OXFORD INSTITUTE FOR ENERGY STUDIES

Justice in Global Warming Negotiations:
How to Obtain a Procedurally Fair Compromise

Benito Müller

Oxford Institute for Energy Studies

EV26

1999

Justice in Global Warming Negotiations

How to Obtain a Procedurally Fair Compromise

by

Benito Müller

EV26
Oxford Institute for Energy Studies
Second, revised Edition
October 1999

ISBN 1 901795 08 X

CONTENTS

EXECUTIVE SUMMARY

In the Kyoto round of the global warming negotiating process, justice was not a major concern. All the relevant parties shared the view that, as the agreement was binding solely for developed countries, the fair solution would be given by some form of 'grandfathering' (i.e. percentage reductions relative to some given emission base line). Future substantive rounds of these negotiations, however, will involve targets not only for developed but also for developing countries, and no such moral consensus is likely to be forthcoming. Considerations of justice will be a key factor in determining the feasibility of any commonly acceptable agreement that has a chance of being ratified. Indeed, given the expected disparity of positions, there will be major obstacles to finding a solution which is likely to be ratified.

The aim of this study is to find a way in which this failure might nonetheless be averted. To this end, a practical method of determining concrete compromise solutions is proposed. The intention is that the procedural fairness and transparency of this method can bring the negotiating parties to see the compromise solution as sufficiently fair to be preferred to a breakdown of negotiations.

In situations where different moral positions can justifiably be upheld, the chances of there being a solution which *all* the (relevant) parties consider to be *completely* fair are negligible. Yet to avert a doomsday scenario we merely need a solution which is commonly regarded as *sufficiently* fair to remain acceptable. The method being proposed to generate such compromises is based on the use of the so-called 'mixed proposals'. These are weighted arithmetic averages over the distribution proposals put forward by the parties as (completely) fair solutions. The problem with the mixed proposals suggested previously is that the weights employed in the aggregation were without exception meant to be determined by way of (strategic) bargaining. In practice, it is argued, this is highly unlikely to generate sufficiently fair proposals. Indeed, if we are dealing with more than two weights, it is unlikely to lead to a proposal at all.

The method advocated in this study, by contrast, determines the weights as reflecting the overall social preference given to the base distributions proposed by the parties: each party ranks all of the proposals according to preference. These rankings are then each expressed in terms of preference scores. The sum total of these scores pertaining to a particular base distribution is a measure (index) of the social desirability of this distribution amongst the parties involved. The weight with which a base distribution is to occur in the mixed proposal is to be given by this index, thus reflecting its social desirability.

This preference score method is presented as a contribution to a debate that may bring about a commonly acceptable compromise solution in future substantive climate change negotiations. As such it deserves to be taken seriously by all involved parties.

INTRODUCTION[1]

The negotiations carried out under the United Nations Framework Convention on Climate Change (FCCC) bear witness to a truly remarkable fact, for it is rare to find the sort of consensus which has emerged about the role of justice (equity[2]) in these global warming negotiations: it appears to have become a self-evident truth that equity issues will be of paramount importance in this context. Thus we are told, to quote just a couple of commentators, that

> It is almost axiomatic that an effective international agreement to limit CO_2 emissions (or other greenhouse gases) will not be undertaken unless it is perceived as fair.[3]
>
> There is now a wide literature on global environmental problems, particularly global warming, which emphasises the need for the emerging regime to be based on the principles of equity or justice.[4]
>
> ...the availability of arrangements that all participants can accept as equitable [...] is necessary for institutional bargaining to succeed...[5]
>
> The key international challenge is ... to find an approach to negotiations which is difficult for any of the major countries or groups to dismiss as unfair.[6]
>
> International co-operation on the scale required will not be achieved without addressing a series of potentially divisive equity issues.[7]

Given this seemingly unconditional support, one could hardly be blamed for being perplexed by the discovery that the role *actually* attributed to justice by Thomas Schelling's dominant focal point theory completely fails to live up to this star-billing. In this focal point account of bargaining, moral properties and considerations are completely coincidental. Clearly justice cannot be critically important and merely epiphenomenal at the same time. But which of the two is its true role in these negotiations? No doubt, as concerns future rounds of global warming negotiations, justice will play a key role in determining the feasibility of a legally binding agreement. But before I go on to describing the argument which led me to this conviction, let me set the scene by giving an introductory account of the relevant issues at stake.

---

[1] I would like to thank and acknowledge the following colleagues for comments and suggestions made on earlier drafts of this paper: Michael Dummett (Oxford), Richard S. Eckaus (MIT, Cambridge Ma.), A. Denny Ellerman (MIT), Meredith Fowlie (OIES), Denis Galligan (Oxford), Henry D. Jacoby (MIT), John Lowe (Natural Resources Canada), Urs Luterbacher (GIIS, Geneva), Hugh Miller III (Goodwin, Procter & Hoar LLP, Boston), John V. Mitchell (RIIA, London).

[2] For stylistic reasons, I shall use 'justice' and 'equity' interchangeably.

[3] Burtraw and Toman (1992:122).

[4] Paterson (1996:181).

[5] Young (1989:368).

[6] Grubb (1990:279).

[7] Nitze (1990:14).

At the heart of these negotiations are, not surprisingly, certain distributive problems, chief of which is the issue of who is going to bear the resulting costs. To achieve the ultimate aim of stabilising the world climate at an acceptable level, there will have to be an agreement which is accepted as legally binding by *all* major greenhouse gas emitters. Assuming, for the sake of argument, that tradable emissions quotas are selected as instrument for achieving economically efficient abatement, the crucial distributive problem becomes that of allocating these quotas in a manner acceptable to all these major players.

There is a general consensus that this allocation should be carried out according to some version of Aristotle's Proportional Allocation Rule. According to this rule, (i) parties are held liable for (or entitled to) different amounts if they differ in some relevant parameter, and (ii) allocation is carried out in proportion to this parameter. The idea being, so Aristotle, that 'What is just then is what is proportional, and what is unjust is what violates the proportion'.[8]

However, seemingly intractable problems emerge as soon as we try to establish what would be the appropriate proportionality parameter. Opinions vary greatly in this respect. Consider, for example, the following listing of candidate parameters which have actually been put forward in submissions to the Intergovernmental Panel on Climate Change (IPCC):

- Per capita emissions (France and Switzerland)
- Per capita economic welfare (Australia)
- Per capita GDP (Poland)
- Emissions intensity of GDP (Norway)
- Relative historical responsibility (Brazil)
- Share of renewables in total energy (Iceland)
- Land area (Russian Federation)

To be sure, there is nothing compelling us to adopt Aristotle's moral interpretation of the proportional allocation rule. But if the rule is chosen as a means to achieve equity, then the chosen proportionality parameter must itself be morally defensible. If the proportional allocation rule is adopted to provide an equitable solution, then morally irrelevant differentiation parameters – such as the average height of inhabitants, or the number of bars in the national anthems – cannot be admitted.

As it turns out all but the last two of the proposals in our list can and have been argued for on grounds of equity.[9] The fact that the Russian proposal has rather been a non-starter suggests that equity considerations have a role in the process of choosing initial allocation proposals, at least if they are to have half a chance of being taken seriously. Moreover, given the remaining number of equity-based parameters in our list, it stands to reason that there is no consensus on what would be a *fair* allocation of quotas.

To give an idea of what it means to justify a differentiation parameter on grounds of equity, consider two particular proposals which, in moral terms, are generally perceived as diametrically opposed to one another. Assume for the sake of argument that there is a maximum quantity T_{max} of anthropogenic greenhouse gases which can be emitted yearly without altering the climate unacceptably. An allocation proposal will be presented as a formula for the yearly distribution of emissions quotas equivalent to this maximum quantity. Perhaps the simplest way of specifying such a formula is to apply the Aristotelian rule with reference to some base line, say the situation in 1990. The two particular proposals I have in mind are

The Simple Grandfathering Distribution: All parties receive yearly permits in proportion to their baseline *emissions*.

The Simple Per-Capita Distribution: All parties receive yearly permits in proportion to their baseline *populations*.

How might these two proposals be justified? It is not difficult to see what would be involved in giving a moral justification for the per capita proposal. All we need to do is to treat our quota distribution problem as something akin to the process of establishing individual property rights for a common good, namely the atmosphere as repository of anthropogenic emissions. Assuming that individual people – as opposed

[8] Aristotle *Nicomachean Ethics*: Bk V: Ch.3.

[9] Given the geographical size of the Russian Federation and the fact that Icelandic electricity generation is overwhelmingly hydroelectric and geothermal, Reiner and Jacoby do seem to be right in thinking that these last two proposals may have been put forward because they are 'near-uniquely favourable to the proposing country'(p.4). But this is not the reason why the parameters involved are morally irrelevant. This irrelevance is rather given by the fact that these geographical and geological attributes of parties are outside their sphere of responsibility, i.e. I take them to be just a matter of good or bad luck. The fact that a proposal favours the proposer *per se* does not necessarily entail that the

to, say, nation states – are taken to be the rightful claimants, the per capita proposal will be justified by arguing on egalitarian grounds that everyone has an equal claim on this common good.

The moral basis of the grandfathering proposal, by comparison, may not seem equally plausible. But this position is not as *prima facie* unethical as its opponents might have us believe, in particular when they argue that it is tantamount to sanctioning the evil of pollution. We must be mindful of the fact that we are *not* allocating *pollution* permits. Anthropogenic emission only becomes a matter of pollution for those quantities over and above the accepted maximum level T_{max} covered by our emissions quotas.[10]

Given there is nothing morally objectionable to emitting *within* these limits, grandfathering can be argued for as follows: (i) The fact is that historically, anthropogenic greenhouse gas emissions have been the by-product of wealth creation processes. (ii) Everyone is entitled to a share of created wealth in proportion to their contribution to the wealth-generating process. *A fortiori*, everyone is entitled to the appropriate proportion in the acceptable use of common amenities in this process. And finally, (iii) the transfer of these entitlements is morally legitimate, i.e. it is neither morally wrong to pass them on, nor is it wrong to accept them.[11]

The theoretical basis of this argument is to be found in what has become known as entitlement theories of justice. This family of theories ranges from the Marxian theory of entitlement derived from labour contributions, to libertarian theories framed in terms of property rights. As things stand, most actual proponents of grandfathering are likely to embrace a theory of just acquisition and transfer of property rights – such

proposal is morally unjustifiable. All it means is that one has to be doubly careful in ascertaining whether it is.

[10] To avoid misunderstandings, let me emphasise that this maximum acceptable level T_{max} is not necessarily the same as the (sum total of) baseline emissions: our simple grandfathering is not necessarily the same as the *status quo* distribution. Baseline emissions are merely used to establish the allocative proportions, the acceptable maximum emissions level may be much lower than global *status quo* (baseline) emissions.

[11] Note that this argument only justifies the simple grandfathering proposal if the proportions in baseline emissions truly correspond to the proportions of wealth creation contributions. I should also mention that there may well be other arguments in support of this position. Indeed, as John Lowe pointed out to me in conversation, a very simple argument which has been used relates to the unfairness of changing the rules after people have made investments of effort and savings.

as the one proposed by Robert Nozick. After all, it would be rather ironic if the developed countries which favour grandfathering were to justify their position in terms of Marxian labour rights, or, for that matter, in terms of their rights as 'eco-squatters'.[12]

This may all be very interesting from a theoretical point of view, but what is the practical 'bottom line'? We have a situation in which different parties can have differing moral positions which are perceived as legitimate by the parties in question. The practical 'bottom line', of course, will only emerge if we consider the economic implications of these positions.

The Second Assessment Report of the IPCC,[13] published in 1996, contains the results of a study appraising the economic effects of our two allocation proposals.[14] The study uses a scenario in which global emissions are permanently held at 1990 levels. Given this scenario, OECD countries will receive an initial allocation of approximately 50 per cent of the emission quotas when the status quo is maintained under a grandfathering scheme, whereas under the equal per capita emissions scheme, their share drops to about 20 per cent.

If *grandfathering* of quotas is adopted (see Fig. 1a) the region of the OECD and that of the so-called economies in transition would be sellers of emission rights, and the resulting wealth transfers would, in the case of the OECD, be sufficient to reduce GDP losses from abatement to a negligible level. Indeed, under grandfathering, Eastern Europe and the FSU would be net beneficiaries. The economic burden on developing countries, however, would be substantial.

[12] In his excellent article on 'Ethical Issues in Income Distribution', Amartya Sen summarises Nozick's position as follows: 'He defines principles of justice in *acquisition* and *transfer*; a person acquiring holdings in accordance with these principles is entitled to them, and no one is entitled to a holding except by repeated applications of these two principles. The principles are so constructed that a person is not only entitled to what he himself produces with his own labour, but also to what is produced by resources owned by him and what he can acquire by free exchange of what he legitimately holds.'(Sen 1984:284f.) This, it seems, does reflect my argument. My reference in (ii) to a 'contribution to a wealth-generating process', however, could quite consistently be interpreted in terms of labour contributions. As concerns 'eco-squatting', it has been suggested in the literature (e.g. Grubb *et al.* 1992:313) that grandfathering is to be justified by reference to the common law doctrine of 'adverse possession', meaning the 'occupation of real property in manner inconsistent with the right of the true owner.'(Walker (1980):34) Note the implicit assumption of pre-existing, presumably common property rights in this suggestion.

[13] Bruce *et al.* (1996).

An equal *per capita* quota allocation scheme would significantly shift the distribution of net benefits and losses (see Fig. 1b): The OECD countries would incur significant net income losses, and the 'rest of world' would be a winner. 'Under either scheme', the IPCC authors tell us, 'China incurs substantial net income losses, except for the very early years. Unless it can greatly increase its energy efficiency, China's rapidly rising demand for energy would far outstrip its allocation of emission rights, even under an equal per capita quota allocation rule.'[15]

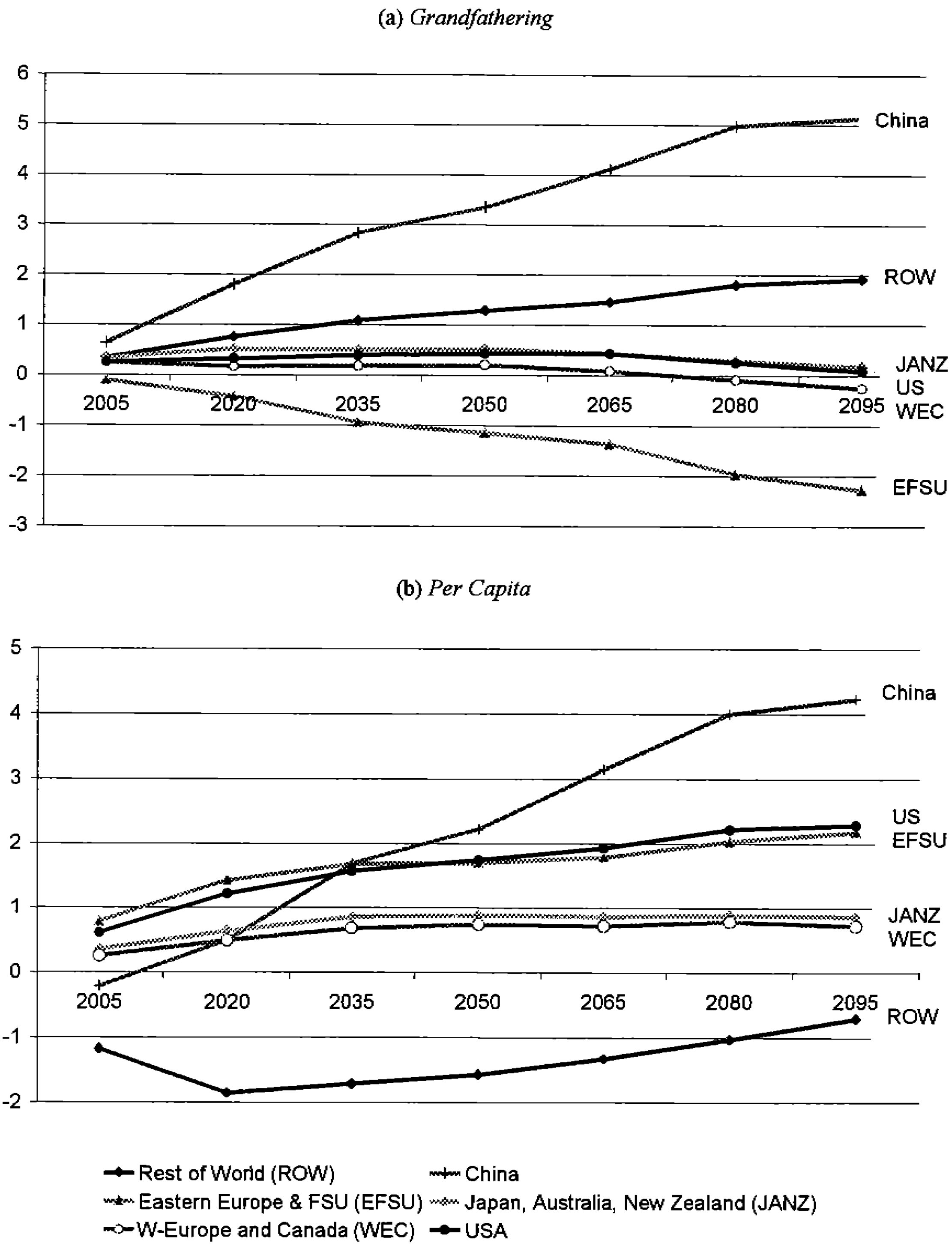
These figures, like all economic predictions, depend on the underlying modelling assumptions and must be taken with a pinch of salt. However, something may safely be concluded, namely (i) the stakes for individual parties can be enormous and (ii) they can vary significantly under the different equity-based allocation proposals.

A sceptic may well grant me this. Indeed he might even concede that equity has a role to play in the selection of *initial* allocation proposals. But surely, he is bound to interject, the *outcome* of the negotiations will be determined by good old-fashioned strategic bargaining, reflecting only the bargaining powers of the parties and the bargaining skills of the negotiators. Unfortunately, I can only say, our sceptic may well be right. The reason I consider this to be unfortunate is my conviction that such a *strategic* agreement is unlikely to be successful. To be successful, an agreement has to be implemented. This, in turn, requires political ratification which normally is beyond the power of mere negotiating agents. Negotiators may be under constraints which make them susceptible to being pressured into accepting a proposal. After all it is their job to hammer out agreements, and not to return home empty-handed. And they can always claim to have obtained the best possible deal given the strategic strengths of their opponents.

---

[14] Edmonds *et al.* (1993).
[15] Bruce *et al.* (1996:341).

Figure 1: The Economic Impacts measured by the Total Cost of Emission Reductions + Net Wealth Transfers from Sales of Permits (as % of GDP)



Source: Bruce *et al.* (1996:340f.)

It would be foolish to assume, however, that bodies such as the US Congress or the Indian Lok Sabha could be similarly bullied into *ratifying* an agreement. Given this, I believe that the perceived equity of an agreement will be a key component in the ratification process. To be sure, I am not claiming that ratification will be endangered by altruistic motives. What I do believe is that parties may refuse to ratify an agreement if they feel it deviates unacceptably from what *they* perceive to be the just solution. In other words, my own scepticism about the ratification prospects of a strategically negotiated agreement is not based on the assumption that ratifying bodies are motivated by the wish that *others* should be treated fairly. All I am saying is that they will be sensitive to what they themselves perceive as unfair treatment by others. Given that we are *prima facie* dealing with a situation where very different positions can justifiably be regarded as fair this does not bode well for the prospects of a *strategically* negotiated global warming agreement.

Things, however, may not be quite as gloomy as this might suggest. For there may be ways to overcome this doomsday scenario by focusing on measures which increase the perceived procedural fairness of the negotiating process. In the second part of my analysis, I shall suggest a procedure that may be capable of generating a successful outcome even when the initial positions are as disparate as they are likely to be in future rounds of the FCCC negotiating process.

SUMMARY

I. Received Views.

I begin my analysis by looking at the way in which justice has been incorporated in existing theories purporting to explain, in particular, negotiations on sharing out (tradable) greenhouse gas emissions permits. Specifically, I consider what appears to be the dominant theoretical interpretation of the role of justice in this context, namely that of providing so-called 'focal points'. I then turn to a dissenting view, where justice is meant to emerge from strategic bargaining and finally discuss some of the equity implications we are likely to be confronted with in the context of negotiating a *global* greenhouse gas emissions accord.

Focal Points: The *Schelling-Barrett Example.* In order to discuss the idea of justice providing 'focal points' – which is by far the most frequently adopted view amongst those authors who actually try to give a theoretical underpinning to their accounts – I make use of an example from the literature in which this idea is explicitly applied to the problem of sharing out emissions quotas between countries. After a brief exposition of how such focal points are *meant* to function, it becomes apparent that justice is merely treated as an epiphenomenon. The fact that the principles are principles of *justice* turns out to be quite irrelevant in the focal point approach. All that is required of them is that they are common knowledge. The role accorded to (principles of) justice in focal point theory is, in essence, that of providing some sort of perspicuous simplicity. This, I argue, will simply not do. More precisely my point is that certain shifts of focal points – induced by a change of background information – cannot be explained without taking into account the *moral* properties of the principles involved. For the sake of its own reputation as an explanatory tool, focal point theory must cease to confuse simplicity with justice and accord the latter a considerably more important role than it actually does.

The Strategic Theory of Justice. After this initial criticism of focal point theory I turn to what may well be the only existing competing view of the role of justice in our negotiating problem. In an article published in 1992, Burtraw and Toman put forward the idea that in negotiations with substantial stakes, the negotiators will, in

the first instance, proceed by way of strategic bargaining. The claim then is that, after the fact, they may agree that the outcome is actually fair, *given* the relative strategic strengths of the parties involved. Again we find justice to be relegated to the realm of epiphenomena as far as negotiations themselves are concerned. However, it also seems clear that this 'might is fair' conception is unlikely to apply to anyone, with the possible exception of parties with strategic advantages.

Doomsday Scenarios. Focal point theory is based on the assumption that amongst all the logically possible solutions there is a range of options which are *commonly acceptable* in the sense that each party would prefer them over not reaching agreement at all. This may well have been the situation at Kyoto, where the parties which undertook commitments were more or less at the same level of economic development.[16] In light of the disparity between, say, the grandfathering and per capita solutions and the history of unbalanced (economic) relations between developed and developing countries, it would not be outlandish to predict that future negotiations intended to achieve a *globally* binding agreement will face what might be called a 'doomsday scenario', i.e. a situation where none of the possible solutions is commonly acceptable in this sense.

In doomsday scenarios, focal point theory cannot even get off the ground. This may not be particularly harmful to the reputation of this theory as tool for explaining the choice of particular agreements, since these scenarios inevitably lead to a breakdown of negotiations. And if there is no agreement, there is nothing in need of this type of explanation. While there is no agreement, there certainly is an outcome (a breakdown) which remains to be explained. The theory put forward for this purpose in Part II explicitly acknowledges justice as a key parameter in this context. It has the particular advantage of showing a way in which doomsday scenarios might be averted, provided equity considerations are taken seriously.

Having criticised some of the received views on the role of justice in global warming negotiations, my own position is laid out in three stages. In Part II, I begin by developing a theoretical framework which explicitly includes parameters reflecting

[16] 'The Kyoto commitments for Annex I are essentially a grandfather system, modified slightly by differentiation and bubbling.'(Mitchell 1998:8)

perceived (in-) equities. In using the resulting theoretical model, I then go on to discuss the way in which procedural fairness may lead to the *creation* of commonly acceptable solutions. Finally, in Part III, I analyse what I consider to be the most likely procedure to overcome doomsday scenarios in this manner: the selection of socially weighted arithmetical averages ('mixed proposals') as compromise solutions.

