Dissertation accepted for the dr. polit. Degree

at the Department of Economics, the University of Oslo:

# Deciding Who does What and When:

## Four Essays on the Economics of Global Climate Change

**Camilla Bretteville Froyn**

CICERO Center for International Climate and Environmental Research - Oslo

## Acknowledgements

Over the years that I have worked on this dissertation I have learned a lot from a great number of people, all of whom in various ways have contributed to the work presented here. Thank you! Some of these people have allowed me to trouble them innumerable times. First of all I would like to express my appreciation to my supervisor, Jon Vislie, for his comments and his flexibility, and also my co-authors, H. Asbjørn Aaheim, Jon Hovi, and Fredric C. Menz. It has been tremendously educational and inspirational working with them. Secondly, I would like to thank the committee; Scott Barrett, Ottar Mæstad, and Geir Asheim for their impressively thorough review. This thesis has benefited a great deal from their comments.

Furthermore I thank Kalle Moene for talking me out of quitting the first couple of years, Tora Skodvin for her pep talks, and Mia Froyn for help with the title. My colleagues at CICERO I thank for all the fun, help and support – in particular I am grateful to Frode Rørvik for helping me with every little computer problem along the way, Lynn Nygaard for excellent editorial assistance, Tone Veiby for helping me with this that and the other thing, and my fellow PhD student, Guri Bang, for her invaluable friendship and support. The work has been financed by the Research Council of Norway and my host institution, CICERO.

I spent one of these years at the University of California at Berkeley, from the summer of 1999 to the summer of 2000. I am very grateful to Michael Hoel and Larry Karp who helped set up the practical arrangement, and to Anthony Fisher, Michael Hanemann, Larry Karp, and the other faculty, as well as my fellow students, at the Department of Agricultural and Resource Economics for interesting discussions, inspiration, and fun. My stay would not have been possible without financial assistance from the Research Council of Norway, CICERO, and the following scholarship funds: Christiania Bank og Kreditkasses Legat, Den Norske Banks Fond til det Samfunnsvitenskapelige Fakultet ved Universitetet i Oslo, August og Mathea Mantheys Legat, Foundation for Economic Research at the Central Bank of Norway, and Stiftelsen Thomas Fearnley, Heddy og Nils Astrup. I am forever grateful to these people and institutions for giving me the opportunity to enjoy a year in the San Francisco Bay Area.

Furthermore I would like to thank my friends and family, in particular Espen Froyn, Carl and Anne Bretteville, Åse and Hans Bjøntegård, Mia and Sven Froyn, Cecilie Mossige, Sigurd Birger Løkeland, Ingvild Strømsheim Wold, Anne Cathrine and Asgeir Ohr, John Bodine, Chris Chouteau, Bob Callahan, and Jana for their love, support, and encouragement. I would not have made this without them.

Lastly I thank my parents, Turid and Christian Bretteville, to whom this work is dedicated and to whom I owe so much. I wish they both could have lived to share this with me.

Camilla Bretteville
Oslo, April 2004

# Table of content:

Dissertation accepted for the dr. polit. degree at

the Department of Economics, the University of Oslo:

# Deciding Who does What and When:

## Four Essays on the Economics of Global Climate Change

## Introduction

**C**limate change has been one of the most widely discussed environmental issues since 1990, following the first report of the Intergovernmental Panel on Climate Change (IPCC). The IPCC's assessment provided evidence that emissions from human activities might be adversely affecting the climate. Subsequent IPCC studies were completed in 1996 and 2001 that supported the initial findings. There is now fairly widespread consensus that climate change is occurring, that it is caused principally by anthropogenic sources, and that some sort of policy response is called for to reduce greenhouse gas emissions from their current levels. Nonetheless, there has been continuing debate in both scientific and policy communities about (a) to what extent, if any, action should be taken (essay no. 1), (b) by whom (essay no. 3 and 4), and (c) over what time period (essay no.2).

These issues are particularly complicated because, due to a combination of several features, climate change is profoundly different from most other environmental problems we have had to deal with. The features include public goods issues that require collective global action, the multiplicity of decision makers ranging from the international governmental level down to the micro level of firms and individuals, and the heterogeneity of emissions and their consequences around the world. Moreover, the long-term nature of climate change due to the fact that it is the concentrations of greenhouse gases that matters, rather than annual emissions, raises the thorny issues of intergenerational transfers of wealth and environmental goods and bads. Next, human activities associated with climate change are widespread, which makes narrowly defined technological solutions impossible, and the interactions of climate policy with other broad socio-economic policies are strong. Finally, large uncertainties, and in some areas even ignorance, characterize many aspects of the problem and require a risk management approach to be adopted in all decision making frameworks that deal with climate change (IPCC 2001).

In response to the IPCC findings, countries at the Second World Climate Conference in 1990 agreed to negotiate a ‘framework treaty’ to control climate change, which culminated in more than 150 nations signing the United Nations Framework Convention on Climate Change (UNFCCC) at the ‘Earth Summit’ in Rio de Janeiro in 1992. Article 2 of this convention states the objective of stabilizing greenhouse gas emissions within a time frame at a level that prevents dangerous anthropogenic interference with the climate system. Signatory nations volunteered to adopt policies to reduce greenhouse gas emissions by the end of the decade to their 1990 levels, but did not specify targets for greenhouse gas reductions beyond 2000. The climate convention was to serve as a framework for future Conferences of the Parties that would specify targets for greenhouse gas emissions reductions and develop mechanisms for achieving these targets.

Procedural negotiations started in Berlin in 1995 with the first Conference of the Parties (COP1), where it was agreed that industrialized countries should set emissions limits with reduction targets to be achieved within specific time limits. The third COP was in 1997 when the Kyoto Protocol was signed, requiring Annex I (industrialized) countries to reduce their combined annual average emissions during the period 2008–2012 to five percent below their 1990 levels. The Kyoto Protocol establishes limits on greenhouse gas emissions for industrialized countries only, differentiated between -8 and +10 percent of the emissions in 1990. It also provides the means to reduce these emissions in cost-effective ways. These so-called flexibility mechanisms include international emissions trading, joint implementation, and the Clean Development Mechanism.

To enter into force, the Kyoto Protocol must be ratified by 55 countries responsible for at least 55 percent of the 1990 carbon dioxide emissions from Annex I countries. At present,[1] 106 parties representing 43.9 percent of these emissions have ratified. Whether Kyoto goes into effect now depends solely on ratification by Russia (with 17.4 % of Annex 1 carbon dioxide emissions). Prospects for implementation of the Kyoto Protocol have been unclear principally because of the United States’ refusal to ratify the accord. Ratification by the United States (36.1 %) was uncertain from the time the Kyoto Protocol was signed, but President Bush eliminated all doubts of the U.S. position when he repudiated the treaty in March 2001. Despite the U.S. rejection of the Kyoto Protocol, there have been subsequent Conferences of the Parties to the UNFCCC in Bonn (July 2001), Marrakech (October/November 2001), and New Delhi (October 2002). These Conferences have resolved some, but not all, of the difficulties

---

[1] April 2004

among signatories regarding emissions trading, carbon sinks, and other implementation issues.

For global environmental protection, the public goods problem of under-provision is difficult to overcome because there is no supranational authority to enforce an agreement. In the absence of such an authority, alternative arrangements are needed. Existing international environmental regimes do not include a centralized equivalent to the World Trade Organization (WTO), but there are a large number of organizations, international treaties and other institutions aimed at resolving transnational environmental conflicts and protecting environmental resources. Currently, the United Nations Environmental Program (UNEP) lists more than 300 international environmental agreements with secretariats or other types of organizational structures. These international environmental agreements can be thought of as institutional structures designed to solve particular international public goods problems.

The existence of international regimes, however, does not ensure that countries will be able to achieve optimal levels of cooperation. Failure to solve the problem of providing international public goods is well known. The incentive for nations to free ride and the costs of detecting and punishing such behavior will be greater the larger the number of nations. Institutions must therefore be designed to reduce these incentives.

One possible solution for providing global environmental goods is the creation of a supranational environmental organization. The advantages of such an institution include the potential for achieving environmental protection levels close to the social optimum, and cost savings from the reduction of administrative duplication and economies of scale in negotiating provisions for coordination of the environmental regime with other international regimes, such as those focusing on trade or development

(Peterson and Wesley 2000). However, costs of negotiation and administration in a large supranational authority could be significant. In addition, large international bureaucracies can themselves be a source of inefficiency and waste. Perhaps most importantly, sovereign nations are generally reluctant to cede national authority to supranational organizations and institutions. Thus, domestic legislation remains the method by which international environmental agreements are implemented.

International environmental treaties face further difficulties because they generally attempt to correct instances of government failure as well as market failure. International contractual solutions to environmental problems are therefore more challenging to achieve than solutions to domestic environmental problems. Those challenges are likely to continue as long as nations remain sovereign (Congleton 2001).

There is a considerable body of literature addressing under-provision of international pollution control. The main problem analyzed by these models is free riding in international pollution control. There are two types of incentive for free riding: the incentive for a country to not sign the agreement and thus benefit from the signatories' abatement efforts, and the incentive for a signatory to violate its commitments in an agreement (non-compliance). Because of these free-rider incentives, there will generally be suboptimal equilibrium coalition structures in global pollution control (Finus and Rundshagen 2001).

A standard game theoretic solution to the free-rider problem is international transfers. The transfer rule redistributes the surplus to be gained from cooperation to compensate the countries that would otherwise have chosen the non-cooperative outcome.[2] The transfer solution might, however, be difficult to implement in the case of

---

[2] For contributions on this issue see for example Carraro (1999), Carraro and Siniscalco (1993) Barrett (1994, 1997, 2003), Finus (2001), and Eykmans and Tulkens (2001).

climate change. Simulation models for climate change mitigation have shown that, in the long run, the gains from cooperation will more than compensate for the initial losses due to abatement efforts. However, the fact that the expected break-even date lies very far into the future makes current cooperation on restricting greenhouse gases difficult. For instance, Eykmans and Tulkens (2001) point out that the transfers in their simulation model are single numbers representing the present value of consumption flows over 320 years, and that these cannot realistically be conceived of as being paid as lump sum transfers at time t = 0. Furthermore, they argue that these transfers also cannot be spread over time because the countries cannot borrow against future gains in order to compensate for early losses, although these kinds of distribution problems might be solvable through some kind of banking system.

Another suggested solution to offset countries’ free-riding incentives is the linkage of environmental negotiations to other economic issues (issue linkage). The idea is to link an issue with excludable benefits (a club good) to the public good provision. It has for example been suggested to link the climate change negotiations with negotiations on trade liberalization (Barrett 1995, 2003) or research and development cooperation (Barrett 2003), (Carraro and Siniscalco 1995, 1997).

By its very nature, a legally binding agreement for climate protection among more than 170 nations – the Kyoto Protocol – will have both supporters and detractors. Many would applaud the notion that such an agreement even exists, given the nature of the problem, the scientific uncertainties, and the long timeframe involved. Others would point to the innovative methods embodied in the Protocol to reduce costs, including emissions trading and the ability of nations to consider changes in land use and new forestry techniques (carbon sinks). Still others support Kyoto on equity grounds, taking the position that since industrialized nations have been the major contributors to the accumulation of greenhouse gases in the atmosphere, it is appropriate for them to bear the major burden of controlling emissions in the future. On the other hand, critics of Kyoto most often point to high costs, uncertain benefits, and lack of commitment from developing countries. Other detractors say that the Protocol left too many specific details to be negotiated at a future date.

For industrialized countries, both damage costs and adaptation costs of climate change are likely to be relatively low. The reason is that the sectors most likely to be affected by climate change – agriculture and forestry – account for a very small share of total output. Thus, the willingness of industrialized countries to pay for climate control is likely to depend on their concern for the situation of developing countries, whether they see climate change mitigation to be a high priority area for aid, and the costs of abatement. Climate change will have a much greater impact in developing countries because of their greater reliance on agriculture and exposure to potential health impacts from waterborne and parasitic diseases. However, developing countries are likely to be less vulnerable to climate change in the future. Furthermore, they might prefer to invest in economic development or climate change adaptation rather than greenhouse gas emission reductions. Finally, the impact on the global climate of reducing emissions in a single country is likely to be negligible. Thus, without a regime that imposes binding commitments on sufficiently many countries, neither industrialized countries nor developing countries have an incentive to voluntarily contribute to climate control (Schelling 1992, 1997).

The climate negotiations began by focusing on the short term, with the industrialized countries agreeing to cut their emissions of greenhouse gases by about five percent relative to 1990 emissions by 2008–2012. Then there was agreement that these cuts should be achieved cost effectively, incorporating ‘flexibility mechanisms.’ Only later did the signatories worry about whether the treaty created incentives for broad participation and full compliance. Thus, the Kyoto agreement offers little incentive for countries to ratify. At the same time, a country may make its participation conditional on easy terms, and easy terms will not protect the climate (Barrett 2001; 2003).

Furthermore, there is little in the Kyoto agreement that ensures that even ratifying countries will actually do what they say they will do. In fact, any compliance mechanism entailing “binding consequences” must be approved by amendment (Article 18), which essentially means a new treaty. Since any party to Kyoto could decline to ratify a subsequent compliance amendment, it can avoid being punished for failing to comply. The compliance mechanism approved in Bonn requires that a country that fails to meet its emission ceiling in the first commitment period (2008–2012) compensate for this deficiency plus an additional 30 percent of this amount in the second commitment period (2013–2017). As additional punishment, the possibility to sell emission permits may be suspended, and a plan for how the non-compliant country intends to get back in compliance must be presented. The legal formulation of the rules in the Kyoto Protocol, however, has been postponed to the first meeting of the parties to the protocol after it has entered into force.[3] The agreement does not address what happens if a country fails to comply in the second period as well. Nor does it provide emissions limits for the

---

[3] For a more detailed description of the compliance mechanism, see Torvanger (2001).

second commitment period. Thus, a country that worries about not being able to comply in the first period may hold out for easy targets in the second period (Barrett 2001; 2003). This way the punishment, even if triggered, does not actually bite and the free-rider problem is not solved.

Game theory thus suggests that if Kyoto enters into force and achieves full compliance, it will be because the treaty achieves very little. A more ambitious version of Kyoto, on the other hand, would likely either fail to enter into force or fail to sustain full compliance. Though Kyoto is considered to be only a first step, if the subsequent stages in the process replicate the Kyoto formula, the outcome is likely to be very close to ‘business as usual.’ Moreover, since many of the proposals for Kyoto alternatives also do not address the fundamental issues of enforcement and participation, they too are likely to fail (Barrett 2001).

Essay no. 1, *‘Decision criteria, scientific uncertainty, and the global warming controversy’* addresses the question of whether to take action. This question is complicated since climate change involves scientific uncertainty on many levels: uncertainties in predicting the timing and magnitude of future climate change caused by greenhouse gas concentration; uncertainties in predicting the ecological, economic, social, and political impacts from this; and uncertainties in predicting the effectiveness and costs of policy options. Because of these uncertainties, policy advisors disagree significantly in their recommendations of how to act.

Two major risks have been put forward. One is the risk of significant human and ecosystem impacts from large-scale climate change in the next century, a potential environmental problem of immense dimensions. The other is the risk of incurring large economic costs now, for policies that might slow global warming or mitigate its impacts, when there is considerable uncertainty about the effectiveness of the policies as well as the severity of the problem. This controversy has been highly visible in the climate negotiations. The U.S. administration has, for example, explained its repudiation of the Kyoto treaty by pointing at the lack of certainty. Essay no. 1 shows, however, several cases where uncertainty regarding climate change impacts indicate implementation of climate policies.

Many economists support the view that we should not impose strong policies before the level of knowledge has been improved (e.g., Manne and Richels 1991; Nordhaus and Popp 1997; Schelling 1992; Kolstad 1996; Ulph and Ulph 1996). Others argue that precisely because of the lack of knowledge, abatement should take place now in order to reduce the possibility of extensive and irreversible damage (e.g., Arrow and Fisher 1974; Henry 1974; Grubb 1997; Tol 1995). Essay no. 1 does not give a direct recommendation of how to act, but rather offers basic insight into what the choice of decision criteria implies for various actions. It illustrates that the complexity of the climate problem makes it nearly impossible to predict the full consequences of choices, and raises the question of whether the use of less information-demanding alternatives to the expected utility model is indicated. The essay concludes that the choice of criterion is a political question, that those in favor of abatement policies might be using one of the alternatives as basis for their advice, and suggests that if the possibility of making irreversible mistakes is of great concern, then the minimax regret criterion might have increased relevance.

One should think that the degree of uncertainty in the climate change area to a large degree would influence the political processes and decisions, and that economic analyses thus would treat uncertainty as a central feature. Although the insights into the basic science have improved substantially over the past decades, there is still uncertainty around almost every aspect of the climate change problem. Nevertheless, the main perceptions about economic issues related to climate change and climate policy, for example as assessed and reported by IPCC (2001), are based on deterministic studies. This does not necessarily mean that the uncertainties have been ignored, but rather that this part of the economics literature is somewhat inconclusive.

An explanation for this may be that most of the climate change studies are based on numerical models. The uncertainties per se are therefore difficult to trace. Results of numerical models hinge on assumptions about economic relationships, such as damage and abatement cost functions, which are far from well known. It is therefore difficult to tell, for example, how a better ability to predict temperature change in the future should affect current policy choices, unless we assume that we know all there is to know about the climate change impacts and the costs of mitigation. This is somewhat of a paradox, because learning is one of the main issues that have been focused on in studies of climate policy under uncertainty.

The purpose of Essay no. 2, '*Option Values and the Timing of Climate Policy*,' is to isolate the effects of future uncertainties on current climate policy choice. The model captures what we perceive as the most important issues, including the irreversibility of climate change due to greenhouse gas emissions and irreversibility in abatement technology investments, both of which will necessarily see a gradual change in uncertainty due to learning.

In the static setting adopted in Essay no. 1, it is found that uncertainty alone, even if large, does not affect present optimal abatement to a large degree, if expected utility is maximized in a risk-neutral framework. This is to some extent due to the complex relationships along the chain of causality from emissions of greenhouse gases to impacts of climate change: We do not have, at the time being, sufficient knowledge about whether the impacts of a possible temperature change ranging from 1.8 to 2.2 ºC in a given future year should be assigned a significantly different expected damage cost than those of a temperature ranging from, for example, 0.2 to 3.8 ºC. There might thus be good reasons to base climate policy on other decision-making criteria than the maximization of expected utility.

The difference between a static and a dynamic analysis under uncertainty is that the decisions are taken sequentially in dynamic analyses. While a static problem gives a solution to the best decision once and for all, a dynamic problem aims at finding the best strategy given the available information. This is why learning becomes relevant: If uncertainty affects the decision, future amendments to the information will change decisions in the future. Whether current decisions should be affected by future learning therefore depends on whether future decisions depend on present decisions, or, in the words of the option value literature, whether present decisions are irreversible.

Irreversibilities and learning in environmental economics were first formalized by Arrow and Fisher (1974) and Henry (1974), and the implication for climate change policy was to introduce mitigation strategies now to keep future options open (‘early action’). However, later contributions (Kolstad 1996a,b; Ulph and Ulph 1995) downplay the importance of the irreversibility of climate change due to greenhouse gas emissions in favor of the irreversibility in abatement technology investments, and thus advocate a ‘wait and see’ strategy. Essay no.2 reconsiders the importance of irreversibilities in climate change policy.

We use a stylized model to highlight three important uncertainty effects on climate change policy. When uncertainty is high, the probability of making a perfect guess is low. Under a process of learning, no matter what strategy is chosen, costs must thus be expected because of adjustments desired due to new information. This is the first effect we study. In addition, there are two opposing option values of policies that maintain flexibilities: one related to climate irreversibility and the other related to investment irreversibility. We show that the climate irreversibility imposes an option value to the ‘early action’ strategies if and only if there is a positive probability of encountering the climate irreversibility constraint, and similarly that the investment irreversibility imposes an option value to the ‘wait and see’ strategies if, and only if, there is a positive probability of encountering the investment irreversibility constraint. We also show that investment irreversibility scales down the future climate effects in the same way as an increase in the discount factor. Furthermore, the effect the climate option value has on current policy choice is reduced the more irreversible the investments, but this effect is smaller, the longer the time horizon.

The preferred policy option thus depends on the relative size of these option values and of the ex ante adjustment costs. We conclude that if the sum of the climate option value and the expected adjustment cost due to too low initial abatement exceeds the sum of the investment option value and the expected adjustment cost due to too high initial abatement, the net option value is positive, the climate irreversibility effect dominates, and ‘early action’ should be preferred over a ‘wait and see’ policy.

The structure and characteristics of international agreements on climate change will have a significant influence on the effectiveness as well as the costs and benefits of mitigation. The effectiveness and the costs and benefits of an international climate change regime (such as the Kyoto Protocol or other possible future agreements) depend on the number of signatories to the agreement and their abatement targets and/or policy commitment. (IPPC 2001). Previous research suggests that, at best, a global climate treaty will achieve very little. At worst, it will fail to enter into force. It is therefore interesting to consider whether multiple treaties can be more successful. In Essay no. 3, *'Regional versus Global Cooperation for Climate Control,'* it is argued that individual nations might have a greater incentive to join a regional coalition than a global one. Regional arrangements for climate control are thus more likely to enter into force. A global agreement like the Kyoto Protocol (was intended to be), that requires agreement among a large number of nations, is therefore not necessarily the best way to effectively control emissions of greenhouse gases.

There is speculation about U.S. interest in a regional climate arrangement that would include NAFTA members (Canada, Mexico, and the United States) and (possibly) Australia. While such an arrangement is complicated since Canada is still within the Kyoto arrangement and Mexico is exempt as a developing country, a western hemispheric continental agreement is possible because an existing institutional framework – the North American Commission on Environmental Cooperation (CEC) – exists to oversee and monitor implementation of a North American emissions trading region should it be established.