II. Alternative Conception

The Inequity-Disutility Model. The basic idea behind the theoretical model I am proposing is that distribution proposals which can be argued for on moral grounds are to be treated as positions of minimum unfairness, relative to the equity principles they are based on. It will be implemented in the envisaged numerical model by way of adopting personal welfare functions which are the sum of two parameters: the utility associated with the economic benefits of an option, on the one hand, and the disutility which the party in question would associate with this option by virtue of a feeling of being unfairly treated, on the other. Conforming with the initially mentioned basic idea, these inequity-disutilities will be minimal for the equity-based proposal which the party in question considers to be the just position, and they may increase in strength for options which deviate from this position, thus diminishing their overall welfare to this party.

One of the attractions of this inequity-disutility model is its ability to represent a classically inexplicable[17] situation in which a breakdown in negotiations or non-ratification (entailing *no* economic benefits) is nonetheless preferred over an option which would clearly bring economic benefits, for the simple reason that the option in question is (perceived to be) too unfair to be acceptable. Accordingly it is easy to represent doomsday scenarios generated by (perceived) inequities. Indeed, the model proposed enables us to give a very simple game-theoretic argument as to why negotiations which end up in such a scenario will inevitably fail to achieve a commonly acceptable agreement.

Making Compromise Possible. Inequity disutilities depend on a host of parameters, not just on the size of shares which the parties are allocated under the distribution in

[17] That is an 'irrational' decision, to those who feel that the traditional theory is infallible.

question. They depend, in particular, on the value of what is at stake and on the manner in which the distribution in question is selected for consideration. Thus, as far as acceptability is concerned, an allocation proposal of, say, one megaton of carbon emissions arrived at under one procedure is not necessarily the same as a megaton proposed under another procedure.[18] The idea is that a judicious choice of these additional parameters may mitigate potential inequity sensitivities sufficiently so as to avert a doomsday scenario.

Given that the size of what is at stake in our allocation problem – namely the global emissions quotas – is a more or less fixed quantity, it will not be feasible to mitigate inequity sensitivities by simply reducing the value of what is at stake. Accordingly I propose that we focus on the parameters associated with the way in which solution proposals are selected.

One factor which *will* be relevant is which options are admitted for consideration in such a selection. As concerns mitigating inequity-disutilities, the best policy, I believe, is to admit only options which are (i) proposed by a party involved in the negotiations and (ii) justifiable in terms of some equity principle. The specification of what I refer to as an *equitable selection base* will include some further conditions in order to ensure that in a *morally ambiguous* situation – where there is no single morally dominant principle (subscribed to by the negotiating parties)[19] – admissible options have the same moral legitimacy.

In addition to this 'substantive' component, inequity sensitivities also depend on certain procedural parameters, chief of which is the (perceived) fairness of the method in which a distribution is selected for consideration, given an equitable selection base. The idea in this case is that the fairer such a procedure, the more likely the mitigation of potential inequity sensitivities, with the aim then being to find a selection procedure which is *sufficiently fair* to create and select a commonly ac-

[18] An allocation figure which is less than what the recipient believes to be his fair share may still be acceptable to him if it was determined by what he considers an acceptable method. But he may equally reject the very same figure, say determined by a random number generator, if he considers this manner of determination to be unacceptable.

[19] Although I believe that there are moral situations which are, if we wish, metaphysically ambiguous, I ought to emphasise that the method I am going to introduce is not tied to this sort of ambiguity. It

ceptable proposal. In pursuing this aim, I first consider the one selection method which seems to be the instinctive favourite, namely to have one's own proposal selected by convincing all parties of its moral superiority. This procedure may be successful in a morally *un*ambiguous situation. In the context of allocating global emissions quotas – which I think is almost certainly morally ambiguous – all that will be achieved is a counter-productive slanging match of mutual recriminations. Indeed, one of the main practical conclusions for global warming negotiations in this study is that – in the context of an equitable selection base – parties should stop accusing each other of taking up morally indefensible positions and acknowledge that in a morally ambiguous situation contradictory positions can be morally legitimate. A cessation of such hostilities actually paves the way for the application of a selection procedure which, I believe, should have a good chance to overcome a global warming doomsday scenario.

Parties	*Individual Preference Scores*		
	D_1	D_2	D_3
A	2	1	0
B	1	2	0
C	0	1	2
***Borda Index*:**	3	4	2

Figure 2: The Borda Preference Score Rule

Having rejected this argumentative selection method, I then briefly turn to another type of selection method involving choices from the selection base, namely the procedures in which one of the 'candidates' in the selection base is *elected* by the parties involved. I shall focus, in particular, on elections by way of the preference score method given by the so-called 'Borda-rule'. In its simplest form, this rule prescribes that each of the voters, say *A*, *B*, and *C* is to rank the candidates in the relevant list ('selection base'), say D_1, D_2 and D_2, according to preference. Each voter will then give a preference score of 0 to their least preferred candidate, a score of 1 to the least preferred but one, and of 2 to their preferred candidate (for an example, see Fig. 2). The Borda election then proceeds by adding up the scores given to each of the candidates, and selecting – with some tie-breaking provision – the one with the maximum 'Borda Index', i.e. the maximum total preference score.

---

will be equally useful if the ambiguity is merely epistemic. In other words, the only case where it will not be useful is if we are dealing with an unambiguous situation which is known to be unambiguous.

While outstanding in its procedural fairness, this method, like all election procedures, is potentially divisive by creating outright winners and losers. In the absence of a way to enforce such an election result (i.e. to *make* parties sign up to/ratify it), this divisiveness may lead to further obstacles to a successful outcome. This is why I will reject the Borda election in favour of procedures of a more inclusive nature.

III Numerical Selection Procedures

In some cases – such as elections of people – the excluding character of choice procedures may be unavoidable. Our allocation problem, however, does admit of more inclusive ways to select proposals for consideration. Our 'candidates', after all, are arrays of (percentage) numbers. As such they can be mathematically aggregated into *compromise* proposals. We are thus in the position to carry out a 'creative election' where the candidate elected for consideration is created from attributes of the candidates on the original list. In the likely case of such an aggregate falling outside the range of options originally put up for consideration, there will be no outright winners. Nor will there be outright losers, since all the original admissible options of the selection base will be reflected in the aggregate. Typical examples of such numerical procedures are those involving 'mixed proposals', such as the one submitted by Norway to the *Ad Hoc Group on the Berlin Mandate* (AGBM) of the FCCC in January 1997, according to which the percentage reduction Y_i of emissions for party i is to be calculated according to the formula:

$$Y_i = A\,[x(B_i/B) + y(C_i/C) + z(D_i/D)]$$

> The relation of B_i to B is CO_2 equivalent emissions per unit of *GDP* for Party i relative to the average in the Annex I Parties. The relation C_i to C is the *GDP* per capita in Party i relative to the average in the Annex I Parties, while the relation of D_i to D is the CO_2 equivalent emissions per capita in Party i relative to the average of Annex I Parties. A is a scale factor to ensure that the desired overall reduction in emissions is achieved. The coefficients x, y and z are weights, which add up to a total of 1.[20]

Another example is the mixture of grandfathering and per capita suggested by Grubb *et al.*:

> if T is the target level and p is the proportion of population in the allocation formula, a country with (modified) population and emissions – which are respectively X% and Y% of the corre-

[20] AGBM (1997:43). Note, incidentally, that the Formula is explicitly meant to determine percentage reductions. In other words, it is used as a means to differentiate between Parties who, in principle, agree on the Grandfathering position. The Formula itself was originally suggested by Ringius *et al.* (1998), which is to be highly recommended, as is Ringius' (1999) analysis of the EU negotiations.

sponding total qualifying population and emissions of the participating countries – would receive an allocation of $[pX + (1-p)Y] \times T$ permits.[21]

This proposal is particularly simple as it takes into account only two base distributions. In general, these numerically mixed compromise procedures involve an application of Aristotle's proportional allocation rule to aggregate claims, themselves generated by taking weighted arithmetic averages over the shares stipulated in the distributions of the selection base. Mixed proposals thus pose the following fundamental questions: (1) Why choose the *Aristotelian* allocation rule as opposed to, say, the 'Contested Garment Rule'[22] advocated in Talmudic tradition? (2) Why use weighted *arithmetic* averages and not, say, geometric or harmonic ones in creating the aggregate claims? And, last but not least: (3) What should be the method of determining these weights?

As far as I am aware, only the last of these questions has been touched in the literature. This is unfortunate. When taken in isolation, the weights in question are nothing but numbers, which means we are left with only one possible answer: they must be determined by way of strategic bargaining, a method which is unlikely to mitigate inequity sensitivities. Grubb and Sebenius' view is indeed that their 'proportion of population' parameter p is to be negotiated, and we can safely concur with Reiner and Jacoby's view that the Norwegian proposal 'offers little guidance as to the weights to be assigned to the three factors, presumably in recognition that ultimately they would be determined in negotiations among the parties'.[23] In light of the inadequacies of strategic methods in overcoming equity-based doomsday scenarios, I fear that this method is rather unlikely to produce a successful outcome. However, by addressing all three questions, we find a type of numerical selection which is much more likely to be sufficiently fair for our purposes.

The crucial issue turns out to be not so much the justification of the Aristotelian rule – given by its being the only 'collusion proof' allocation rule – but the question how

[21] Grubb and Sebenius (1992:209), and Grubb *et al.* (1992:321).
[22] See Young, H.P. (1994:67 ff.).
[23] Reiner and Jacoby (1997):8. Indeed, in a recent publication, one of the 'fathers' of the Norwegian Formula, Asbjørn Torvanger, concedes that 'the weight of each criterion [in the Norwegian Formula] was in the original proposal subject to negotiation, but in later proposals they were equalised to one third'[Torvanger and Godal (1999):40]

to justify the use of (weighted) *arithmetic* averages. The justification which I propose is rather lengthy and, in parts, somewhat technical. But since I have appended my technical justification with its own informal summary, I shall omit details here and simply sketch the core idea: If we are dealing with an equitable selection base then we can justify the use of the weighted arithmetic aggregates by way of identifying the resulting *mixed shares* as *value entitlements*. In other words, we can interpret the weights as distribution-relative *unit-values* for the contended good. In doing so we find that the share allocated to a party under the arithmetically mixed proposal amounts to precisely the *total value* it is entitled to.

This sort of justification obviously only works if we can find a satisfactory procedure for determining these unit-values. My own proposal is based on the idea of interpreting them in terms of the *social* value of the different base distributions, as expressed in the indices generated by the Borda rule: a unit distributed under a distribution which is socially (i.e. amongst the parties) preferred over another one will accordingly have a higher (social) value than one distributed according to the less preferred distribution. The use of weighted arithmetic averages is thus to be justified in terms of allocating to the parties the *total social value* they are due in accordance with the Borda valuation of the base distributions.

In view of the acknowledged procedural fairness of both the Aristotelian and the Borda rule and given the inclusive nature of mixed proposals, my contention is that the suggested preference score compromise procedure – where the base proposals are aggregated with weights identified as the relevant normalised Borda indices – should have a good chance of being regarded as sufficiently fair to generate a commonly acceptable compromise proposal.

In the simplest case of just two base distributions, the Borda weights, as it turns out, are given by the proportion of the parties which prefer the distribution in question. In other words, the mixed proposal selected by our preference score procedure – namely $[p^{*}X+(1-p^{*})Y]\times T$, to use the Grubb-Sebenius example – corresponds to the creative election of a mixed 'candidate' in accordance with the rules of proportional representation, since the preference score parameter p^{*} turns out to be nothing but the fraction of the parties which would have voted for the per capita distribution.

This preference score procedure is, of course, quite general as concerns the number of the involved base proposals. Indeed, one of its principal merits is its ability to suggest appropriate weights for formulae such as the Norwegian proposal with more than two weights. Take, for example, the case of just three parties *A*, *B*, *C*, and assume that

$$D_1 = \langle d_A^1, d_B^1, d_C^1 \rangle,\ D_2 = \langle d_A^2, d_B^2, d_C^2 \rangle, \text{ and } D_3 = \langle d_A^3, d_B^3, d_C^3 \rangle$$

are the base distributions according to emissions per GDP, GDP per capita, and emissions per capita, respectively. The Norwegian proposal will then be determined by

$$D^p = \langle d_A^p, d_B^p, d_C^p \rangle = p_1 D_1 + p_2 D_2 + p_3 D_3, \text{ where}$$

$$p_k \geq 0 \text{ and } p_1 + p_2 + p_3 = 1 \text{ (see Fig 3).}^{24}$$

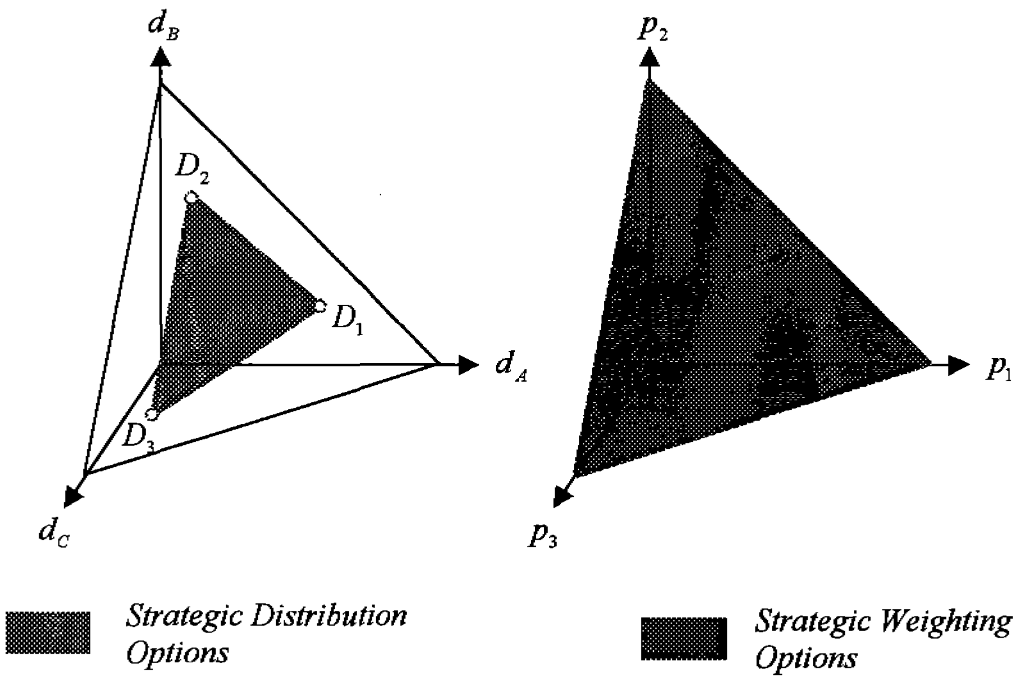


Figure 3: The Norwegian Proposal – Strategic Options

[24] To be clear, in the case of the Norwegian Formula (*NF*), the 'mixtures' D^p are *not* themselves meant to distribution proposals. They are to be interpreted as percentage differentiations from the Grandfathering distribution. In other words, if g_k is the amount allocated to *k* under Grandfathering, then *k* would receive $\alpha g_k d_k^p$ under the *NF* (with α a scale factor to ensure the desired global emission level is reached). Of course, this is just a peculiarity of the *NF* which does not diminish the generality of the Preference Score method.

In the absence of a procedure such as the one proposed above, the (strategic) negotiating task would thus be to determine two parameters, each between 0 and 1,[25] such as to obtain a commonly acceptable mixed solution. Such a task may just be humanly feasible in the context of a single parameter leading to a simple mixed solution. For two or more parameters, however, it does seem to take on truly Herculean proportions.

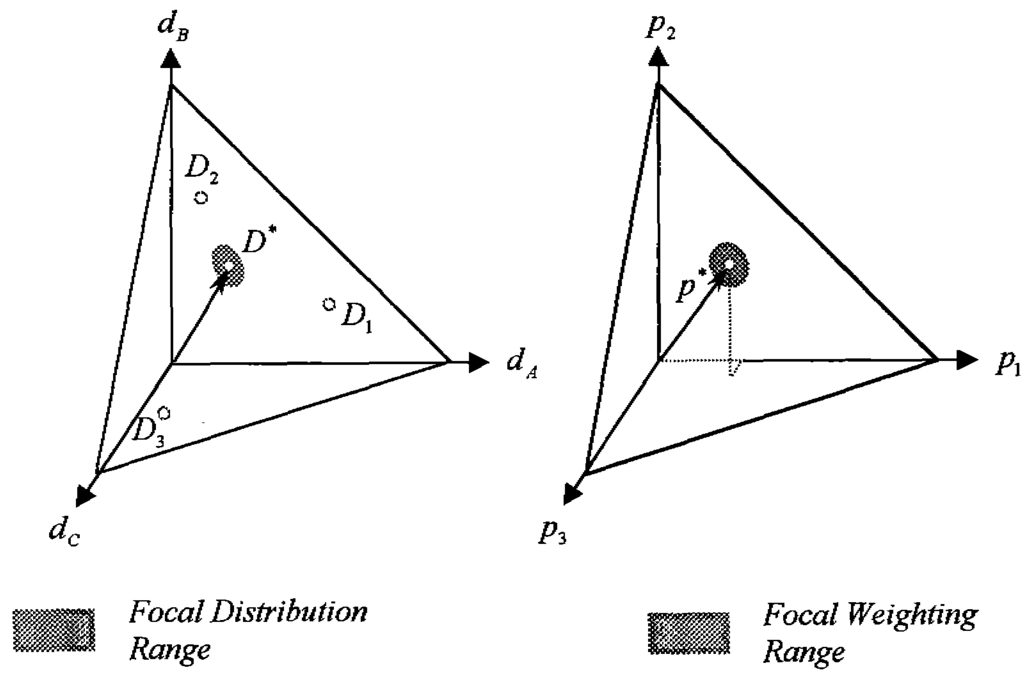


Figure 4: The Norwegian Proposal – Preference Score Options

In applying the preference score method, by contrast, we determine a particular *preference score mixture*:

$$D^* = \langle d_A^*, d_B^*, d_C^* \rangle = p_1^* D_1 + p_2^* D_2 + p_3^* D_3, \text{ i.e. } d_k^* = p_1^* d_k^1 + p_2^* d_k^2 + p_3^* d_k^3$$

by establishing the relevant (normalised) *Borda weighting*: $p^* = \langle p_1^*, p_2^*, p_3^* \rangle$,

which, in the case illustrated in Fig. 2, is:

$$p_1^* = 3/(3+4+2) = 1/3; \ \ p_2^* = 4/9; \ \ p_3^* = 2/9.$$

To be sure, I am not claiming that this preference score proposal will invariably be successful (commonly acceptable). However, *if* the base proposals submitted for consideration constitute an equitable selection base, and *if* the parties have

[25] The third weight being determined by the fact that $p_1 + p_2 + p_3 = 1$.

come to accept the equal legitimacy of the base proposals, then they will hopefully be enticed by the fairness of the suggested preference score procedure to accept that the mixed proposal with the Borda weighting does indeed demarcate a region in which *acceptably fair* solutions can be found, namely a neighbourhood of the preference score mixture D^* (see Fig. 4). Once this stage is reached, the parties can then happily engage in focal point bargaining and conclude their negotiations with a successful outcome somewhere in this neighbourhood.

In adopting the suggested preference score method we remove, so to speak, the conflict from one playing field with irreconcilable differences to another - namely the specification of admissible base distributions - where there may be room for consensus. But is this really all there is to it? Are we not actually complicating the issue by creating a whole new set of decision problems to do with the scoring procedures used in determining the social welfare weightings. Indeed, two pertinent issues do spring to mind. First, why should we go along with the 'single party – single scores' provision adopted in our procedure? Why should the preferences of, say, Switzerland be counted on equal terms as those of India or the USA? Second, who is to be treated as a party? Should, say, the fifteen members of the EU be counted as a single party, or as fifteen individual ones? Or why should the USA not be allowed to be counted as fifty parties?

As it happens, both of these 'representative equity' problems can be resolved by switching from *single national* ('single party – single scores') to *global* preference scores, where each party is permitted to multiply its scores by the number of people it represent at the negotiations.[26] Indeed, this global preference score procedure has the advantage of being collusion proof in the sense that a party, taken on its own, will receive precisely the sum total of what its parts could realise if instead *they* were treated as the eligible individuals.[27]

[26] It is important not to confuse these demographic weights for preference scores with population parameters used to differentiate allocations of emission quotas. The former are used to achieve representative equity at the negotiations and are thus tied to the situation at the time the negotiations take place. The latter, by contrast, are not tied to the status quo in this manner: they can refer to some prior base-line situation or they can be intertemporal aggregates.

[27] The assumption being that the 'parts' would share their preferences concerning the base distributions with the 'whole' (see Appendix 1).

The proposed preference score method of generating socially weighted mixed allocations is obviously applicable to many allocation problems other than those arising in global warming negotiations. And yet, given the strong likelihood of a doomsday scenario if this particular negotiating process proceeds by using the prevailing strategic methods, this preference score procedure ought to be of particular interest to anyone wishing these negotiations to succeed in arriving at a commonly acceptable agreement.

I

RECEIVED VIEWS:

FOCAL POINTS AND STRATEGIC ACCEPTANCE

Focal Points: The Schelling-Barrett (*SB*-) Example

To illustrate the supposed focal point role of equity principles, consider the following (slightly adapted) example due to Scott Barrett,[1] itself based on examples originally put forward by Thomas Schelling in his seminal work on *The Strategy of Conflict*.[2] Suppose that *n* parties are engaged in negotiations concerning the distribution of a quantity *T* of a (homogeneous, divisible) good – such as greenhouse gas emissions permits/quotas – between themselves.

Scenario A. Assume to begin with that the parties have no information about one-another bar the fact that they are *n* in number (call this information 'I^A'). It is then likely, says Barrett, that the parties would agree on an '*egalitarian*' distribution given by the distribution rule, say R^A, stipulating that $d_1^A : d_2^A : \cdots : d_n^A = n_1 : n_2 : \cdots : n_n$ – where n_k = the number of parties represented by *k* – i.e. the *equal distribution*:[3]

$$D^A = \langle \tau^A n_1, \tau^A n_2, \ldots, \tau^A n_n \rangle = \langle \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n} \rangle.^4$$

Scenario B. Suppose now that the parties know nothing of one another except (I^B) that there are *n* parties, and party *k* has (represents) a population of size P_k. In this case, Barrett tells us, it is likely that amongst the many possible outcomes, the parties would choose an allocation 'on the basis of population', that is to say they would mutually consent to a rule, say R^B, stipulating that that $d_1^B : d_2^B : \cdots : d_n^B$ = $P_1 : P_2 : \cdots : P_n$, and thus adopt the *per capita distribution*:

---

[1] Barrett (1992:87ff).

[2] Schelling (1960:61ff).

[3] A 'distribution', in the technical sense used here, is simply a sequence $D = \langle d_1, d_2, \ldots, d_n \rangle$ of non-negative real numbers adding up to 1: $d_k \geq 0$, $1 = \sum d_k$ representing the shares allocated to the *n* parties involved: $d_k T$ is the size of the part of *T* allocated to party *k* under distribution *D*. Specific distributions will be denoted by means of superscript indices: $D^A, D^B, D^C, \ldots$ or $D^1, D^2, D^3, \ldots$, and, whenever required, these indices will also be attached to their components.