The European Union (EU) is already dealing with climate change in a regional arrangement. In March 2000, the European Commission launched the European Climate

Change Programme (ECCP) to coordinate the EU response. The EU response includes ratification of Kyoto by the EU Commission in May 2002 and the establishment of an emissions trading scheme within the European community in 2005. The EU-wide emissions trading system is intended not only to ensure that the EU achieves the eight percent cut in emissions by 2008–2012 to which it is committed under the Kyoto Protocol, but also to reduce member-nation abatement costs and lessen competitive impacts of achieving regional compliance.

There is thus a possibility of at least two regional arrangements (NAFTA + Australia, and the EU), and a likelihood of more as developing nations make arrangements to reduce greenhouse gas emissions in the future. Using a simple dynamic game-theoretic model, Essay no.4 shows that a regime based on regional agreements can sustain a larger number of cooperating parties than a global treaty. This is true even when the cost of reducing emissions is the same for both types of regime. The model provides upper and lower bounds on the number of parties under each regime. It is shown that a system with two agreements can Pareto-dominate a regime based on a global treaty. We conclude that, should the Kyoto Protocol not enter into force, cooperation based on regional agreements could be an even better substitute. And even if Kyoto *does* enter into force, regional cooperation might still be an option for future commitment periods.

Decision making frameworks related to climate change involve multiple levels ranging from global negotiations to individual choices and a diversity of actors with different resource endowments, and diverging values and aspirations. This explains why it is difficult to arrive at a management strategy that is acceptable for all. The dynamic interaction among economic sectors and related social interest groups makes it difficult to arrive at a national position to be presented at international forums in the first place. The intricacies of international climate negotiations result from the manifold, often ambiguous, national positions as well as from the linkage of climate change policy with other socio-economic objectives (IPCC 2001).

Conflicts of interests represent a major obstacle for the achievement of a common strategy to combat global climate change. The parties' positions in the climate negotiations are results of political solutions to internal conflicts due to different preferences among people, uneven perceptions of fairness, and various economic effects of policy measures. The performance of a party is closely associated with the causes of these conflicts, how serious they are, and the required means to mitigate them.

It is a matter of common knowledge that the existing climate regime will not be sufficient to keep greenhouse gas concentrations under control. First, the Kyoto Protocol has not yet entered into effect. Second, important issues around compliance and sanctions have yet to be decided on. Third, the targets beyond the year 2012 have to be negotiated. Forth, no global agreement will be fully effective without the involvement of the countries which have not yet committed to reducing emission levels under the Kyoto framework, like China, other developing countries, the United States, and Australia (the enlargement issue).

For a region like the European Union (EU), with cross-national institutions and political infrastructure, the prospects of solving severe conflicts are brighter than for the global society. Therefore, the means to develop a common strategy for climate policy for the EU also represents an opportunity to demonstrate directional leadership in the global climate regime. Hence, it is crucial to analyze whether the EU can play a leadership role with respect to the enlargement issue by proposing strategies, measures, and institutions that can help expand the number of countries that commit to controlling their emissions (Carraro 2000).

To explore possible internal means for the EU, and to evaluate their applicability for the global society, we need to understand the sources and the extent of the underlying internal conflicts.

Climate policy touches every economic sector and thus matters to most interest groups in the economy (Michaelowa 2000). This is why conflicts of interest among and within EU countries have represented a major obstacle for the achievement of a common climate policy. Organized interests have resisted increased costs and have successfully used arguments like that of competitiveness. National positions in the policy negotiations are results of political solutions to internal conflicts, due to different preferences among people, perceptions of fairness, and various economic effects of policy measures. The stance of each nation is closely associated with the causes of these conflicts, how serious they are, and the means required mitigating them.

More than ten years after the EU Commission originally proposed a directive on carbon taxes they reached agreement on a proposed framework for energy taxation. The 1992 proposal faced so much domestic resistance that agreement was not reached until the directive had been considerably watered down. The lobbying succeeded in having the tax altered significantly, with an energy component added, energy-intensive sectors exempted, and the entire tax package made conditional on other OECD countries undertaking precisely the same tax.

Essay no. 4, *‘Sectoral Opposition to Carbon Taxes in the EU – a Myopic Economic Approach’*, study economic reasons for the political infeasibility of extensive carbon taxes. More specifically, it points out reasons for why the opposition to carbon taxes was so great. The study is based on estimating and comparing the costs of expected emissions cuts across sectors and across countries in the EU. This illustrated how different economic sectors might have anticipated the impacts from an expected carbon tax. This focus shows exactly how what seems to be cost-effective and to the best for EU as a region on paper may, because of the myopic vision of the affected sectors, turn out too controversial to be politically feasible.

From the analysis it is clear that the selection of measures is vital, and that different countries have different needs when it comes to policy design. While certain measures may be acceptable to some countries, they are likely to generate opposition in others. The conflicts of interest are largely the result of dependency on fossil fuels, which is higher in the three northern countries and lower in the three southern EU countries in this study. In general, the greater the interest conflicts, the greater are the opportunities to find ways to reduce emissions because of the relative low efficiency in these countries' electricity production.

Three main conclusions can be drawn. First, common measures across countries are generally less attractive since a particular measure may be advantageous to one country in order to keep the national cost of climate policy down, but may in another country spur opposition that could be avoided if exceptions were allowed. Second, the electricity sector plays a key role in climate policy in every EU country and in the potential opposition to any policy measure. Conditions in the electricity sector vary greatly among countries, and a long-term strategy to reduce opposition due to internal as well as external conflicts of interest would, therefore, be to introduce a common electricity market. Third, before a common electricity market can be established, further differentiation of national emissions targets could represent an important opportunity for mitigating conflicts of interests across countries.

**References:**

Arrow, K.J. and A.C. Fisher (1974), ‘Environmental Preservation, Uncertainty, and Irreversibility’. *Quarterly Journal of Economics* **88** (2), 312–319.

Barrett, S. (1994), ‘Self Enforcing International Environmental Agreements’, *Oxford Economic Papers* **46**, 804-878.

Barrett, S. (1995), ‘Trade restrictions in International Environmental Agreements’, London Business School, London.

Barrett, S. (1997), ‘A Theory of International Cooperation’, MIMEO, London Business School, London.

Barrett, S. (2001), ‘Towards a Better Climate Treaty’. AEI-Brookings Joint Center for Regulatory Studies. Policy Matters 01-29. November 2001.

Barrett, S. (2003), Environment & Statecraft. The Strategy of Environmental Treaty-Making, Oxford University Press, New York, 2003.

Carraro, C. (1999), ‘The Structure of International Agreements on Climate Change’, in C. Carraro (Ed.), International Environmental Agreements on Climate Change, Kluwer Academic Pub, Dordrecht.

Carraro, C. (2000). ‘The economics of coalition formation – Potential for structural and instrumental leadership’. Chapter 8 in J. Gupta and M. Grubb (eds.), Climate Change and Eropean Leadership, 135-156. Kluwer Academic Publishers, the Netherlands.

Carraro, C. and D. Siniscalco (1993), ‘Strategies for the International Protection of the Environment’, *Journal of Public Economics* **52**, 309-328.

Carraro, C. and D. Siniscalco (1995), ‘Policy Coordination for Sustainability: Commitments, Transfers, and Linked Negotiations’, in I. Godin and A. Winters (Eds.), The Economics of Sustainable Development, Cambridge University Press, Cambridge.

Carraro, C. and D. Siniscalco (1997), ‘R&D Cooperation and the Stability of International Environmental Agreements’, in C. Carraro (Ed.), International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Congleton, R.D. (2001), ‘Governing the Global Environmental Commons: The Political Economy of International Environmental Treaties and Institutions’, in Shultze and Ursprung (Eds.), Globalization and the Environment, Oxford University Press, NY.

Eyckmans, J. and H. Tulkens (2001), ‘Simulating Coalition Stable Burden Sharing Agreements for the Climate Change Problem’, FEEM Working Paper No. 75.

Finus, M. (2001), Game Theory and International Environmental Cooperation, E. Elgar, Cheltenham.

Finus, M. and B. Rundshagen (2001), ‘Endogenous Coalition Formation in Global Pollution Conrol’, FEEM Working Paper No. 43.

Henry, C. (1974): ‘Option Value in the Economics of Irreplaceable Assets’, *The Review of Economic Studies* **41**, Issue Symposium on the Economics of Exhaustible Resources, 89–104.

Grubb, M. (1997), ‘Technologies, Energy Systems, and the Timing of $CO_2$ Abatement: An Overview of Economic Issues’, *Energy Policy* **25** (2), 159–172.

IPCC (2001). ‘Climate Change 2001: Synthesis Report. Contribution of Working Groups I, II, and III to the Third Assessment Report of the Intergovernmental Panel on Climate Change.’ R.T. Watson and the Core Writing Team (Eds.). Cambridge University Press, Cambridge.

Kolstad, C.D. (1996a), ‘Fundamantal Irreversibilities in Stock Externalities’, *Journal of Public Economics* **60**, 221-233.

Kolstad, C.D. (1996b), ‘Learning and stock effects in environmental regulation: The case of greenhouse gas emissions’, Journal of Environmental Economics ans Management **31**, 1-18.

Manne, A.S. and R.G. Richels (1991), ‘Buying Greenhouse Gas Insurance’, *Energy Policy* **19** 543–552.

Michaelowa, A. (2000), ‘The strength of different economic interests in shaping EU climate policy’, *Energy and Environment,* **11** (3) 277-292.

Nordhaus, W.D. and D. Popp (1997), ‘What is the Value of Scientific Knowledge? An Application to Global Warming Using the PRICE Model’, *The Energy Journal* **18** (1), 1–45.

Peterson, W. and F. Wesley (2000), ‘The Design of Supranational Organizations for the Provision of International Public Goods: The Case of Global Environmental Protection’, *Review of Agricultural Econom*ics **22-2**, 352-366.

Schelling, T.C. (1992), ‘Some Economics of Global Warming’, *American Economic Review* **82**, 1-14.

Schelling, T.C. (1993), ‘Intergenerational Discounting,’ Mimeo. Maryland: School of Public Affairs, University of Maryland.

Schelling, T.C. (1997), ‘The Costs of Controlling Global Warming’, *Foreign Affairs* November/December (1997), 8-14.

Tol, R.S. (1995), ‘The Damage Costs of Climate Change. Toward more Comprehensive Calculations’, *Environmental and Resource Economics* **5,** 353–374.

Torvanger, A. (2001),’ An Evaluation of Business Implications of the Kyoto Protocol’, Report 2001:5, Center for International Climate and Environmental Research, Oslo, Norway.

Ulph, A. and D. Ulph (1995): ‘Global Warming, Irreversibility and Learning’, *Discussion Paper in Economics and Econometrics* **9601**, Department of Economics, University of Southampton.

A former version of this essay is accepted for publication in
*Mitigation and Adaptation Strategies for Global Change*

# Essay no. 1:

# Decision Criteria, Scientific Uncertainty, and the Global Warming Controversy

CAMILLA BRETTEVILLE FROYN[4]
CICERO - Center for International Climate and Environmental Research - Oslo

## Abstract

This paper applies several well-known decision criteria to the climate change problem. The policy process is represented by a simple game against nature with two possible choices: abate or no action. The outcome is considered a compound lottery, with one representing emissions and another representing damages. Assuming that costs exceed benefits of abatement for the participant, the paper analyzes how different decision criteria affect the decision to abate. The role of expert opinion and quality of information in climate change decisions are also considered.

The complexity of global warming makes it impossible to completely overlook the consequences of alternative choices. The paper discusses the question of whether the use of less information demanding alternatives to expected utility theory is indicated. It concludes that the choice of criterion is a political question, and that those in favor of abatement policies might be using one of the alternatives as basis for their advice, and suggests that if the possibility of making irreversible mistakes is of great concern, then the minimax regret criterion might have increased relevance.

**Key words**: climate policy, decision criteria, global warming, greenhouse gas abatement, irreversibility, scientific uncertainty, uncertain choice

---

[4] E-mail: camilla.bretteville@cicero.uio.no, web: www.cicero.uio.no

## 1. Introduction

The global warming controversy has to a large degree been a debate among scientists, policy analysts, and politicians about how to deal with uncertainty. Two major risks have been put forward. One is the risk of significant human and ecosystem impacts from large-scale climate change in the next century, a potential environmental problem of immense dimensions. The other is the risk of incurring large economic costs now, for policies that might slow global warming or mitigate its impacts, when there is considerable uncertainty about the effectiveness of the policies as well as the severity of the problem. This controversy has been highly visible in the climate negotiations. The US administration has for example explained their repudiation of the Kyoto treaty by pointing at the lack of certainty. This paper shows, however, several cases where uncertainty regarding climate impacts and policy effects could be reason for policy-makers to implement climate policies.

Regardless of which strategy policy-makers choose when faced with the threats of global warming, there is considerable uncertainty with respect to the outcome. On one hand, they risk undertaking costly abatement measures to avoid what may turn out to be a minor problem. On the other hand, refusing actions to mitigate climate change is also risky business. Although most of the debate has been over the scientific evidence, the real issue is rather which risk is perceived to be the greater threat (Colgazier 1991).

The complexity of global warming makes it impossible to completely overlook the consequences of alternative choices. A question thus arises of whether this problem, that exhibits such severe forms of uncertainty, should be analyzed in a framework of ignorance, or at least partial ignorance. Theories of rational behavior under complete ignorance can be found for example in Arrow and Hurwicz (1972). Non-probabilistic criteria build on such a notion of ignorance. Critics of these criteria have put forward that the decision maker must at least have some vague partial information about the true state of nature (Luce and Raiffa 1957). The question remains, however, if this vague partial information is sufficient to assign subjective probabilities to the possible states of the world.

Theories on decision-making under uncertainty have suggested various alternatives to expected utility maximization. This paper examines some of these well-known principles in the context of the global warming problem. The most well-known non-probabilistic criterion is perhaps the *maximin* principle (Rawls 1971). Other ones are *the principle of insufficient reason,* and minimax regret, suggested by Savage (1951) as an improvement on the maximin criterion. Discussion of one or more of these criteria can be found for instance in Arrow (1951), Chernoff (1954), Luce and Raiffa (1957), Cohen and Jaffray (1980), Sinn (1980), Fishburn (1987), Dobbs (1991), Chrisholm and Clark (1993), Bouglet and Vergnaud (2000), and Chevé and Congar (2002). Combinations of probabilistic and non-probabilistic criteria have also been suggested. Those studied here are the *limited degree of confidence* criterion, recently treated in Stigum (1990), Eichberger and Spaniels (1998), Chichilnisky (2000), and Lange (2003), and g*eneralized maximin/maximax* criterion, also known as *the pessimism-optimism index criterion* of Hurwicz (1951). Lastly this paper gives a brief discussion of a completely different class of decision criteria known as safety first criteria (see Rawls 1971), where one example is the highest constant consumption criterion, treated for instance in Solow (1974).

When evaluating a decision criterion it is possible to look at the axiomatic foundation, the practical implications, or both aspects of the problem. Generally speaking, if we accept the axiomatic foundation underlying the criterion, we must also accept their implications. However, axioms are abstract and often difficult to fully understand. Therefore it can be fruitful to look at their implications when applied to a specific problem which is exactly what is done here.

Applying the criteria to a practical problem that needs solving demonstrates the implications of using different decision criteria in a frame of uncertainty. As an illustration, this paper uses a very simplified example of a game against nature, with two possible policy options: action to try to prevent global warming (abatement), and business as usual (no action). The framing of the example is such that the decision-maker runs the risk of investing in vain, because the probability of the cost of abatement exceeding the damage avoided is more than 50 percent. In other words, it is likely that the abatement measures will eventually result in a loss. The exercise shows that applying the abstract principles to the problem of global warming illustrates how different decision criteria affect the decision to abate, the role of expert opinion, and the implications of information quality. The paper builds heavily on Aaheim and Bretteville (2001) and Bretteville (1999), and the object is to focus on the lack of certainty regarding impacts of greenhouse gas emissions and the effectiveness of policy in order to compare implications of the use of alternative criteria with those of expected utility maximization. One finding is that those arguing in favor of abatement policies might use one of the alternatives as basis for their advice. The choice of criterion is thus a political question.

The global warming controversy is briefly discussed before sketchily attending to the structure of standard decision theory under uncertainty, presenting various decision criteria, and exploring their implications for the simple game example. Conclusions are presented in the final section.

## 2. The Global Warming Controversy

The question of human-induced global warming definitely involves scientific uncertainty: uncertainties in predicting the timing and magnitude of future climate change caused by greenhouse gas emissions; uncertainties in predicting the ecological, economic, social, and political impacts; and uncertainties in predicting the effectiveness and costs of policy options (Colgazier 1991). Because of these uncertainties, policy advisors disagree significantly in their recommendations of how to act[5].

Referring to the precautionary principle, some argue that we should cut emissions of greenhouse gases now. The precautionary principle has many definitions. Among the common themes are the undesirability of irreversible damage, the need to prevent and anticipate damage, and the argument that lack of complete scientific uncertainty should not be used as an excuse for inaction (Harding and Fisher 1992). The Bergen Ministerial Declaration from 1990 states that the precautionary principle implies that "Where there are threats of serious or irreversible damage, lack of full scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation."

---

[5] This is discussed further in Aaheim and Bretteville 2001.

The precautionary principle thus implies that there is a case for reducing carbon dioxide ($CO_2$) and other greenhouse gas emissions now, even if it is uncertain whether these emissions contribute to the greenhouse problem, and what the economic costs of climate change impacts might be (Chrisholm and Clarke 1993). This interpretation is, of course, dependent on the belief that the climate problem represents “threats of serious or irreversible damage”. Those who share this view argue that uncertainty should dictate abatement now to reduce the possibility of extensive and irreversible damage. They look upon abatement as insurance against catastrophic events and claim that uncertainty alone gives rise to a willingness to pay for insurance (Schelling 1995). A parallel would be fire insurance, which is considered wise even if one is not realistically expecting one's house to burn down.

Those who question the view of the climate problem as representing “threats of serious or irreversible damage” argue that one should not impose stringent policies before the level of knowledge has improved significantly. By postponing action, they argue, we may learn more about the effects of climate change, and new and cleaner technologies may be discovered. Meanwhile, we can invest in alternative projects, which will better prepare us to both abate greenhouse gases and adapt to changes in the future (Kolstad 1994). Additional knowledge about the impacts of climate change will allow the design of better climate policies. Yet, we do not know whether improved knowledge implies less uncertainty. This is why we need more discussion about framing, what information is needed, and how to use available knowledge about the uncertainty in order to make better decisions.

Much of the climate change literature gives recommendations of how to act. Many economists support the view that we should not impose strong policies before the level of knowledge has been improved (i.e. Manne and Richels 1991; Nordhaus and Popp 1997; Shelling 1992; Kolstad 1996; Ulph and Ulph 1996). Others argue that precisely because of the lack of knowledge, abatement should take place now in order to reduce the possibility of extensive and irreversible damage (i.e. Arrow and Fisher 1974; Henry 1974; Grubb 1997; Tol 1995). This paper does not give a direct recommendation of how to act, but rather offers basic insight to the implications for action of the choice of other decision criteria compared to expected utility maximization.

## 3. The Stylized Climate Problem

What is the best way to approach the climate problem? When making policy choices under uncertainty in areas characterized by disagreement among experts, a value judgment will have to be made about what counts as evidence. The question thus becomes: what scientific evidence is sufficient and admissible to justify a policy decision under these conditions? Where to set these standards of proof - how sure is sure enough - is a value judgment. People's values are affected by both their general principles of rights and responsibilities, and their self-interest. They might be primarily concerned either about a fair decision-making process or about the fairness of outcomes. Values play a significant role in judging the fairness in allocation of the costs, benefits and risks to initial stakeholders, and also with respect to the outcome for society as a whole (Colgazier 1991).

The disagreement among expert recommendations on climate policy thus may originate in different value judgments. The discussion of how to deal with the scientific uncertainty is mixed with discussions on other complicated questions such as whether to discount future costs and benefits, the best method for valuing environmental damages, and how to use the precautionary principle. This paper looks beyond both global and intergenerational political problems by introducing a benevolent decision-maker, planner, or principal as a theoretical abstraction whose goal is to decide the best climate policy. Political feasibility and the problems attached to reaching a global climate agreement is thus not touched upon. In other words the decision-maker is assumed to have the `right' preferences, from the social point of view, and will take all relevant factors into consideration in taking action, focusing on the lack of certainty regarding impacts of greenhouse gas emissions and the effectiveness of policy. This rather abstract perspective is taken to make the implications of choosing different decision criteria more transparent and thus comparable.

## 4. Decision Criteria

A decision problem is defined by the acts or options among which one must choose, the possible outcomes or consequences of these acts, and the contingencies or conditional probabilities that relate outcomes to acts (Tversky and Kahneman 1987). Theories of decision-making under risk and under uncertainty have attempted to formalize the way a decision-maker chooses among alternative courses of action when the consequences of each course of action are unknown at the time the choice is made. A situation is said to involve *risk* if the attached randomness can be expressed in terms of specific numerical probabilities (roulette lotteries). These probabilities are objectively specified, as with lottery tickets, or a dice: The numbers are known. However, situations where one cannot (or does not) assign actual probabilities to the alternative possible occurrences are said to involve *uncertainty* (horse lotteries). In this case probabilities reflect the individuals own subjective beliefs: The numbers are unknown (Knight 1971). The two main theories on this field is the von Neumann and Morgenstern (1944) expected utility theory with risk and Savage's (1954) theory of expected utility with uncertainty.

The essence of the von Neumann-Morgenstern theory is a set of restrictions imposed on the preference relations over roulette lotteries. It determines a von Neumann-Morgenstern utility function that can be used to range different roulette lotteries. The Savage theory is a theory of decisions in a situation of horse lotteries. Here the outcomes are uncertain (which horse that wins), while *acts* are the subject of decision (which horse, if any, to place the bet on). Acts are defined as functions from the set of states to the set of outcomes. The theory suggests that the decision-maker has preferences over acts and that these preferences can be represented by a utility function assigning subjective probabilities to states, and furthermore, that choice between acts is equivalent to choice between lotteries. The Anscombe and Aumann (1963) theory combines this. They define *acts* as functions from states to randomized outcomes. The objective randomization over outcomes determines a von Neumann-Morgenstern utility function while the preferences over acts determine subjective probabilities over states[6].