[4] $n_k = 1$, $\tau^A = 1/\sum_k n_k = n$.

$$D^B = \langle \tau^B P_1, \tau^B P_2, \ldots, \tau^B P_N \rangle = \langle \frac{P_1}{P}, \frac{P_2}{P}, \ldots, \frac{P_n}{P} \rangle.^5$$

Scenario C. Finally, suppose that each negotiator not only knows the number and the total population but also the GNP (G_k) of the parties (call this information I^C). 'The problem', according to Barrett, 'is now more complicated, and the decision harder.' He thinks the negotiators might continue to split the emissions permits evenly (i.e. adopt D^A)[6] or allocate them on the basis of population (D^B). 'Alternatively, they might now allocate [the permits] on the basis of GNP', by adopting the 'inverse GNP'-rule ($R^{C'}$), which stipulates that $d_1^{C'} : d_2^{C'} : \cdots : d_n^{C'} = 1/G_1 : 1/G_2 : \cdots : 1/G_n$, and consequently agree on the *inverse GNP distribution*

$$D^{C'} = \langle \tau^{C'} \frac{1}{G_1}, \tau^{C'} \frac{1}{G_2}, \ldots, \tau^{C'} \frac{1}{G_n} \rangle.^7$$

'Perhaps the most likely outcome', in Barrett's view, 'would involve the allocation of [permits] on the basis of GNP per capita' according to a rule ($R^{C''}$) stipulating that $d_1^{C''} : d_2^{C''} : \cdots : d_n^{C''} = P_1/G_1 : P_2/G_2 : \cdots : P_n/G_n$. In other words, the distribution most likely to be chosen according to him is the *inverse GNP-per-capita distribution*:

$$D^{C''} = \langle \tau^{C''} \frac{P_1}{G_1}, \tau^{C''} \frac{P_2}{G_2}, \ldots, \tau^{C''} \frac{P_n}{G_n} \rangle.^8$$

What is this example *meant* to illustrate? To explain this, we need to consider Barrett's not as yet mentioned fundamental assumption that the negotiations in question are 'tacit', in Schelling's sense, i.e. (in game-theoretic jargon) that the parties are playing a 'non-cooperative coordination game'. In such games, the parties are forbidden to communicate with one another. At the same time they are generally assumed to share certain information as common knowledge.[9] The aim of the game then is for all of the parties to make the same choice from a range of alternatives under these conditions. To give a very simple example, the parties could be told that

5 The normalising factor, in this case, is $\tau^B = 1/P$, with $P = \sum_k P_k$

6 Note, incidentally, that the same could be said in the context of *Scenario A*.

7 $\tau^{C'} = 1/[\sum_{k=1}^{n} (1/G_k)]$.

8 $\tau^{C''} = 1/[\sum_{k=1}^{n} (P_k/G_k)]$.

9 'Information is common knowledge in a game if it is known to all of the players, and if every player knows that all the players know, and that every player knows the others know that he knows, and so forth.'(Roth 1985:262)

each of them would benefit in some desired way provided only they all manage to choose (say in some limited time frame) the same object amongst the alternatives given in Fig. 5a.

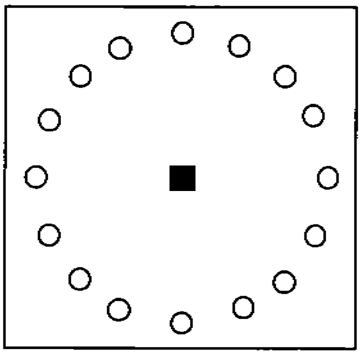


Figure 5a

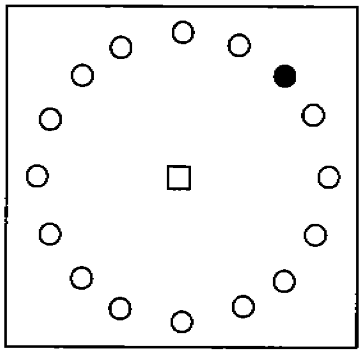


Figure 5b

The framework adopted presupposes that for each of the parties, making *some* choice is preferable over 'walking away,' i.e. over a breakdown of negotiations. Everyone is thus presupposed to be content if a common choice is achieved, regardless of which particular alternative is chosen.

In *The Strategy of Conflict*, Schelling pointed out that there are many such non-cooperative coordination games in which winning choices occur much more frequently than could be explained probabilistically, and he argued convincingly that this phenomenon is to be explained by the fact that the parties manage to coordinate their choices by virtue of common knowledge and the presence of some 'obvious' choice, which he terms a 'focal point'. Indeed, in the case of choosing from Fig. 5a, there is one particular option which is distinguished from all the others in a way which quite literally makes it a visual focal point, namely the black square located equidistantly from all the white circles. Given the common knowledge that all parties are trying to coordinate their choices, the choice of this focal point by any one (rational) party becomes not only likely but almost inevitable. To be noted is the fact that different characteristics (colour, location, shape) can single out one and the same option as the 'obvious' choice under the circumstances. They may, of course, reinforce the focal point 'signal' of this solution, but none of them is essential to a successful co-ordination. All that matters to raise the success-rate over the probabilistic level is that some (preferably just one) of the options have some commonly known characteristics which distinguish them from all the others.

As illustrated in Fig. 5b, there may well be different focal points, not necessarily with equal 'focal signal intensity': given that the central square is distinguished from the majority of the options by two characteristics, namely location and shape, and the black circle only by one, it is not unreasonable to assume that the former would be the stronger focal point of the two. To use Barrett's phrase, the central square would perhaps be the most likely outcome in this game.[10]

But why, one might well wonder, should all this be relevant to non-tacit ('explicit') negotiations where the parties are not barred from communicating? Schelling's answer to this clearly legitimate question is as follows:

> Most bargaining situations ultimately involve some range of possible outcomes within which each party would rather make a concession than fail to reach agreement at all. In such a situation any potential outcome is one from which at least one of the parties, and probably both, would have been willing to retreat for the sake of agreement, and very often the other party knows it. Any potential outcome is therefore one that either party could have improved by insisting; yet he may have no basis for insisting, since the other knows or suspects that he would rather concede than do without agreement. Each party's strategy is guided mainly by what he expects the other to accept or insist on; yet each knows that the other is guided by reciprocal thoughts. The final outcome must be a point from which neither expects the other to retreat; yet the main ingredient of this expectation is what one thinks the other expects the first to expect, and so on. Somehow, out of this fluid and indeterminate situation that seemingly provides no logical reason for anybody to expect anything except what he expects to be expected to expect, a decision is reached. These infinitely reflexive expectations must somehow converge on a single point, at which each expects the other not to expect to be expected to retreat.[11]

To reach an agreement, explicit bargaining is thus meant to involve a coordination of the participants' expectations, and since parties will, in many cases, not divulge their true positions in this respect, it seems not unreasonable to interpret this part of negotiations as an instance of tacit bargaining. While I do not wish to deny that principles of equity can be involved in this sort of tacit second-guessing of expectations, I reject

[10] Note, incidentally, the crucial role of the common background knowledge in these deliberations: if the options presented in Fig. 1a had been coloured in red and green, and if the possibility of colour blindness were common knowledge, then there would only be a single focal point in this game.

[11] Schelling (1960:69f.)

the view that the function assigned to them in these coordination games exhausts the actual role they play in negotiations of the sort we are considering here.

Let me begin to explain my objections by looking more closely at the role actually attributed to equity principles in the focal point theory. The sole role of focal point signals is to single out one of the available options and so fill 'the vacuum of indeterminacy that would otherwise exist'.[12] The way in which they achieve this and the fact that they are principles *of equity* is completely immaterial. Indeed, the equal distribution D^4 could equally well have acquired its focal point nature, say, from the common background knowledge that all parties are mathematicians with some predilection for symmetry. As it happens, they would not even need to have a predilection for symmetry, as long as one could be sure that to a mathematician, the said distribution somehow stands out from all the others of the 'continuum of possible alternatives'.[13] In precisely the same vein can parties easily not have a predilection for equity without diminishing the focal point signal strength of equity principles. Indeed, both Barrett and Shelling emphasise this very possibility: Barrett, on the one hand, proclaims that, in the examples he was considering,

> the outcomes that seemed compelling did not derive their attraction from their ethical properties. Rather, the ethical rules were known to each party, and were known by each party to be known by the other party, and so on. The ethical rules thus served as focal points.[14]

Schelling, on the other, claims that, in a two-party context,

> 50-50 seems a plausible division, but it may seem so for too many reasons. It may seem 'fair'; it may seem to balance bargaining powers; or it may, as suggested in this paper, simply have the power to communicate its own inevitability to the two parties in such fashion that each appreciates that they both appreciate it.'[15]

And he concludes that

> when the pressure of public opinion seems to force the participants to the obviously 'fair' or 'reasonable' solution, we may exaggerate the 'pressure' or at least misunderstand the way it works on the participants unless we give credit to its power to coordinate the participants' expectations. It may, to put it differently, be the power of suggestion, working through the mechanism described in this paper, that makes public opinion or precedent or ethical standards so effective. [...] Finally, even if it is truly the force of moral responsibility or sensitivity to public opinion that constrains the participants, and not the 'signal' they get, we must still look to the source of the public's own opinion; and there, the writer suggests, the need for a simple, qualitative rationale often reflects the mechanism discussed in this paper.[16]

---

[12] Schelling (1960:73).
[13] Schelling (1960:70).
[14] .Barrett (1992:87ff).
[15] Schelling (1960:72).
[16] Schelling (1960:73).

The picture of the principles' moral force emerging from this analysis is clearly less than impressive. The only reason, it appears, they have any role to play at all in negotiations is that they are assumed to be part of the common background knowledge and that they are associated with some sort of 'crude simplicity'.[17] To be sure, the simplicity in question is inherent in the distribution rules determined by the principles and not in the equity-based distributions themselves, but the crucial point is that, by virtue of this simplicity (and not because of their ethical nature), equity principles are meant to single out certain solutions as positions from which none of the parties expects the others to retreat. The role of simplicity in this is meant to be that of round figures in, say, wage negotiations, where there can be no doubt that a party is more likely to retreat from a demand of say 2.0437892 per cent than it would from a nice 'round' figure such as 2 per cent. The role of equity in negotiations is thus, in essence, meant to be exhausted in the provision of a commonly appreciated measure of simplicity, allowing the negotiators to dig in their heels and to be expected to do so.[18] Moral properties, in particular, are purely epiphenomenal in this context.

A second point which was meant to be illustrated by the *SB*-example is the dependence of equity-based focal points on the presupposed background information. The fact that the applicability of principles of justice is tied to the presence of a particular kind of information has, of course, been long established in moral philosophy. Indeed, Amartya Sen has produced the following interesting example of how informational limitation can lead to a disambiguation of what would otherwise be morally ambiguous situations:[19]

> The point can be illustrated by considering the problem of a person who is asked by three boys to arbitrate who should get a flute (made of bamboo) about which the boys are quar-

---

[17] 'More impressive, perhaps, is the remarkable frequency with which long negotiations over complicated quantitative formulas or *ad hoc* shares in some costs or benefits converge ultimately on something as crudely simple as equal shares, shares proportionate to some common magnitude (gross national product, population, foreign exchange deficit, and so forth), or the shares agreed on in some previous but logically irrelevant negotiation.'(Schelling 1960:67)

[18] Indeed, Schelling seems to hold the view that the only reason why the public may subscribe to certain equity principles is because they are associated with this sort of simplicity. I find this simplistic (not to say patronising) view of public opinion astonishing, to say the least, especially since it is not confined to this author, as we shall see later on.

[19] By 'ambiguous' I mean essentially the situations which Sen refers to as 'complex': 'When more than one moral claim is accepted and there are several non-compulsive principles competing for attention, we have a complex moral structure.'(Sen 1984:288) I myself shall use 'complex' to refer to situations where a plurality of principles is applicable, regardless of what their moral relations may or may not be.

relling. Consider first three alternative scenarios. In the first case, it is known that boy *A* plays the flute well and with very great pleasure, while boys *B* and *C* are less musical. It is clear to the arbitrator that *A* will get more happiness out of the flute than the other two. The arbitrator knows nothing else about the three boys, and decides to give the flute to *A*, in conformity with utilitarianism. In the second case, the arbitrator knows that boy *B* is much more deprived than the other two and has very few toys and other sources of pleasure and that he is generally much less happy than the other two. Nothing else is known about the boys, including who plays the flute well; the arbitrator decides, in this case, to give the flute to *B* on grounds of leximin or difference principle. In the third case, the arbitrator gathers that boy *C* made the flute with his own labour starting from a bamboo belonging to no one, while the others not only did not contribute anything to this effort, but wanted to take the flute away from him. She knows nothing else about the boys, for example, who is how well off, or who enjoys playing the flute more. In this case, the arbitrator decides to give the flute to *C* because of his labour, or as part of an entitlement structure incorporating the right to what one has produced, or on libertarian grounds (see, for example, Nozick, 1974).[20]

This example is of particular interest in the present context because it may, on the one hand, shed some light on the way in which the shifts in focal point in the *SB*-scenarios might come about. On the other, it contains certain (hidden) features which raise questions about the role assigned to equity in focal point theories. The fact is that, strictly speaking, none of the three alternative scenarios presented by Sen are morally simple. The information provided in each of them contains the sort of information which in the *SB*-example was denoted by 'I^A', namely that there are *three* competing boys. This data, when taken by itself, clearly could have suggested an application of an egalitarian principle, leading to a 'time-share' ruling by the arbitrator. The question thus has to be why we can nonetheless confidently follow Sen in assuming that 'in each case the arbitrator may feel that an unambiguously correct decision has been made'.[21] What seems to be clear is that this is not an issue of simplicity of outcome: if anything, the egalitarian solution seems to be at least as simple as any of the ones actually put forward by the arbitrator. The only possible answer, in my mind, is that – in the context of the information actually provided in Sen's scenarios – the 'moral force' of the principles which the arbitrator *did* evoke was, without exception, perceived by her as stronger than that of the egalitarian principle.

And the same applies, *mutatis mutandis*, to the *SB*-example. The 'official' reason why we are likely to find a shift of focal point from the equal distribution D^A to, say, the *per capita* distribution D^B if the initial background information I^A is amplified by including population figures is that this is achieved 'just by suggesting their rele-

[20] Sen (1984:290).

vance and making them prominent in the problem'.[22] Yet there are limits to the effectiveness of this sort of external suggestion: although I have not carried out an empirical study, it is clear to me that no one in their right mind would think that, say, the number of bars in the parties' national anthems would be of any relevance whatsoever, no matter *how* prominent the presentation. There may be many reasons why certain parameters are considered to be relevant to the distribution problem at hand, such as the fact that they occur in an equity-based distribution rule. Mere prominent presentation, however, cannot be one of them.

What may happen, of course, is that by emphasising a parameter commonly known to be relevant, say population number, a subject's attention might be drawn to the relevant distribution rule (R^B). But this, I submit, is not sufficient to explain the 'obviousness' of the *per capita* distribution in this context. In trying to co-ordinate expectations, a rational party must take into account the fact that there are two potential focal points (D^A and D^B).[23] The question then becomes what sort of differentiating features between the two (if any) makes one of them more 'obvious' than the other. The answer, according to the received focal point theory, would presumably be twofold, namely (i) some sort of simplicity and (ii) the suggestive power given by the emphasis of the population parameter. Note, however, that the focal point signals imparted by these properties point in opposite directions: while the prominence of the population figures is meant to enhance the focal point strength of the *per capita* distribution, there can be no doubt that, as far as simplicity is concerned, the equal distribution (rule) carries the prize. In view of these conflicting signals, I find it difficult to see how, in the absence of any other contributing factors, anything short of an explicit external endorsement of the *per capita* distribution could explain why subjects would most likely regard this distribution to be the 'obvious' one. Given then that according to both Schelling and Barrett, coordination is meant to be generally accomplished even without such explicit endorsements, focal point theory, as it stands, faces an explanatory impasse which impairs its credibility in this context.

[21] Sen (1984:290).
[22] Schelling (1960:65).
[23] Similar to the game involving Figure 1b, a successful outcome here involves a co-ordination between different competing focal point signals. Yet while it was possible to 'disambiguate' the scene presented in (a red-green version of) Figure 1b by reference to some common knowledge that some of the parties might simply not see one of the potential focal points, there is no analogous 'moral colourblindness' which could be evoked to achieve the same result in the present case.

The only way I can see to remedy this situation is to take seriously the 'moral properties' of equity-based distribution proposals. In doing so, we can evoke a focal point signal which actually reinforces the parameter signal (ii), namely the perception of a generally acknowledged moral superiority of the *per capita* rule R^B over the egalitarian rule R^A in the context described in *Scenario B*. In other words, the co-ordination can be explained by reference to the parties believing that, in the situation described, the *per capita* rule would generally be regarded as the more equitable one.[24]

This is not to say that non-cooperative coordination can *only* be achieved by listening to these moral signals. Indeed, given a certain moral complexity of the scenario, it seems highly unlikely that any one of the possible equity-based distributions could appear to have the generally accepted status of being morally superior to all the others, forcing the players to search for some other individuating feature in their striving to coordinate their choices.[25] The point here is simply that focal point theory, even within the narrow confines of tacit bargaining as exemplified in non-cooperative games, can be a credible explanatory theory only if 'moral properties' are admitted as decisive factors in their own right and the role of equity is acknowledged to go beyond the mere provision of simplicity.

And yet, I am afraid, focal point theory – even in such a 'morally sensitive' version – faces further problems as an explanatory tool for future global warming negotiations. But before I turn to discuss these shortcomings, let us briefly consider an account which has been put forward as an alternative to the focal point view, namely the *realpolitik* solution endorsed by Dallas Burtraw and Michael Toman (1992).

[24] Note that this does not mean that the parties have to subscribe to this view, let alone that they must be assumed to be driven by equitable motives. All they must believe in is that there is their said social norm.

[25] Schelling, incidentally, describes an experiment which, in essence, corresponds to the situation described in Scenario 2. Interestingly, and contrary to the likelihoods expressed by Barrett, co-ordination was dominantly achieved by choosing the equal distribution. Schelling himself puts this down to the 'refined signal' – in our case, the '*per capita* signal' of *Scenario B* – being 'drowned out by "noise"'(Schelling 1960:65) in the sense of the population figures losing their prominence in the field of other parameters mentioned in the Scenario. My own conjecture would be that the information provided in *Scenario C* entails a morally complex set-up incapable of sending a sufficiently unique focal point signal, thus forcing the players to focus on the signal associated with simplicity.

The Strategic Theory of Justice

In the first half of their study, Burtraw and Toman analyse the distributional consequences of four 'fairness rules' in the context of CO_2 mitigation, namely

1. *Ability to pay*: Set each country's emission reduction so that abatement cost relative to precontrol income is the same for all countries, subject to some overall level of cost to be borne by all nations.
2. *Polluter pays*: Set each country's emission reduction so that abatement cost relative to precontrol emissions is the same for all countries, subject to some overall level of cost to be borne by all nations.
3. *Equal percentage cuts*: Set each country's emission reduction equal to the same percentage of pre-emission levels.
4. *Tradable, population-based emission rights*: Endow each country with a right to emit CO_2 in proportion to its population, subject to global emissions being cut by some percentage; then countries either reduce emissions to meet their national quotas or purchase emission rights from other nations.

On the basis of this analysis they conclude that the disparities among the allocations of burden prescribed by these four criteria are such as to 'raise doubts about whether focal points can effectively guide CO_2-containment negotiations'.[26] There is no doubt in my mind that, as far as future negotiations are concerned,[27] Burtraw and Toman have got it right, but, I am afraid, for quite the wrong reason. Remaining firmly within the non-cooperative bargaining paradigm, they reject the focal point analysis chiefly because of a 'lack of precedent favoring particular rules'.[28] Indeed, they credit their alternative 'endogenous emergence' conception to what they regard as an 'important lesson from noncooperative bargaining theory'[29] which, they claim,

> suggests that when negotiations have significant stakes, negotiators are more inclined to turn from precedent and struggle to achieve incremental advantage in the new setting (Binmore et al. 1985, 1988). After the fact, all parties may agree that the outcome was fair, based on the relative strengths of the parties rather than on the invocation of focal points (Binmore et al. 1989).[30]

Having read the three mentioned articles by Binmore *et al.*, I can see how their empirical evidence could be used to suggest that negotiating parties, through a learning process, could come to accept the strategic solution as the best they could hope for given the relative bargaining strengths.[31] But I was unable to find any evidence that

[26] Burtraw and Toman (1992:127).
[27] That is, negotiations involving both developed and developing (non-Annex I) countries.
[28] Burtraw and Toman (1992:127).
[29] Burtraw and Toman (1992:127).
[30] Burtraw and Toman (1992: 128).
[31] Just to give an example, Binmore et al. have carried out laboratory experiments involving non-cooperative coordination games such as the following two-stage game: 'Stage I: The cake is of size 100 pence. Player 1 makes a proposal (X); Player 2 accepts (1 receives X, 2 receives $100 - X$) or

they see this outcome as being fair. Indeed, I find it difficult to believe – in particular, given the strategic elements explicitly mentioned by Burtraw and Toman (*viz.* relative rate of time preference, relative risk aversion, coercive use of unrelated issues) – that the use of these bargaining instruments should somehow lead to generate new standards of equity. If I pressure someone at knifepoint to accept an offer which, under normal circumstances, he would consider to be unfair, he may or may not accept it, but it is highly unlikely that he will regard it as being fair. And this will remain the case if I repeat the 'offer he cannot refuse' long enough for him to learn that it would be wise to accept.

Although I find it difficult to believe that Burtraw and Toman would wish to embrace this 'might is right' conception of equity, on reflection[32] I cannot but conclude that they do. Even though they find themselves in distinguished company – such as David Lloyd George who in the 1918 election campaign proclaimed on the topic of war debts:

> Who is to foot the bill? By the jurisprudence of any civilised country the loser pays. It is not a question of vengeance, it is a question of justice.'[33]

– I find it difficult to believe that anyone – with the possible exception of those with strategic advantages – could accept this sort of strategic criterion ('if it is the best we can get, it has to be fair').

The reason why Burtraw and Toman put forward their strategic theory is ultimately their realisation that

> There remains the possibility that standards of equity held by citizens in different countries are mutually exclusive, and that they do not allow room for agreement. Without modifying

rejects (game continues). Stage II: The cake is of size 25 pence. Player 2 makes a proposal (X'); Player 1 accepts (1 receives X', 2 receives $25-X'$) or rejects (1 receives 0, 2 receives 0). A game-theoretic analysis requires that Player 1 makes an opening demand in the range 74 – 76 pence, and Player 2 accepts any opening demand of 74 pence or less (for he cannot do better by refusing, even if he obtains the entire cake in the second stage)'(Binmore *et al.* 1985:1178). They found that while first time players most frequently started with a 50:50 opening demand, experienced players most frequently chose 75:25 (the strategic solution, not a focal point!). And they conclude that 'subjects, faced with a new problem, simply choose "equal division" as an "obvious" and "acceptable" compromise – an idea familiar from the seminal work of Thomas Schelling (1960). We suspect, on the basis of the present experiments, that such considerations are easily displaced by calculations of strategic advantage, once players fully appreciate the structure of the game'(Binmore *et al.* 1985: 1180).

[32] Indeed, in an earlier version of this paper, I tried to convince myself that the authors might have been the victims of a conceptual confusion. However, I no longer find this convincing.

[33] As quoted in Smith (1974):715.

these standards, then the prognosis for an international agreement that can be meaningfully implemented is non-existent.[34]

While fully agreeing with their initial statement – in particular when interpreted as a diagnosis of potential globally comprehensive negotiations – I have no faith at all in the efficacy of the proposed prescription: to think that a comprehensive international agreement could be achieved by trying to move the equity bench-marks through strategic negotiating techniques is simply a non-starter.

To be fair, Burtraw and Toman do concede that conventional equity principles ('as expressed by potential focal points'[35]) may retain some relevance to negotiations, but only as 'constraints' in a second phase of the bargaining problem:

> The first phase involves the actual negotiations among various agents from different nations; the second involves the ratification and implementation of the agreement by the body politic. We conjecture that in the first phase, the strategic elements of the bargaining problem will be dominant. [...] The outcome of this phase of the problem, taken in isolation, might depend little on fixed standards of allocational equity. [...] and new standards would emerge that rationalize the bargaining outcome.
>
> However, negotiations will take place in the shadow of the influence of opinions held by the body politic. Any outcome of the negotiations must be credible with regard to ratification by the different nations. Any proposed outcome that cannot credibly be implemented in the second phase of the game also cannot be credible during the first phase. Through this constraint the prior beliefs of the principals manifested in simple fairness rules may take on strategic importance in the negotiations.[...]
>
> This framework emphasizes the asymmetry between principals and their agents. Principals will not have immediate access to all the information and discussions in the negotiation that can promote learning and lead to a convergence of expectations.'[36]

Burtraw and Toman are quite right in emphasising the relevance of an asymmetry between principals and their agents, but the real relevance is not the one they have in mind. It may be true that the principals are ignorant of the 'technical issues of the negotiations' *including* the strategic elements, but it is wrong to assume that they cling to a belief-system guided by some equity-principle or other because of this ignorance, and that, if properly educated, they would see the light and recognise the strategic outcome negotiated by their agents, 'based on the relative strengths of the parties'[37] as being fair. To think that re-educating the populations of all the 'strategically challenged' countries would ultimately lead them to switch to the 'might is right' doctrine is simply to delude oneself about the actual force of the traditional principles of equity in people's decision-making processes.