## 4.1 MAXIMIZATION OF EXPECTED UTILITY

The theory of *maximization of expected utility* have a set of underlying axioms. In short, these axioms require that individuals be able to rank actions (completeness), make

---

[6] See e.g. Fishburn (1987), Karni and Schmeidler (1991) or Kreps (1990) for a more comprehensive treatment of standard decision theory under uncertaity.

consistent choices between them (transitivity), and evaluate proximity choices correctly (continuity) (see e.g. Arrow 1965). This implies that the choices of an individual, whose actions are consistent with the axioms, can be described in terms of the utilities of various outcomes for that individual. The utility of an act is equal to the expected utility of its outcomes, obtained by weighing the utility of each possible outcome by its probability. When faced with a choice, the model says that a rational decision-maker prefers the act that offers the highest expected utility.

This theory has been challenged from several perspectives since it was formalized by von Neumann and Morgenstern (1944) and later Savage (1954)[7]. A common basis for much of the criticism has been that people either *does not* act according to expected utility theory in practice, that they *cannot* behave according to expected utility theory because it is too demanding in terms of informational requirements, or that rational people *should not* behave according to expected utility theory because the axioms on which the theory is based are not sound or appealing.

Maximization of expected utility is based on the assumption that the decision-maker can attach probabilities, $p_s$, and outcomes, $x_s$, to each possible state of nature, $s$, and that utility, $u(x_s)$, can be assigned to each of the possible outcomes. The decision is made by maximizing expected utility or welfare:

$$\mathrm{Eu} = \sum_{s} p_s u(x_s)$$

Machina (1989) points out two critical properties of expected utility. First, the comparison between two pairs of outcome and probability is made independent of all other pairs of outcome and probability. If a given combination of outcome and

---

[7] Frank P. Ramsey developed basically the same decision-making framework, published in Ramsey (1931), before his death in 1928.

probability ($x_1$, $p_1$) is preferred to another combination ($x_2$, $p_2$) in one lottery, a replacement of ($x_2$,$p_2$) by ($x_1$,$p_1$) will be preferred in any other lottery as well. This is called the replacement property. Second, if one outcome is preferred to another outcome, i.e. u($x_1$)>u($x_2$), any combination, $px_1$ will be preferred to $px_2$. That is, one may mix outcomes with any probability without affecting the preferences. This is called the mixture property. These two properties can be summarized in the independence axiom.

*The independence axiom*: Lottery X is preferred to ($\succ$) or indifferent to (~) to lottery Y if and only if

$$(\mathrm{X}, \mathrm{p}, \mathrm{Z}, 1-\mathrm{p}) \succeq (\mathrm{Y}, \mathrm{p}, \mathrm{Z}, 1-\mathrm{p})$$

for all lotteries Z and all $p>0$.

Replacement follows directly from the independence axiom because the preference of X over Y also means that a compound lottery of X and Z is preferred over a compound lottery of Y and Z if the probability for Z is the same in both compound lotteries. By setting $p$ = 1 in the independence axiom we get $X \succeq Y$. The axiom states that this preference applies for all $p$, which is the mixture property.

Experimental studies have revealed systematic violations of the expected utility hypothesis[8]. The evidence indicates in particular that decision-makers do not satisfy the independence axiom. In the choice between acts Ellsberg (1961) concludes that decision-makers preferences are inconsistent with expected utility theory and in particular with the existence of additive subjective probabilities. This is referred to as the Ellsberg-paradox. In one version of the experiment subjects are presented with a jar that holds 90 balls. They are told that 30 of these balls are red and that each of the

---

[8] See e.g. Fishburn (1988, Chapter 3) or Machina (1982) for a review of experimental evidence.

remaining 60 balls is either white or black and a bet involves guessing the color of a ball randomly drawn. Subjects are asked to rank the three possible bets. The finding reported is that the bet on red typically is preferred over the other two, who are considered equivalent. Furthermore, he finds that subjects also prefer a bet on either black or white to black or red and to white or red, while they are indifferent between the latter two bets. An explanation for this is that people dislike the ambiguity that comes with choice under uncertainty; they dislike the possibility that they may have the odds wrong and so make the wrong choice. Hence, they go with the bet where they know the odds.

In the choice between prospects the most well-known violation is referred to as the Allais-paradox. One version of this (taken from Kreps 1990) goes as follows:

a) Choose between two gambles. The first gives a 0.33 chance of $27,500, a 0.66 chance of $24,000, and a 0.01 chance of nothing. The second gives $24,000 for sure.
b) Choose between two gambles. The first gives a 0.33 chance of $27,500, a 0.67 chance of nothing. The second gives a 0.34 chance of $24,000 0.66 chance of nothing.

The typical response pattern is to choose the sure thing in a) and the first gamble in b). This violation to expected utility theory has been explained by that individuals rescale probabilities, with more weight (proportionally) given to small probability events.

Driven by the experimental results, several alternative theories have been proposed that depart from expected utility theory. A full review is beyond the scope of this paper[9]. One direction, however, discussed in Aumann (1962) is to weaken the assumption of complete preferences. According to his theory, one can still rank roulette

---

[9] See i.g. Karni and Schmeidler (1991) for an overview.

lotteries, but it opens up for the possibility of not knowing whether one will place a bet, and if one does, which horse to bet on. In this case one cannot assign a specific subjective probability distribution over states. Aumann shows that the preferences over acts in stead can be characterized by a set of subjective probability distributions, such that one act is preferred to another if it is considered better under all subjective probability distributions in the set. Another direction motivated by the Ellsberg paradox and discussed for example in Gilboa and Schmeidler (1989) and Schmeidler (1989), also characterizes preferences by a set of subjective probability distributions, but an act is evaluated by using the subjective probability distribution that gives the lowest expected utility.

Violations of the axioms of expected utility theory is not the subject of this paper, nor is the axiom’s soundness or appeal. The analysis below is instead motivated by the claim that maximization of expected utility assumes a higher ability to consistently handle large amounts of complex information than one can possibly expect from decision-makers. Whether this criticism applies depends, however, on the goal of the decision analysis. An argument in support of the expected utility model has been that people's inability to handle complex information consistently confirms the need for analysis, because the analysis may contribute to better decisions.

Savage (1954) holds the view that by processing partial information we can generate an a priori (subjective) probability distribution over the states of nature sufficient for making decisions. However, to transform vague information concerning the states of nature into an explicit a priori probability distribution, the decision-maker has to register consistent choices in a series of simple hypothetical problems involving these states (Luce and Raiffa 1957). Expected utility theory is thus particularly useful when dealing with situations where probabilities and possible outcomes are within the normal range of human experience. The climate problem is not within this range. We simply do not have the experience to calculate proper weights to aggregate utility over states. The requirement of consistent treatment of a large amount of information represents a problem and alternative decision rules, where the amount of information needed might be less, have therefore been suggested. The aim of this paper is to illustrate and discuss the implication of some of these decision criteria within a specific numerical example of greenhouse gas abatement. The example is not chosen to demonstrate the problems with expected utility maximization, but is simply to compare the differences in implications for the abatement decision. In fact, there is nothing in this example indicating that there cannot be a stable assignment of utilities to outcomes.

### *4.1.1 The numerical example*

Consider a strongly simplified numerical example of abatement with uncertain benefits where the decision-maker has to choose between abatement and no action at all. The problem is interpreted as a game against nature, a non-strategic player[10]. The uncertainty about the final outcome is related partly to uncertainty regarding the magnitude of future emissions, and partly to the uncertainty regarding the future damage of global warming. Future emissions are also uncertain due to factors beyond the control of the policy-maker, such as population growth, economic growth, and technological innovations. The outcome of a given choice can be considered a compound lottery, which consists of two sub-lotteries. The first lottery concerns

---

[10] The example used is the same as in Aaheim and Bretteville (2001).

emissions, and the second concerns the damage from an increase in the concentrations of greenhouse gases.

The numerical example is illustrated in Figure 1. With a business as usual strategy (no abatement) there are two scenarios, alternative 1 and 2, and a policy choice of abatement is assumed to give the same decision tree (the bottom one in Figure 1). In alternative 1, the decision maker replaces the top left scenario with the bottom scenario by choosing to implement climate policy. In this case abatement reduces the damage regardless of what happens, without affecting the probabilities, whereas in alternative 2, abatement affects both the spread of outcomes and the probabilities. For alternative 2, a choice of implementing climate policy will change the relevant scenario from the upper right hand tree to the bottom one. The main difference between the two alternative scenarios, however, is that a possible catastrophe is avoided through abatement in alternative 2.

The specific numbers are chosen to construct these differences between the two alternatives, to simplify calculations, and to generate the same expected damage costs for both alternatives; in fact, the expected damage costs of each branch are the same for the two alternatives[11]. A risk neutral decision-maker will thus view them as identical.

It is reasonable to assume that the considerable difference in the spread of possible outcomes in alternative 1 and alternative 2 is important to a decision-maker. However, applying risk neutrality, expected utility theory does not separate between the two alternatives because the mixture of probabilities from alternative 1 to alternative 2 does not alter the ‘contribution’ to the expectations from each branch of the decision trees. For example, for the upper branch of the two alternatives, we have 0.5 * 0.77 *

---

[11] For benchmark estimates of the damages of climate change, see e.g. Tol (2002).

302.4 = 0.5 * 0.025 * 9313.9 = 116.3. This equality applies for all the branches in alternative 1 and 2. Hence, due to the independence axiom, the two alternatives in the upper part of Figure 1 are identical to a risk neutral expected utility maximizer. A risk averse or risk loving expected utility maximizer will however discriminate both alternatives. Attitude towards risk thus play a key role in decision making under uncertainty (IPCC 2001)[12].

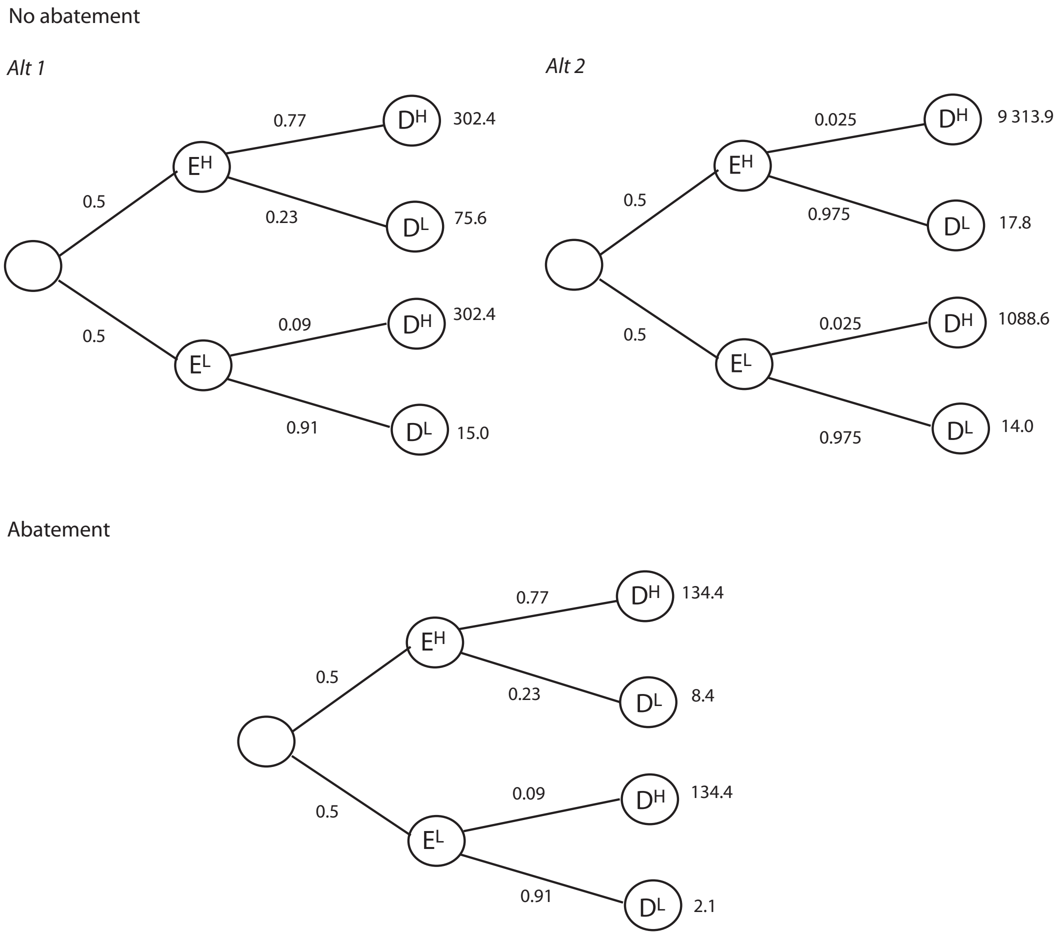


Figure 1: *Alternative scenarios for uncertain choice with equal expected damage costs and two policy options*

---

[12] Confer with the brief treatment of the Ellsberg paradox and uncertainty aversion above.

First, let the policy choice of no abatement be described by alternative 1 in Figure 1. This decision tree indicates a fifty-fifty chance of `high' (H) and `low' (L) emissions. The expected value of damage costs with no action is assumed to be USD 146 billion. If emissions turn out to be high, there is a considerable chance (77 percent) for a relatively high damage of USD 302.4 billion, and there is a probability of 23 percent for relatively low damages of USD 75.6 billion. If emissions instead turn out to be low, there is a 91 percent probability of an even lower damage cost of USD 15 billion, and only a 9 percent chance of the same high damage cost as with high emissions.

The results from a policy choice of abatement are described by the bottom decision tree in Figure 1. In this example, abatement reduces the damage regardless of what happens, without affecting the probabilities. The result is that the expected damage is reduced to 60 US billion and the spread of possible outcomes is less. Without abatement, the possible range of outcomes is between USD 15 and 302 billion, while abatement narrows the range to be between USD 2 and 134 billion.

A risk-neutral decision-maker will not abate if the cost of abatement measures exceeds the expected gains, USD 146 - 60 = 86 billion. Note that even the risk-neutral decision-maker runs the risk of investing in vain because the probability that damage turns out lower than USD 86 billion is 57 percent. In other words, it is likely that the abatement measures eventually result in a loss. If risk averse, the decision-maker is actually willing to pay more than USD 86 billion as an insurance against significant loss, should damages turn out to be high. This is because abatement also narrows the range of possible outcomes.

Now consider alternative 2 in Figure 1, which is very different from alternative 1, but where the expected damage costs are the same: USD 146 billion. There is still a fifty-fifty chance of high and low emissions, but now both cases have the outcome of low damage costs from global warming as the most likely (with a probability of 97 percent). This can be due to an expectation of people being able to adapt to climate change without severe problems. Nevertheless, there is also a small chance that the damage costs become severe, for example because it is more or less impossible to adapt. In this case the damage costs become extreme if future emissions turn out to be high and severe even if emissions are low.

Assume, moreover, that abatement leads to the same probability tree as in the former alternative. This means that the effect of abatement is different from the first alternative, because now abatement affects both the spread of outcomes and the probabilities (endogenous risk). Hence, for alternative 2, abatement implies that a catastrophe is avoided, but the abatement costs will most likely turn out to be a waste of resources. The question is now what to do in the second alternative. Since abatement affects the expected cost of climate change equally in alternative 1 and alternative 2, a risk-neutral decision-maker will again choose action if abatement costs are lower than USD 86 billion.

As mentioned above, a large number of experiments have shown that people do not consider compound lotteries with identical expected utility values to be identical (Machina 1989; Camerer 1995). Machina (1989) identifies two effects. First, people become more risk averse if the spread of outcomes increases. The above decision tree was tested in a small audience. The results showed a the tendency to accept higher abatement costs in alternative 2 than in alternative 1. Second, a linear transformation of the probabilities of a pair of prospects tends to change the preferences from the `high probability, low gain' to `low probability, high gain' if the high probability is reduced. This effect is consistent with the first effect, but relates to a change in probabilities rather than a change in outcomes.

It is important to be aware of the axioms on which expected utility builds and that they may be violated when tested against behavior. However, this is not the same as claiming that maximization of expected utility is useless or ‘wrong’. It may limit the applicability of the expected utility model, but on the other hand the expected utility model may also help decision-makers make better decisions. People often make systematic errors. Experiments with the Monty Hall Game[13] show one example. If that is what is going on here, then policy should not reinforce intuitive but destructive judgments, but rather strive to be rational. It is important to emphasize that systematic violations are usually identified in extreme cases, where the probabilities are small and outcomes are large. In many cases, this weakens the criticism of expected utility, although probably not in the case of climate change, where it is difficult to discard the possibility of catastrophic events. The issue in this discussion, however, is not about preferences regarding uncertainty, but rather about how to process information about uncertainty. Let us thus return to the discussion of decision criteria.

## 4.2 THE PRINCIPLE OF INSUFFICIENT REASON

The principle of insufficient reason, first formulated by Bernoulli in the 17th century, states that if there is no evidence leading us to believe that one event is more likely to

---

[13] see for example http://www.stat.sc.edu/˜west/javahtml/LetsMakeaDeal.html

occur than another, then the events should be judged equally probable. Applied to the example presented in Figure 1, the probability of each outcome would be 0.25, and the decision rule would be abate greenhouse gases if the abatement costs $y \leq 104$ for alternative 1 and abate if $y \leq 2549$ for alternative 2, which of course in both cases is less restrictive than for expected utility maximization, and the tendency to abate is thus higher.

### 4.3 MAXIMIN (MINIMAX) PRINCIPLE

The *maximin* (maximum minimorum) criterion ranks alternatives by their worst possible outcomes, and adopts the alternative for which the worst possible outcome is preferable to all other worst outcomes (Rawls 1971). Similarly, the *minimax* principle applies when disutility is considered; that is when the decision-maker attempts to minimize the maximum loss from taking action.

The maximin principle implies maximization of the welfare in the worst possible case, and essentially it allows risk aversion to become infinite. Applied to the problem of climate change, it can be interpreted as the following. The decision-maker chooses the level of abatement, A, that maximizes the social welfare, w, in the worst possible state of nature, chosen from all possible states, s:

$$\max_A \left\{ \min_s w(A, s) \right\}$$

The example presented above allows only two alternative levels of abatement: the decision-maker decides either to use resources to try to prevent global warming (abatement) or to do nothing (no action). The choice between the two will naturally depend on the magnitude of abatement costs, y, and on the decision-maker's beliefs about possible future damages[14]. Generally, the decision-maker will choose abatement if the costs are less than the benefits, namely the damage prevented. If he believes in alternative 1 (see Figure 1), where choosing abatement will lower the damage in all cases without affecting the probabilities, he will choose abatement if $302.4-134.4-y \geq 0$, that is, if $y \leq 168$ billion USD. If he believes in alternative 2, where choosing abatement affects the spread of both outcomes and probabilities, he will choose abatement if $y \leq 9180$ billion USD. Thus the condition for imposing measures to reduce emissions is less restrictive with the maximin criterion, compared to the expected utility criterion where the critical value was 86 billion USD. This reflects the criterion's emphasis on the worst case scenario, and it is particularly obvious in alternative 2 where there is a small chance of a catastrophic outcome under a no action policy.[15]

### 4.4 MINIMAX REGRET

Minimax regret (risk/loss), suggested by Savage (1951) as an improvement on the maximin criterion, aims at minimizing the difference between the best that could happen and what actually does happen (see also Fishburn 1987). The decision-maker tries to minimize regrets for not having, in hindsight, made the superior choice. In the global warming context, this could be interpreted as choosing the option which the decision-maker believes future generations would least regret:

$$\min_{y}\left\{\max_{s}\left[w\left(A^{*}(s),s\right)-w(A,s)\right]\right\}$$

---

[14] It is important to note that abatement costs are encumbered with just as much uncertainty as the damages, although they are treated in the example as fixed for the sake of simplicity.

[15] For resent treatments of maximin, see e.g. Bouglet and Vergnaud (2000), and Chevé and Congar (2002).

Let $A^*(s)$ be the optimal choice of abatement if we knew for certain that s would occur. The term A(s) represents the level of abatement that would have been chosen if we knew in advance what would happen in the future. Since s is unknown ex ante, we will most likely choose a different level of abatement, denoted A. The minimax regret criterion says that the ex ante choice of A should be chosen in order to minimize the future loss in the case where your decision is as bad as it can get. This averts the chance of making a major mistake. In order to find the preferred policy, the decision-maker must predict and compare the maximal regrets for all possible policy choices.

To more easily grasp the concept of this criterion, consider first the following simple utility payoff and regret matrix:

| payoffs | $s_1$ | $s_2$ | | regrets | $s_1$ | $s_2$ |
|---|---|---|---|---|---|---|
| $a_1$ | 0 | 100 | | $a_1$ | 1 | 0 |
| $a_2$ | 1 | 1 | | $a_2$ | 0 | 99 |

In terms of payoffs, the result is 0 if $s_1$ occurs and 100 if $s_2$ occurs when choosing $a_{1,}$ and the payoff is 1 in both states if $a_2$ is chosen. In terms of regrets, choosing alternative $a_1$ has the possible maximum regret of 1 because if $a_1$ is chosen and $s_1$ is realized the payoff is 0, but it could have been 1 if $a_2$ was chosen instead. If $a_1$ is chosen and $s_2$ is realized the result is the maximum payoff and there is nothing to regret. Choosing alternative $a_2$ has the possible maximum regret of 99. This is because if $a_2$ is chosen, there is nothing to regret if $s_1$ occurs, but a regret of 99 if $s_2$ is realized. Consequently, alternative $a_1$ minimizes the maximum regret, and should therefore be chosen according to this criterion.

For the numerical global warming example assumed in this paper, the minimax regret cost matrix is shown in Table 1. The numbers in the matrix are found in the following way: Take for instance the possible state of high emissions and high increase in temperature (HH) in Figure 1. The maximum regret will depend on the abatement cost y. For alternative 1, the regret will be the net benefit from abatement; the damage when no action is taken, 302.4, minus the damage that is faced when action is taken, 134.4, and minus the cost of abatement, y, namely 168-y. Or, if y is sufficiently large, the regret will be zero. For sufficiently large costs, abatement should not be chosen and the decision-maker should have no regrets when choosing no action. The regret given alternative 1 with no abatement is thus the maximum of 168-y and 0. If instead the state of high emissions and low increase in temperature (HL) occurs, the regret would be the maximum of 67-y and 0, etc. The maximum regret for a strategy is found by comparing the regrets in each state. If we believe in alternative 1 and consider choosing no action, the maximum regret would thus be the maximum of 168-y and 0.

| States/ Beliefs + Strat. | HH | HL | LH | LL | Max Regret |
|---|---|---|---|---|---|
| Alt. 1, NA | max(168-y,0) | max(67-y,0) | max(168-y,0) | max(13-y,0) | max(168-y,0) |
| Alt. 1, A | max(0,y-168) | max(0,y-67) | max(0,y-168) | max(0,y-13) | max(0,y-13) |
| Alt. 2, NA | max(9180-y,0) | max(9.4-y,0) | max(924-y,0) | max(12-y,0) | max(9180-y,0) |
| Alt. 2, A | max(0,y-9180) | max(0,y-9.4) | max(0,y-924) | max(0,y-12) | max(0,y-12) |

Table I: *The game's maximum regret matrix*

Given alternative 1 and abatement, the regret in the state of HH is the maximum of 0 and y-168; for sufficiently high costs the decision-maker will regret choosing abatement, and the regret will be the abatement cost minus the avoided damage. Thus, if we believe in alternative 1 and consider choosing abatement, the maximal regret would be the maximum of 0 and y-13. The numbers for alternative 2 are found in a similar fashion.

The preferred policy is found by comparing the maximal regrets and choosing the minimum. The result will naturally depend on y, and for alternative 1 in the example the strategy has to satisfy:

$$\min_A \left\{ \max_s (168 - y, 0), \max_s (0, y - 13) \right\}$$

When applied to our example the criterion says that if y ≤ 168, the decision-maker should choose abatement. For alternative 2, the critical value for choosing abatement is y ≤ 9180. These are the same critical values as for maximin.[16]

### 4.5 GENERALIZED MAXIMIN/MAXIMAX

The maximin and the minimax regret criteria are each very conservative in the sense that they put emphasis on the worst case scenario. What if we look at the best state, or a combination of the best and the worst? This is the essence of *the pessimism-optimism index criterion* of Hurwicz (1951), also known as generalized maximin/maximax. This criterion states that the level of abatement should be chosen in order to maximize a weighted average of the social welfare in the best and the worst state:

$$\max_A \left\{ \alpha \min_s [w(A, s)] + (1 - \alpha) \max_s [w(A, s)] \right\}$$

One interpretation of the size of the pessimism-optimism index, α, is that it reflects the decision-maker's beliefs about the likelihood of facing the worst case in the future. The generalized maximin/maximax is thus a simplification of the expected utility criterion. When determining α, the decision-maker can either find the best and the worst case with

---

[16] For other discussions on this criterion, see e.g. Chernoff (1954), Crisholm and Clark (1993), and Chevé and Congar (2002).

appurtenant subjective probabilities, or divide all possible states into two groups: the bad states and the better states, and α will be the probability of realizing one of the bad states. For the numerical global warming example assumed in this paper this means picking out the best and the worst case for both alternatives.

| α | 0.0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|
| Alt. 1: Abatement if y ≤ | 13 | 38 | 59 | 90 | 121 | 152 | 168 |
| Alt. 2: Abatement if y ≤ | 12 | 929 | 2762 | 4596 | 6429 | 8263 | 9180 |

Table II: *Abatement criteria for alternative choices of* α under the generalized maximin/maximax criterion

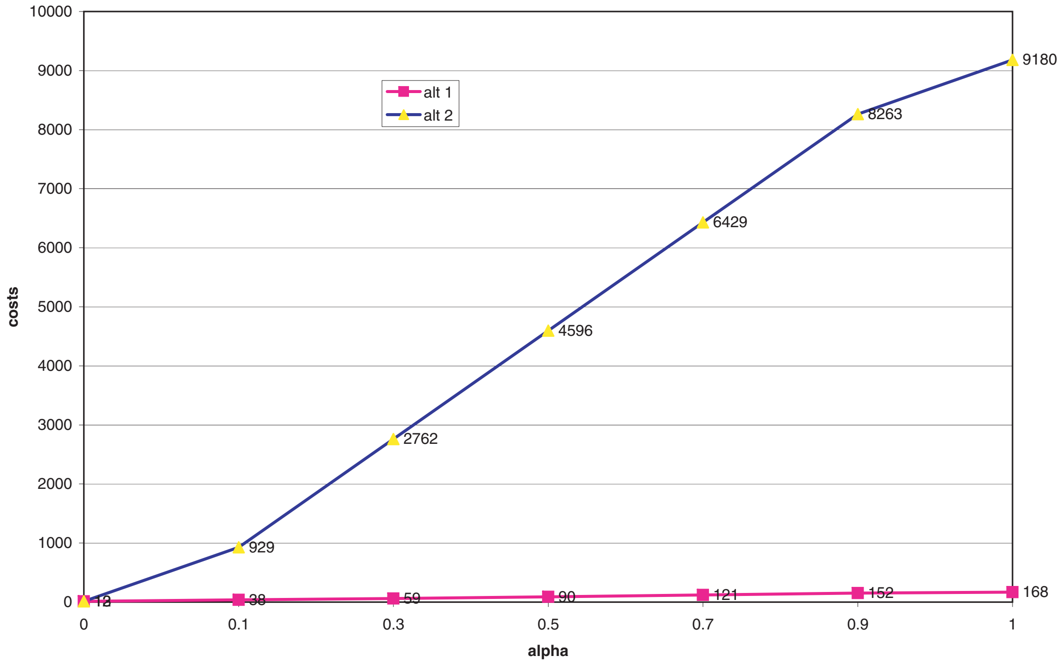


*Figure 2: Critical values for the generalized maximin/maximaxcriterion*

The critical value of abatement costs will obviously depend on the size of α, and is shown in Table II and in Figure 2. As can be seen, for alternative 1, the remedial policy will be implemented if y < 13 billion USD if α =0 (no weight is put on the worst case), and if y < 168 billion USD if α =1 (all weight is put on the worst case, which coincides with maximin). The critical value of y in alternative 2 ranges from 12 to 9180.

Choosing the value of α that corresponds to the probability distribution for alternative 2 in Figure 1 yields α = 0.5 * 0.025 = 0.013. The generalized maximin/maximax criterion then prescribes abatement if y < 126 billion USD, which again is less restrictive than the expected utility criterion's 86 billion USD. If we use alternative 1's corresponding weight, however, α = 0.43, we get the slightly stricter critical value y < 80. Obtaining the expected utility level for the critical y requires a belief about α of 0,47 for alternative 1, and 0,008 for alternative 2. Thus, if alternative 2 applies, the critical value of imposing climate change mitigation is less restrictive than with expected utility if the probability of a catastrophe is less than 0.008. This illustrates how sensitive the generalized maximin/maximax criterion is to the possibility of catastrophic events. Recall that the probability of the worst case in alternative 2 is 0.0125, which is more than one and a half times 0.008.

### 4.6 LIMITED DEGREE OF CONFIDENCE

The limited degree of confidence criterion implies that the decision-maker maximizes a weighted sum of the expected utility criterion and the maximin criterion:

$$\max_A \left\{ \gamma E_s \left[ w(A,s) \right] + (1-\gamma) \min_s \left[ w(A,s) \right] \right\}$$

The decision-maker's degree of confidence in the probability distribution underlying $E_s w$ is measured by $\gamma \in [0,1]$ and for $\gamma < 1$ more weight is put on the worst case than a under standard expected utility maximization. In the case of full confidence, γ = 1, the expected utility criterion is used, whereas under complete uncertainty, $\gamma = 0$, the maximin decision rule is applied. The discussion above showed that the conditions for imposing measures to reduce emissions are less restrictive using the maximin criterion, compared to the expected utility criterion, due to the maximin criterion's emphasis on the worst case scenario. Thus, $\gamma$ represents the uncertainty about the uncertainty, or alternatively the degree of ignorance. In this interpretation, ignorance is simply the counterpart of confidence in the probabilistic assessment underlying the expected utility calculation (Eichberger and Spaniels 1998; Stigum 1990).

| $\gamma$ | 0.0 | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|
| Alt. 1: Abatement if $y \leq$ | 168 | 160 | 143 | 127 | 110 | 94 | 86 |
| Alt. 2: Abatement if $y \leq$ | 9180 | 8270 | 6451 | 4633 | 2814 | 995 | 86 |

Table III: *Abatement criteria for alternative choices of* $\gamma$ under the *limited degree of confidence criterion*

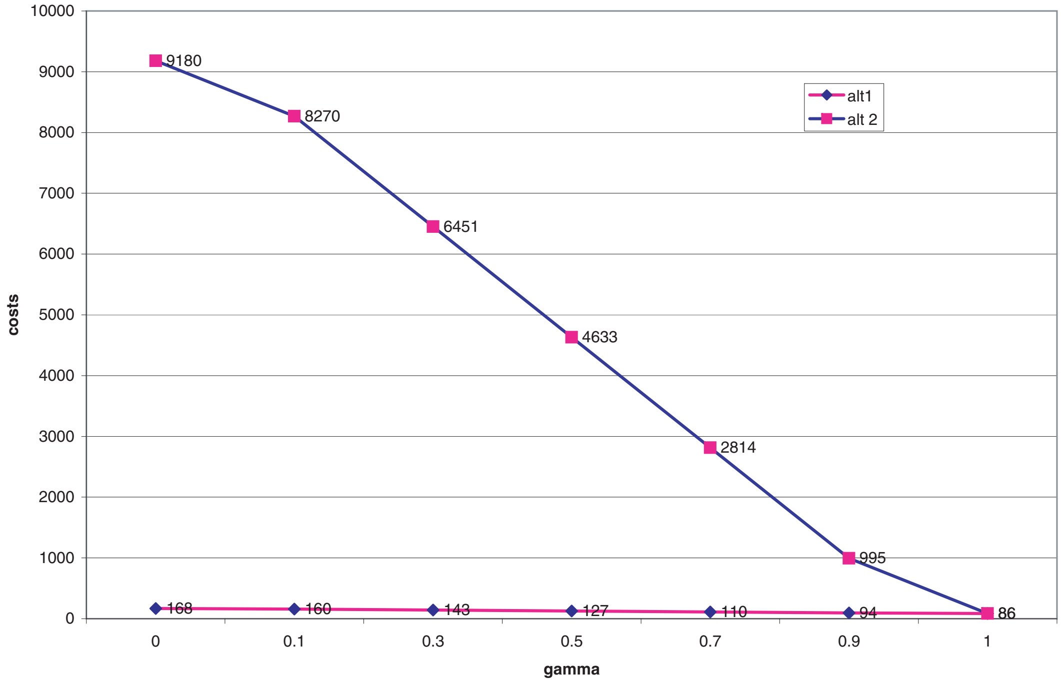


Figure 3: *Critical values for the limited degree of confidence criterion*

Returning to the example, it can be seen from Table III that the critical value of abatement costs spans from the maximin value to the expected utility value and the lower the confidence in the probability distribution, the higher the willingness to pay for emission reductions. This is also illustrated in the graph in Figure 3.

In Lange (2003) the limited degree of confidence criterion is compared to expected utility maximization in a two period model with learning.[17]

## 4.7 SAFETY FIRST

Another way of showing precaution is by lowering the probability that the welfare level in the future will be too low (Rawls 1971). Two alternatives are:

$$\min_A \left\{ \Pr_s \left[ w(A,s) \leq k \right] \right\}$$

$$\max_A \left\{ E_s w(A,s) \right\} s.t. \left\{ \Pr_s \left[ w(A,s) \leq k \right] \leq \beta \right\}$$

The first alternative, a probabilistic non-expected utility criterion, minimizes the probability of the welfare being less than some constant k. The second alternative maximizes the expected social welfare, subject to a constraint, which says that the probability of welfare being less than level k should be less than β.

Another safety-first approach is to have constraints on the welfare of the worst off group, land or region. One example is:

$$\max_A \left\{ E_s w(A,s) \right\} s.t. \left\{ \Pr_s \left\lfloor \min_i w_i(A,s) \leq k \right\rfloor \leq \beta \right\}$$

Now the decision-maker maximizes the expected future social welfare subject to the probability of the welfare of the worst off, being less than β. If β = 0 the constraint can

---

[17] Chichilnisky (2000) discusses this criterion in a static setting.

be interpreted as a floor. If k is chosen properly, this criterion will then ensure expected intergenerational equity and sustainable development. Loosely put, sustainable development demands that each generation use no more than their legitimate share of the world's resources (Asheim 1993). If one or several generations consume more than their legitimate share, or their consumption generates too much pollution, the consumption possibilities of later generations will be undermined. Later generations might then not be able to reach the social welfare level k, and the criterion is violated. This is also known as the highest constant-consumption criterion (Solow 1974).

### 4.8 APPLICABILITY OF ALTERNATIVE DECISION CRITERIA

The use of any of these criteria as a basis for decision-making requires ideally the decision-maker to have complete awareness of the possible states of nature. In a way they require what Savage would call ‘a closed universe’. Nothing genuinely surprising can be expected to happen. Given the uncertainty structure of the climate problem, this is a very restrictive assumption. Most of these criteria, however, focus on one or a few of the possible outcomes, so the information on the rest need not be as accurate. The theory of *expected utility* demands, in a way, more detailed information than most of the other criteria discussed here, and also an opinion on the probabilities. Yet, expected utility theory may provide an adequate framework for explaining and predicting social choice in uncertain situations. The information problems associated with this approach are not necessarily avoided by using other approaches. Furthermore, even if individuals place higher weights on low-probability extreme events, this does not mean that the decision-maker necessarily will do the same.

It is important to note that the conclusions presented above are highly dependent on the framing of the problem, i.e. how states of nature are defined, and in the case of expected utility, on the subjective beliefs about the probability distribution over states. An obvious framing problem with the *principle of insufficient reason* for example is that the number of states considered will influence the decision: Consider a case where the policy choice depends on the possibilities that a) there will be no damage from global warming, or b) there will be severe damage. Since what will happen is unknown, the principle of insufficient reason says that we should use a probability of 1/2 on both states. Second, consider the case where three possibilities are considered, namely a) no damage from global warming, b1) severe damage from sea-level rise, and b2) severe damage from extreme weather events. If the principle of insufficient reason is applied, there is a probability of 1/3 on each state. Compared with the first framing of the problem, however, the probability of state a) has decreased from 1/2 to 1/3, and the probability of b) has increased from 1/2 to 2/3. Many such listings are possible, and in general these different abstractions of the same problem will, when resolved by the principle of insufficient reason, yield different solutions (Luce and Raiffa 1957).

Using *maximin* to decide upon abatement policy reflects an extreme fear of the worst case. Its conservatism, however, might make good sense in some contexts. Chevé and Congar (2002), for instance, argue that maximin complies with the precautionary principle, however, knowing the actual nature of the worst case is problematic. What *is* the worst that can happen as a consequence of anthropogenic climate change? Answering this question is difficult because of uncertainties about the strength of climate change impacts, and whether and how they will materialize. Another problem is that the worst case might be a catastrophe of such dimensions that deciding between two policy options might have no significant impact on the outcome. Furthermore, if there is a tiny chance that the abatement strategy is ineffective, the worst case scenario would be to implement a costly remedial policy that fails to avert severe damages. This means that unless there is complete certainty regarding the policy effectiveness (which is assumed in this example), abatement cannot be rationalized as the appropriate maximin strategy in any game against nature.

The following simple illustration exhibits an obvious objection to the maximin criterion:

Matrix I

| payoffs | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 100 |
| $a_2$ | 1 | 1 |

Matrix II

| payoffs | $s_1$ | $s_2$ |
|---|---|---|
| $a_1$ | 0 | 1,000,000 |
| $a_2$ | 0.000001 | 0.000001 |

For matrix I, the maximin criterion favors $a_2$ over $a_1$. Some consider this unreasonable, because it would still be true if 1 were reduced to 0.000001 and the 100 increased to 1,000,000 as is the case in matrix II. The critics agree that since nature is considered a non-strategic player, clearly $a_1$ must be chosen when matrix II is the problem (Luce and Raiffa 1957).

*Minimax regret* has been criticized first on the grounds that it has never been clearly demonstrated that differences in utility do in fact provide a measurement of regret. In other words, it is not clear that the `regret' of going from a state of utility 5 to a state of utility 3 is equivalent to that of going from 15 to 13. Second, one can construct examples where an arbitrarily small advantage in one state of nature outweighs a considerable advantage in another state. Such examples tend to produce the same feeling of uneasiness that is explained in the previous paragraph on maximin. A third objection is that in some examples the minimax regret criterion may select a strategy $a_3$ among the possible strategies $a_1$, $a_2$, $a_3$, and $a_4$. If, for some reason, $a_4$ was made unavailable, this might lead to the choice of $a_2$ instead. And the criticism is thus that the presence of an undesired strategy should not influence the choice among the remaining strategies (Chernoff 1954).

The minimax regret rule is also criticized on the grounds that regret consists of "crying over spilt milk", which may not be the best way to optimize. On the other hand, applying the notion of regret to the maximin rule gives some weight to the relationship between the costs of implementing a remedial policy and the loss of doing nothing when the damages caused by global warming turn out to be significant. This holds even when there is uncertainty about whether the abatement policy will be effective (Chrisholm and Clark 1993). However, the uncertainties regarding impacts and the results of abatement efforts make it difficult to define and measure the regrets. When using this criterion, the decision-maker has the same need for information as with the expected utility criterion, but does not need to assign probabilities to the various outcomes.

Minimax regret might be particularly relevant when policies serve dual purposes. In many instances, actions to combat local environmental problems, protect biodiversity, and so forth will simultaneously have a desirable impact on global warming (Ekins 1996; Aunan et al. 1998). Conversely, policies aimed at combating an enhanced greenhouse effect, such as reducing forms of air pollution caused by the use of fossil fuels, and reducing the clear-felling of forests, will also often help alleviate local environmental problems. These dual purpose policies are particularly attractive if there is a significant chance that either a greenhouse problem will not occur, or that human preventive action will be ineffective because *if* the policy should fail to remedy global warming, it will still reduce local environmental problems, and the abatement effort has not been a sheer waste. If the ancillary benefits are sufficiently large, the costs of implementing the abatement policy will be outweighed. If this is the case, we will have $-y > 0$ and the remedial policy will always be preferred, regardless of its effect on climate change, and regardless of the criterion chosen. No-regret policies might thus provide a way for the politicians to minimize the ex post critique.

The *safety-first* criteria pose questions about how to determine welfare levels, $k$ and probability, $\beta$. What determines the acceptable size of $\beta$, the probability of the future welfare being less than $k$? How likely can it be? This question might depend on the size of $k$. How low a welfare can be accepted? This is again dependent on whether or not $k$ measures the welfare on average. If so, is this a weighted average, and how then are the weights determined? And what level of aggregation is applied? Is $k$ a weighted sum of each group, land or region? In addition, even if $k$ and $\beta$ are determined in the face of the uncertainty, how can the decision-maker know that a certain policy decision will satisfy the constraints? The typical economics approach to climate change focuses on total or average effects. Nordhaus for example aggregates impacts in his analyses of what to do about this problem, but people who are worried about climate change focus on specific outcomes like flooding in Bangladesh or loss of coral reefs. Is it right to average over total effects? Or does it matter that some things are being harmed even while other things are being helped by climate change?

The class of safety-first criteria has information challenges in addition to the problems of deciding upon the strictness of restrictions touched upon above. The idea of safeguarding the worst off from too low a level of welfare seems sensible. However, in extreme cases, this could mean that one would have to lower the welfare of the majority substantially in order to marginally benefit the worst off, and no politician could do that without losing popularity. One possible solution to this problem is side payments. The worst off could get transfers from the majority which would increase the utility for all parties in comparison with strict safeguarding.

Using the *highest constant-consumption* criterion, the decision-maker will seek to pick an efficient path. He will ask for the largest steady per capita consumption that can be maintained indefinitely (the maximum steady state). However, predicting and finding such a path demands pretty much the same amount of information as the calculation of expected future welfare, and will thus be highly uncertain.

A common criticism of criteria such as maximin, minimax regret, generalized maximin/maximax, and the principle of insufficient reason is that they are rationalized on some notion of ignorance. In practice, however, the critics claim that the decision-maker usually has some vague partial information concerning the true state of nature

(Luce and Raiffa 1957). The question remains, however, if this vague partial information is enough to assign subjective probabilities to all possible, or at least sufficiently many, states of nature - and for a case like global warming, whether the available information is enough even to apply non-probabilistic criteria.

The principle of insufficient reason, maximin, generalized maximin/maximax, and the limited degree of confidence criteria can be looked upon as simplifications of the expected utility criterion. What separates them is which states are given the larger weights, and also the type and extent of information they demand. They all, however, demand less information than the expected utility criterion, since probabilities are not required. For the limited degree of confidence criterion, an unsatisfactory level of information about the probability distribution over states is compensated for by putting extra weight on the worst case scenario. These criteria, as well as the expected utility criterion, all attempt to control the outcome. The safety-first criteria have a different goal: to control the uncertainties. The minimax regret criterion's concern is that of making mistakes, and may therefore be the criterion that best reflects the political process. In fact, the tendency of people to systematically violate the independence axiom in experimental settings might indicate that they are using a form of the minimax regret criterion when the stakes are high. In the example above, this implies that people are willing to accept higher abatement costs in alternative 2, where a catastrophe is avoided with abatement, than they do when comparing alternative 1 with the abatement case[18].