[34] Burtraw and Toman (1992: 130).
[35] Burtraw and Toman (1992: 128).
[36] Burtraw and Toman (1992: 129).
[37] Burtraw and Toman (1992: 128).

The true asymmetry between principals and agents, in particular between those who lack these advantages, is that agents are more likely to succumb to strategic pressures for the simple reason that they have a considerable additional incentive to come to a conclusion. After all, they are sent to negotiate, and indeed paid to come back with the best agreement they can get, and not to return empty handed. The principal, by contrast, will generally have no such incentive. Indeed, I very much doubt that an 'immediate access to all the information and discussions [including threats of coercive use of unrelated issues] in the negotiations'[38] would lead the populations in question to modify their equity standards and to accept the strategic outcome as being fair. If anything, such knowledge is likely to lead to an outright rejection of the settlement whatever it may be.

Doomsday Scenarios

In my initial critique of the generally adopted focal point theory I argued that – due to an equivocation of equity with simplicity – it fails in explaining even some tacit bargainings. This shortcoming, I suggested, might be remedied if 'moral properties' were admitted as focal point signals in their own right.[39] Unfortunately for the focal point approach, this is not where the problems end, at least not as far as post-Kyoto climate change negotiations are concerned.

The crucial point here is that, even in this morally sensitive guise, focal point theory relies on the fundamental assumption of there being a range of possible negotiating outcomes which are commonly acceptable, in the sense that for any one of them 'each party would rather make a concession than fail to reach agreement at all.'[40] Indeed, focal point theory considers only options of this generally acceptable type. In order to be a focal point, a (logically) possible solution must therefore be a generally acceptable option to begin with.

[38] Burtraw and Toman (1992: 129).
[39] Let me emphasise again that this is not the same as the assumption that the parties involved are necessarily motivated to act morally. All that is needed is a general awareness of the fact that in certain situations there may be an equity principle which, by virtue of its moral force, is generally acknowledged to be the right one for the situation in question.
[40] Schelling (1960:69).

Given that the parties who were meant to take on emission targets at the negotiations in Kyoto were all more or less on the same (high) level of economic development,[41] it may well be that there was indeed such a range of commonly acceptable possible solutions from the outset of the negotiations. If so, focal point theory might be able to provide an explanation of the outcome of the Kyoto negotiations. The fatal obstacle for a focal point explanation of such negotiations arises when they become truly global, i.e. when they are meant to bring about an agreement which is binding for both developed and developing countries.

From the evidence available, we can safely predict that in this global negotiating context, both 'grandfathering' and 'per capita' distribution proposals will be put forward as the (only) just solution to the claims problem. Given their disparity, it does seem unlikely in the extreme that either of these proposals is commonly acceptable, implying that neither of them could serve as a focal point. Indeed, in view of this disparity, it is questionable whether there will be *any* possible outcomes of this commonly acceptable nature at all, a situation which, for obvious reasons, might be called a '*doomsday scenario*'.

In the context of such a doomsday scenario, focal point theory simply does not get off the ground. This, it might be argued, is not surprising for there would be no negotiating outcome to be explained in the first place, as doomsday scenarios must end in a breakdown of negotiations. While saving focal point theory from being branded inadequate, such a defeatist view does nothing to save us from the effects of global warming. But are we really forced to adopt such a defeatist position? Are there no ways in which we might be able to avert a breakdown even if we are faced with such disparate initial negotiating positions?

The answer, I believe, is 'no'. There is room for guarded optimism for there are ways in which a doomsday scenario might be averted even if the initial positions are as *prima facie* irreconcilable as the ones predicted for future rounds of the FCCC negoti-

[41] The exception being the 'economies in transition', whose economies had collapsed after 1990. And indeed, it seems unlikely that the Kyoto negotiations could have avoided an equity discussion if the base line for the targets had been set later than 1990.

ating process. To explain how this could be achieved, we will however have to incorporate justice (equity) into the basic fabric of our theories. We will have to acknowledge its proper (fundamental) role and provide the theoretical substance to its commonly proclaimed star-billing, which otherwise remains mere lip-service.

II

ALTERNATIVE CONCEPTION:

INEQUITY-DISUTILITY AND PROCEDURALLY FAIR SOLUTIONS

The alternative conception I have in mind is based on the idea that equity-based proposals are modelled as *positions* ('fix points') *of minimum unfairness*, relative to the equity standards which they are grounded on:[1] If '$D^{(k)}$' denotes the distribution which party k considers to be the 'just' solution, then k is likely to regard certain deviations from this position as diminishing his welfare because of a (perceived) inequity. In other words, the overall welfare-level which k associates with (accepting) some distribution D may, if we wish, 'contain a negative component,' a feeling of being unfairly treated to a degree tied to D's position *relative* to $D^{(k)}$. This is not to say that other welfare parameters associated with D – such as the welfare which k would derive from the economic benefits of D – will necessarily be outweighed by this negative component. What is crucial here is simply that this *may* happen, in which case we could expect a breakdown even though an acceptance would have meant the (*ceteris paribus*) acquisition of a good, a state of affairs which is inexplicable in the traditional framework.[2]

The Inequity-Disutility (*ID*) Model

To flesh out this idea, let me introduce a rudimentary mathematical model framed in terms of a conceptual scheme which has proven to be useful in this context.[3]

[1] I am consciously using the term 'equity *standard*' (as opposed to 'equity *principle*') in order to allow for the possibility that distributions are perceived as being minimally unfair on the basis of some form of moral intuition.

[2] The reason for this Schopenhauer style interpretation of fairness as the removal of unfairness is to make room for a phenomenon which is inexplicable in the traditional framework: if, as presupposed, the commodity to be distributed is indeed an economic 'good' (an assumption which seems to be justifiable in the case of tradable emissions permits), then, according to the traditional framework, any (*ceteris paribus*) acquisition of this commodity is meant to *increase* the level of welfare. Given furthermore the traditional welfare-maximising conception of an agent, the choice to forego a possible acquisition becomes inexplicable (or 'irrational' to those who feel uncomfortable with the thought that there could be something which the traditional theory might not be able to explain). As far as this traditional conception is concerned, the framework proposed here amounts to a 'hidden variable explanation', i.e. it puts down the shortcoming of the former to its ignoring certain relevant parameters.

[3] I must emphasise that the mathematical model I am going to construct has to be taken with a pinch of salt, for I fully agree with Kreps's warning that 'the units in a utility scale, or even the size of relative cont.

Assuming that T is a quantity of a divisible, homogeneous good to be distributed amongst, say, two parties A and B, a *distribution* D of T is determined by the fractions: $d_A \cdot T$ and $d_B \cdot T$ (with $0 \leq d_A, d_B \leq 1$, $d_A + d_B \leq 1$) allocated to A and B, respectively. It can thus be represented by a pair of positive real numbers: $D = \langle d_A, d_B \rangle$ - or $D^m = \langle d_A^m, d_B^m \rangle$, if we wish to differentiate between distributions by way of using indices.

If *all* of T is allocated, we have a *total* distribution: $d_A + d_B = 1$. For modelling purposes, it is advantageous to include *partial* distributions - such as the one where A receives three-fifths and B one-fifth: $\langle 3/5, 1/5 \rangle$ (see Fig. 6) - in the domain of our model, if only to avail ourselves of the null-distribution: $\langle 0, 0 \rangle$ as a representation of the state of affairs both prior to the negotiations and after a possible breakdown.[4] Accordingly the domain of our model will be the class: $\mathfrak{D}_0$ of all the logically possible partial and total distributions.

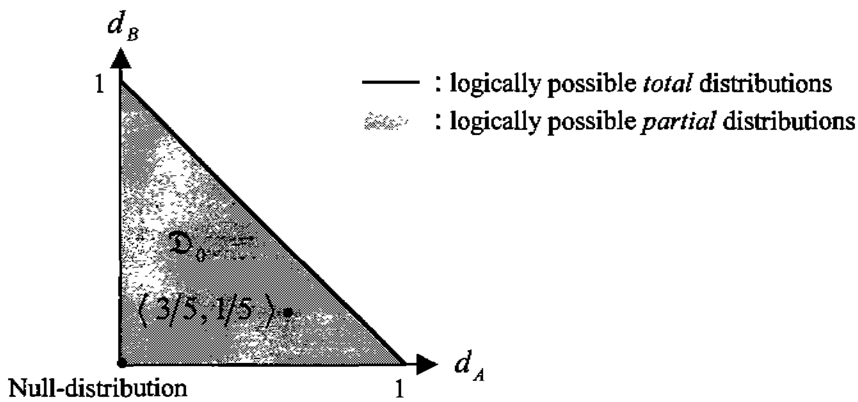


Figure 6: The Domain of the *ID*-Model

differences, have no particular meaning. We can't, in looking at a change from x to y, say that the consumer is better off by the amount $U(x) - U(y)$ or by anything like this. At this point [...] the utility function is introduced as an analytical convenience. It has no particular cardinal significance. In particular, the "level of utility" is unobservable, and anything that requires us to know the "level of utility" will be untestable. This is important as we go through demand theory; we'll want to be careful to note which of the many constructions we make are based on observables, and which are things that exist (if at all) only in the mind of the economist.'(Kreps 1990:32)

[4] To simplify matters, it is assumed that if one of the parties rejects a final offer then we are back to square one, i.e. no one gets any part of T. Note, however, that this is not an essential presupposition for the model suggested.

In some cases, it will be useful to interpret (finite) subsets of $\mathfrak{D}_0$ as '*distribution matrices*', where the distributions are taken to make up, say, the matrix columns. With this, we can turn to numerically representing the parties' welfare in terms of welfare functions: Let $W_k(D)$ represent the welfare which k would experience if D were the outcome of the negotiations, and let $u_k(t)$ be the '*acquisition utility*' (or short '*utility*') which k would derive from obtaining t units of the goods in question.[5] Since we shall focus on comparisons with the *status quo ante*/breakdown levels, we can furthermore assume that these functions are appropriately calibrated:

$$W_k(\langle 0, 0 \rangle) = 0 \text{ and } u_k(0) = 0 \quad (\text{for } k = A, B).$$

From a traditional point of view, this formulation would seem to be somewhat redundant, since no distinction is made between the two: the individual welfare associated with a distribution is simply taken to *be* the utility derived from obtaining the amount allocated under this distribution, i.e. the traditional understanding is that if $D = \langle d_A, d_B \rangle$ then $W_k(D) = u_k(d_k T)$. In the *ID*-model, however, the two concepts are

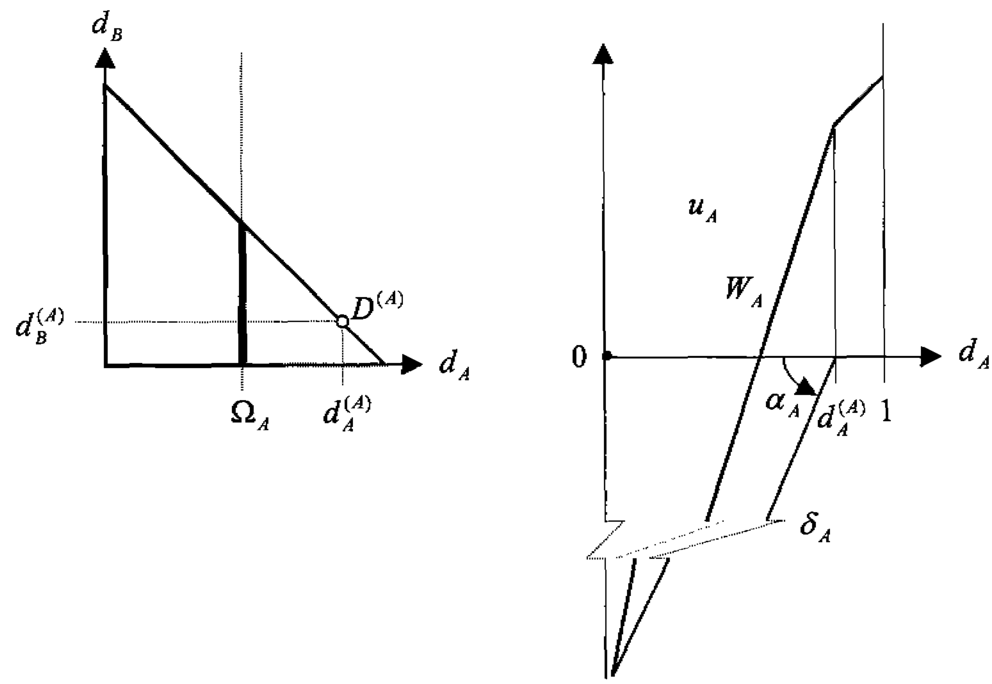


Figure 7: A Possible Set of Welfare Component Functions for Party *A*

[5] This acquisition utility is simply meant to represent the welfare which the party in question would derive purely from acquiring the good, i.e. in ignorance of factors such as how much the other parties get, etc.

distinct due to equity effects on the overall well-being of the parties. In applications,[6] u_k is traditionally also assumed to be linear in the amount of the good received. As we are engaged in a general conceptual exposition, I propose to follow suit by stipulating that: $u_k = u_k(d_k T) = d_k T$ (see Fig. 7). Indeed, to simplify things even further, let us assume that $T = 1$, which leaves us with: $u_k(d_k) = d_k$.

To model the mentioned equity effects, *inequity-disutilities* (or short: '*disutilities*') reflecting say again A's aggravation from being unfairly treated, is represented by a function δ_A into the negative real numbers. In a first instance, these disutilities depend on the distributions up for consideration: $\delta_A = \delta_A(D) \leq 0$.[7] Since $D^{(A)} = \langle d_A^{(A)}, d_B^{(A)} \rangle$ is the equity-based distribution which A regards as minimally unfair,[8] it is meant to impart minimal inequity disutility to A: $\delta_A(D^{(A)}) = 0$. Moreover, in keeping with the traditional pessimistic view on the moral fabric of economic agents, I propose to adopt initially what might be called the 'toddler conception'[9] of fairness according to which (i) A's disutility depends only on his share: $\delta_A(D) = \delta_A(d_A)$, and (ii) minimal disutility is felt if he gets more than his 'fair share': $d_A^{(A)}$,[10] while the disutility increases in strength when the share moves (in opposite direction) below $d_A^{(A)}$:

(i) $\delta_A(d_A) = 0$ for $d_A \geq d_A^{(A)}$;
(ii) $\delta_A(d'_A) \leq \delta_A(d_A) \leq 0$ for $d'_A \leq d_A \leq d_A^{(A)}$.

To simplify matters even further, we might as well use the linear disutilities depicted in Fig. 7:

(i) $\delta_A(d_A) = 0$ for $d_A \geq d_A^{(A)}$;
(ii) $\delta_A(d_A) = (d_A^{(A)} - d_A)\tan\alpha_A$ for $d_A < d_A^{(A)}$, $0 \leq \alpha_A < \pi/2$.

---

[6] See Binmore et al. (1989):753.

[7] As before, the universal calibration of minimal disutility at 0 is a convention for the sake of arithmetic simplicity and, as such, not essential to the model.

[8] It is, of course, possible that some parties do not have any views on this matter, meaning that for some *k*s there may not be such a $D^{(k)}$. Yet, I believe that, in their case, no harm is done to our model if they are simply assigned an arbitrary $D^{(k)}$, provided that their disutility is then presumed to be nil throughout.

[9] I find it striking, from personal experience, how often the first usage of 'It's not fair!' is tied to this sort of conception. Moreover, I believe that this instinctive view stays very much with us even when we are meant to have developed into mature moral agents. The only reason for adopting it here, however, is simplicity of exposition.

[10] If we did wish to make room in our model for somewhat more altruistic players, we would, presumably, have to allow for non-minimal disutilities in this case.

This use of linear disutilities enables us, in particular, to interpret A's '*inequity sensitivity*' (or short: '*sensitivity*') in a very simple manner as the angle α_A, ranging from complete insensitivity at $\alpha_A = 0$ to unsurpassable hypersensitivity at $\alpha_A = \pi/2$. A's total welfare from a particular outcome D is then represented as the sum of the utility and the disutility he derives from this outcome:

$$W_A(D) = W_A(d_A) = u_A(d_A) + \delta_A(d_A)$$

These disutilities thus force A's welfare at $d_A = \Omega_A$ below the welfare he associates with the breakdown option.[11]

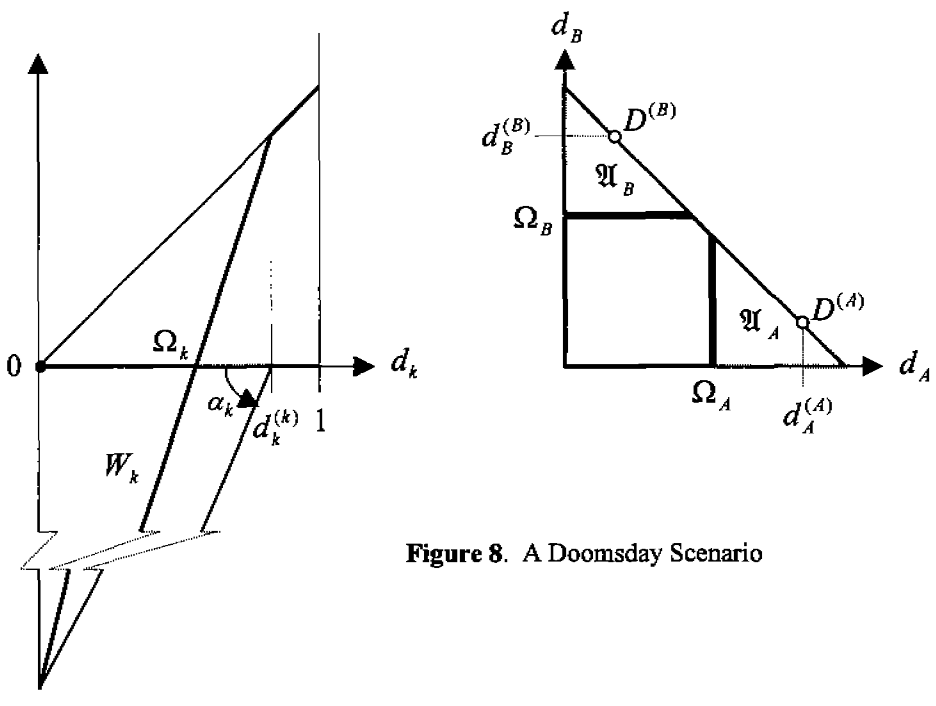


Figure 8. A Doomsday Scenario

Consider now the case where *both* parties conform to the situation depicted in Fig. 7 (and 8). In other words, assume they both have the same sensitivity to being treated unfairly, and they both think that the 'cake' should rightly be divided in a specific, uneven proportion; all they disagree about is who should get the bigger piece.

[11] The use of linear disutilities may be advantageous as far as representing sensitivities is concerned. However, it also has the less desirable effect of introducing a discontinuity into the welfare functions at the null-distribution. In Appendix 4, I shall introduce a more palatable continuous version of these disutilities. At this stage all I can suggest is that anyone offended by the said discontinuity, ignore partial distribution options (apart from the null-distribution) and interpret our welfare function diagrams as pertaining to the relevant total distributions alone.
cont.

If by k's '*individual* acceptability range': $\mathfrak{A}_k$ we mean the range of distributions which, if accepted, would make k better off than at the outset (after a breakdown): $\mathfrak{A}_k = \{ D \in \mathfrak{D}_0 : W_k(D) > 0 \}$, we find that in the situation depicted in Fig. 8, the '*common* acceptability range': $\mathfrak{A} = \mathfrak{A}_A \cap \mathfrak{A}_B$ is empty. There are no possible solutions which both parties would prefer over a breakdown. This, I contend, corresponds precisely to the sort of situation envisaged by Burtraw and Toman where the standards of equity (together with the sensitivities to inequity) 'do not allow room for agreement'.[12]

The conjecture that an occurrence of the scenario depicted in Fig. 8 in the final negotiating stage will inevitably lead to a breakdown can now be supported by a very simple game-theoretic argument. Consider the following 'take-it-or-leave-it' game: player A tables a final proposal; player B then either signs up or walks away. To be more precise, the game is meant to involve the following four strategies:

Player A:	*Player B*:
s_1 = tables $D \in \mathfrak{A}_A$ (*cf. Fig.* 8)	t_1 = accepts proposal.
s_2 = tables $D \notin \mathfrak{A}_A$ (i.e. $D \in \overline{\mathfrak{A}}_A$).	t_2 = walks away.

The specific pay-offs in this game will, naturally, depend on the particular choice of D. All we need to know, however, is that – according to Fig. 8 – it will have a bimatrix of the type

	t_1	t_2
s_1	p, q	0, 0
s_2	$\bar{p}, \bar{q}$	0, 0

where $p > 0$, $q < 0$ and $\bar{p} \leq 0$. In this case the strategy profile (s_1, t_2) – i.e. a breakdown – is a Nash equilibrium whichever D is actually proposed (in fact, in most cases the only one), and there is a story making it the obvious solution of this game:[13] by opting for s_1, player A can ensure that he gets a (utility-) pay-off which is greater or

[12] Burtraw and Toman (1992:130).
[13] I emphasise the existence of such a story simply because being a Nash equilibrium is only a necessary but not a sufficient condition for being a solution.

equal to anything he can expect when choosing s_2.[14] If, however, he does choose s_1, player B will adopt t_2.

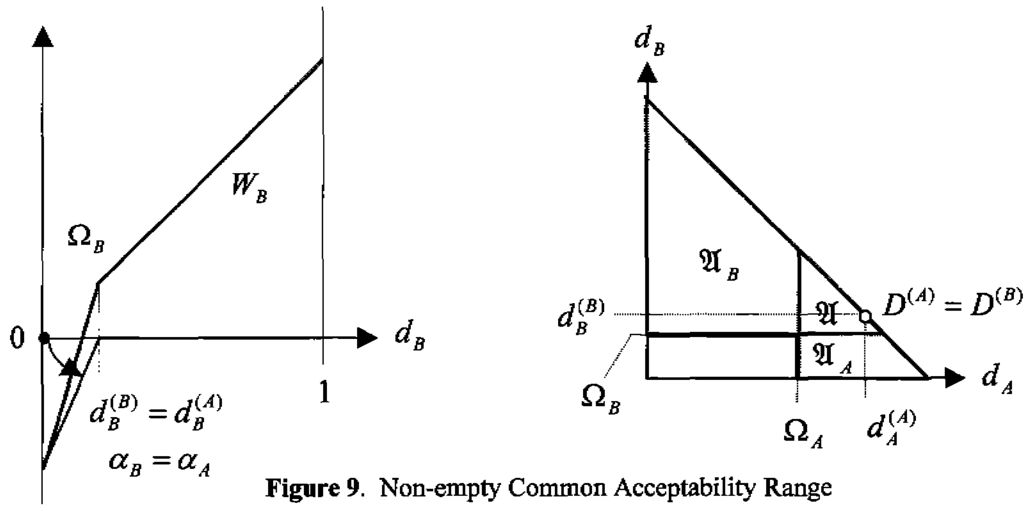


Figure 9. Non-empty Common Acceptability Range

Consider now the situation depicted in Fig. 9, where all is exactly as before, but for the fact that B actually concurs with A's view of what would be the fairest distribution. In this case, player A has four possible strategies, namely to table a proposal D with

(i) $s_1 = D \in \mathfrak{A}_A$ and $D \in \mathfrak{A}_B$ (i.e. $D \in \mathfrak{A} = \mathfrak{A}_A \cap \mathfrak{A}_B$),
(ii) $s_2 = D \in \mathfrak{A}_A$ and $D \notin \mathfrak{A}_B$,
(iii) $s_3 = D \notin \mathfrak{A}_A$ and $D \in \mathfrak{A}_B$,
(iv) $s_4 = D \notin \mathfrak{A}_A$ and $D \notin \mathfrak{A}_B$.