---

[18] This is consistent with the result we got when it was tested in a small audience.

## 4.9 THE TIME DIMENSION

Unlike most other political issues, climate policy has an extremely long time horizon. Nearly half of the greenhouse gases will remain in the atmosphere for a hundred years and some of them for thousands of years. In addition it is likely that estimates of impacts and benefits will improve considerably over the next century, but we will never reach full knowledge. In the example discussed above, we make a decision ‘today’ and know ‘tomorrow’. In a more realistic description of climate policy, the decision would be a process where adjustments were made according to new information. A possible way to model the process would be that the outcomes of the lottery we consider today are tickets to new lotteries with different probabilities.

The long-term nature of climate change due to the fact that it is the concentrations of greenhouse gases that matters, rather than annual emissions, raises the thorny issues of intergenerational transfers of wealth and environmental goods and bads. This again raises the question of discounting[19]. One argument for discounting over time is that consumers are impatient; they have a preference for immediate over postponed consumption. A second is that the marginal utility of consumption decline with growing per capita consumption. One could question, however, the validity of the first argument, when discussing discounting in the light of the climate problem. In this case we are talking about consumption hundred or several hundred years ahead. It might be difficult to justify the impatience argument, considering the time horizon. It is not the same generation consuming today and a hundred years from now, and some might

---

[19] This is a widely discussed issue. Some examples are d'Arge et al. (1982), Azar and Sterner (1996), Howarth (1996, 2004), Jones-Lee and Loomes (1995), Khanna and Chapman (1996), Pearce et al. (2003) Schelling (1993), Toth (1995), and Weitzman (1994, 1998, 2001).

consider it immoral to state that our consumption today is more important, or more valuable, than the consumption of the generations to follow.

The second argument in favor of dsicounting follows from an assumption about continuing future growth, which implies increased future consumption. Assuming decreasing marginal utility, this future consumption will be of less value, on the margin. This means that, given continued growth, the last unit consumed in the future has less utility value than the last unit consumed today. The generation that undertakes the abatement, however, is not the one who will consume the possible long-term benefits from the reduction in emissions of greenhouse gases. Thus, if the economic sacrifices are carried by the countries that can best afford it; the transfers will tend to be from people in the developed part of the world, to people in the less developed parts. Hopefully the residents of the developing countries will be far better off a century from now than they are today. However, they may not yet be as well off then, as people in the developed part of the world is today (Schelling 1993).

### *4.9.1 Time inconsistency*

One of the responsibilities of present policy makers is to take care of the interests of future generations. It is a common experience, however, that people tend to postpone unpleasant tasks, preferring to have them done in the next period, even if it might be optimal to complete the task now. This might also apply to policy makers, and will probably reflect time-inconsistent preferences.

A decision rule is said to be time-consistent if at each decision node reached when the rule is followed, the decision rule is still optimal in the sense of maximizing the welfare as evaluated at the reached node (Asheim 1997). Considering the climate issue, policy makers' preferences (and they might also reflect the preferences of the present generation) seem to give extra weight to current welfare over future welfare. These present-biased-preferences might lead to procrastination of climate change remedial action, a decision that might be time-inconsistent. The incentives to procrastinate stem from the fact that the policy makers plan to do a task based on its long-run benefits, and these benefits are strongly dependent on whether they accept taking on the costs of greenhouse gas abatement today[20].

## 4.10 SCOPE FOR FURTHER WORK

The simplicity of a presentation like the above always raises the question of whether the results would stand up if the analysis introduced greater complexity. A natural direction to go would be to expand the analysis to take into account change (or lack of change) in certain variables over time. How will the irreversibilities affect decision-making in a dynamic framework? Would we abate now or would we wait and see what happens? Taking into account the possibility of change (or lack of change) over time will allow us to consider how future learning might influence decision-making. Then the problem of climate change evolves as a subject of risk management in which strategies should be reformulated as new knowledge arises (IPPC 2001). This will give an opportunity to analytically address robustness. Robust strategies[21] are closely tied to the idea of adaptive decision strategies, that is, strategies that can change over time in response to observations of the climate and economic systems. (Adaptive decision strategies differ from sequential strategies in that in the former information is endogenous, that is, the

---

[20] For a contribution on incentives for procrastination see O'Donoghue and Rabin (1999).
[21] See for example Lempert and Schlesinger (2000).

type, rate, and quality of information gained depends on both the unfolding scenario and policy choices whereas information is exogenous in the sequential strategies.) Adaptive decision strategies are closely tied to the concept of robustness, because such strategies are most useful in situations of deep uncertainty–where robustness, as opposed to optimization, appears to be the best decision-making criterion (IPCC 2001).

A second possible direction could be to check how the results are affected by the ability to make continuous choices, i.e. abate 30%, 31% etc. Today the policy question is not so much about whether to abate, but rather how much to abate.

A third direction could be to check how the results depend on an expanded strategy set. For example, instead of only looking at a decision about whether to abate, the analysis could consider a choice of investing in research and development that would lower the cost of abating in the future, or alternatively, investing in expanding future consumption to compensate for future climate change damages.

## 5. Conclusions

The information needed to make better climate policy decisions depend, on the one hand, on which decision-making criterion is chosen. On the other hand, the amount and type of information available may help determine which decision criterion should be used. When selecting a decision criterion, one of the very first things that needs to be determined is the target of control. Then the preferred policy choice depends very much on the framing of the problem: how states of nature are defined, subjective beliefs about the probability distribution over states, and the magnitude of costs. Furthermore, the answer to whether more uncertainty should imply more or less abatement to a large extent depends on the specification of relationships. Different functional forms, i.e. linear vs. non-linear, will give different results[22]. The outmost consequence of this is, of course, that the skilled modeler can get any result he desires. It is therefore of great importance that the framework for analysis of climate policy is as realistic as possible.

Focusing on the lack of certainty regarding impacts of greenhouse gas emissions and the effectiveness of policy made a comparison of alternative criteria with the maximization of expected utility model possible: for maximin and minimax regret the condition for imposing measures to reduce emissions of greenhouse gases are less restrictive, that is, the costs of abatement (y) can be higher, compared with the critical value for y under maximization of expected utility; the generalized maximin/maximax criterion is less restrictive under alternative 2 where a possible catastrophe is avoided, but slightly more restrictive under alternative 1; and for the limited degree of confidence criterion the conditions are equal to or less restrictive depending on the degree of ignorance (See Table IV).

| | Decision Criterion | Alt. 1 Choose abatement if | Alt. 2 Choose abatement if |
|---|---|---|---|
| EU (risk neutral) | $\mathrm{Ew} = \sum_{\mathrm{s}} \mathrm{p_s w(A_s)}$ | $\mathrm{y} \leq 86$ | $\mathrm{y} \leq 86$ |
| Minimax | $\max_A \left\{ \min_s w(A,s) \right\}$ | $\mathrm{y} \leq 168$ | $\mathrm{y} \leq 9180$ |
| Minimax Regr. | $\min_y \left\{ \max_s \left[ w\left(A^*(s), s\right) - w(A,s) \right] \right\}$ | $\mathrm{y} \leq 168$ | $\mathrm{y} \leq 9180$ |
| Gen. Minimax | $\max_A \left\{ \alpha \min_s \left[ w(A,s) \right] + (1-\alpha) \max_s \left[ w(A,s) \right] \right\}$ | $\mathrm{y} \leq 80$<br>$(\alpha = 0.43)$ | $\mathrm{y} \leq 126$<br>$(\alpha = 0.013)$ |
| Lim. D. of Conf. | $\max_A \left\{ \gamma E_s \left[ w(A,s) \right] + (1-\gamma) \min_s \left[ w(A,s) \right] \right\}$ | $\mathrm{y} \leq [168,86]$<br>$\gamma = [0,1]$ | $\mathrm{y} \leq [9180,86]$<br>$\gamma = [0,1]$ |

---

[22] This is discussed further in Aaheim and Bretteville (2001).

Table IV. *Critical cost values for choosing abatement*

Differences in recommendations on climate policy might thus be due to the use of different criteria as basis for analysis. Those arguing in favor of abatement policies might use one of the alternative criteria as basis for their analysis, rather than maximization of expected utility. Furthermore, unless there is complete certainty regarding the policy effectiveness, maximin can be used to back up a no-action strategy. This is because if there is a tiny chance that the abatement policy is ineffective, the worst case scenario would be to implement a costly remedial policy that fails to avert severe damages.

The choice of a given decision criterion is not independent of the attitude towards climate risks and the amount, type of, and confidence in the available scientific information. The complexity of the climate problem makes it very difficult, if not impossible, to completely overlook the consequences of alternative choices and how they affect uncertainty. The result may be that people cannot say whether they prefer maximization of expected utility or an alternative decision rule as a basis for climate policy making. It is therefore hard to determine whether decision criteria such as the maximization of expected utility model, which demands high levels of information, represent decision-makers' preferences properly. This may be an argument for the use of alternative rules that demand less information and enable the decision-makers to better predict the consequences of their choices. It is important to emphasize that the tendency to violate the expected utility axioms is particularly strong in extreme cases, where there are small probabilities of extreme outcomes. For the case of global warming, it is difficult to discard the possibility that at least one disastrous event will occur, so expected utility may fail to reflect the preferences of decision-makers. However, whether the benevolent decision-maker *should* use the expected utility criterion is another discussion.

The choice of decision criterion is thus a political question. The decisions of today will influence the consumption possibilities of future generations, and long-term problems in particular should be treated with this in mind. Colgazier (1991) claims that a good guide for making decisions under uncertainty is to hedge your bets. This requires finding a course of action that is reasonably robust in leading to positive outcomes and avoiding negative ones. This might point towards using the principle of expected utility maximization. However, when faced with a problem such as global warming, which is very long term and outside the range of human experience, the possibility of making huge mistakes should be of great concern, and hedging one's bets might require that the choice of climate policy is, at least partly, in accordance with the minimax regret criterion.

## Acknowledgements

I am grateful to Geir Asheim, Scott Barrett, Leif Helland, Larry Karp, Fredrik C. Menz, Kalle Moene, Ottar Mæstad, Jon Vislie, participants at the Second World Congress of Environmental and Resource Economists 2002, various seminar participants, anonymous referees, the editor of *Mitigation and Adaptation Strategies for Global Change*, and my colleagues at CICERO, in particular H. Asbjørn Aaheim for very helpful discussions and comments on earlier drafts. I also thank Lynn Nygaard for excellent editorial help, Tone Veiby for figure drawing assistance, Frode Rørvik for computer support, and the Research Council of Norway for funding. Part of this work was carried out during a years visit to the Department of Agricultural and Resource Economics at UC Berkeley who's hospitality is gratefully acknowledged. Responsibility for errors and opinions is my own.

**References**

Aaheim, H.A. and C. Bretteville (2001), 'Decision-Making Frameworks for Climate Policy under Uncertainty'. CICERO Working Paper 2001:2. Oslo: Center for International Climate and Environmental Research – Oslo

Aaheim, H.A., K. Aunan and H.M. Seip (2000), 'The value of the environment: Is it a matter of approach?' *Integrated Assessment* **1** (1), 49-61.

Aumann, R. (1962), 'Utility theory without the completeness axiom'. *Econometrica* **30**, 445-462.

Aunan, K., G. Pátsay, H.A. Aaheim, and H.M. Seip (1998), 'Health and Environmental Benefits from Air Pollution Reductions in Hungary'. *The Science of the Total Environment* **212,** 245–268.

d'Arge, R.C., W.D. Schulze and D.S. Brookshire (1982) 'Carbon Dioxide and Intergenerational Choice'*The American Economic Review*, Vol. 72, No. 2, Papers and Proceedings of the Ninety-Fourth Annual Meeting of the American Economic Association. (May, 1982), pp. 251-256.

Arrow, K.J. (1951), 'Alternative Approaches to the Theory of Choice in Risk-Taking Situations, *Econometrica* **19,** No. 4, 404-437

Arrow, K.J. (1965), 'Aspects of the Theory of Risk-Bearing'. Helsinki: Yrjö Jahnsson's Lectures.

Arrow, K.J. and A.C. Fisher (1974), ‘Environmental Preservation, Uncertainty, and Irreversibility’. *Quarterly Journal of Economics* **88** (2), 312–319.

Asheim, G.B. (1997), ‘Individual and Collective Time-Consistency’. *Review of Economic Studies* **64,** 427–443.

Asheim, G.B. (1993), ‘Hvor bærer det hen? Utvikling av teori for en bærekraftig økonomi’. *Sosialøkonomen* (4).

Azar, C. and T. Sterner. ‘Discounting and Distributional Considerations in the Context of Global Warming.’ *Ecological Economics*, Vol. 19 (1996), pp. 169-184.

Bretteville, C. (1999), ‘Decision Criteria under Uncertainty and the Climate Problem’. CICERO Working Paper 1999:10. Oslo: Center for International Climate and Environmental Research – Oslo.

Camerer, C. (1995), ‘Individual Decision Making’ in J. Kagel and A. Roth, eds, *The Handbook of Experimental Economics*. Princeton: Princeton University Press.

Chernoff, H. (1954), ‘Rational Selection of Decision Functions,’ *Econometrica* **22,** 422–443.

Chichilnisky, G. (2000), ‘An Axiomatic Approach to Choice under Uncertainty with Catastrophic risks’, *Resource and Energy Economics* **22**, 221-231.

Chrisholm, A.H. and H.R. Clark (1993), ‘'Natural Resource Management and the Precautionary Principle’, in E. Dommen, ed., Fair Principles for Sustainable Development: Essays on Environmental Policy and Developing Countries. Hants, UK: Edward Elgar Publishing.

Colgazier, E.W. (1991), ‘Scientific Uncertainties, Public Policy, and Global Warming: How Sure is Sure Enough?’ *Policy Studies Journal* **19** (2), 61–72.

Eichberger, J. and W. Spaniels (1998), ‘Liquidity and Uncertainty: Banks or Asset Markets?’ Mimeo. Germany: University of Saarland.

Ekins, P. (1996), ‘How Large a Carbon Tax is Justified by the Secondary Benefits of $CO_2$ Abatement?’ *Resource and Energy Economics* **18,** 161–187.

Ellsberg, D. (1961), ‘Risk, ambiguity, and the Savage axioms’. *Quarterly Journal of Economics* **75**, 643-669.

Elster, J. (1989). *Nuts and Bolts*. Cambridge: Cambridge University Press.

Fishburn, P.C. (1987), ‘Utility Theory and Decision Theory’, in *The New Palgrave: A Dictionary of Economics.* London and Basingstoke: Macmillian Press Limited.

Gilboa, I. and D. Schmeidler (1989), ‘Maximin expected utility with non-unique prior’. *Journal of Mathematical Economics* **18**, 141-153.

Grubb, M. (1997), ‘Technologies, Energy Systems, and the Timing of $CO_2$ Abatement: An Overview of Economic Issues’, *Energy Policy* **25** (2), 159–172.

Hanley, N., J.F. Shogren, and B. White (1997), *Environmental Economics in Theory and Practice*, New York: Oxford University Press.

Harding, R. and L. Fisher (1992), ‘The Precautionary Principle’, Paper for Ecopolitics VI Conference 1992.

Henry, C. (1974), ‘Option Values in the Economics of Irreplaceable Assets’. *The Review of Economic Studies* **41**, Issue Symposium on the Economics of Exhaustible Resources, 89–104.

Howarth, R.B. (1996), ‘Climate Change and Overlapping Generations’. *Contemporary Economic Policy*; 14(4), October 1996, 100-111.

Howarth, R.B. (2004). ‘Discounting and uncertainty in climate change policy analysis’. Forthcoming *Land Economics*

Hurwicz, L. (1951). ‘Optimality Criteria for Decision-Making under Ignorance’, Cowles Commission Discussion Paper, Statistics, No. 370.

IPCC (2001). Third Assessment Report: Climate Change 2001

Jones-Lee, M.W. and G. Loomes (1995). **‘**Discounting and Safety’.*Oxford Economic Papers*, New Series, Vol. 47, No. 3. (Jul., 1995), pp. 501-512.

Khanna, N. and D. Chapman (1996), ‘Time Preference, Abatement Costs, and International Climate Policy: An Appraisal of IPCC 1995’. *Contemporary Economic Policy*; 14(2) (April 1996) pp. 56-66.

Kahneman, D. and A. Tversky (1979), ‘Prospect Theory: An Analysis of Decision under Risk’. *Econometrica* **47** (2).

Karni, E. and D. Schmeidler (1991), ‘Utility theory with uncertainty’, Chapter 33 in *Handbook of Mathematical Economics, Vol. IV*, Edited by W. Hildenbrand and H. Sonnenschein. Elsevier Science Publisher B.V.

Knight, F.H. (1971), *Risk, Uncertainty, and Profit*. Chicago: University of Chicago Press.

Kolstad, C.D. (1994), ‘George Bush versus Al Gore’, *Energy Policy* **22** (9), 771–778.

Kolstad, C.D. (1996), ‘Learning and Stock Effects in Environmental Regulation: The Case of Greenhouse Gas Emissions’, *Journal of Environmental Economics and Management* **31** (1), 1–8.

Kreps, D.M. (1990), *A course in microeconomic theory*. Harvester Wheatsheaf, London.

Lange, A. (2003), ‘Climate Change and the Irreversibility Effect – Combining Expected Utility and MaxiMin’, *Environmental & Resource Economic*s **25**, 417-434

Lempert, R.J. and M.E. Schlesinger (2000), ‘Robust Strategies for Abating Climate Change’, *Climate Change **45**, 387-401*

Luce, R.D. and H. Raiffa (1957), ‘Games and Decisions: Introduction and Critical Survey’. New York: Dover Publications Inc.

Machina, M.J. (1982), “Expecte utility’analysis without the independence axiom’. *Econometrica* **50**, 277-323.

Machina, M.J. (1989), ‘Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty’. *Journal of Economic Literature* **XXVIII** (December), 1622–1668.

Malnes, R. (1997), *Filosofi for statsvitere*. Oslo: Tano Aschehough.

Manne, A.S. and R.G. Richels (1991), ‘Buying Greenhouse Gas Insurance’, *Energy Policy* **19** 543–552.

Von Neumann, J. and O. Morgenstern (1944), *Theory of games and economic behavior*. Princeton University Press, Princeton.

Nordhaus, W.D. and D. Popp (1997), 'What is the Value of Scientific Knowledge? An Application to Global Warming Using the PRICE Model', *The Energy Journal* **18** (1), 1–45.

O'Donoghue, T. and M. Rabin (1999). 'Incentives for Procastination'. *Quarterly Journal of Economics,* **114** (3), August 1999, pp. 769-816.

Pearce, D., B. Groom, C. Hepburn, and P. Koundouri (2003), 'Valuing the Future'. *World Economics*, Vol. 4, No. 2 (April-June 2003), pp. 121-141.

Ramsey, F. (1931), *Truth and Probability*. The Foundation of Mathematics and Other Logical Essays, London: K. Paul, Trench, Trubner and Co.

Rawls, J. (1971), *A Theory of Justice*. Cambridge, Ma.: Harward University Press.

Ringius, L. (1999), 'Differentiation, Leaders, and Fairness: Negotiating Climate Commitments in the European Community'. *International Negotiation* **4,** 133–166.

Savage, L.J. (1951), 'The Theory of Statistical Decision'. *Journal of the American Statistical Association* **46,** 55–67.

Schelling, T.C. (1992), 'Some Economics of Global Warming', *American Economic Review*, **82** (1): 1-14.

Schelling, T.C. (1995), 'Intergenerational discounting', *Energy Policy* **23**(4/5):395-401.

Schmeidler, D. (1989), ‘Subjective probability and expected utility without additivity’. *Econometrica*, Vol. 57, No. 3 (May 1989), pp. 571-587.

Solow, R.M. (1974), ‘The Economics of Resources or the Resources of Economics – Richard T. Ely Lecture,’ *American Economic Review*, Papers and Proceedings, **64**, 1-14 (Reprinted in Dorfman and Dorfman (Eds.) (1993)).

Sinn, H.-W. (1980), ‘A Rehabilitation of the Principle of Insufficient Reason’, *The Quarterly Journal of Economics* **94** (3): 493-506.

Stigum, B. (1990), *Toward a Formal Science of Economics*. Cambridge, Ma.: MIT Press.

Toth, F.L. (1995). "Discounting in Integrated Assessments of Climate Change". *Energy Policy* 23:403.

Tol, R.S. (1995), ‘The Damage Costs of Climate Change. Toward more Comprehensive Calculations’, *Environmental and Resource Economics* **5,** 353–374.

Tol, R.S. (2002), ‘Estimates of the Damage Costs of Climate Change, Part I: Benchmark Estimates’, *Environmental and Resource Economics* **21**, 47-73.

Ulph, A. and D. Ulph (1996), ‘Global Warming, Irreversibility and Learning.’ Discussion Paper No. 9601, Department of Economics, University of Southampton.

Weitzman, M.L. (1994). ‘On the “environmental” discount rate’. Journal of Environmental Economics and Management, **26**, 1, pp. 200-209.

Weitzman, M.L. (1998). ‘Why the far distant future should be discounted at the lowest possible rate’. Journal of Environmental Economics and Management, **36**, pp. 201-208.

Weitzman, M.L. (2001). “Gamma Discounting.” *American Economic Review,* Vol. 91 (March, 2001), pp. 260-271.

West, R.W.(web) ‘The Let's Make a Deal Applet’, http://www.stat.sc.edu/~west/javahtml/LetsMakeaDeal.html

“All things are so very uncertain,
and that's exactly what makes me feel reassured.”
Too-Ticky in *Moominland Midwinter* by Tove Jansson

# Essay no. 2:

# Option Values and the Timing of Climate Policy

CAMILLA BRETTEVILLE FROYN[23] and H. ASBJØRN AAHEIM[24],
CICERO[25]

## Abstract

This paper reconsiders the importance of irreversibilities in climate change policy. The model presented captures the irreversibility of both climate change due to greenhouse gas emissions and abatement technology investments, both of which are subject to uncertainty that will necessarily change gradually over time as a result of learning. The climate irreversibility adds an option value to the ‘early action’ strategies as long as there is a positive probability of encountering the climate irreversibility constraint. In contrast, irreversibility in abatement investments scales down the future climate effects in the same way as an increase in the discount factor would. The effect the climate option value has on policy making is reduced the more irreversible the investments, but this effect decreases with the length of the time periods. The preferred policy option depends on the relative size of these option values and the possible policy adjustment costs. The framework allows a clear distinction between costs caused by desired adjustments due to new information and the two opposing option values. It also allows graphic illustrations that increase the understanding of the different effects on optimal climate policy. If the sum of the climate option value plus the adjustment costs resulting from too low initial abatement exceeds the sum of the investment option value plus the adjustment costs resulting from too high initial abatement, the net option value is positive, the climate irreversibility effect dominates, and ‘early action’ should be preferred over a ‘wait-and-see’ policy[26].

**Key words:** climate policy, decision criteria, global warming, irreversibility, learning, option value, timing, uncertainty

---

[23] camilla.bretteville@cicero.uio.no
[24] aaheim@cicero.uio.no
[25] Center for International Climate and Environmental Research – Oslo: www.cicero.uio.no
[26] The paper was originally prepared for the 12TH Annual Conference of EAERE, Bilbao, Spain, June 2003

## 1. Introduction

Intuitively it seems very plausible that environmental irreversibilities would play an important role in questions concerning global climate change. In the economics literature, however, this is widely disputed. Due to the long time horizon and the uncertainty aspects of the problem, climate change policy decisions made today may influence welfare hundreds of years into the future. It is also likely that many of the changes caused by global warming will be irreversible. One example is the possible change in the path of the Gulf Stream; it can hardly be assumed that it could be moved back to its original basin once change has occurred. Another example is that changes in the climate might destroy unique ecosystems or drive certain species to extinction if their habitats are destroyed. Also, since greenhouse gases accumulate in the atmosphere, the emissions themselves are irreversible.[27]

The issue of irreversibility and learning in environmental economics was introduced in the seminal papers by Arrow and Fisher (1974) and Henry (1974). Focusing on a one-time development decision, their basic idea was that making an irreversible decision induces additional costs because the current decision restricts future decision possibilities. This implies that a flexibility premium or an option value, adds to the reversible alternatives. Hanemann (1989) defines the option value as the value of flexibility. This extra value is the value of retaining the option to choose any of the alternatives in the light of new information – an option that is lost if an irreversible alternative was chosen in the first place. Thus, if there is a chance of learning, they

---

[27] Every gas has a specific atmospheric lifetime – that is, the time it takes before *1/e* of an emitted quantity of the gas is left in the atmosphere, eg. 114 years for $N_2O$, 12 years for $CH_4$, and up to 200 years for $CO_2$. No single lifetime can be defined for $CO_2$ because of the different rates of uptake by different removal processes (IPPC 2001).

argue, it becomes more important to keep future options open. With regard to climate change policy, the implication is that the current level of greenhouse gas emissions should be lower if there is a possibility of learning more about irreversible damages in the future. Arrow and Fisher (1974) refer to this as the ‘irreversibility effect’, and the implication for climate change policy is to introduce mitigation strategies now to keep future options open (‘early action’). However, later contributions (Kolstad 1996a,b; Ulph and Ulph 1995) downplay the importance of the irreversibility of climate change due to greenhouse gas emissions (henceforth referred to as *climate irreversibility*) in favor of the irreversibility in abatement technology investments (henceforth referred to as *investment irreversibility*), and thus advocate a ‘wait and see’ strategy[28].

Brekke and Lystad (2000), however, observe that Kolstad is not depreciating the capital: his capital lasts forever. They show that with a small depreciation rate on capital, Kolstad’s conclusion is turned around.  Fisher and Narain (2003) points out that in Kolstad (1996b) the obvious reason why there are no findings of impacts from climate irreversibility is that there is no scenario were it would be optimal to emit negatively in the future to compensate for too high emissions today. Furthermore, in the parameterization of the model, the non-negativity restriction on emissions, used in the model to define emissions irreversibility, is never binding.

A different example where the question of timing is addressed is Pindyck (2002), here analyzed as an optimal stopping problem. He looks at when (if ever) climate policy should be adopted. So adopting a policy today competes not only with never adopting the policy, but also with adopting it next year, in two years, and so on. He focuses largely on a one-time adoption of an emission reducing policy.

---

[28] See Fisher (2001); Fisher and Narrain (2003); or Heal and Kriström (2002) for more details.

All in all, the rather sparse economic literature on the implications for the timing of climate policy of uncertainty and irreversibility seems to favor the ‘wait and see’ policy over ‘early action.’ Fisher (2001) argues that the reason for this result, which may appear counterintuitive to most, is that, given the assumptions or parameter values built into the models, there is relatively little cost in deferring abatement while waiting to learn more about the benefit. The climate irreversibility, therefore, plays less of a role in driving current policy recommendation on controlling emissions.

Contrary to the standard assumptions of economic climate change models, Fisher and Narain (2003) assume that the probability of a catastrophic impact at some point in the future (the next period), although very small, may be positively related to the level of greenhouse gas concentration in the atmosphere. Their model also allows for the possibility of regrettable first period investments. The introduction of endogenous risk has a positive effect on abatement investments, but they find the effect of investment irreversibility to dominate the effect of climate irreversibility essentially because the concentration of greenhouse gases in the atmosphere does not change much over a 10 year period. However, Torvanger (1997), using a three-period stochastic dynamic programming model, shows that in the case of endogenous probability of irreversible climate change the climate irreversibility effect dominates. Aaheim (2004), using a stochastic model with Brownian motion, shows that both investment irreversibility and climate irreversibility may affect optimal abatement significantly, but if the policy is updated frequently in accordance with new information, the two effects tend to weigh each other out.

The conflicting irreversibilities were first treated in Kolstad (1996a). He finds that, if the learning process is sufficiently slow, compared with the rates of pollution decay, and capital depreciation, learning does not influence the decision. However, if learning is significant, the two irreversibilities can affect the desired first-period level of emissions in opposite directions. The dominant effect is determined by the relative sizes of the rates of pollution decay, and capital depreciation, as well as the expectations about damages. It thus seems pertinent to study further the conditions for the irreversibility effect to hold, that is: Under what conditions will the implications for the timing of climate policy of uncertainty and irreversibility favor an ‘early action’ type policy over ‘wait and see’? In an effort to make this complicated issue more transparent we use a stylized model to highlight three important uncertainty effects on climate change policy. The simplicity of the model is chosen to enable an explicit illustration of these effects. It is important to note that the choice in our model is continuous. However, we operate with an interval in which the policy choice is expected to lie. Hence, when we refer to a ‘wait and see’ policy we refer to an abatement choice that is relatively closer to the lower bound for abatement investment (low initial abatement), and when we refer to ‘early action’ we mean a choice that is relatively closer to the upper bound (high initial abatement). The question addressed here is thus *how much* to do, rather than whether to take action.

When uncertainty is high, the likelihood of making a perfect guess is low. Under a process of learning, no matter what strategy is chosen, costs must thus be expected because of adjustments desired due to new information (henceforth referred to as ‘adjustments costs’). This is the first effect we study. In addition, we study the two opposing option values of policies that maintain flexibilities: one related to climate irreversibility and the other related to investment irreversibility. We show that the climate irreversibility imposes an option value to the ‘early action’ strategies if, and only if, there is a positive probability of encountering the climate irreversibility constraint, and similarly that the investment irreversibility imposes an option value to the ‘wait and see’ strategies if, and only if, there is a positive probability of encountering the investment irreversibility constraint. We also show that irreversibility in abatement investments scale down the future climate effects in the same way as an increase in the discount factor. Furthermore, the effect the option value of early abatement has on current policy choice is reduced the more irreversible the investments, but this effect decreases as the time horizon increases.

Our framework allows a clear distinction between costs caused by desired adjustments due to new information and the two opposing option values, which to our knowledge is a new approach. It also allows graphic illustrations that increase the understanding of how the adjustment costs and option values affect optimal climate policy. This clarification is our main contribution to the debate, and it shows clearly that the preferred policy option depends on the relative size of the two option values and of the ex ante adjustment costs. We conclude that if the sum of the climate option value plus the expected adjustment cost resulting from too low initial abatement exceeds the sum of the investment option value plus the expected adjustment cost resulting from too high initial abatement, the net option value is positive, the climate irreversibility effect dominates, and ‘early action’ should be preferred over a ‘wait and see’ policy.

A short discussion on changes in uncertainty and learning is given in Section 2. Section 3 starts with a presentation of the general two-period model with continuous policy choice, uncertainty, and learning, followed first by an analysis of climate irreversibility, which corresponds to the traditional option value discussion, and second by an analysis of how investment irreversibility may affect the solution. Section 4 considers briefly implications of the use of alternative decision criteria, while concluding remarks are given in Section 5.

## 2. Uncertainty and learning

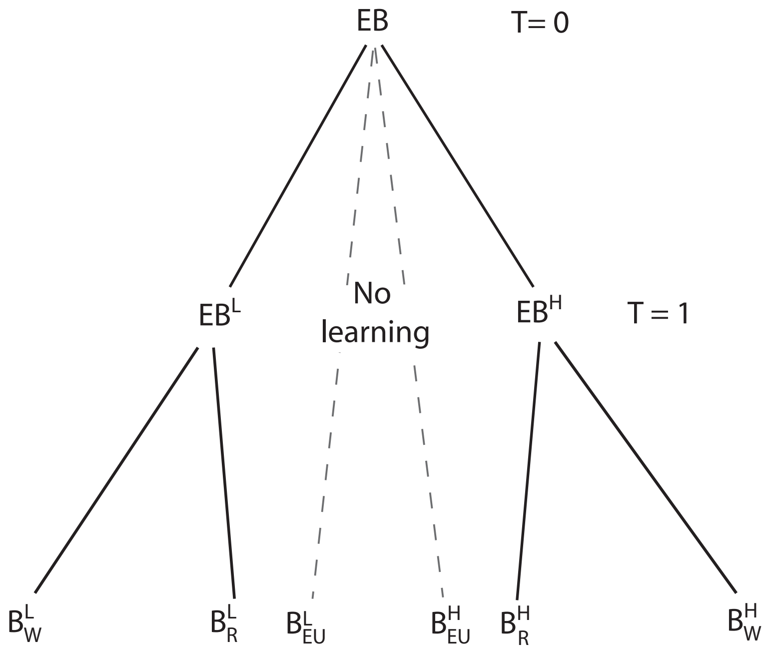
Although insights into the basic science of climate change have improved substantially over the past decades, there is still uncertainty around almost every aspect of the problem – including future levels of emissions, temperature changes resulting from greenhouse gas concentrations, impacts from changes in global mean temperature, technological developments, and costs of abatement and adaptation. One might think that the degree of uncertainty would significantly influence the political processes and decisions, and that economic analyses thus would treat uncertainty as a central feature. Nevertheless, the main perceptions about economic issues related to climate change and climate policy, for example as assessed and reported by IPCC (2001), are based on deterministic studies. This does not necessarily mean that the uncertainties have been ignored, but rather that this part of the economics literature is inconclusive.

One possible explanation is that most of the climate change studies are based on numerical models. The uncertainties per se are therefore difficult to trace. Results derived from numerical models depend on assumptions about economic relationships such as damage and abatement cost functions, which are far from well known. It is therefore difficult to tell, for example, how better ability to predict temperature change in the future will affect current policy choices, unless we assume that we know all there is to know about possible climate change impacts and the costs of mitigation. This is somewhat of a paradox, because the impact of learning is one of the main issues that have been studied in analyses of climate policy choice under uncertainty.

Impacts from learning must be studied in a dynamic framework where decisions are taken sequentially. While a static problem gives a best solution once and for all, a dynamic problem aims at finding the best strategy under a given set of information. This is why learning becomes relevant: If uncertainty affects the decision, future amendments to information will change decisions in the future. Whether decisions of today are affected by future learning therefore depends on whether future decisions depend on present decisions, or, in the words of the option value literature, whether present decisions are irreversible.

The alternative strategies can be illustrated schematically by Figure 1. This is a two-period decision tree with alternative strategy paths. The dotted paths in the middle represent static expected utility maximization without learning. According to this strategy, at time t = 0 (the first decision node) we choose abatement levels based on a weighted sum of the probabilities of states (our expectations). The probabilities are formed in accordance with our beliefs. If damages turn out to be high, we get the net benefit of $B_{EU}^{H}$; and if they turn out to be low, we get $B_{EU}^{L}$. The outer branches represent sequential decision paths: At time t = 0 we have to decide which state we perceive as the most likely, and choose our first period abatement level accordingly. Either way we risk making the wrong guess. The state is, however, now revealed at time t = 1. This means that if we initially made the wrong guess we can now adjust our abatement effort according to the new available information. If we guess the state of damages to be low at t = 0 (the left hand side in figure 1), the net benefit is $B_{R}^{L}$ if we are correct and $B_{W}^{L}$ if

*Figure 1: Alternative decision strategy paths*



we are wrong. If we guess the state of damages to be high, the net benefit is $B_R^H$ if we are right and $B_W^H$ if we are wrong. Regardless of our choice of path, we get a lower benefit if we make the wrong guess because the level of first period abatement is not chosen optimally.[29]

Such a sequential decision-making process aims to identify short-term strategies in the face of long-term uncertainties. The next several decades will offer many opportunities for learning and mid-course corrections. The relevant question is not

---

[29] The benefits in Figure 1 are in arbitrary order. Generally we can only say that $B_R^H > B_W^H$ and $B_R^L > B_W^L$.

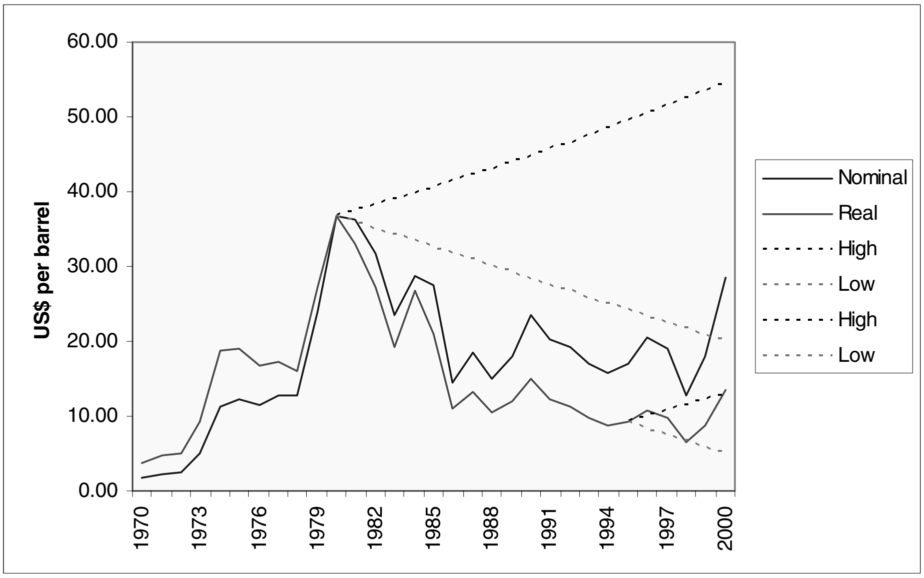
"What is the best course of action for the next 100 years?", but rather "What is the best course for the near-term given the long-term objective?" (IPPC 2001).

There is a vast amount of ongoing research aimed at improving our understanding of the implications of anthropogenic emissions of greenhouse gases. We will therefore learn more about the problem in the future, and most likely the scientific background for climate policy will improve. To the extent that learning has an influence on decision making, it should therefore be taken into account also in current climate policy choice.

However, a better understanding of climate change issues does not necessarily mean less uncertainty. The connection between uncertainty and the state of knowledge is illustrated by comparing reports from the UN Intergovernmental Panel of Climate Change (IPCC). The Second Assessment Report (IPCC 1995) estimates that a doubling of $CO_2$ concentrations in the atmosphere (2 x $CO_2$) will lead to a global mean temperature increase between 1.8 ºC and 4.2 ºC. In the Third Assessment Report (IPCC 2001), the possible range of the future increase in global mean temperature increased to between 1.4 ºC and 5.6 ºC. The reason for the increase in uncertainty is improved knowledge in the underlying science linking greenhouse gas emissions to atmospheric accumulation and global temperature change. A better and more precise description of the relationships between policy, economic and human development and emissions, on the one hand, and atmospheric processes, on the other, improved the consistency between the scenarios in the Third Assessment Report (TAR), and resulted in a wider range of possible outcomes.

The Third Assessment Report was the first to also indicate the degree of uncertainty associated with its predictions by stating the degree of confidence the leading authorities in the field of climate change report with their forecasts, ranging from ‘very high’ (≥95%) to ‘very low’ (≤5%). Most of its findings are assigned a probability ranking (see the TAR or e.g. Heal and Kriström (2002) for more details). Even though improved knowledge may not reduce the total uncertainty in the future, the aspect of learning is still important, simply because some of the uncertainty is resolved as time goes by. Just as it is easier today to predict the climate in 2010 than in 2100, it will be easier in 2090 than it is today to predict the climate in 2100.

*Figure 2. Spot prices for Saudi light (1970–1979) and Brent Blend (1980–2000) in nominal prices and real 1990 prices*[30].



To illustrate this more or less passive way of learning, consider the price of crude oil over the past 30 years. In the wake of oil price hikes in the 1970s, significant effort was put into forecasting the oil price around 1980. Forecasts were based on

---

[30] Sources: Aslaksen et al. (1990) and Statistics Norway (2003)

uncertainty and Hotelling's rule,[31] which predicted an increase in the price of exhaustible resources, such as petroleum.

Figure 2 shows the actual development of oil prices from 1970 to 2000. Most forecasts of the trend made in the early 1980s were found in the upper half of the area between the dotted lines, typically in the range between US$ 35 and US$ 55, in real terms.[32] Thus, most pessimistic (low price) forecasts turned out to be far too optimistic. However, one cannot claim that oil price changes are better understood today than they were 20 years ago. It may be granted that the importance of the Hotelling rule was exaggerated, but this is also a result of past observations. Therefore, a forecast for the year 2000 made the year before with the same uncertainty had, of course, a much higher probability of being correct than a forecast made in 1980. The main point is that short-term predictions are more accurate than long-term predictions, since learning occurs along the way, regardless of whether the uncertainty has decreased.

One may expect that gradual resolution of uncertainty will be considered the main source of learning in the climate policy process, at least in the foreseeable future. Some of the uncertainty may be irreducible in principle, and hence decision makers will have to continue to take action under significant degrees of uncertainty. Thus, the problem of climate change evolves as a subject of risk management in which strategies should be reformulated as new knowledge arises (IPPC 2001).

---

[31] See e.g. Hanley et al (1997) for a description of Hotelling's rule.

[32] Lorentsen et al. (1985) denotes by US$ 20 a 'collapse price scenario.'

## 3. Irreversibilities and option values

This section analyses the abatement decision taken by an agent subject to uncertainty both in benefits of mitigating climate change (damages prevented) and in abatement technology investment costs. We will concentrate on irreversibility in one of these variables at a time to see how both types of irreversibility affect the optimal solution. We start with a presentation of the general two-period model with uncertainty and learning. Then, we analyze the case of climate irreversibility, which corresponds to the traditional option value discussion (Arrow and Fisher, and Henry). Next, we look at how investment irreversibility may affect the solution.

### *3.1 A two-period model with a binary distribution of outcomes*

We consider climate policy over two periods, 'the present' and 'the future.' We do not discount 'the present', hence this is referred to as period 0, and 'the future' is referred to as period 1. Without abatement, emissions of greenhouse gases are fixed in both periods and denoted $e_0$ and $e_{1,}$ in the initial and future time periods, respectively. Abatement may be invoked in either period at costs of $a_0$ and $a_1$, respectively. The abatement cost per unit of emissions cut, $c_{t,}$ is assumed to be independent of scale in each period. Hence, the actual abatement is $a_0/c_0$ in period 0 and $a_1/c_1$ in period 1.

The emissions of greenhouse gases in period $t$ can then be written as $[e_t - a_t/c_t]$.

These add to the previous level of concentrations of greenhouse gases, denoted by $y_{t-1}$. Then, the development in atmospheric concentration[33] can be written $y_t = y_{t-1} + e_t - a_t/c_t$. For the two specific periods we have:

$$y_0 = \hat{y} + e_0 - \frac{a_0}{c_0}, \tag{1}$$

$$y_1 = y_0 + e_1 - \frac{a_1}{c_1}, \tag{2}$$

where the historical level of concentrations, $\hat{y}$, is assumed known.

Assume, moreover, that the relationship between concentrations and economic damages has a logarithmic form, with an exponent larger than one.[34] Without loss of generality, we choose a quadratic damage function for the sake of simplicity, $f(y_t) = \alpha y_t^2$, where α is a constant.

Recall that the periodic emissions levels are exogenously given. The benefits of abatement can thus be expressed by the damage avoided. For the initial period, this is

$$b(y_0) = q_0(f(\hat{y} + e_0) - f(y_0)) = q_0(\alpha(\hat{y} + e_0)^2 - \alpha(\hat{y} + e_0 - \frac{a_0}{c_0})^2) \tag{3}$$

where $q_0$ is the price on damages or the willingness to pay for avoiding them in period 0. Similarly, the benefits of abatement in the future can be written,

---

[33] Note that concentrations are usually measured in terms of parts per million or billion by volume (ppmv and ppbv respectively), whereas emissions are measured in tons. We therefore convert emissions to concentration units and calculate costs in corresponding terms.