Clearly this game has a solution (an 'obvious way to play'[15]). Moreover, interpreting (i) - (iv) independently of the situation depicted in Fig. 9 as a type-characterisation of our games, it also follows quite generally that acceptance (i.e. t_1) features in a solution if and only if s_1 is a viable strategy (i.e. iff $\mathfrak{A} \neq \varnothing$), in which case the solution is (s_1, t_1).

In practice, however, this solution is not quite as obvious, for the strategy profile (s_1, t_1) can only be implemented if the first player is sufficiently familiar with the common acceptability range: he has to be able to decide whether the proposal he is

[14] In the case of $\bar{p} = 0$ there may be no danger for 1 of losing anything by choosing the second strategy. But if he were to do so he would deprive himself of the chance of gaining something, namely in case 2 acts 'irrationally'.
[15] See Kreps (1990:404).

about to table is indeed acceptable to his opposite number(s). Even if – from a 'God's eye' point of view – there are possible solutions which would lead to a mutually beneficial conclusion of the negotiations, there is no guarantee that mere human negotiators would be able to see them and thus avoid a breakdown.

As far as equity, or rather inequity, is concerned, certain contexts may facilitate the negotiating task. In some circumstances negotiators can safely assume inequity effects to be negligible. If, say, the quantity of the good to be distributed is of very little value to all parties involved, chances are that their inequity sensitivities are close to zero. In

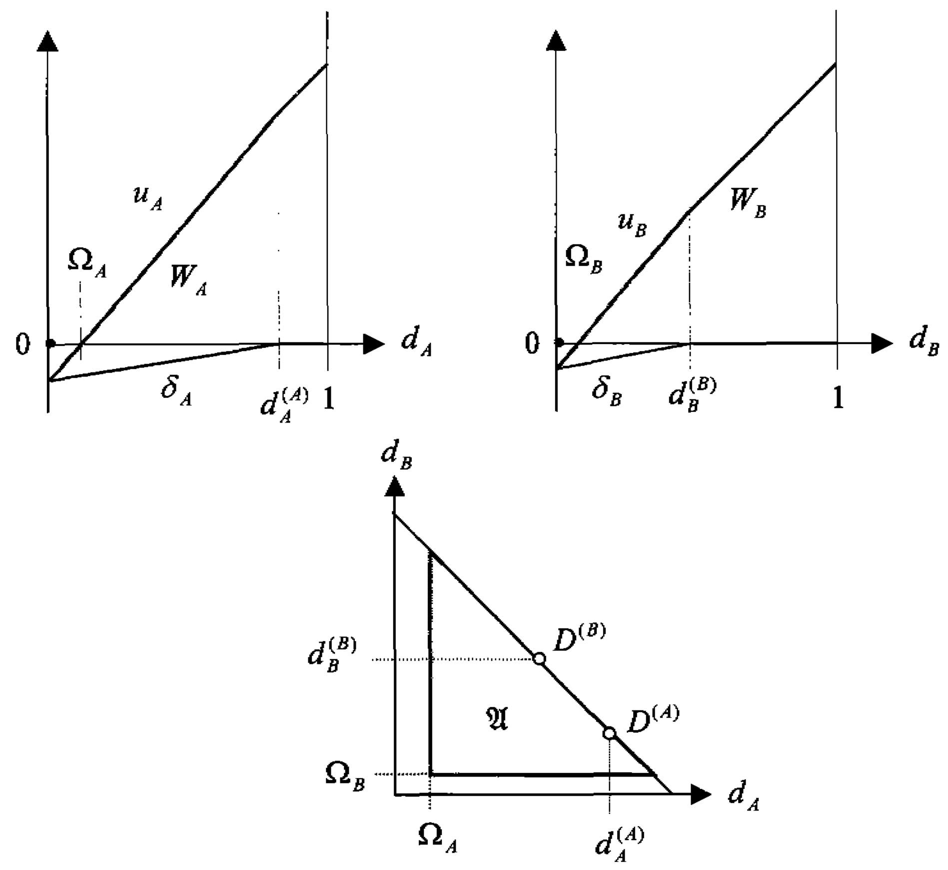


Figure 10: Weak Sensitivities $\alpha_A, \alpha_B << \pi/2$

this case it would not be unreasonable to assume there to be a common acceptability range, indeed one which extends to *most* of the possible solutions, apart from the most extreme cases (see Fig. 10).[16]

[16] Note that, by ignoring equity effects, $\mathfrak{A} = \mathfrak{D}_0$ is arguably one of the presuppositions of Schelling's focal point theory. Yet, if this is so, this theory needs further scrutiny for those cases where the common acceptability range, although non-empty, is a proper sub-set of the domain of all possible cont.

Alternatively, the equity positions adopted by the parties may be sufficiently close to warrant the assumption that they are contained in a neighbourhood of commonly acceptable possible solutions. But what is to be done if both the stakes and the inequity sensitivities are high and the stated equity positions are at 'extreme ends of the spectrum'? As this is more than likely to involve a doomsday scenario, chances for a successful outcome – judging from our game-theoretic conclusions – may seem very bleak, indeed.

Making Compromise Possible: The Role of Procedural Fairness

Could there not be a way of avoiding such a doomsday scenario? Fortunately, I believe, there is. Contrary to what might be concluded from the conception used so far, the inequity-disutilities associated with some possible outcome *D* depend on an array of factors, say $\overline{X} = \langle X, X', X'', \cdots \rangle$, other than just the shares: d_k allocated under *D*. And variations in these other factors may mitigate potential sensitivities sufficiently so as to ensure the creation of commonly acceptable options. In other words, disutilities: $\delta_k(d_k, \overline{X}_0)$ which would generate a doomsday scenario may be mitigated by switching from $\overline{X}_0$ to some other parameter set $\overline{X}_1$: $\delta_k(d_k, \overline{X}_0) < \delta_k(d_k, \overline{X}_1)$. This mitigation, in turn, may be sufficient to allow for commonly acceptable solutions in the scenario involving the mitigated disutilities: $\delta_k(d_k, \overline{X}_1)$.

As mentioned earlier, one parameter which might mitigate sensitivities is the relative value of what is at stake.[17] Yet it would be wrong to pin our hopes for resolving the

distributions, an exercise which, I believe, has to go beyond carrying out laboratory experiments with relatively negligible sums of money to be divided.

[17] Given that we granted ourselves the expository luxury of cardinal utilities in our *ID*-model, the natural way of representing numerically the value which *k* attaches to the acquisition of a share: $d_k T$, is to identify it with what we referred to as *k*'s acquisition utility: $u_k(d_k T) = d_k T$. With this representation, it will be apparent from Fig. 7 that δ_k must be a function of u_k, at least as far as our model is concerned. Otherwise our model will be incapable of accommodating the phenomenon of common acceptability increasing under *ceteris paribus* decreases of the total stake *T*, for it will actually 'predict' a decrease of the individual acceptability range $\mathfrak{A}_k$, if the involved acquisition utility changes *ceteris paribus*, say from $u_k(d_k) = d_k$ to $u_k(d_k) = d_k / 100$ (corresponding to a reduction from $T = 1$ to $T = 1/100$).

global warming doomsday scenario on simply lowering the stakes: the fact is that the size T of total emissions permits is actually, by and large, a fixed quantity.

Fortunately, there are yet other factors which may have a mitigating effect on inequity sensitivities. What I have in mind are parameters relating to properties of what might be called the 'selection mechanism' employed in the negotiations. By this I mean the procedure π under which – based only on information contained in some presupposed *selection base* $\mathfrak{D} \subseteq \mathfrak{D}_0$ – a particular option: $\pi[\mathfrak{D}] \in \mathfrak{D}_0$ is selected for consideration. A *selection based on* $\mathfrak{D}$, it has to be emphasised, is not quite the same as a *choice from* $\mathfrak{D}$ (in the well-known technical sense): selection mechanisms, for one, are always meant to select a single option (as opposed to a 'choice *set*') and, more importantly, this selected option can be any of the logically possible alternatives and is not restricted to being a member of the selection base.[18]

One of the factors which, I believe, *is* relevant in this context is the nature of the adopted selection base $\mathfrak{D} \subseteq \mathfrak{D}_0$. There are several reasons why the composition of the selection base may have an effect on inequity disutilities. For one, we can reasonably expect that if the equity-based option $D^{(k)}$ adopted by k is left out, k will feel less fairly treated than if it is not, whatever the selection procedure might be: if $D^{(k)} \in \mathfrak{D}$ and $\mathfrak{D}' = \mathfrak{D} \setminus \{D^{(k)}\}$, then $\delta_k(\pi[\mathfrak{D}], \mathfrak{D}) \geq \delta_k(\pi[\mathfrak{D}'], \mathfrak{D}')$. Alternatively, if some party puts forward a morally unjustifiable proposal D for inclusion in the selection-base (in particular a *purely* self-serving one), then it would not be surprising if other parties felt less happy about a selection involving D, than one which does not.[19] One might furthermore object on grounds of 'double counting' to the inclusion of a proposal

[18] Just to give a very simple example: if the range of possible alternatives is the set of real numbers in the unit interval: $\mathfrak{D}_0 = [0, 1]$ and $\mathfrak{D}$ is a finite sub-set thereof, say $\mathfrak{D} = \{0.2, 0.8\}$, then π = 'take the arithmetic mean of the elements of $\mathfrak{D}$' is a selection procedure grounded in the selection base $\mathfrak{D}$ which – given that $\pi[\mathfrak{D}] = 0.5 \notin \mathfrak{D}$ – is clearly not a choice from $\mathfrak{D}$.

[19] Consider, for example, the proportional procedure of allocating divisible goods under conflicting claims: if, say, party 1 claims half of the cake (an inheritance, say) and party 2 claims all: d_1=1/2, d_2=1, then the cake is to be distributed in proportion to the claims: $d_1 : d_2$=1/3:2/3. There is no doubt in my mind that either party would perceive this solution to be fairer if both claims are known to be legitimate than if it was known that the claim of the other party is illegitimate (say based on a forged will). And this quite independently of what they think about the fairness of the procedure.

which is essentially the same as one already contained in the selection base.[20] And last, but not least, one might feel that if a distribution proposal D is based on an equity principle which in the given context is clearly morally inferior to another principle, then D should not be included in the selection base.[21]

This and the fact that the parties in international negotiations generally act as agents for collective bodies (nations, governments),[22] has lead me to conclude that the chances of overcoming our doomsday negotiating scenarios are enhanced – indeed, maximised as far as selection bases are concerned – by adopting what I shall refer to as an '*equitable*' selection base: $\mathfrak{E} \subset \mathfrak{D}_0$ including all (and only) the proposals put forward by the parties concerned,[23] provided (i) they can be justified by an equity principle, (ii) they do not give rise to double counting, and (iii) the principles involved are 'morally independent'. I am aware that, for practical purposes, this characterisation is in need of further elaboration. But since it is equally clear that the scale of such an undertaking would warrant a separate research project, this general description will have to suffice for the present purposes.

The property I believe to be of particular importance in mitigating sensitivities and the one I wish to focus on instead is the (perceived) *procedural* fairness of the involved selection procedures: if two procedures π and π' select the same outcome on the basis of the same selection base $\mathfrak{D} \subseteq \mathfrak{D}_0$: $D = \pi[\mathfrak{D}] = \pi'[\mathfrak{D}]$, and if k perceives π' to be fairer than π, then it stands to reason that k would associate less injustice with D

[20] If, for example, GNP is indeed causally tied to CO_2 emissions, then one might well object to including a distribution proposal which is justified in terms of grandfathering current GNP as well as one which is based on the grandfathering of current emission levels.
[21] Thus if we were in a morally complex yet unambiguous situation, only the proposal based on the morally dominant principle would be acceptable for inclusion in the selection base.
[22] If the parties were agents for themselves or other *individuals*, we might want to include proposals based on *intuitive* equity standards. But in the case of collective principals, we can reasonably restrict ourselves to the more easily recognisable proposals based on general principles of equity. Whether a particular proposal can be justified by a principle of equity may, in practice, be a matter of controversy, but this is infinitely easier to decide than whether it genuinely conforms to an ethical intuition of the proponent, and not merely to purely selfish motives.
[23] Note that this is also meant to imply that each and every one of the admitted proposals is seen by at least one of the parties in question as being the just solution (i.e. as being minimally unfair).

qua outcome of the fairer π' than if it had been selected by means of π: $\delta_k(\pi'[\mathfrak{D}],\ \pi')) \geq \delta_k(\pi[\mathfrak{D}], \pi)$.[24]

A simple way of incorporating this dependency into our elementary *ID*-model is by introducing *conditional* inequity-disutilities: $\delta_k^{\pi}(d_k)$,[25] designating the disutility which k would associate with the share d_k *if* it were selected by means of procedure π: $\delta_k^{\pi}(d_k) = \delta_k(d_k, \pi)$. Retaining the linearity assumption – i.e.

$$\delta_k^{\pi}(d_k) = \delta_k^{\pi'}(d_k) = 0, \text{ for } d_k \geq d_k^{(k)};$$
$$\delta_k^{\pi}(d_k) = (d_k^{(k)} - d_k)\tan\alpha_k, \text{ and}$$
$$\delta_k^{\pi'}(d_k) = (d_k^{(k)} - d_k)\tan\alpha'_k \text{ for } d_k < d_k^{(k)},\ 0 \leq \alpha_k, \alpha'_k < \pi/2.$$

– π and π' will involve sensitivity parameters with $\alpha'_k \leq \alpha_k$. This makes it self-evident (see Fig. 11) how our model manages to reflect the idea that, say, A's individual acceptability range: $\mathfrak{A}'_A$ determined by the fairer procedure (with α'_k) is an extension of the one: $\mathfrak{A}_A$ that would be determined by the other procedure (with α_k). And the greater these individual ranges, the greater, of course, the chances of a non-empty common acceptability range.[26]

[24] Note that an adopted equity-based distribution – i.e. $D^{(k)}$ in the case of k – is likely to be a fix-point also in the sense that it always has minimal disutility, regardless of the procedure involved: $\delta_k(D^{(k)}, \pi) = 0$ for all π.

[25] Given that *all* inequity disutilities are conditional in this sense, the use of this term should not be misunderstood as introducing a subdivision amongst them, but merely as stressing this conditional nature: there are no 'absolute' inequity disutilities.

[26] In terms of the *ID*-model, Burtraw and Toman's 'strategic emergence' idea translates roughly as follows: a strategic bargaining procedure π_s is applied to select a final proposal D_s. In the process of learning about the strategic limitations and strengths, $D^{(A)}$ and $D^{(B)}$ (here taken to be variables) converge on this strategic solution such that in the signing stage of the negotiations, we find ourselves in a situation very much akin to the one described in Fig. 9, namely one where $D^{(A)} = D^{(B)} = D_s$. My argument, in turn, was essentially (i) that '$D^{(A)}$' and '$D^{(B)}$' are not variables but proper names designating positions fixed by equity considerations, and (ii) that by adopting π_s as selection procedure, what is likely to happen is an increase in the inequity sensitivity of the strategically weaker parties, thus potentially even transforming situations with commonly acceptable outcomes at the outset into doomsday scenarios.

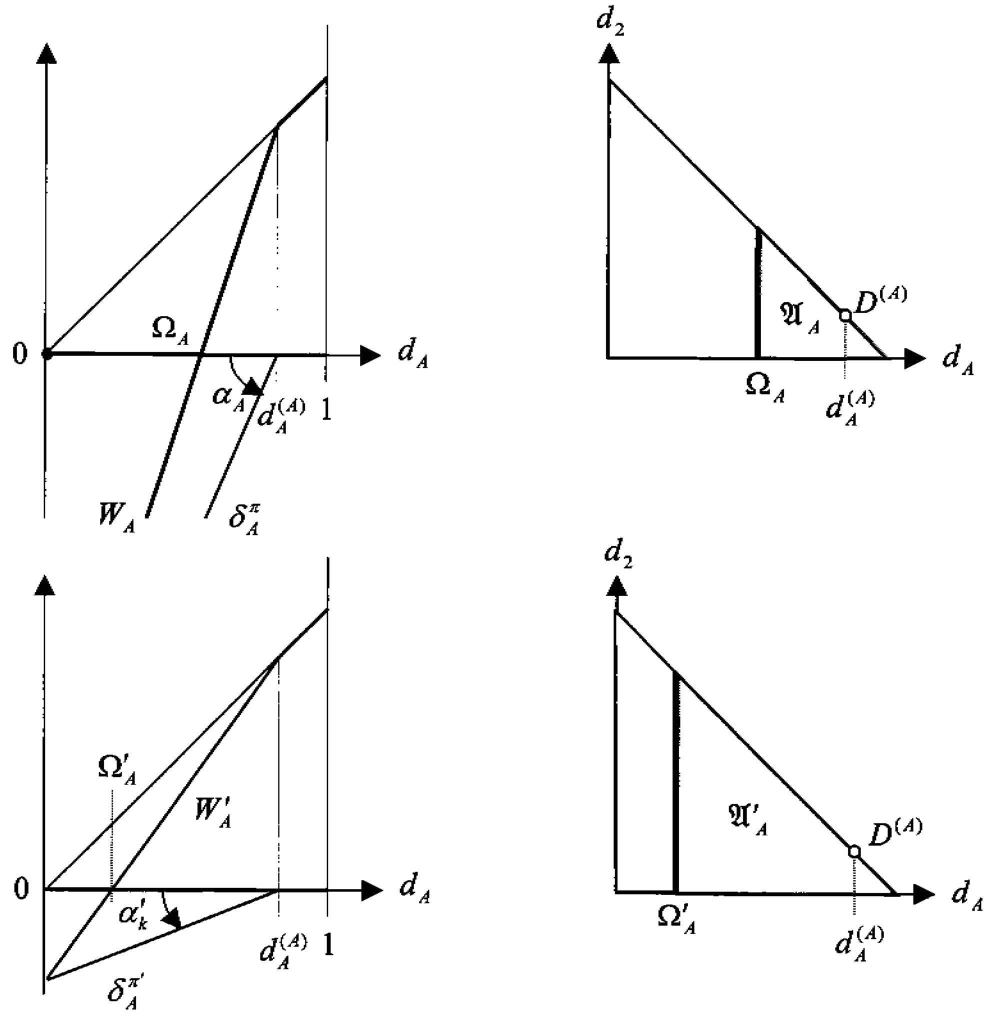


Figure 11: Increasing Procedural Fairness

Can we make any recommendation about a selection procedure π which would be *sufficiently fair* for our purposes, i.e. one which – particularly when based on $\mathfrak{E}$ – not only (i) mitigates any potential inequity sensitivities sufficiently so as to ensure commonly acceptable possible solutions: $\mathfrak{A} \neq \varnothing$, but (ii) actually selects one of them: $\pi[\mathfrak{E}] \in \mathfrak{A}$? I am afraid, all I will be able to do in the remaining pages of this study is to argue the case of a particular selection procedure which I believe has the potential of being sufficiently fair in this sense. The issue whether there might be others which are more likely to satisfy (i) and (ii) will simply have to remain open for future analysis.

In the search for a sufficiently fair selection procedure, we might reasonably be lead to consider the sort of mechanisms we referred to as 'choice procedures', i.e. selection procedures restricted to options from the selection base in question. Arguably the most tried and tested procedure of this type is the one where each party is out to convince the others about the moral superiority of their own position, thus trying to make the others see 'the error of their ways'. This may be successful in situations – such as Sen's 'bamboo-flute-scenarios'[27] – which are morally *un*ambiguous in that amongst the several relevant equity principles there is one which is morally superior to all the others. In morally ambiguous situations, where none of the competing equity principles can justifiably be said to morally dominate all others, it is unlikely that the individual negotiating agents (let alone their governments/parliaments) would be persuaded by this sort of 'reproach tactics'. Trying to apply this 'righteousness' approach to situations where the proposals collectively constitute an equitable selection base will only create a lot of ill feeling which may be hard to overcome, even if a different procedure is adopted at a later stage. After all, one of the ideas behind the conception of such equitable selection bases was that in morally ambiguous situations there may well be *different* equity-based solutions with the *same* moral legitimacy. Indeed, for the parties involved in such a situation to acknowledge this fact – and consequently to desist from accusing each other! – will, I shall argue, open up the possibility of a fair compromise.

Anyone sharing my belief that the context of our global emission allocation problem is indeed a morally ambiguous one will therefore have to search for a more appropriate selection procedure. Another type of choice mechanisms which might well be worth considering in this search are *voting* procedures used to (s)elect one of the 'candidates' put up in $\mathfrak{E}$. Take, for example, the *preference score method* originally (1781) recommended by Jean-Charles de Borda for use in elections to the French Royal Academy of Sciences. In its simplest form, the Borda procedure prescribes that each voter *k* is to order the set of 'candidates':[28] $\mathfrak{E} = \{ D^1, D^2, \ldots, D^l \}$ according to preference and to communicate this order in terms of an (order-preserving) scoring-function s_k from $\mathfrak{E}$ into the natural numbers $\{ 0, 1, 2, \ldots \}$, such that if D^m is least preferred by *k*, then $s_k(D^m) = 0$, if it is least preferred but one, then $s_k(D^m) = 1$ etc.[29] The Borda

[27] Recall that Sen's scenarios were found to be complex but unambiguous, unlike the Schelling-Barrett *Scenario A* which was genuinely morally simple (and thus *a fortiori* unambiguous).
[28] In order to simplify the notational requirements, let us for the time being assume that each party has adopted a different equity-based distribution.
[29] For simplicity's sake, let us henceforth assume that the order imposed by *k* is one of strict preference. If it were not, we would have to introduce some additional 'bracketing rules' – see, e.g., Dummett (1984).

selection then proceeds on the basis of these individual scores by adding up the scores given to each of the 'candidates': $\beta_k = \sum_m s_m(D^k)$, and selecting – with some tie-breaking provision – the one with the maximum 'Borda index', i.e. the maximum total score β_k.[30]

According to Dummett, the Borda procedure reflects 'how far the support for a candidate outweighs the opposition to him'.[31] It generates what Sen refers to as a social welfare functional.[32] As such it can lay claim to a degree of 'social legitimacy,' in particular given Dummett's convincing argument that it is the fairest possible voting mechanism in the sense of reflecting 'as accurately as possible the preferences of the voters'.[33]

So what could possibly be wrong with using this procedure to select one of the equity-based distribution proposals for our purposes? The one feature which I can see might be detrimental to its inequity mitigating capacity is something it actually shares with all $\mathfrak{C}$-*choice* procedures. While taking into account all the options in the selection base, the outcome will inevitably have an excluding, divisive character by creating outright losers and winners. In the context of elections proper, with the possibility of enforcing the outcome amongst those whose candidates have lost the election, this will normally create no problems. If, however, acceptance of the selection outcome cannot be enforced, then it would be unwise to ignore the possibility that those who lose out in a *choice* from the (equitable) selection base may feel excluded by the out-

[30] There is an extensive literature on voting procedures in general, and on the Borda rule in particular. Amongst the most noteworthy pieces are Dummett (1984) and the relevant papers in Sen (1982), in particular Sen (1977).
[31] Dummett (1984:177).
[32] That is to say a functional which (i) specifies exactly one social ordering of the selection base for any *n*-tuple of personal welfare functions, and which (ii) satisfies Arrow's well-known specification bar, obviously, the 'Independence of irrelevant Alternatives' (*IA*) [see Sen 1977]. The fact that the Borda functional does not satisfy *IA* is, as far as Dummett is concerned, of no real importance, for he maintains that *IA* 'lacks complete intuitive justification, since it conflicts with the more compelling principle that whether *x* would be a fairer outcome than *y* depends not only on how many (or which) voters prefer *x* to *y*, and how many prefer *y* to *x*, but on how strong their preferences are'(Dummett (1984):54).
[33] Dummett (1984:29).

come no matter how fair the procedure.[34] Indeed, the very fact that there are outright winners may increase the inequity sensitivity of some of the losers, thus increasing the probability of a breakdown. What we need to find are selection procedures which are more 'inclusive' from this point of view.