[34] The usual choice in economic studies is somewhere between 2 and 3. See e.g. Aslaksen (1990) for a discussion on the choice of exponent.

$$b(y_1) = q_1(f(\hat{y} + e_0 + e_1) - f(y_1)) = q_1(\alpha(\hat{y} + e_0 + e_1)^2 - \alpha(\hat{y} + e_0 + e_1 - \frac{a_0}{c_0} - \frac{a_1}{c_1})^2) \quad (4)$$

To simplify the formalization, we will in the following normalize the value of damages and thus set $q_0 = q_1 = 1$.

In the initial period, there is certainty about present costs and benefits but uncertainty about the state in period 1 (the future). Furthermore, assume that learning takes place during the initial period, such that the decisions in the future are adjusted according to this new information. Since we simplify the problem to look at two periods, this is modeled as if the future decisions are made under full certainty. However, when we assume that the uncertainties resolve in the future, we refer only to the information revealed since we made the initial policy decision. In real life there will be substantial uncertainty remaining also in the future.

When deciding upon abatements, we take into account that our initial decisions will influence future benefits. We simplify the uncertainty by considering only two future states of the world: low damages (state A) and high damages (state B). Denote by $\pi$ the probability of state A.

A standard expected utility maximizing agent determines

$$W = \max_{a_0, a_1} E\left\{\sum_{t=0}^{1} \frac{b(y_t) - a_t}{(1+r)^t}\right\}, \quad (5)$$

where $E$ is the expectations operator over possible outcomes, and $r$ is the discount rate. Note that, so far, we have not taken any kind of irreversibility into account. That is, the solution to this problem gives the optimal policy, from the initial-period perspective, when there are no other restrictions.

Inserting for *b*($y_t$) we can write the welfare function as

$$W = \max_{a_0,a_1} E\left\{ \begin{array}{l} \left[ f(\bar{y}+e_0) - f(\bar{y}+e_0-\frac{a_0}{c_0}) \right] - a_0 \\ +\frac{1}{(1+r)}\left[ \left( f(\bar{y}+e_0+e_1) - f(\bar{y}+e_0+e_1-\frac{a_0}{c_0}-\frac{a_1}{c_1}) \right) - a_1 \right] \end{array} \right\} \qquad (6)$$

Assume that the damage cost function has a quadratic form $f(y_t) = \alpha y_t^2$ and that uncertainty in period 1 is characterized by two states, *A* and *B* with a vector of outcomes, $x = (\alpha_1, c_1)$.such that

$$x = \left\{ \begin{array}{l} P(x^A) = \pi \\ P(x^B) = 1-\pi \end{array} \right\}$$

Then, the first order condition for abatement in period 0 can be written as:

$$1 = \frac{2\alpha_0}{c_0}(\hat{y}+e_0-\frac{a_0}{c_0}) + \frac{1}{1+r}\left[ \pi\frac{2\alpha_1^A}{c_0}(\hat{y}+e_0+e_1-\frac{a_0}{c_0}-\frac{a_1}{c_1^A}) + (1-\pi)\frac{2\alpha_1^B}{c_0}(\hat{y}+e_0+e_1-\frac{a_0}{c_0}-\frac{a_1}{c_1^B}) \right]$$

(7)

where $\alpha^i$ is the initial damage per unit concentration, that is, the damage of an increase in concentrations by, for example, 1 ppmv carbon from "today" if the present damage is zero. The term $\frac{2\alpha_t}{c_t} y_t$ is the marginal damage per euro of abatement when the concentration level is $y_t$. This term may be interpreted as the shadow price of concentrations at level $y_t$, and because of the quadratic damage function it is linear in the level of concentrations.

Similarly, the first order condition for abatement in the future period can be written as:

$$1 = (\hat{y} + e_0 + e_1 - \frac{a_0}{c_0})(\pi \frac{2\alpha_1^A}{c_1^A} + (1-\pi)\frac{2\alpha_1^B}{c_1^B}) - a_1(\pi \frac{2\alpha_1^A}{c_1^A}\frac{1}{c_1^A} + (1-\pi)\frac{2\alpha_1^B}{c_1^B}\frac{1}{c_1^B}) \tag{8}$$

Equations (7) and (8) imply that the marginal expected value of the concentrations in each period should be equal to 1, which is the marginal cost of abatement in value terms.

In order to simplify the expressions further, define the following constants:

$$C_0 = (\hat{y} + e_0)\frac{2\alpha_0}{c_0} + \frac{\hat{y} + e_0 + e_1}{1+r}\left[\pi \frac{2\alpha_1^A}{c_0} + (1-\pi)\frac{2\alpha_1^B}{c_0}\right] - 1,$$

$$C_1 = (\hat{y} + e_0 + e_1)\left[\pi \frac{2\alpha_1^A}{c_1^A} + (1-\pi)\frac{2\alpha_1^B}{c_1^B}\right] - 1.$$

$C_0$ and $C_1$ are both expressions for the expected marginal damage per euro abatement minus the unit cost of abatement (= 1) at a business as usual concentration. The difference is that $C_0$ take in hand both periods, while $C_1$ only applies to period *1*.

Moreover, define

$$\Phi_0 = \frac{2\alpha_0}{c_0}\frac{1}{c_0} + \frac{1}{1+r}\left[\pi\frac{2\alpha_1^A}{c_0} + (1-\pi)\frac{2\alpha_1^B}{c_0}\right]\frac{1}{c_0},$$

$$\Phi_1 = \pi\frac{2\alpha_1^A}{c_1^A}\frac{1}{c_1^A} + (1-\pi)\frac{2\alpha_1^B}{c_1^B}\frac{1}{c_1^B},$$

$$\Psi_0 = \frac{1}{1+r}\left[\pi\frac{2\alpha_1^A}{c_0}\frac{1}{c_1^A} + (1-\pi)\frac{2\alpha_1^B}{c_0}\frac{1}{c_1^B}\right],$$

$$\Psi_1 = \pi\frac{2\alpha_1^A}{c_1^A}\frac{1}{c_0} + (1-\pi)\frac{2\alpha_1^B}{c_1^B}\frac{1}{c_0}.$$

$\Phi_t$ *(t = 0,1 )* can be interpreted as the contribution, at the margin, to total benefits from abatement in period *t* whereas $\Psi_t$ can be interpreted as the contribution, at the margin, to total benefits at time t from abatement at time *s* (*s = 0,1; s ≠ t*). Thus, $\Psi_0$ denotes the marginal present value in period *0* of the abatement that is planned carried out next period. This value includes both damages avoided and saved abatement costs due to undertaken investments. In the same manner $\Psi_1$ denotes the marginal value in period *1* of the abatement undertaken in period *0*.

The first order conditions can now be written as the linear equation system:

$$C_0 = \Phi_0 a_0 + \Psi_0 a_1, \tag{9}$$

$$C_1 = \Psi_1 a_0 + \Phi_1 a_1 . \tag{10}$$

Equations (9) and (10) require that in optimum, the marginal damage should be equal to the marginal value of abatement in period *i* of the abatement scheme planned over both periods.

The level and the distribution of abatement in the two periods can now be written as two linear functions in $a_0$ and $a_1$. We rewrite (9) and (10) and get

Period 0:
$$a_0 = \frac{C_0}{\Phi_0} - \frac{\Psi_0}{\Phi_0} a_1 \tag{11}$$

Period 1:
$$a_0 = \frac{C_1}{\Psi_1} - \frac{\Phi_1}{\Psi_1} a_1 \tag{12}$$

These two linear functions express the trade-off between abatement costs in the two periods. Thus, for period 0, a reduction of 1 € in present abatement cost, $a_{0,}$ must be replaced by an increase of $\psi_0/\Phi_0$ € in period 1 in order to keep marginal costs equal to marginal benefits. For period 1, a 1 € reduction in period 0 must be replaced by increased abatement cost of $\Phi_1/\psi_1$ € in period 1, if marginal costs are to be equal to marginal benefits.

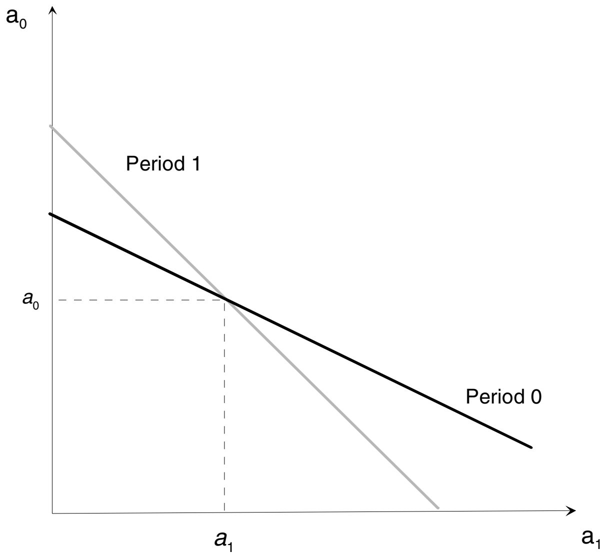
Note that the linearity between $a_0$ and $a_1$ applies regardless of the uncertainties involved. How the degree of uncertainty affects optimal abatement, or how uncertainty in the damage costs affect the solution compared to the uncertainty in abatement costs can therefore be analysed by corresponding shifts in the parameters $C_t$, $\Phi_t$ and $\Psi_t$. In Appendix 1, we show how these parameters are affected by an increase in the uncertainties, and what the consequences are for the intersections and the slopes of the $a_1 = f(a_0)$ lines. This provides a background for the discussions about the implications for the importance of the irreversibilities below.

Figure 3 illustrates the optimal allocation of abatement costs in the two periods. The curves correspond to (11) and (12) and represent all abatement allocations whereby marginal costs equal marginal benefits. That is, the period 0 curve shows all combinations of $a_0$ and $a_1$ that satisfy the first order condition for period 0, and the period 1 curve shows all combinations that satisfy the first order condition for period 1. Where the curves intercept, both conditions are satisfied simultaneously.

Abatement costs for period 0 start at a lower level and increase at a lower rate compared to the line for period 1. From the definitions of the constants we have that $\psi_0/\Phi_0 < \Phi_1/\psi_1$, unless there is large uncertainty in abatement costs (see Appendix 1). If so, the trade-off curve is steeper in period 1 than in period 0. Then, a necessary, but not sufficient, condition for interior solution is that $C_1/\psi_1 > C_0/\Phi_0$. Alternatively, if $\Phi_1/\psi_1 < \psi_0/\Phi_0$ we must have $C_1/\psi_1 < C_0/\Phi_0$ to obtain an interior solution. It follows from the definition of $C_1$ and $C_0$ that this depends on the relationship between damage costs and abatement costs in period 0 and period 1. ‘Corner solutions’ with abatement in only one period is possible, but the following discussion will be based on the case of interior solutions. Then a reduction in present abatement may be compensated by more abatement in the future.

The optimal solution is found where the two lines intersect. Consider an allocation with abatement in period 0 only. Then the total benefit of abatement would be $C_0/\Phi_0$. This value, however, would increase if some abatement were postponed to period 1. Reallocation from today until the future would, in fact, increase the total benefit of abatement until the two lines intersect. From this point on, further reallocation of abatement between the two periods will give lower total benefit. Intertemporal optimality is therefore attained at the intersection.

*Figure 3. Optimal allocation of abatement between the present and the future*



Changes in the parameters will affect total abatement, as well as the allocation between the periods. If $q_1\alpha_1$ increases, the constants for abatement assigned to both periods increase, but more so in equation (12), which applies for the future (period 1), than in equation (11), which applies for the present (period *0*). Also the multiplicative term in equation (11) increases, while the multiplicative term in equation (12) is unchanged. In sum, both total and future abatement increase compared with present abatement. An increase in the expected unit cost of future abatement does not affect the constant term, but reduces the multiplicative term for period *0*. The constant term for period 1 decreases, while the multiplicative term in period 1 increases. All these three shifts thus contribute to an allocation of abatement from the future to the present.

An increase in the discount rate affects the abatement assigned to period *0* only in this model. The effect is that the present value of abatement in period 1 is reduced. This contributes to a reduction of the abatement in period *0*. At the same time, abatement in period 1 becomes relatively less expensive. This tends to change the trade-off between abatement in the two periods. Both effects reduce abatement in period *0*. Some may be allocated to period 1, but the total amount is likely to be unchanged.[35]

Changes in uncertainty on future beliefs will also affect optimal abatement in the two periods. One exception is the case where there is uncertainty in future damage costs only, while the future abatement costs are certain. Appendix 1 shows that all the above-defined constants remain unaffected if the expected damage remains unchanged. However, if future damage costs are certain, but abatement costs for this period are not, we show in the appendix that the constants will increase. Except for the constant term in equation (11), the terms in equations (11) and (12) will also be affected. The multiplicative term for period 0 increases, while both terms for period 1 decrease. All changes contribute to a reallocation of abatement from the present to the future.

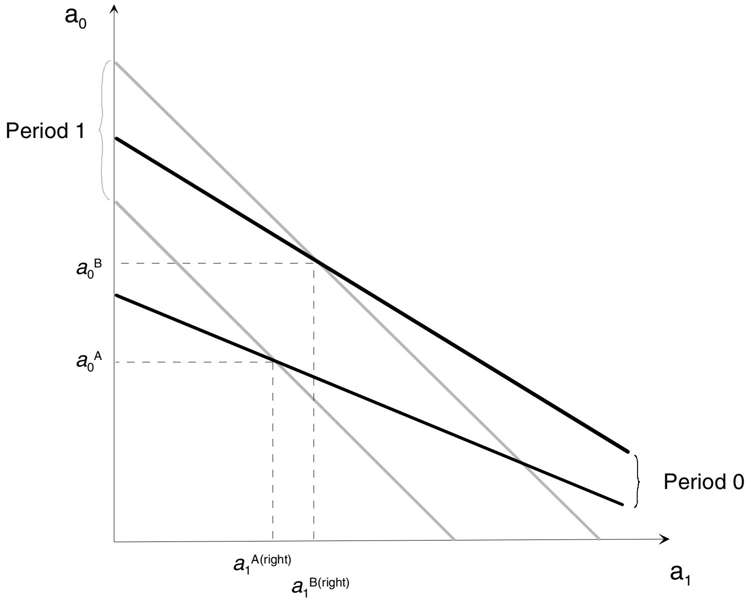
An illustration of the case with uncertain damages is given in Figure 4A. This is simply a more complicated version of Figure 3, where there is uncertainty regarding the position of the trade-off curves. Figure 4A displays two alternative trade-off curves between the abatement in period 0 (the present) and period 1 (the future): one pair representing state A (low future damage costs) where the optimal solution is $a_0^A$ and

---

[35] Appendix 2 shows the changes in these parameters for a specific numerical example.

$a_1^{A(right)}$, and one pair representing state B (high future damage costs), where the optimal choice is $a_0^B$ and $a_1^{B(right)}$. These two solutions are to be considered the boundaries for the abatement choices. The rational decision maker would choose an allocation between the two extremes, based on his subjective probability distribution over states. For example, insertion of the parameters in equations (11) and (12) gives the best ex ante allocation of abatement investments for the expected utility maximizer.

*Figure 4A. Optimal abatement allocation when damage costs are uncertain*



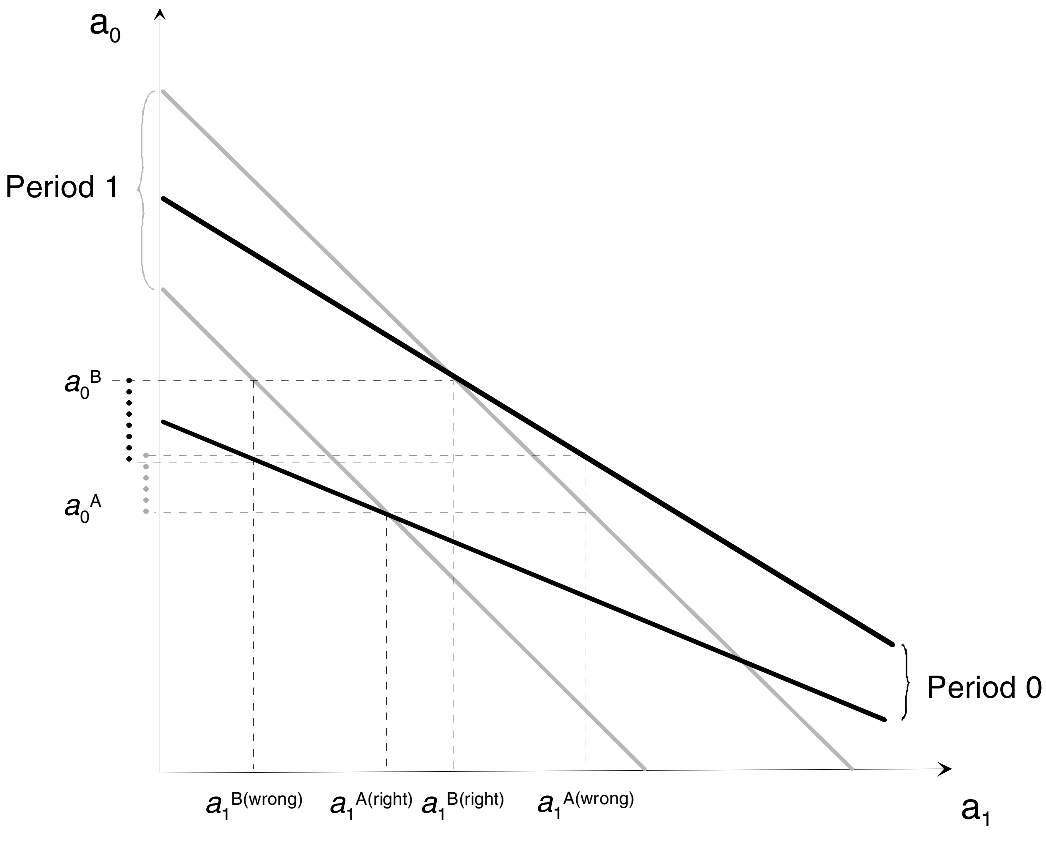
Because of uncertainty, the initial period decision is based on guesses about the future state of the world and the amount of next period abatement. If learning takes place (uncertainty is lessened) before the final decision for the next period, the decision maker may want to make adjustments to her beliefs in period 0. Irreversibilities place a constraint upon these adjustments which will be addressed below, but let us first take a closer look simply at the costs of policy adjustments.

As a benchmark, assume that the initial period decision is based exclusively on beliefs about either state A (the lower pair of trade-off curves) or B (the upper pair of trade-off curves) in Figure 4A, and not on maximization of expected utility. Furthermore, assume that one of these states is realized in period 1. If the guess is right, future abatement ($a_1$) will be chosen as predicted. If the guess is wrong, however, period 1 abatement should be adjusted according to the new information. Since the choice of initial period abatement ($a_0$) was conditioned upon a guess about $a_1$, the benefits will be higher when the guess is right. The situation for when the guess is as wrong as it can get is illustrated in Figure 4B. Here we put all our eggs in one basket, but after learning we wish we had put them in the other. The best abatement solutions after learning that the guess was wrong are marked $a_1^{A(wrong)}$ and $a_1^{B(wrong\,)}$, respectively (see below for explanations).

Recall that each curve represents abatement allocations that satisfy the period's first order condition for optimal abatement, and that the social loss of guessing wrong can be measured in present value terms. According to the objective function (6), this can be measured by differences in initial (period 0) abatement costs. In Figure 4B these losses are indicated by the two vertical lines to the left of the $a_0$-axis. The black dotted line represents the loss incurred when the wrong guess was the high damage cost scenario, and the grey one the loss for when the wrong guess was the low cost alternative, and in both cases the other extreme was learned to be the real state of the world. The explanation goes as follows:

Start with supposing we based our decision solely on state B; the high damage cost scenario. Then the initial period choice of abatement cost was $a_0^B$. Suppose further that at the start of the next period we discover that damage costs are instead low. Thus, now we wish we had chosen $a_0^A$ initially, but since we cannot move back in time, this is no longer a possible option. For the optimality condition to hold for period 1, abatement cost must be on the period 1 trade-off curve for state A, and the best move is to choose $a_1^{B(wrong)}$. This is a reduction compared with our initial guess for period 1 abatement, and the social loss of guessing wrong can be measured by the difference between the $a_0$ value where $a_1^{B(right)}$ crosses the trade-off curve for period 0 in scenario A and the $a_0$ value that was chosen, namely $a_0^B$. This social loss is thus the difference between what we thought our initial abatement was worth and what it turned out to be worth according to the information received. Since we abated too much in period 0, the initial abatement

*Figure 4B. Cost of guessing wrong under damage cost uncertainty*



is worth less than we thought. The loss is illustrated by the black dotted line along the $a_0$-axis.

Suppose instead that the initial guess is state a; the low damage cost scenario. The period 0 abatement cost choice is then $a_0^A$. Suppose further that at the start of the period 1, we discover that costs instead are high. Since the optimal allocation for state B is no longer attainable, the best we can do is to satisfy the optimality condition for period 1, given this new information. In this situation we abated too little initially and have to compensate by abating more than we thought we would in period 1. The best choice is $a_1^{A(wrong)}$. The social loss of guessing wrong, in this case, can be measured as the difference between $a_0^A$ and the $a_0$ –value where $a_1^{A(wrong)}$ intersects with the trade-off curve for period 0 in state B. The loss is illustrated by the grey dotted line along the $a_0$-axis in Figure 4B.