[34] Majority voting procedures, in particular, may create resentment in those parties which lose out in the chosen outcome, which is why they are often complemented with some minority protection measures. Note, incidentally, that the Borda rule is *not* a majority procedure.

III

NUMERICAL SELECTION PROCEDURES:

SOCIALLY WEIGHTED MIXED PROPOSALS

Simple-Claims Allocation vs Aggregate-Claims Allocation

In electing people, it is difficult to envisage how anything but a choice procedure could be employed. While one might possibly try some sort of time-share arrangement, one thing is clearly – some might say unfortunately – out of the question: to create a completely new ('mixed') candidate out of bits and pieces of the ones up for election. However, if the issue at hand is to find a commonly acceptable compromise between a number of 'candidate' distributions of a homogeneous divisible good, the situation changes markedly, for these 'candidates' are arrays of numbers which *can* be mathematically manipulated to produce new distributions as 'compromise candidates'.

Take, for example, what might be called the '*simple proportional procedure*' (π^P) where each claimant k – in conformity with Aristotle's proportionality principle – receives an amount proportional to his share of the total amount T under what *he* considers to be the fair distribution amongst the proposals in the underlying equitable selection base $\mathfrak{E} = \{ D^1, D^2, \ldots, D^l \}$, i.e. proportional to $d_k^{(k)}T$.[1] The outcomes of this procedure – under which the proportional distribution rule R^P is applied with reference to these 'simple' claims: $d_1^P : d_2^P : \cdots : d_n^P \ = d_1^{(1)} : d_2^{(2)} : \ldots : d_n^{(n)}$ – are distributions of the form:

$$D^P = \langle \tau^P d_1^{(1)}, \tau^P d_2^{(2)}, \ldots, \tau^P d_n^{(n)} \rangle,^2$$

and there is no question as to the manner in which they are constructed out of components of the distributions in the underlying selection base.

---

[1] Note that the superscript '(k)' in our notation can be interpreted as a function from the collection of parties involved: $\{1, 2, \ldots, n\}$ *onto* the class of indices: $\{1, 2, \ldots, l\}$ used to individuate the legitimate equity-based distribution proposals in the distribution matrix: $\mathfrak{E} = (D^1, D^2, \ldots, D^l)$, the implicit assumption being, of course, that each equity-based distribution which is being put forward is also seen by at least one of the parties as the just outcome.

[2] $\tau^P = 1/[\sum_{k=1}^{n} d_k^{(k)}]$.

The question is only whether these constructs could really serve as acceptable 'compromise candidates'? In approaching this question, we must remember that the Aristotelian proportionality rule is *not* the only allocation rule which could have been applied in the context of this simple second-order claims problem.[3] According to ancient Talmudic tradition, for example, the solution would generally be quite distinct from the Aristotelian compromise solution D^P.[4] There is, however, one feature that sets the proportional rule apart from all other claims allocation rules. The fact – proven by O'Neill (1982) – is that *it alone* (i) only depends on the particular claims and the total amount to be distributed, and (ii) is collusion proof, in the sense that a consolidation of the claims of several claimants into one claim will not change the total amount that these claimants receive.

> This result does not say that the proportional rule is completely immune from manipulation. For example, if the claims merely represent assertions by the claimants about how much they deserve, then under the proportional rule it is clearly desirable to inflate one's claim as much as possible. When the claims are both verifiable and transferable, however, it makes sense to use a rule that does not encourage the splitting or consolidation of claims among various groups of claimants. Under these circumstances the proportional rule is the most appropriate solution.[5]

If we are thus engaged in negotiating an allocation of, say, *tradable* (transferable) emissions permits, and if we follow our earlier recommendation to choose $\mathfrak{E}$ as selection base, then the proportional rule recommends itself for reasons which go beyond culture-relative precedent.[6]

In light of this, it would not be unreasonable to think that our simple proportional procedure might be able to generate acceptable compromise distributions. The predicament is that by focusing on a *single* claim per claimant, the procedure ignores a large proportion of the information given in our compromise problem, information

[3] That is to say the problem of how to distribute the total amount T if each has a legitimate claim given by his share under his preferred equity-based distribution.

[4] According to this tradition, we first establish each claimant's *uncontested* portion, i.e. (in the case of *two* claimants: A and B), $m_A = \max\{T - d_A^{(A)}T, 0\}$ for claimant A, and $m_B = \max\{T - d_B^{(B)}T, 0\}$ for claimant B. The Talmudic solution $D^T = (d_1^T, d_2^T)$ gives each claimant his uncontested portion plus half of the excess over and above the sum of uncontested portions: $d_A^T = (m_A + s/2)/T$ and $d_B^T = (m_B + s/2)/T$, where $s = T - (m_A + m_B)$. For a more detailed discussion of this 'contested garment rule' (and a generalisation for more than two claimants), see Young (1994:67ff).

[5] Young (1994:79).

[6] Note, incidentally, that all the allocation rules used in the Schelling-Barrett example (i.e. R^A, R^B, $R^{C'}$, and $R^{C''}$) are proportional rules, albeit involving different yard-sticks (population size, national cont.

which the parties might well regard as relevant to finding an acceptable solution. After all, why should the size of a claimant's allocation depend only on what *he* thinks is due to him, given that all the base proposals are meant to be (mutually recognised as) equally legitimate? Say we are dealing with three claimants, *A*, *B*, *C*, and two equity-based distribution proposals $\mathfrak{E} = \{ D^1, D^2 \}$, and say that *A* is allocated 45% under both proposals: $d_A^1 = d_A^2 = 45\%$, while the allocations to *B* (*C*) are $d_B^1 = 10\%$ ($d_C^1 = 45\%$) and $d_B^2 = 45\%$ ($d_C^2 = 10\%$), respectively. Assuming, as always, that the two distributions form an equitable selection base, *A* could clearly not be blamed if he thought that his fair share, even under a compromise, should be 45%. After all, each of the claimants is assumed to regard one of the distributions as the fairest one, and therefore it would be difficult for *B* or *C* to object to *A*'s 45% claim as being unfair. Yet if, which again does not seem implausible, *A* together with *C* prefer the first and *B* the second of the base-distributions, we have that $d_A^{(A)} = d_B^{(B)} = d_C^{(C)} = 45\%$, and consequently obtain a simple proportional distribution which allocates the disputed good in exactly equal shares, i.e. 33.⅓% each. *A* could therefore hardly be blamed for rejecting this compromise proposal as unacceptable, since everyone agrees that, in fairness, he should get more.

The input to our simple proportional procedure – i.e. the selection base (distribution matrix) $\mathfrak{E} = \{ D^1, D^2, \ldots, D^l \}$ – thus contains information which may be relevant to the acceptability of the generated compromise proposal, but which cannot be accommodated by this procedure. Indeed, it cannot be (directly) accommodated by any of the traditional allocation rules, insofar as they are 'simple' in the sense of relying on the assumption of there being a single (legitimate) claim for each of the claimants involved. The difficulty with our compromise problem lies in the fact that, in general, any claimant *k* will have a whole array of equally legitimate (usually) conflicting claims, given by the '*claims vector*': $\overline{C}_k = \langle d_k^1, d_k^2, \ldots, d_k^l \rangle$, i.e. the distribution matrix row for *k*. There are two natural ways in which one might proceed if one wishes to retain the advantages of the Aristotelian rule. One the one hand, one might simply tell the claimants to go away and make up their mind and apply the

wealth, etc.). It would therefore not be unreasonable to expect proponents of the resulting equity-based distributions to admit R^P in the second-order claims problem.

Aristotelian rule once they have done so. Since it can safely be assumed that (in most cases) the outcome would be the simple proportional distribution encountered above, this option is unlikely to be promising as far as acceptable compromise is concerned. The second option I have in mind is to try and aggregate each claims vector $\overline{C}_k$ mathematically into a single number (claim) $\mu(\overline{C}_k)$ in a 'sufficiently fair manner', which then allows us to adopt an *Aristotelian Aggregate-claims Allocation* rule $R^{P\mu}$, according to which:

$$\text{(AAA)} \qquad d_1^{P\mu} : d_2^{P\mu} : \cdots : d_n^{P\mu} \; = \mu(\overline{C}_1) : \mu(\overline{C}_2) : \ldots : \mu(\overline{C}_n).$$

In order to pre-empt objections, let me re-emphasise at this point our assumption that each and every of the options contained in an equitable selection base is meant to be the just solution for at least some of the involved parties. Given this, it is possible to justify the occurrence of *all* the options in the presupposed equitable selection base in our aggregate claims by our argument as to why each party's favoured proposal ought to be included.[7]

There are a number of ways which can, and have been used to bring about such aggregations (each of which corresponding to a type of mean value): Assuming $\overline{w} = \langle w_1, \ldots, w_l \rangle$ to be a *weighting* – i.e. an array of positive real numbers – we have, for example (for $m = 1 \ldots l$):

1) $\mu_a(\overline{w}, \overline{C}_k) = \sum_m w_m d_k^m$ — Weighted Arithmetic Mean
2) $\mu_g(\overline{w}, \overline{C}_k) = (\prod_m w_m d_k^m)^{1/l}$ — Weighted Geometric Mean
3) $\mu_h(\overline{w}, \overline{C}_k) = 1/(\sum_m 1/w_m d_k^m)$ — Weighted Harmonic Mean
4) $\mu_q(\overline{w}, \overline{C}_k) = (\sum_m (w_m d_k^m)^2)^{1/2}$ — Weighted Quadratic Mean

[7] What will have been excluded by this assumption are certain situations where an application of our AAA-rule would clearly lead to anomalous results. Just to give an example: If all the parties actually agree on what would be the just solution, say D^J then it would clearly be wrong to aggregate D^J with other proposals, even if these were also morally defensible. Thus in the Kyoto negotiations it would probably have been counter-productive to insist on introducing a mixture involving a per capita distribution, assuming that the parties required to take on commitments believed some sort of grandfathering to be the just solution.

In the natural sciences, the preferred way of aggregating a 'vector quantity' $\bar{V} = \langle V_x, V_y, V_z \rangle$ for the purpose of comparing magnitudes is by reference to its norm $\| \bar{V} \| = \mu_q(\bar{1}, \bar{V})$, where $\bar{1} = \langle 1, 1, \dots, 1 \rangle$ is the *trivial weighting*. A velocity $\bar{V}$, for example, is greater in magnitude than another one $\bar{V}'$ (both measured within the same frame of reference) if and only if the speed of the former is greater that that of the latter: $\| \bar{V} \| > \| \bar{V}' \|$. In the social sciences, the situation may well generally be more diverse,[8] but as far as greenhouse gas abatement proposals are concerned, attention is equally focused on a single aggregation method, namely the weighted *arithmetic* aggregation μ_a which, together with the proportional aggregate-claims allocation rule, gives rise to what has come to be known in the literature as 'mixed proposals'. Considering that using different aggregation functions will generally create different aggregate distributions for the same weightings and base distributions, the choice of μ_a for our purposes is clearly in need of some justification.

Before we turn to discuss this choice, it is important to stress that all procedures which might be considered for the purpose of generating an acceptable compromise between the relevant equity-based distribution proposals (i.e. the proposals in $\mathfrak{E}$) are, if we wish, 'second-order' distribution procedures. They differ from their 'first-order' counterparts used to generate justifiable equity-based distributions in several relevant ways, the most obvious of which being that, in contrast to the first-order cases, their 'input' is not a state of the world but the 'output' of first-order procedures. More importantly, given our assumption that all the elements of this 'input' are equally justified, these (second-order) compromise procedures should not discriminate between the proposals by appeal to parameters on 'first-order' grounds. To use an analogy: an adjudication between equally legitimate but conflicting claims should not discriminate between these claims on ground pertaining to their legitimacy. In other words, once the legitimacy of the claims is established, such adjudications should be impartial as to the different grounds for legitimacy.

[8] Indeed, in his discussion of real-income measures (indices), Samuelson (1974:597) refers to all the aggregation functions in our list, bar the quadratic one.

One way of ensuring the analogous impartiality for our second-order compromise proposals would be simply to prohibit their using the 'standards of comparison'[9] applied in our first-order procedures – such as base-line (or diachronic) population figures, base-line GDP, or base-line (diachronic) emission figures – and stipulate that compromise procedures should only make reference to properties and components of the relevant base-distributions.[10] However, this sort of regimentation – while effective in ensuring the said impartiality – would really be too restrictive. After all, the required impartiality is not endangered by, say, population figures *per se*, but only by population figures introduced on 'first-order grounds'. In other words, population data may well figure in a compromise procedure, but their use has to be justified in a manner other than by reference to the sort of justifications applicable in the context of the original distribution problem.[11] And the same goes for all the other first-order standards of comparison. This 'second-order caveat' will therefore have to be borne in mind in the following discussion.

Value Interpretations of Weighted Arithmetic Aggregate Distributions

Arguably one of the most prominent – certainly one of the most publicised[12] – of these mixed proposals is the one put forward by Michael Grubb and James Sebenius. Their allocation formula is 'simple' in the sense of involving just two equity-based distribution proposals, namely the per capita distribution D^B and the distribution proportional to current emissions, say D^C.[13] Using the weighted arithmetic aggregation function μ_a (with normalised weights: $\sum_m w_m = 1$), they employ the AAA-rule $R^{P\mu_a}$ which leads them to propose the simple aggregate ('mixed') distribution

[9] For a formal definition of this notion see Young (1994:76).

[10] Assuming that preferences between these base distributions by the parties in question can be loosely interpreted as properties of these distributions, our simple proportional compromise procedure would satisfy this regimentation.

[11] In particular, it would not be legitimate to demand that compromise allocations should be weighted according to, say, emissions (population) figures *on the grounds* that grandfathering (per capita) is the fair allocation principle. Nota bene: this is not to say that (as we shall see below) some form of population weighting could not legitimately be introduced, but merely that it has to be on other grounds than the ones pertinent to the original distribution problem.

[12] See, for example, Grubb and Sebenius (1992), Grubb *et al.* (1992), Banuri *et al.* (1996).

[13] Since GNP may be regarded as a rough proxy for actual emissions, a substitute for D^C could indeed by introduced in *Scenario C* of the Barrett-Schelling example, namely the distribution in proportion to GNP.

$D^* = \{d_1^*, d_2^*, \dots, d_n^*\}$ given by: $d_k^* = w_B d_k^B + w_C d_k^C$, or, to conform with their own formulation:

$$d_k^* = w_B d_k^B + (1 - w_B)\, d_k^C .$$

In general, a *WAA* (*'weighted arithmetic aggregate'*) *distribution* D^* – generated by the use of (AAA) with $\mu = \mu_a$ – is given by

$$\text{(WAA)} \qquad d_k^* = \tau^* \mu_a(\overline{w}, \overline{C}_k) = \tau^* \sum_m w_m d_k^m .$$[14]

So how could the use of this sort of aggregate distribution in the context of our claims problem be justified? One way in which this might be achieved is by way of an analogy with the traditional theoretical conception of an 'allocation':[15] whereas our claims vectors represent different claims (by some person) on a *single* good, allocations in this traditional technical sense are arrays of some person's claims for *different goods*, one per claim. Thus if we were to assume (i) that $\boldsymbol{T}$ is partitioned into equal parts, one for each of the distributions in $\mathfrak{C} = \{D^1, D^2, \dots, D^l\}$: $\boldsymbol{T} = \bigcup \boldsymbol{T}^m$ and $|\boldsymbol{T}^m| = T/l$, and (ii) that D^m pertains exclusively to $\boldsymbol{T}^m$,[16] we would no longer have a claims problem, for k would then indisputably receive $\sum_m d_k^m T/l$ units of $\boldsymbol{T}$. As this is precisely his share $d_k^*(\overline{w}, \overline{C}_k)\, T$ under the *pure* arithmetic aggregate distribution with $\overline{w} = \overline{1}$, it stands to reason that the arithmetic aggregation function is indeed the appropriate choice in this context, at least if we are dealing with trivial weights. This interpretation of a party's aggregate share has the added advantage of providing a clear idea of the way in which the underlying equity-based distributions can be seen as *being part of* the resulting aggregate distribution; something which, as I maintained earlier, is important as far as the mitigation of inequity sensitivities is concerned. But what about the case of *WAA*-aggregations with non-trivial weights?

[14] With $\overline{w} = \langle w_1, \dots, w_n \rangle$ and $\tau^* = 1/\sum_m w_m$. Note that for the trivial weighting $\overline{w} = \overline{1}$, we have that $\tau^* = 1/\sum_{m=1}^{l} 1 = 1/l$.

[15] See Young (1994): Appendix 1.

[16] Note that this initial assumption amounts to an initial egalitarian treatment between the alternative base distributions.

There is, as far as I can envisage, no justification if we retain our present conception of weightings as arrays of pure numbers. However, if we decide to give these numbers an ('empirical') *interpretation*, then it becomes possible to extend our justification to the case of non-trivial weightings. What I have in mind, in particular, is an interpretation of our weights as some sort of *values* (per unit of T) which, in turn, will allow us to interpret the quantities allocated in the above-mentioned manner as differently *valued* quantities.

The basic idea behind this 'valuation strategy' emerges if our numerical weights are interpreted as different *prices* p_m (in some common currency unit), one for each of the segments T^m of our partition, since under such a monetary interpretation it is clearly defined what the 'total value' of the amounts allocated to k in the above-mentioned allocation model is meant to be, namely the total *Monetary Value* $MV_k(\overline{p}, \overline{C}_k) = \sum_m p_m d_k^m T / l$.[17] Moreover, it is not difficult to see that k's aggregated claim $d_k^*(\overline{w}, \overline{C}_k) \cdot T$ – when valued at the average price $\hat{p} = (1/l) \cdot \sum p_m$ – is precisely the same as this total monetary value:

$$d_k^*(\overline{p}, \overline{C}_k) \cdot T \cdot \hat{p} = MV_k(\overline{p}, \overline{C}_k).$$

Needless to say that, in the absence of a 'price setting formula' or 'mechanism', this monetary interpretation is useful only for conceptual purposes. If we wish to proceed towards an interpretation which can actually be applied, we will have to focus our attention on the issue of how these unit values could be fixed in a 'sufficiently fair' manner for them to be acceptable as aggregation weights. The reason for using 'unit *value*', instead of '*price*' in this context is simply that I cannot envisage how this could be done in monetary terms. My proposal instead is to consider an interpretation in terms of certain 'social values'.

[17] With $\overline{p} = \langle p_1, \dots, p_l \rangle$.

Borda Indices as Aggregation Weights

The idea is to make use of the possibility to interpret the *Borda index* β_m – i.e. the sum total of preference scores given to a base-distribution D^m under the Borda rule – as a measure of the 'social desirability' of D^m amongst the negotiating Parties. Given the one-one correlation between these base distributions and the partitions of our interpretative allocation model, this measure is then 'transferred' to the units distributed under D^m in this model. Applying these 'social unit values' in the context of the conception introduced in the preceding section then leaves us with a total *Social Value* of the amounts allocated to a claimant k in our interpretative allocation model of $SV_k(\overline{\beta}, \overline{C}_k) = \sum_m \beta_m d_k^m T / l$. Given that $\sum_m \beta_m = n \cdot \sum_m \sigma_{m-1}$ we consequently find this total social value to be the same as k's *WWA*-share multiplied by n (the number of claimants) times the index base average:

$$d_k^* \cdot T \cdot n\hat{\sigma} \quad = \quad SV_k .$$

It goes without saying that this equation has to be taken with some caution. For one, it is questionable whether and to what extent these 'social values' can be meaningfully manipulated in the same manner as monetary values. Indeed, there are good reasons to believe that they are really not meaningful as absolute quantities. The most we ought to conclude from the correlation established above is that our *WWA*-allocations are proportional to these total social values:

$$d_1^* : d_2^* : \ldots : d_n^* = SV_1 : SV_2 : \ldots : SV_n .$$

Provided that this 'proportional' use of total social values is itself coherent, this conclusion should, however, be sufficient to justify the use of Borda indices as aggregation weights on equity grounds. In light of their central role in these constructions, it may at this stage be useful to take a somewhat closer look at Borda indices, and indeed at the Borda rule itself.

Borda Functionals. It was mentioned above that the Borda rule generates a Social Welfare Functional (*SWFL*), say *BF*, within the ordinal non-comparability (*ONC*) framework. In other words, it determines (within the *ONC*-framework) a particular 'social' ordering $\prec$ of the relevant set of social states – our equitable selection base

$\mathfrak{E} = \{D^1, D^2, \ldots, D^l\}$ – for any '*welfare situation*', i.e. any n-tuple $\overline{W} = \langle W_1, \ldots, W_n \rangle$ of real valued personal welfare functions, each defined over $\mathfrak{E}$.

This is achieved, first and foremost, by what might be called a '*Preference Score Functional*' (*PS*) which, for any given welfare situation $\overline{W}$, determines an n-tuple $\overline{s}$ of *preference score functions* (on $\mathfrak{E}$); i.e. an n-tuple of (order preserving) functions s_k from $\mathfrak{E}$ into a chosen '*index base*' $\overline{\sigma} = \langle \sigma_0, \sigma_1 \ldots \sigma_{l-1} \rangle$, each of which constructed in accordance with the Borda rule. In a second step, these preference score functions are used to create the Borda index for each of the distributions (the 'social states') in $\mathfrak{E}$. Given any n-tuple of preference score functions $\overline{s}$, a '*Summation Functional*' (*SF*) generates an l-tuple $\overline{\beta} = \langle \beta_1, \ldots, \beta_l \rangle$ of cumulative Borda scores by summing up the different preference scores for each of the alternative base distributions D^m, thus generating its *Borda-index*: $\beta_m = \beta(D^m) = \sum_k s_k(D^m)$. Finally, the Borda ordering is determined in terms of the natural ordering $<_{\overline{\sigma}}$ of the index base: $x_i \prec x_j$ iff$_{df}$ $\beta_i <_{\overline{\sigma}} \beta_j$, a process which, in function terms, can be described as $\Omega(\overline{\beta}) = \prec$.

The sole reason why the functional $BF = \Omega \circ SF \circ PS$ is a *SWFL* – in the *ONC* framework, where welfare situations are taken to be the same: $\overline{W} \approx_{ONC} \overline{W}'$ iff$_{df}$ there is an n-tuple $\langle \Psi_k \rangle$ of positive monotonic transformations, such that $W_k' = \Psi_k(W_k)$ for all k – is the nature of the underlying preference score functions: the Borda rule is defined in terms of preference *orderings* between the social states and, *given an index base*, it will determine a specific preference score attribution for any preference ordering between the social states in question. In other words, we have that if $\overline{W} \approx_{ONC} \overline{W}'$ then $PS(\overline{W}) = PS(\overline{W}')$. This, of course, is sufficient to ensure that *BF* satisfies Sen's *ONC-Invariance Requirement* for *SWFL*s, according to which any $\overline{W}$ and $\overline{W}'$ with $\overline{W}' \approx_{ONC} \overline{W}$ must give rise to the same social ordering: $BF(\overline{W}') = BF(\overline{W})$.

A point worth noting in this context is that the 'determinacy' of the Borda rule – arising from the prescription of a specific set of numbers for the preference scores (the 'index base') – which is sometimes criticised as an unnecessary restriction, is actually a *sine qua non* within the *ONC*-framework. Consider, for example, the alternative rule, under which the scores can be chosen at will, as long as they reflect the underlying preferences, and assume we are dealing with two individuals *A* and *B*, and three base distributions D^1, D^2, D^3 – which *A* prefers in ascending and *B* in descending order. Under this alternative rule, both parties would clearly have correctly revealed their preferences by choosing the preference scores $s_A(D^1) = 0$, $s_A(D^2) = 1$, $s_A(D^3) = 2$, and $s_B(D^1) = 2$, $s_B(D^2) = 1$, $s_B(D^3) = 0$, respectively. But they could equally well have chosen, say, $s'_A(D^1) = 0$, $s'_A(D^2) = 1$, $s'_A(D^3) = 4$, and $s'_B(D^1) = 4$, $s'_B(D^2) = 1$, $s'_B(D^3) = 0$. Under the assumptions of the *ONC* framework, there are no reasons for discriminating between these two contexts and consequently both should give rise to the same social welfare ordering. The problem is that if we follow the Borda rule by using the Summation Functional *SF* to aggregate these scores, we find that $\bar{s}$ and $\bar{s}'$ generate *different* orders between the three base distributions, namely $D^1 \sim D^2 \sim D^3$ and $D^1 \sim D^3 \succ D^2$, respectively.[18]

The example just discussed is also useful in revealing certain connections arising from applications of the Borda rule itself. After all, the said differences between the preference score functions would also occur under the Borda rule if one were to switch from the 'natural' index base $\bar{\sigma} = \langle\, 0, 1, 2 \,\rangle$ to $\bar{\sigma}' = \langle\, 0, 1, 4 \,\rangle$. Accordingly it is really misleading to talk of '*the* Borda functional' since we can only speak of functional relationships, in this context, if we divide the relation between welfare situations and social orderings established by the Borda rule according to the index bases used. This leaves us with a *plurality* of Borda functionals, distinguished by the chosen index base, a distinction to be reflected by way of an index-notation: $BF_{\bar{\sigma}}$, $BF_{\bar{\sigma}'}$. We should, however, not jump to the conclusion that a difference in index base inevitably signifies a difference in Borda functional. After all, we know that for $\bar{\sigma}' = c\bar{\sigma}$ (with

[18] As this point would be equally true if our social ordering procedure were based on, say, multiplication as aggregation function – as opposed to the Borda rule's summations – it stands to reason that within the *ONC*-framework the choice of a pre-determined index-base for preference scores is a pre-requisite for arithmetical social ordering procedure (based on preference scores) in general. The use of such a pre-determined set of possible scores thus cannot be objected to on *ONC* grounds. The only real *ONC*-objection I can envisage could be against a justification of the summation functional on grounds of reflecting 'strengths of preferences'.

$c > 0$),[19] the Borda indices based on $\overline{\sigma}'$ will be c times the ones based on $\overline{\sigma}$, i.e. $(\forall \overline{W})\, BW_{\overline{\sigma}'}(\overline{W}) = c \cdot BW_{\overline{\sigma}}(\overline{W})$, and consequently that the orderings they generate will be the same: $(\forall \overline{W})\, BF_{\overline{\sigma}'}(\overline{W}) = BF_{\overline{\sigma}}(\overline{W})$. Since this, in turn, is precisely what we mean by $BF_{\overline{\sigma}'} = BF_{\overline{\sigma}}$, it now stands to reason that – in addition to the equivalence between welfare situation ($\overline{W}' \approx_{ONC} \overline{W}$) – Borda functionals involve a second fundamental equivalence, namely an 'essential sameness' between index bases given by:

$$\overline{\sigma}' \approx_{BF} \overline{\sigma} \text{ iff } BF_{\overline{\sigma}'} = BF_{\overline{\sigma}}.$$

The practical significance of this relation is that, as concerns creating a *social welfare ordering* by way of the Borda rule, it is quite irrelevant which of the two index bases one employs. The use of such 'Borda orderings' would obviously be much simpler if all index bases were equivalent in this sense, for this would dispense one from having to justify a choice of index base. As things stand, the world is not so simple, but at least it is not as complex as it would have been if Borda functionals had turned out to be different for each and every index base.

Our use of Borda weightings in *WWA-distributions* gives rise to an analogous equivalence between index bases, defined in terms of whether they generate the same '(Borda) Preference Score Distribution' for any given welfare situation and any selection base, i.e.

$$\overline{\sigma} \approx_{PSD} \overline{\sigma}' \text{ iff } (\forall k)(\forall \mathfrak{E})(\forall \overline{W})\, d_k^*(BW_{\overline{\sigma}}(\overline{W}), \mathfrak{E}) = d_k^*(BW_{\overline{\sigma}'}(\overline{W}), \mathfrak{E}).$$

The unfortunate fact is that this equivalence too fails to be universal, which is why we shall shortly have to turn to the question of justifying a choice of base distribution. But before turning to this point, let me raise a point of interest for further investigation. Both of the mentioned equivalence relations between index bases are defined in terms of the sameness of the generated construct, i.e. social orderings, on the one hand, and *WWA*-distributions, on the other. They are, if we wish, 'product specific' and could be treated quite independently of one another. However, since

[19] $c\overline{\sigma} = \langle c\sigma_i \rangle$ is the 'similarity transform' of $\overline{\sigma} = \langle \sigma_i \rangle$.

both types of constructions crucially involve a use of Borda indices, there might be links between the two procedures which might prove to be useful, in particular, to the quest for a justification of Borda indices as weights in *WWA*-distributions. Specifically, there might be reasons for inferring that if two Borda weightings give rise to the same social welfare ordering, then they should also give rise to the same preference score distribution. Taken to its logical conclusion, such an inference would generate a (necessary) condition for the adequacy of our socially weighted distribution method: if two index bases are the same as concerns welfare orderings: $\bar{\sigma} \approx_{BF} \bar{\sigma}'$ then they must also be equivalent in their distributive use: $\bar{\sigma} \approx_{PSD} \bar{\sigma}'$. Whether or not this is the case, or indeed required, is an issue which will have to be left for further investigation.[20] One connection between the two relations, however, can easily be ascertained, namely that both of them identify similarity transforms, meaning that for any index base $\bar{\sigma}$ and any number $c > 0$ both $\bar{\sigma} \approx_{BF} c\bar{\sigma}$ and $\bar{\sigma} \approx_{PSD} c\bar{\sigma}$. In other words, regardless of whether we are engaged in constructing a social welfare ordering or a preference score distribution, any index base will be as good as any of its similarity transforms.

Choosing the Index Base. The fact that preference score distributions can change solely due to a change of index base means that the choice of index base becomes a non-trivial matter which has to be justified. Given the envisaged purpose of these distributions, the obvious 'parameter' for such a justification would seem to be the overall fairness of the resulting procedure. In other words, the obvious approach to the 'index base problem' is to ask ourselves: is it possible to identify an index base which would commend itself on grounds of the fairness of the resulting distributions?

As far as generating Borda welfare orderings is concerned, Michael Dummett clearly believes this to be the case for the 'natural' index base $\bar{\sigma}_N = \{0, 1, 2, \ldots, l-1\}$, and it is illuminating to consider his reasons for this choice. The first one, focusing on the potential for strategic manipulations of the procedure, is summarised in the following quotation:

[20] If it were the case, then one might be able to justify our preference score distributions independently of our 'value interpretation' in terms of the justification for using Borda orderings (assuming, of course, that such a justification exists).

> The preference score procedure lends itself easily to adaptation when the voters desire to skew the scores in one direction or the reverse, which they are particularly likely to do – misguidedly, in my opinion – when electing a candidate to a post. If they agree in preferring a candidate with strong support, though strong opposition, to one generally ranked as rather better than middling, they can give extra weight to a candidate's having a high place on a voter's list; for example, when there are seven candidates, by assigning scores of 9, 6, 4, 3, 2, 1, and 0 points to the candidates on a voter's list. If, conversely, they agree in preferring a generally acceptable candidate to one who rouses strong opposition, they can skew the scores in the reverse direction, say by assigning the scores 9, 8, 7, 5, 2, 1, and 0. The more the preference scores are skewed, the greater will be the temptation to strategic voting, but those who wish to skew them may accept this as the lesser evil; anyone who understands that there can be no ideal voting procedure is prepared to choose between evils.[21]

The objection against the use of skewed index bases put forward here is thus based on the idea that an increase in the potential for strategic voting amounts to a decrease in (perceived) procedural fairness. Indeed, in conversation, Dummett has used the same line to object to preference score methods without a fixed index base. His view – in my understanding – is that *if* we could assume that voters did assign numbers of their own choosing, say between 0 and 10, which reflected the strength of their preferences as truthfully as they could be reflected, then we would have a method which is superior to the Borda rule. However, since the premise is unlikely ever to be satisfied, this method has to be rejected in view of its unlimited potential for strategic manipulation.

Yet if we do reject this sort of individual indexation, and instead adopt a predetermined index base, then the adoption of an unskewed index basis, Dummett contends, becomes a matter of fairness for egalitarian reasons, and not just due to a trade-off with strategic potential. An index base like $\langle 0, 1, 10 \rangle$ may reflect the strengths of preference of, say A, more accurately than the unskewed base $\langle 0, 5, 10 \rangle$. But why should A be given this sort of preferential treatment over someone whose preference strengths are better reflected by an index base like $\langle 0, 9, 10 \rangle$? The choice of unskewed index bases as fairest option can thus be grounded by a 'veil of ignorance' argument.

While this reasoning is partly based in an informational framework which presumably would have to be categorised as '*cardinal* non-comparable' – a fact which might

[21] Dummett (1984:241).

worry adherents of the *ONC* creed[22] – it is important to emphasise that Dummett's initial 'strategic manipulation argument' is quite independent of preference measurement questions, and it is, in my view, quite sufficient to justify the use of 'unskewed' index bases. As a matter of fact, the need to restrict the possibilities of strategic manoeuvring in order to increase procedural fairness arises in yet another context, namely the choice of who is meant to be a party to the procedures in question.

[22] Dummett's use of preference 'strengths' is incompatible with a purely ordinal conception of preferences. However, it is not self-evident that his arguments rely, if we wish, on a 'crude' cardinal conception of these strengths according to which they are measured in terms of a preference strength unit such as the (in)famous 'util'. In particular, Dummett's arguments would seem to remain valid if preference strengths were merely 'affine quantities,' reflecting only proportions and not absolute measures. This might be able to placate at least those who object to the cardinal conception of preferences on grounds of there not being a meaningful concatenation operation (the assumption being, of course, that these 'affine quantities' are indeed coherent in the absence of such an operation).

Choosing the 'Constituency': Simple National or Global Preference Scores

Consider the following hypothetical scenario (Scenario 1): The Parties to the FCCC will come together in 2010 in order to decide on a distribution (formula) for global carbon dioxide emissions, to be applied in a commitment period after the one adopted at Kyoto. Let us assume, for argument's sake, they decide to use the preference score procedure on the bases of the Per Capita and the (emissions) Grandfathering distribution, both with, say, 1995 base-lines. In order to carry this scenario further, we need to make some assumptions about the relevant preferences of the Parties. Given the figures listed in Table 1, it is not implausible to assume that Annex I Parties would give preference to Grandfathering, while non-Annex I Parties would prefer the Per Capita solution. Accordingly, the United States would therefore give a score of one to Grandfathering and a zero-score to the Per Capita option; and each of the 15 EU member states and the 22 remaining Annex I Parties would do the same. China, India, and the other 126 non-Annex I Parties to the Convention would do the reverse. This would leave a mixture of approximately ¼ Grandfathering and ¾ Per Capita as Preference Score Distribution (Table 2).

Table 1: Scenario Parameters

	Population			*CO_2 emissions*		*Per Capita Allocations*	*Grandfathering Allocations*
	2010[a]	*1995*		*1995*			
	million	million	%	Gt	%	%	%
Annex I							
USA	304	271	*4.8*	5.1	*24.5*	*4.8*	*24.5*
EU	378	374	*6.6*	3.0	*14.5*	6.6	14.5
Rest	589	578	*10.2*	5.3	*25.3*	10.2	25.3
non-Annex I							
India	1347	929	*16.3*	0.9	*4.1*	16.3	4.1
China	1152	1205	*21.2*	2.8	*13.3*	21.2	13.3
Rest	3121	2330	*41.0*	3.8	*18.2*	41.0	18.2
Total	6891	5687	*100.0*	21.0	*100.0*	100.0	100.0

[a] UN medium estimate.

Now consider an alternative scenario (Scenario 2), which is exactly the same as the one above, bar the fact that the United States have managed to be counted as a regional grouping of 50 separate Parties (i.e. its 50 States). After all, the US delegation could argue, say, that since they represent almost the same number of people as the EU, it is unfair that the EU's preferences should be given 15 times more weight than their own. Assuming these new Parties to have the same preferences as the Federal US government means that Grandfathering will now get 49 more scores than in Scenario 1, thus increasing its preference score weight from 0.23 to 0.4 (see Table 3). More importantly, the USA (*qua* region) will have benefited from this 'regionalisation'. Indeed, all the Parties of the same 'persuasion' as the USA (i.e. Annex I) benefit at the expense of those of the opposing view (non-Annex I). The preference score method with 'single national' scores is therefore clearly *not* collusion proof, at least not in the strong technical sense.[23]

Table 2: Single National Preference Scores. Scenario 1

	Preference Scores		*Preference Score Distribution*	
	Grandfathering	*Per Capita*		%
USA	1	0	*USA*	9.3
EU (15 Parties)[a]	15	0	*EU*	8.4
Rest Annex I (22 Parties)	22	0	*Rest Annex I*	13.6
India	0	1	*India*	13.5
China	0	1	*China*	19.4
Rest non-Annex I (126 Parties)	0	126	*Rest non-Annex I*	35.8
Total Preference Scores	38	128		
Preference Score Weights	**0.23**	**0.77**		

Table 3: Single National Preference Scores. Scenario 2

	Preference Scores		*Preference Score Distribution*	
	Grandfathering	*Per Capita*		%
USA (50 Parties)	50	0	*USA*	12.7
EU (15 Parties)	15	0	*EU*	9.8
Rest A.I (22 Parties)	22	0	*Rest Annex I*	16.3
India	0	1	*India*	11.4
China	0	1	*China*	18.0
Rest non-A.I (126Parties)	0	126	*Rest non-Annex I*	31.8
Total Preference Scores	87	128		
Preference Score Weights	**0.40**	**0.60**		

[a] Strictly speaking, the EU should be counted as 16 Parties, namely the 15 member states + the EU itself. But this (politically motivated) anomaly cannot be admitted in our preference score procedures.

If – as seems plausible – this sort of 'representative equity' argument is indeed tenable, then clearly China could equally demand a recognition of its 23 provinces. Given that at these negotiations China represents 34 thousand times the population of Liechtenstein, it could not be condemned if it were to insist on a further-reaching regionalisation. This opens up the question what would happen if all Parties were allowed to regionalise to their hearts' content? The first thing to be noted is that even though the term used to describe this process of splitting Parties has geographical connotations, its limits have to be demographic if the representative equity argument is to retain its validity. In other words, the division has to be essentially a partition of the

[23] Given that Parties will generally not benefit from colluding in the sense of forming a union, it is at least 'weakly' collusion proof, in the sense that this kind of collusion is unlikely to be attempted.

(represented) population, which means that the limit of these regionalisations is reached when each individual inhabitant is identified as a regional unit.

Table 4: Scenario 3 Parameters

Regions / Parties	*95 Population (million)*[a] people	n%[b]	*95 CO_2 emissions (Gt CO_2)* t CO_2	n%	*Grandfathering (%)*	n%	*Per Capita (%)*	n%
USA (304m Parties)	*(271)*		*(5.1)*		*(24.5)*		*(4.8)*	
SAM	271/304		5.1 Gt/304 m	*80.5*		80.5		15.7
JANE	0.89	*15.7*	16.9	*80.5*		80.5		15.7
EU (378m)	*(374)*		*(3.0)*		*(14.5)*		*(6.6)*	
ISABELLE	374/378		3.0 Gt/378 m	*38.4*		38.4		17.4
WOLFRAM	0.99	*17.4*	8.1	*38.4*		38.4		17.4
Rest of Annex I (589m)	*(578)*		*(5.3)*		*(25.3)*		*(10.2)*	
BORIS	0.98	*17.3*	9.0	*43.0*		43.0		17.3
India (1152m)	*(929)*		*(0.9)*		*(4.1)*		*(16.3)*	
ANISHA	0.81	*14.2*	0.8	*3.6*		3.6		14.2
China (1347m)	*(1205)*		*(2.8)*		*(13.3)*		*(21.2)*	
LI	0.89	*15.7*	2.1	*9.9*		9.9		15.7
Rest of non-Annex I (3121m)	*(2330)*		*(3.8)*		*(18.2)*		*(41.0)*	
JUAN	0.75	*13.1*	1.2	*5.8*		5.8		13.1
Total *(6891m)*	*(5687)*	*10^{11}*	*(21.0)*	*10^{11}*	*(100)*	10^{11}	*(100)*	10^{11}

[a] Figures in brackets are regional aggregates. [b] n% = nano percent (100% = 10^{11}n%).

Consider thus Scenario 3 under which all Parties to the FCCC have managed to gain recognition as groupings of this extreme type of regionalisation. India, for example, will then be a regional grouping of 1.15 billion Parties, such as ANISHA (see Table 4), a Party 'inhabited' in 2010 by exactly one person, namely Anisha. Given that in this case all the Parties would have exactly the same population (namely 1) there could no longer be any justifiable objections to the use of the egalitarian single national preference scores. Before we can contemplate even a hypothetical application of our preference score procedure in this maximally regionalised context, we must make some assumptions concerning the underlying base-distributions. While it may in theory be possible to define these base-distributions with intra-regional differentiations, it is for the present purposes quite sufficient to treat the Parties of a region as equals. In other words, in order to arrive at a 1995 emission figure for, say again, ANISHA, we shall simply divide the 1995 emissions of India (0.9Gt) by the number of Parties in India (1.15b), which leaves us with an absolute figure of 0.8t, or 3.6n% of the 1995 world total. And the same procedure sets ANISHA's portion of the 1995 world population at 15.7n%. Assuming that ANISHA, together with all the other Indian Parties, prefer getting more rather than less, the Per Capita distribution would hence receive a total of 1.15b scores from the Indian region alone (while Grandfathering would receive zero). Under analogous assumptions for all the regions (Table 5), we are thus left with preference score weights of 0.82 for the Per Capita option and 0.18 for Grandfathering.

Table 5: Single National Preference Scores. Scenario 3

Region	*Regional Totals of Preference Scores (million)*		*Preference Score Distribution for Regional Parties*	*Aggregate Per Region*
	Grandf.	*Per Capita*	n%	%
USA (304m Parties)	304	0	27.6	8.4
EU (378m)	378	0	21.3	8.0
Rest of Annex I (589m)	589	0	22.0	13.0
India (1152m)	0	1152	12.2	14.1
China (1347m)	0	1347	14.7	19.7
Rest of non-Annex I (3121m)	0	3121	11.8	36.8
Total *(6891m)*	1271	5620	10^{11}	100.0
Preference Score Weights	**0.18**	**0.82**		

While the single national preference score procedure in this scenario could not be blamed for reasons of representative inequity, the scenario itself is obviously completely impracticable. Nonetheless, it is useful if only because it points towards, and indeed justifies, a modification to our preference score procedure making it equally immune to objections on grounds of representational inequity. All we need to do is to switch from single national scores to 'global' preference scores, under which each Party is allowed to multiply its scores by the number of individuals it represents. From our assumptions about the preferences of Scenario 3 Parties within a given region, it follows that the global preference score weighting by our initial (Scenario 1) Parties has to be the same as the one arrived at in Scenario 3. Moreover, we find that under this global preference score procedure, the Parties of our first scenario will receive precisely the sum total of shares which they are allocated as regions under Scenario 3 (Table 6). Indeed, the global preference score method not only recommends itself by virtue of its representational equity, but also because it turns out to be strictly collusion proof, at least for those cases where the underlying first-order differentiation parameters are additive.[24]

[24] The additivity of say the population size parameter is given by the fact that the population of the sum of two countries is equal to the sum of their populations. The same is obviously true for emissions, but not for, say, emission intensities (emissions per GDP): The emissions of, say, the OECD is the sum of emissions of its member countries, but the OECD emission intensity is not the sum of its member states' intensities. (See also Appendix 1.)

Table 6: Global (2010) Preference Scores.

	Preference Scores		*Preference Score Distribution*	
	Grandfathering	*Per Capita*		*%*
USA	304	0	*USA*	8.4
EU	378	0	*EU*	8.0
Rest Annex I	589	0	*Rest Annex I*	13.0
India	0	1152	*India*	14.1
China	0	1347	*China*	19.7
Rest non-Annex I	0	3121	*Rest non-Annex I*	36.8
Total Preference Scores	1271	5620		
Preference Score Weights	**0.18**	**0.82**		

The reference to 'first-order differentiation' is a useful reminder that something needs to be said about the proposed use of global preference scores with respect to our previous 'second-order caveat'. After all, population figures are paradigmatically used as first-order differentiation parameters. Are we therefore entitled to use them again as 'demographic weights' for the preference scores of the Parties involved? The reason why our use of population figures does not endanger what we referred to as the 'impartiality' of the resulting global preference score procedure is that this use is motivated by considerations of representative equity which are quite independent of the first-order question as to what would be a fair allocation of emission quotas. The difference between these two contexts becomes particularly perspicuous if we consider the potential of using inter-temporal aggregate population figures. It is well-known that in the first-order context such inter-temporal aggregates have been used, in particular under the so-called 'historic per capita' proposals. To argue that such inter-temporal figures should be applied in order to rectify representative inequities, however, would be unconventional, if not absurd (in particular if the representative mandate is taken to be given through an electoral process).[25]

In sum, as far as procedural equity is concerned, global preference scores are preferable over their single national counterparts for reasons of both representative fairness and avoidance of potential strategic manipulations.

[25] It may just be conceivable that the domain of representation could be extended beyond the present generation (the electorate) to future generations, say by reference to representing 'interests'. But it seems absurd in the context of representative equity that the same could be applied to past generations. The notion of (political) representative equity has to be restricted to individuals whose life might be affected by the policies framed by the representative agents.

Simple Weighted Arithmetic Aggregate Distributions

To conclude, let me briefly recapitulate some of the effects of this analysis on the sort of simple 'mixed' proposals advocated, amongst others, by Grubb and Sebenius. In other words, let us assume that we are dealing with just two base distributions $\mathfrak{E} = \{D^1, D^2\}$. Following Grubb and Sebenius, the negotiators should focus their attention on *WAA*-solutions given by

$$d_k^*(p) = p\,d_k^1 + (1-p)\,d_k^2 \quad (\text{with } 0 \le p \le 1),$$

and try to come to a settlement as to what the parameter p should be. Due to the purely numerical character of this parameter, there is indeed no other way than strategic bargaining in which this parameter could be determined. Being generally pessimistic about the inequity mitigation properties of such strategic procedures, I cannot see how this type of selection procedure could save us from the doomsday scenario of a break-down in negotiations, even though it has the inclusive nature of numerical selections. If, however, we were to apply the socially weighted procedure suggested above, the chances of reaching a generally acceptable solution might well look better. Assuming Dummett is right in thinking that amongst the different Borda rules, the one with an 'unskewed' index base is the procedurally fairest, let us adopt $\overline{\sigma}_c = \langle\, 0, c, 2c, 3c, \cdots \,\rangle$ $(c > 0)$.

Single Preference Scores. If n is the total number of parties, n_1 the number of parties which prefer D^1 over D^2 and n_2 the one for which the reverse is true, then we have – still assuming strict preferences throughout for the sake of simplicity – that (i) $n_2 = n - n_1$, and (ii) that the aggregated single preference scores for the two distributions are $\beta_1 = \beta(D^1) = cn_1$ and $\beta_2 = \beta(D^2) = cn_2 = cn - \beta_1$, respectively. According to (WAA), k's share under the Borda weighted arithmetic aggregate distribution is $d_k^*(\overline{\beta}) = \tau^* \sum_m \beta_m d_k^m$, which – given that $\tau^* = 1/\sum_m \beta_m = 1/cn$ – means

$$d_k^*(\overline{\beta}) = \frac{n_1}{n} d_k^1 + \frac{n_2}{n} d_k^2 .$$

Global Preference Scores. If instead of using single preference scores we allow the Parties to use the 'global' variety, i.e. if we allow them to multiply their scores by their population size, then we find that $\beta_1 = c\,p_1$, $\beta_2 = c\,p_2$, and $\tau^* = 1/cp_w$, with p_k = sum of the populations of the parties preferring D^k (k = 1, 2) and p_w = world population. Accordingly we have that, under global preference scores:

$$d_k^*(\overline{\beta}) = \frac{p_1}{p_w} d_k^1 + \frac{p_2}{p_w} d_k^2$$

In the context of a simple ('two-candidate') claims problem, the linear (unskewed) preference score procedure thus generates a 'mixed-candidate' in accordance with the principles of proportional representation. Indeed, in the case of just two parties: A and B, – where the strategic mixed proposals would be given by:

$$D^p = \langle d_A^p, d_B^p \rangle = pD^1 + (1-p)D^0, \text{ i.e. } d_k^p = p\,d_k^1 + (1-p)d_k^0, \quad 0 \le p \le 1$$

– the single (national) preference score mixture:

$$D^* = \langle d_A^*, d_B^* \rangle = \tau^*(\beta_0 D^0 + \beta_1 D^1), \text{ i.e. } d_k^* = \tau^*(\beta_0 d_k^0 + \beta_1 d_k^1),$$

will be the simple arithmetic mean of the two base distributions: $D^* = D^{1/2}$, since $\tau^* = 1/(\beta_0 + \beta_1) = 1/2$ and (without loss of generality):

Parties	*Individual Preference Scores*	
	D^0	D^1
A	0	1
B	1	0
Borda Index:	$\beta_0 = 1$	$\beta_1 = 1$

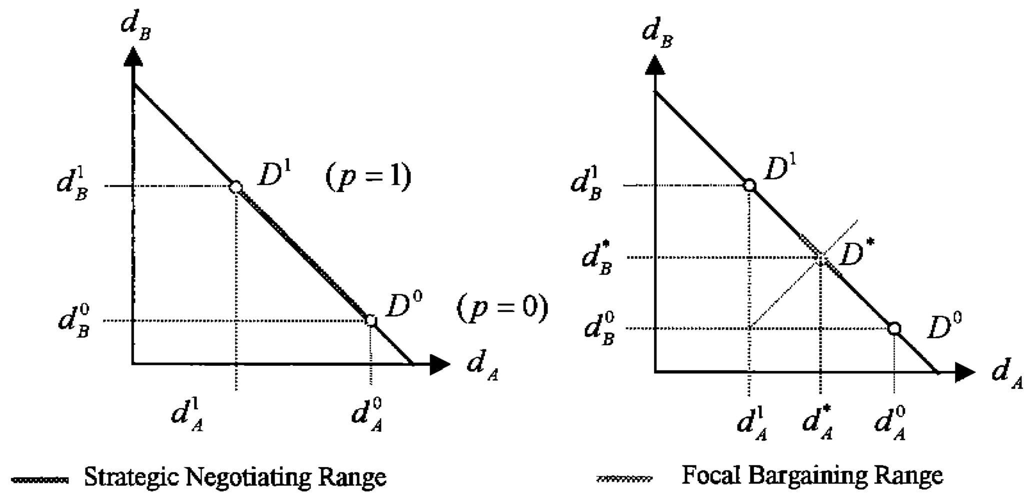


Figure 12: Simple Preference Score Mixtures

This is not to say that this will automatically be the solution which will be signed up to. After all, if $d_A^* = 0.6203$ and $d_B^* = 0.3797$, then it would not be surprising if the negotiations were to end up with the focal point solution $\langle 0.6, 0.4 \rangle$. The intended primary function of this preference score procedure, after all, was not to replace the bargaining process as modelled in focal point theory, but simply to transform potential doomsday scenarios into situations where these bargaining processes can get off the ground.

The inclusiveness and the fairness of the preference score procedure is meant to mitigate potential sensitivities in such a manner that the selected preference score mixture D^* could reasonably be regarded as an anchor point for subsequent (focal point) negotiations, i.e. a position where the parties involved could agree that – as far as equity considerations are concerned – the outcome ought to be somewhere in the neighbourhood. Whether or not the preference score procedure actually manages to generate a non-empty common acceptability range, let alone one which contains D^*, obviously depends on the initial sensitivities of the parties involved. What I hope to have shown in this study is merely that if there is a serious interest in a *world-wide* settlement on greenhouse gas emissions quotas, then one, if not the only manner in which such a settlement might be attainable is by way of determining the total social values of what is due to the individual parties – as allocated under the preference score mixture – as (hopefully) a commonly acceptable anchor-point for the bargaining process.

Informal Synopsis of Part III

I began this discussion of numerical selection procedures by considering a particularly simple interpretation of the claims problem at hand, in which each claimant was taken to have a single claim, namely the percentage which he would get under his favoured base-distribution. The advantage of this sort of simplified situation was that all the claims in question were directly comparable in magnitude, which, in turn, meant that certain traditional allocation rules could be brought into play. Amongst the variety of these rules, the most appropriate for our purposes turned out to be the Aristotelian proportional allocation.

Having singled out the Aristotelian rule in this manner, I then argued that the outcome achieved can be made still fairer, not because of a short-coming in the Aristotelian rule, but because of the *over*-simplified nature of the claims which were taken into account. This was due to the fact that in the context of our claims problem, a claimant could justifiably put forward not just one, but a whole array of claims for the good in question. The problem with this, however, turned out to be that such claims vectors failed to be directly comparable in magnitude, thus automatically excluding the application of any traditional allocation rule, Aristotelian or not.

Wishing to retain the benefits of the Aristotelian rule, the only way around this problem was to transform these claims vectors into aggregate claims which *are* comparable in magnitude. Since such an aggregation could be achieved by different methods – each of which leading to different outcomes under the Aristotelian rule (AAA) – we were, yet again, confronted with making a justifiable choice in answer to the following questions: (i) Are there aggregation functions which could justifiably be applied in the context of our claims problem, and, if yes, (ii) which of them would – in conjunction with (AAA) – be most likely to constitute what we referred to as a sufficiently fair selection procedure?

Conforming with the existing literature I chose to focus on the slightly narrower issue whether we can justify the choice of the weighted arithmetic average function as used in the so-called 'mixed proposals'. The general idea of how such a justification might be given was quite simple. Based on the fact that all the base distributions were taken to have equal legitimacy, an 'allocation model' was introduced under which the total disputed good was equally allocated to these base distributions. Moreover, each of these equally sized parts was meant to have a specific value-per-unit of the good it contains. A justification of weighted arithmetic aggregations – and consequently of the resulting weighted arithmetic aggregate distributions – was then seen as given if the *WAA*-distribution-share of each claimant is equal to the total value due to him in this partition model (under an interpretation of the aggregation weights as these unit values).

The main problem was to find a suitable mechanism to fix these 'unit values', and the solution was to use a type of 'stated social value', represented by what I referred to as Borda indices. This type of welfare index is given by the sum of preference scores allocated to a base distribution under the Borda rule and reflects the social acceptability (or, to be more precise, the acceptability amongst the claimants) of the base distribution in question. By interpreting our aggregation weights as Borda indices, it was thus possible to justify the use of weighted arithmetic aggregate distributions as compromise solutions for our claims problem on the grounds that, in doing so, the Parties receive their *total social dues*, as determined by the Borda rule.

Unfortunately, these 'total social dues' turned out to be sensitive to the chosen 'index base' – i.e. the set of numbers to be used as preference scores – which meant that yet another argument for a particular choice was required. Following Michael Dummett, it was argued that, as concerns procedural fairness, the 'unskewed' index-base (0, 1, 2, 3, and so forth) is the fairest option. This, in turn, left one more equity problem to be dealt with, namely the fact that the 'Single Party – Single Score' principle inherent in the ordinary Borda rule might justifiably be objected to on grounds of 'representational equity': why should the views of a small number of people (such as the population represented by the delegation from Liechtenstein) carry the same weight as those of a very large number (China)? The solution to this problem was to switch from 'Single National' to 'Global' preference scores, i.e. to permit each Party to multiply their scores by the number of people they represent. This switch had the additional advantage of turning the preference score procedure into a collusion proof mechanism.

Given the inclusive nature of these preference score distributions (no outright losers) and the procedural fairness of both the Borda rule as a voting mechanism and the Aristotelian allocation rule (AAA), it stands to reason that this sort of selection procedure may have a better chance than most others of being sufficiently fair to overcome potential doomsday scenarios in our allocation problem. In the absence of such a fair compromise proposal, all that we might end up with in the UNFCC process is a flourishing conference tourism service industry. As side-effect, this might be acceptable, as sole effect, it clearly is not.[26]

[26] As a postscript, let me also highlight the fact that the preference score method proposed here in the context of global warming negotiations is likely to have a much greater range of applications, one of which may indeed be the aggregation of otherwise unrelated indices. But this will have to be left for further investigation.

APPENDIX 1

Global preference score mixtures of additivebase distributions are collusion-proof (assuming the collusion is between parties of identical preferences).

Consider, on the one hand, a scenario with three parties A, B_1, B_2, and, on the other, the situation where two of the parties: B_1, B_2 collude. The collusion scenario thus has just two recognised parties: A, B, and the 'joint' party B is assumed to share the preferences with B_1 and B_2 regarding the elements of the relevant, say two-element selection base: $\{ \widetilde{D}^1, \widetilde{D}^2 \}$ for the collusion and $\{ D^1, D^2 \}$ for the initial scenario. The main difference between the distributions of the two selection bases is, of course, that the former are pairs: $\widetilde{D}^m = \langle \widetilde{d}_A^m, \widetilde{d}_B^m \rangle$ and the latter triples: $D^m = \langle d_A^m, d_{B_1}^m, d_{B_2}^m \rangle$ of distribution numbers.

The assumption that we are dealing with additive base-distributions simply means that:

$$\text{(i) } \widetilde{d}_A^m = d_A^m \text{ and (ii) } \widetilde{d}_B^m = d_{B_1}^m + d_{B_2}^m \text{ (for } m = 1, 2).$$

If $\widetilde{s}_k^m$ is the Borda score of k (= A, B) for $\widetilde{D}^m$, and s_k^m is the Borda score of k (= A, B_1, and B_2) for D^m, then collusion itself is meant to imply that (iii) $\widetilde{s}_B^m = s_{B_1}^m = s_{B_2}^m$, and (iv) $\widetilde{s}_A^m = s_A^m$. If p_X is the population of Party X, then the demographically weighted Borda indices are

$$\text{(v) } \widetilde{\beta}^m = p_A \widetilde{s}_A^m + p_B \widetilde{s}_B^m \text{ and (vi) } \beta^m = p_A s_A^m + p_{B_1} s_{B_1}^m + p_{B_2} s_{B_2}^m$$

Given (iii) and (iv) it follows from (vi) that

$$\beta^m = p_A \widetilde{s}_A^m + (p_{B_1} + p_{B_2}) \widetilde{s}_B^m$$

This, together with (iii) and the fact that $p_B = p_{B_1} + p_{B_2}$ then implies that $\beta^m = \widetilde{\beta}^m$.

.

According to the general formula for weighted arithmetic aggregate distributions:

$$d_k^* = \tau^* \sum_m w_m d_k^m \text{ with } \tau^* = 1 / \sum_m w_m$$

the Global preference score shares of the parties consequently are

$$\tilde{d}_k^* = \tau^*(\beta^1 d_k^1 + \beta^2 d_l^2) \quad \text{with } \tau^* = 1/(\beta^1 + \beta^2)$$
$$\tilde{d}_k^* = \tilde{\tau}^*(\tilde{\beta}^1 \tilde{d}_k^1 + \tilde{\beta}^2 d_k^2) \quad \text{with } \tilde{\tau}^* = 1/(\tilde{\beta}^1 + \tilde{\beta}^2)$$

Thus we have not only $\tau^* = \tilde{\tau}^*$, but also, by virtue of (ii) that

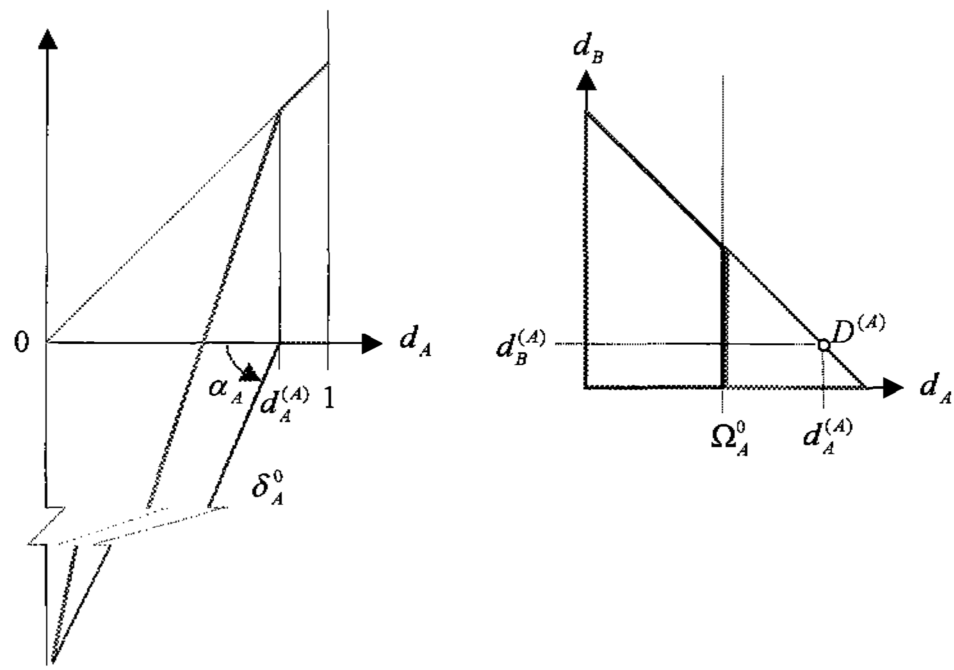
$$d_{B_1}^* + d_{B_2}^* = \tau^*[\beta^1 (d_{B_1}^1 + d_{B_2}^1) + \beta^2 (d_{B_1}^2 + d_{B_2}^2)] = \tilde{\tau}^*[\tilde{\beta}^1 \tilde{d}_B^1 + \tilde{\beta}^2 \tilde{d}_B^2] = \tilde{d}_B^*.$$

APPENDIX 2

Linear Base Disutilities

(i) $$\delta_A^0(d_A; d_A^{(A)}, \alpha_A) = \begin{cases} 0 \quad \text{if} \quad d_A \geq d_A^{(A)}, \\ (d_A^{(A)} - d_A)\tan\alpha_A, \text{ else} \end{cases}$$

(ii) $$W_A^0(D; D^{(A)}, T, \alpha_A) = u_A(d_A T) - \delta_A^0(d_A; d_A^{(A)}, \alpha_A), \text{ where}$$
$$u_A(d_A T) = d_A T.$$

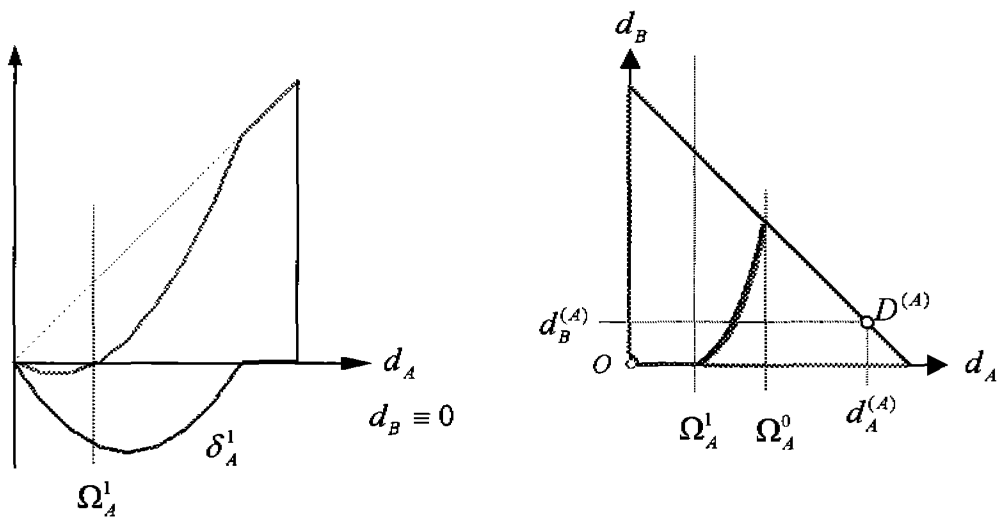


As mentioned earlier, this purely linear model has the advantage of providing a simple representation for sensitivities, namely the angle α_A. At the same time it has at least two disadvantages. First it involves a discontinuity at the null-distribution, *if* we were to apply (i) and (ii) also to partial distributions (as suggested in the right-hand figure). Second, it actually implies that A's break-off point: $\Omega_A(T)$ moves to the right with diminishing T, thus decreasing the chances for commonly acceptable solutions. There are, however, ways in which both these shortcomings can be remedied, albeit at the cost of a loss of simplicity of exposition. Let me begin by suggesting a numerical formula for disutilities which would remedy the said discontinuity, by introducing

Value Dependent Disutility Representations

(i) $\delta_A^1(D; d_A^{(A)}, \alpha_A, T) = \delta_A^0(d_A)\,(d_A + d_B)\,T$, where
$(d_A + d_B)\,T$: total value at stake

(ii) $W_A^1(D; D^{(A)}, T, \alpha_A) = u_A(d_A T) - \delta_A^1(d_A; d_A^{(A)}, \alpha_A)$,



These modified disutilities for A clearly are no longer linear, nor are dependent only on the share of A as determined by d_A. Indeed, the left-hand diagram pertains only to (total and partial) options in which $d_B = 0$. What is equally clear, however, is that A's welfare function W_A^1 is continuous for all the possible options. Unfortunately, this modified welfare function still fails to reflect the idea that a lowering of the stakes will increase the chances of an agreement (the break-off points are stationary under variations of T). This, however, can easily be remedied by 'strengthening the influence' of T on disutilities, by say multiplying our base disutilities with a power of T: $\delta_A^2 = \delta_A^1\, T^n$. In this case the break-off point, say for total distributions, is at $\delta_A^0\, T^n$, and thus moves in the desired direction (towards 0) with decreasing T.

REFERENCES

AGBM (1997), *Framework Compilation of Proposals from the Parties for the Elements of a Protocol or Another Legal Instrument.* FCCC/AGBM/1997/2, 31 January.

Banuri, T., K. Göran-Mäler, M. Grubb, H. K. Jacobson, and F. Yamin. (1996) 'Equity and Social Considerations.' in Bruce *et al.* (1996): 79 – 124.

Barrett, Scott (1992), ' "Acceptable" Allocation of Tradable Carbon Emission Entitlements in a Global Warming Treaty,' Chap. VI in *Tradable Entitlements for Carbon Emission Abatement*, Geneva: UNCTAD, INT/91/A29: 85–113.

Binmore, K., A. Shaked, and J. Sutton (1985), 'Testing Noncooperative Bargaining Theory: A Preliminary Study,' *American Economic Review* 75:1178–80.

—. (1988) 'A Further Test of Noncooperative Bargaining Theory: Reply,' *American Economic Review* 78: 837–39.

—. (1989) 'An Outside Option Experiment,' *Quarterly Journal of Economics* 104, no. 4: 753–70.

Bruce, J. P., H. Lee, and E. F. Haites (eds.) (1996), *Climate Change 1995: Economic and Social Dimensions of Climate Change* (Contributions of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change), Cambridge: C.U.P. for IPCC.

Burtraw, Dallas, and Michael A. Toman (1992), 'Equity and International Agreements for CO_2 Constraints.' *Journal of Energy Engineering* 118, no. 2:122–35.

Dummett, Michael A. E. (1984), *Voting Procedures,* Oxford: Clarendon Press.

Edmonds, Jae, Marshall Wise, and David W. Barns (1995) 'Carbon Coalitions: The Cost and Effectiveness of Energy Agreements to alter Trajectories of Atmospheric Carbondioxide Emissions' *Energy Policy* 23: 309–35.

Grubb, Michael (1990), *Energy Policies and the Greenhouse Effect,* Vol. 1. London: RIIA.

Grubb, Michael, and James Sebenius (1992), 'Participation, Allocation and Adaptability in International Tradable Emission Permit Systems for Greenhouse Gas Control.' Chap. 9 in *Climate Change: Designing a Tradable Permit System*, Paris: OECD:185–225.

Grubb, Michael, James Sebenius, Antonia Magalhaes, and Susan Subak (1992), 'Sharing the Burden,' in *Confronting Climate Change: Risks, Implications, and Responses*, edited by Irvin M. Mintzer. Cambridge: C.U.P.: 305–22.

Kreps, David M. (1990), *A Course in Microeconomic Theory,* New York: Harvester Wheatsheaf.
Mitchell, John V. (forthcoming), 'Post-Kyoto Implications for Energy Markets: The Geopolitical Perspective,' Proceedings of the IAEE/GEE Conference on *Energy Markets: What's New?,* Berlin, 9 September 1998.

Nitze, W. (1990), *The Greenhouse Effect: Formulating a Convention,* London: RIIA.

Nozick, R. (1974), *Anarchy, State and Utopia,* Oxford: Blackwell.

O'Neill, B. (1982), 'A Problem of Rights Arbitration from the Talmud,' *Mathematical Social Sciences* 2: 345–71.

Paterson, Mathew (1996), 'International Justice and Global Warming.' in *The Ethical Dimensions of Global Change*, edited by Barry Holden, London: Macmillan.

Reiner, David M., and Henry D. Jacoby (1997), *Annex I Differentiation Proposals: Implications for Welfare, Equity and Policy,* Report No. 27, Cambridge MA: MIT Joint Program on the Science and Policy of Global Change.

Lasse Ringius, Asbjørn Torvanger and Bjart Holtsmark (1998), 'Can Multi-Criteria Rules Fairly Distribute Climate Burdens? OECD Results from Three Burden Sharing Rules,' Energy Policy 26(10) (August 1998), pp. 777-793.

Lasse Ringius (1999), 'Differentiation, Leaders, and Fairness: Negotiating Climate Commitments in the European Community'. International Negotiation vol. 4, no. 2 (Summer 1999).

Roth, A. (1985), 'Towards a focal Point theory of Bargaining,' in *Game-theoretic Models of Bargaining,* edited by Alvin E. Roth. Cambridge: CUP: 259–68.

Samuelson, Paul A. (1974), 'Analytical Notes on International Real-Income Measures,' *Economic Journal* 84: 595–608.

Schelling, Thomas (1960), *The Strategy of Conflict,* 1980 ed., Cambridge Mass.: Harvard University Press.

Sen, Amartya (1977), 'On Weights and Measures: Informal Constraints in Social Welfare Analysis,' *Econometrica* 45 (September):1539–72. Reprinted in Sen (1982):226–63.

— (1982), *Choice, Welfare and Measurement,* Oxford: Blackwell.

— (1984), *Resources, Values and Development,* Oxford: Basil Blackwell.

Smith, Goldwin (1974), *A History of England,* 4th ed. (1st ed. 1949), New york: Charles Scribner's Sons

Torvanger, Asbjørn and Odd Godal (1999),'A survey of differentiation methods for national greenhouse gas reduction targets', Oslo: CICERO, Report 1999:5.

Walker, David M. (1980), *Oxford Companion to Law,* Oxford: Clarendon Press.

Young, H. P. (1990), 'Sharing the Burden of Global Warming,' *Report from the Institute for Philosophy and Public Policy* 10, no. 3/4: 6–10.

— (1994), *Equity in Theory and Practice,* Princeton, NJ: Princeton University Press.

Young, Oran R. (1989), 'The Politics of International Regime Formation: Managing Natural Resources and the Environment,' *International Organization* 43, Summer:349–75.

OXFORD INSTITUTE FOR ENERGY STUDIES
57 WOODSTOCK ROAD, OXFORD OX2 6FA ENGLAND
TELEPHONE (01865) 311377
FAX (01865) 310527
E-mail: publications@oxfordenergy.org
http://www.oxfordenergy.org