*3.2 Irreversible emissions of greenhouse gases.*

So far, we have imposed no restriction on possible abatement in each period. Since emissions of greenhouse gases are considered irreversible, however, abatement in each period cannot exceed the emissions in that period. This gives the constraint

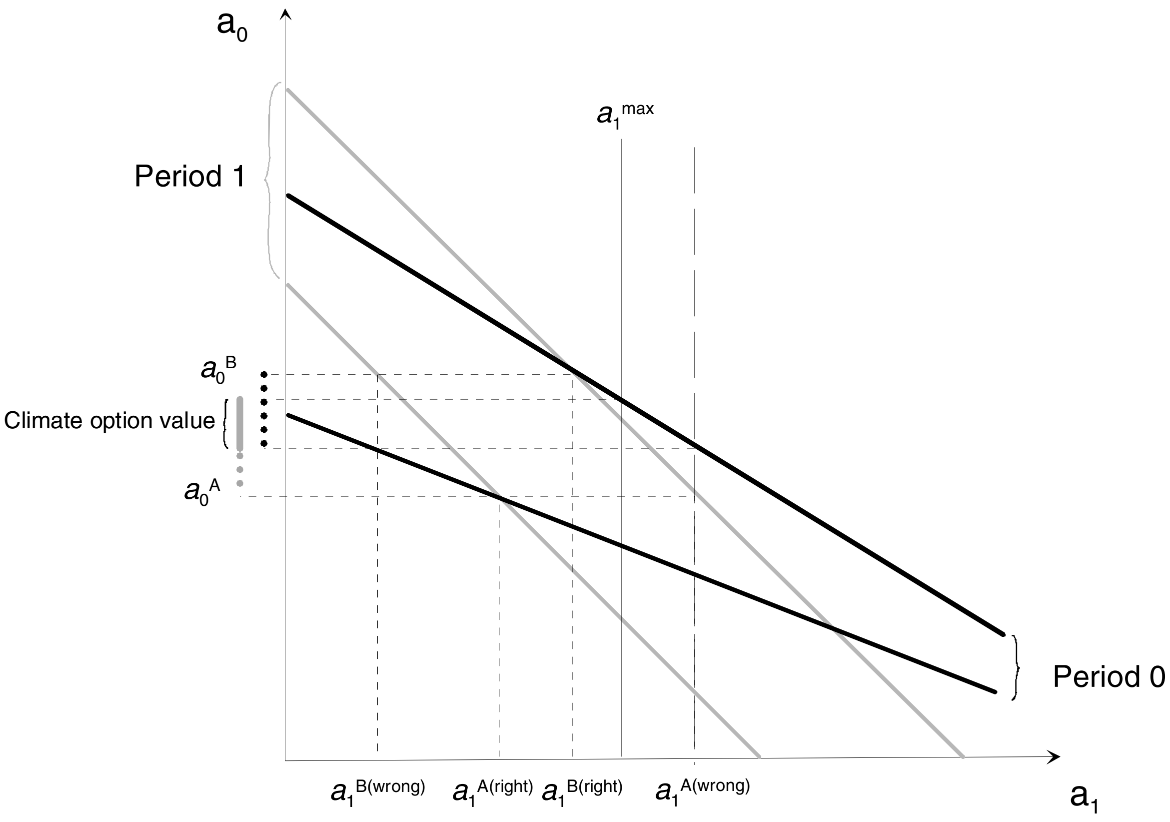
$$a_1^* \leq c_1^* e_1 \tag{13}$$

where $c^*_1$ is the actual unit costs in period 1 and $a_1^*$ is the optimal abatement cost in this period. (Note that both $c_1$ and $e_1$ are exogenous.) If the decision in the initial period was based on a prediction of higher future abatement than $c_1^* e_1$, the optimal solution cannot be achieved. If future emissions are low or uncertainty is large, or both, there will be a positive probability for encountering this irreversibility constraint. The lower we set present abatement, the higher is this probability, and an extra value thus attaches to the choice of high initial abatement. This is the option value of early abatement.

Assume that there is uncertainty in the damage costs only and that these turn out to be higher than expected. Initial abatement was then lower than optimal, when evaluated in hindsight. A relative loss has therefore occurred and adjustments are needed. In this case abatement will have to be adjusted upwards, something that can cause a violation of the maximum abatement constraint (13). If (13) is violated, the optimal level of abatement for period 1 is not attainable, the climate irreversibility constraint is binding, and the best we can do is to choose the maximum level of abatement.

The option value of early abatement is illustrated in Figure 4C. The maximum period 1 abatement ($a_1^{max}$) is illustrated by the vertical solid line. Suppose, as a benchmark, we choose the most conservative 'wait-and-see' strategy and base our initial period decision solely on low damages (state A). Suppose further that we learn that damages instead are high (sate B). We then would prefer to choose abatement costs equal to $a_1^{A(wrong)}$ . This point is, however, to the right of $a_1^{max}$. Since we cannot choose an abatement cost level that is higher than the maximum, the climate irreversibility constraint is binding. The second best solution is thus not reachable, and the best we can do under these circumstances is to choose $a_1$ equal to $a_1^{max}$. The social loss from guessing wrong is represented by the grey line along the $a_0$-axis, where the solid segment represents the option value, and the dotted segment represents adjustment costs[36]. Thus, the loss imposed if the climate irreversibility constraint is encountered adds to the costs of the ‘ordinary’ adjustments to resolved uncertainty. The black dotted line to the left of the $a_0$-axis is the same as in Figure 4B and represents the loss for the benchmark case when the most expansive ‘early action’ policy is chosen and we learn that we should have chosen the most conservative ‘wait-and-see’ strategy.

*Figure 4C. Cost of guessing wrong under climate irreversibility*



In the special case of figure 4C, if wrong, the total cost of assuming low damages exceeds the total cost of assuming high damages, but this depends on the dimension of the option value of early abatement. The option value depends on the

---

[36] The graphic representations shows the relative, rather than the actual, size of the different costs linked to the three uncertainty effects studied.

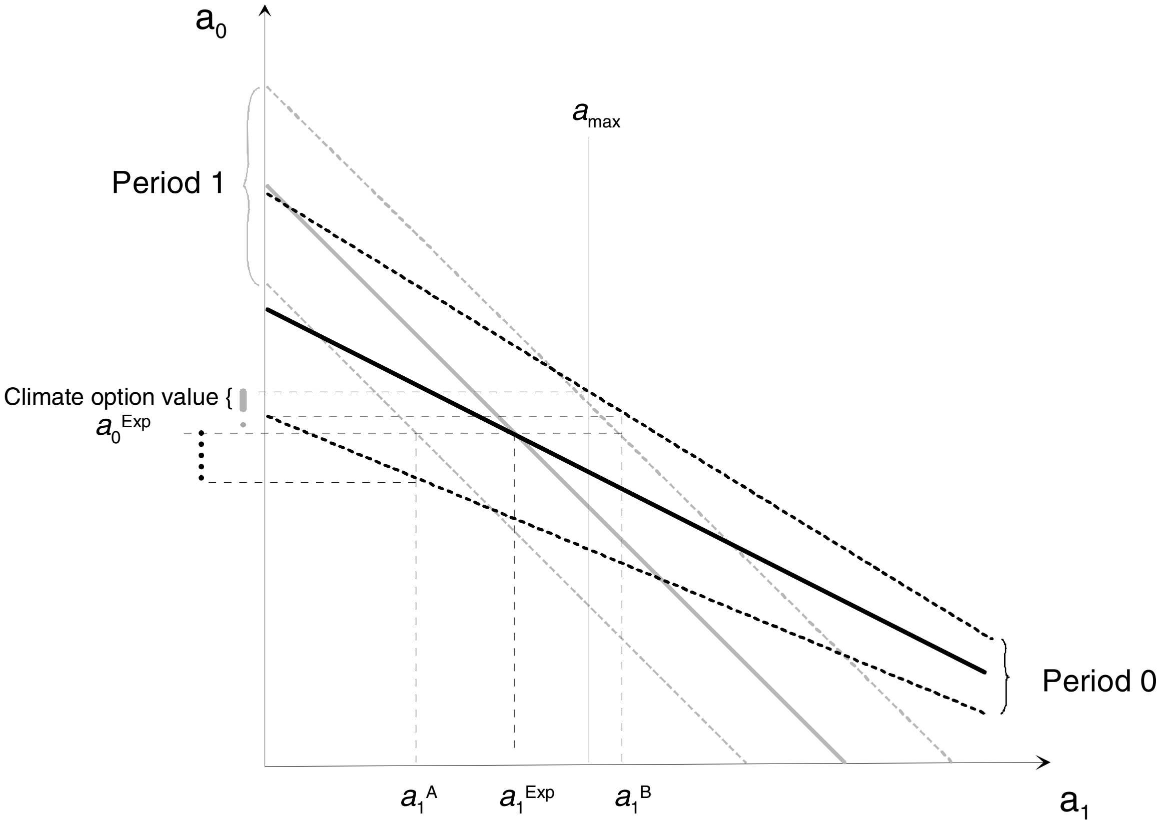
probability of encountering the climate irreversibility constraint, which depends on $a_1^{max}$, and the maximum possible abatement depends on the periodic emissions, which are exogenously given. In the special case of Figure 4C, the irreversibility constraint will be encountered under high damage costs if the policy in period 0 is based on low damage costs. Note that $a_1^{A(wrong)}$ represents the maximum possible abatement that the decision maker might want to choose in period 1. Hence, if $a_1^{max}$ is to the right of $a_1^{A(wrong)}$, there is no option value of early abatement. As pointed out in the introduction this is the case in Kolstad (1996b), where he does not have a binding climate irreversibility constraint in the parameterization of his model. This means that there is no option value of early abatement, and the climate irreversibility therefore, by assumption, has no impact on decisions, something that makes his conclusion rather obvious.

If instead the damage costs turn out to be lower than initially expected, one could think of a situation where it is optimal to emit more than $e_1$. This constraint does not depend on the initial period decision, but the likelihood of encountering it is of course higher with higher initial levels of abatement.

Recall that the rational decision maker will choose an abatement allocation between the two extremes discussed above. Figure 4D illustrates the option value when the decision is based on expected utility maximization. Compared to Figure 4C, the trade-off curves for state A and B are now dotted. The solid lines in between are the ones that apply for expected utility maximization with equal probabilities ($\pi = 0.5$). The optimal ex ante solution is $a_0^{Exp}$ and $a_1^{Exp}$.

The social losses of guessing wrong are represented by the vertical lines to the left of the $a_0$-axis. Again we look at the most serious mistakes possible. First, take the case where we learn at the end of period 0 that the state of the world is low damage costs (state A). Since we cannot do anything about our period 0 abatement cost choice, we must compensate by lowering the period 1 abatement cost until the first order condition for this period is satisfied, namely to $a_1^A$ in Figure 4D. The loss can be measured as the difference between $a_0^{Exp}$ and the $a_0$ value at the point where $a_1^A$ crosses the period 0 trade-off curve for state B. It is represented by the black dotted line to the left of the $a_0$-axis.

*Figure 4D. Cost of guessing wrong under expected utility maximization*



Second, suppose instead that we learn that the state of the world is high damage costs (state B). We then wish we had done more initially, and this must be compensated for by increasing future abatement. To satisfy the first order condition for period 1, we must choose $a_1^B$. This level is, however, not attainable since it is to the right of $a_1^{max}$. The best we can do then is to choose period 1 abatement equal to $a_1^{max}$. The cost of guessing wrong is represented in the figure by the grey line to the left of the $a_0$-axis, where the solid segment represents the option value and the dotted segment represents the adjustment cost.

How the option value of climate irreversibility is affected by the degree of uncertainty in damages and abatement costs is found by investigating the resulting shifts in the trade-off curves for state B (the upper pair in Figure 4D). In Appendix 1 we show that increased uncertainty in damage costs leads to an upward shift it the state B curves for both periods. This leads to an increase in the climate option value. The increase is, however, to some extent counteracted by the fact that the state B trade-off curve for period 0 becomes steeper, implying that the adjustment costs contribute a larger part of the total cost of guessing wrong, when damage costs increase. More uncertainty in the abatement costs also implies that the state B trade-off curve for period 0 and becomes gentler. The shift in the trade-off curve for period 1 is ambiguous. However, since the strongest shift is in the trade-off curve for period 0, the cost of encountering the climate irreversibility constraint, the climate option value, decreases.

To conclude so far, the option value of early abatement relates only to the costs of meeting the climate irreversibility constraint. Under a process of learning, additional costs will occur because of adjustments desired as a result of new information. In the model discussed here, the costs related to these adjustments depend on the slope and the position of the trade-off curves for the two periods. These positions depend on the choice of damage functions and abatement costs, and to which of these that are being subject to uncertainty. For example, Figure 4 is restricted to uncertainty in damage costs. If abatement costs are uncertain, the slope of the trade-off curve for period 0 becomes less steep for high costs without changing the intersection with the $a_1$-axis. The trade-off curve for period 1 shifts downwards and becomes steeper. Therefore, whether a policy close to the ‘wait and see’ strategy, i.e. a policy that puts relatively more weight on state A, should be preferred to a policy close to ‘early action’, i.e. a policy that put relatively more weight on state B, or vice versa, or whether the optimal strategy is exactly in the middle, depends on the specification of the model. The only conclusion we can draw so far, unless a numerical study is carried out, is that the climate irreversibility imposes an option value to ‘early action’ strategies as long as there is a positive probability of encountering the climate irreversibility constraint. This is, however, not sufficient for recommending such a strategy. Compared to the case with equal probabilities, such a recommendation also requires that the climate option value exceeds the loss resulting from deviating from this strategy in period 0. On the other hand, a ‘wait-and-see’ policy can never be expected to do better than the equal probability strategy when there is climate irreversibility only. The expected cost of encountering the climate irreversibility constraint always is higher in a ‘wait-and-see’ policy option, unless additional restrictions are imposed.

### *3.3 Irreversible abatement technology investments*

The model (1) to (4) assumes that the abatement level is set in each period, independent of previous actions. As Kolstad (1996a,b) and Ulph and Ulph (1995) point out, when focusing on the importance of irreversibility, it would be particularly inappropriate to disregard the fact that capital costs contribute to a large share of the costs of mitigating climate change. Furthermore, these investments are to a large degree sunk and lock the economy into a particular use of resources, thereby representing an irreversible cost. Capital irreversibility is therefore modeled by letting the initial investment costs, $a_0/c_0$, adhere to the next period. Thus, compared to above the capital investments are 100 percent irreversible which is a rather strong assumption, but chosen here to underscore the effects of capital irreversibility. Note that the interpretation of abatement investments is now somewhat different. One way of thinking about it is that the investment costs undertaken in period 0 only covers part of the total investment costs; that some of the costs, for instance in terms of interests and repayments, have to be paid in future periods. Equation (2) is replaced with

$$y_1 = y_0 + e_1 - \frac{a_0}{c_0} - \frac{a_1}{c_1}. \tag{14}$$

The capital irreversibility has two implications for future options, as long as $a_0 > 0$, because the initial abatement investment adheres to the next period. The maximum level of concentrations in the future is now reduced to $y_0 + e_1 - a_0/c_0$ compared with $y_0 + e_1$ in the previous sections. Consequently, the climate option value, the value of keeping open the option of further increasing the abatement in the future, is reduced because the initial investment will lower the abatement also in the future. Second, this means that the maximum level of abatement in the future now also depends on the investment decision taken initially. In other words, a wrong guess in period 0 may lead to negative optimal abatement in period 1. Since negative abatement is not possible, an extra cost may occur in period 1 because of the positive probability of encountering this capital constraint. Thus an extra value attaches to the 'wait-and-see' strategies that we choose to call *the option value of late investments* (investment option value). Inserting for (5) in the benefit function yields

$$b(y_1) = f(\hat{y}+e_0+e_1) - f(y_1) = \alpha(\hat{y}+e_0+e_1)^2 - \alpha(\hat{y}+e_0+e_1-2\frac{a_0}{c_0}-\frac{a_1}{c_1})^2 \tag{15}$$

which corresponds to equation (4) above.

The introduction of capital irreversibility thus implies that the initial abatement becomes 'twice as important' compared to the case with climate irreversibility only. Apart from this, the model is unchanged. The optimal solution can, in principle, thus be discussed within the same framework as in section 4.1 and 4.2. However, with capital irreversibility, the constants $C_0$, $\Phi_0$ and $\Psi_0$ change to:

$$C_0^* = (\hat{y}+e_0)\frac{2\alpha_0}{c_0} + \frac{2(\hat{y}+e_0+e_1)}{1+r}\left[\pi\frac{2\alpha_1^A}{c_0} + (1-\pi)\frac{2\alpha_1^B}{c_0}\right] - 1$$

$$\Phi_0^* = \frac{2\alpha_0}{c_0}\frac{1}{c_0} + \frac{2}{1+r}\left[\pi\frac{2\alpha_1^A}{c_0} + (1-\pi)\frac{2\alpha_1^B}{c_0}\right]\frac{1}{c_0}$$

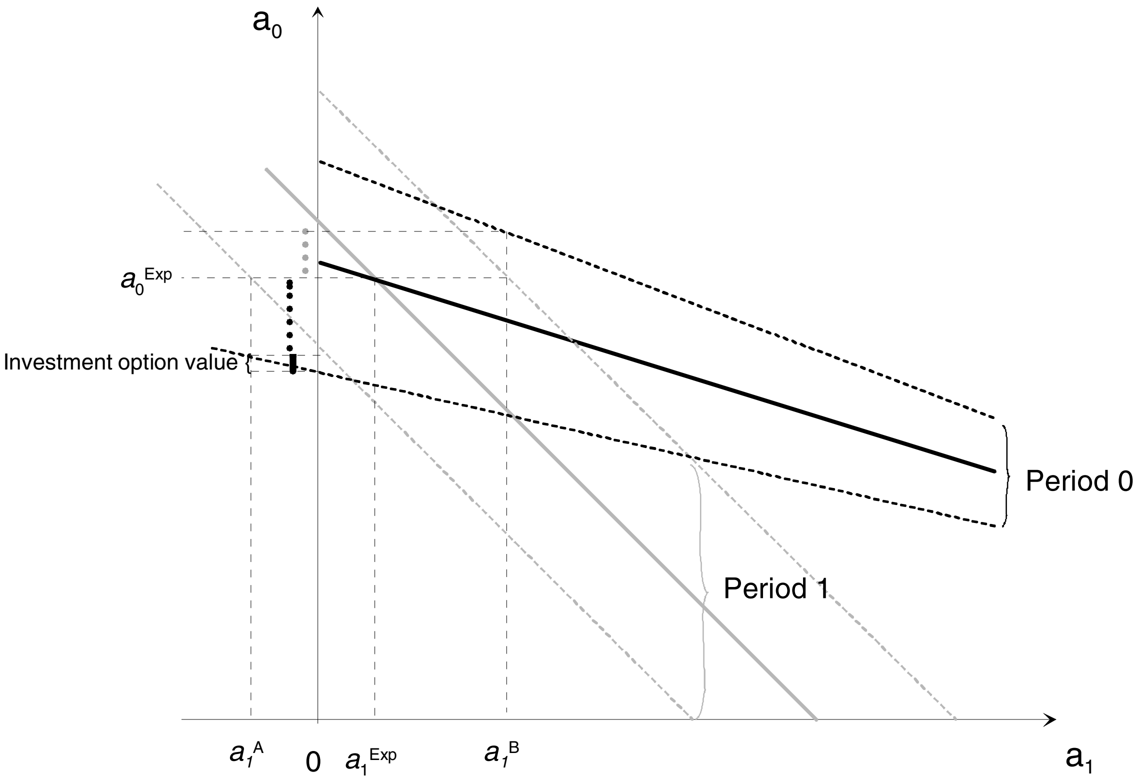
$$\Psi_0^* = \frac{2}{1+r}\left[\pi\frac{2\alpha_1^A}{c_0}\frac{1}{c_1^A} + (1-\pi)\frac{2\alpha_1^B}{c_0}\frac{1}{c_1^B}\right]$$

while the expressions for $C_1$, $\Phi_1$ and $\Psi_1$ are unaltered. Note that in this model, capital irreversibility with no depreciation implies only that the discount term $1/(1+r)$ is doubled. Ceteris paribus this leads to an allocation of abatement from period 1 to period 0. Total abatement is, however, likely to decrease. (See the discussion in section 4.1.)

Compared to Figure 4, the introduction of capital irreversibility leads to a positive shift in the trade-off curves for period 0, and they also become steeper, since $\Psi_0/\Phi_0 < \Psi_0^*/\Phi_0^*$. The probability of encountering the climate irreversibility constraint is thus lowered. Recall that this constraint is due to the restriction $a_1 \geq 0$; we cannot have negative abatement. The likelihood of encountering this constraint depends on the emissions under 'business as usual' in the future ($e_1 - a_0/c_0$).

The option value of late investment is illustrated in Figure 5. If the decision in period 0 is based on expected utility maximization with equal probabilities, we choose $a_0^{Exp}$, expecting to choose $a_1^{Exp}$ in period 1. If the low damage state occurs in period 1, it turns out that $a_0^{Exp}$ was more than enough investment for both periods. The optimal choice, according to the new information, is to lower $a_1$ until the low damage cost trade-off curve for period 1 is reached (the state A curve), that is to $a_1^A$. In Figure 5 this point is to the left of the $a_0$ axis and is not attainable due to the requirement of non-negative investments. The loss due to this overinvestment can be measured by the difference between the chosen policy, $a_0^{Exp}$, and the value of $a_0$ where $a_1 = 0$ crosses the low damage cost trade-off curve for period 0. The loss is indicated by the black bar along the $a_0$ axis in Figure 5, where the option value of late investment represents the solid segment.

*Figure 5. Cost of guessing wrong under irreversible investments and expected utility maximization*



Correspondingly, the loss when the high damage cost scenario is realized is the difference between the $a_0$ –value at the point were $a_1^B$ crosses the trade-off curve for period 0 in the high damage cost state and $a_0^{Exp}$. This is indicated by the grey dotted line to the left of the $a_0$ axis in Figure 5.

The effects from changes in the uncertainties in damage costs and abatement costs on the interception point of the state A trade-off curves with the $a_0$-axis affect the investment option value. We show in Appendix 1 that an increase in the uncertainty related to damages, implies that the interception point of the state A curve with the $a_0$ -axis for both periods shift downwards. This leaves open the question of whether it increases the investment option value. If the uncertainty in abatement costs increases, however, the interception of the period 0 curve shifts downwards, while the change in the interception point for period 1 is ambiguous. The investment option value is therefore likely to increase.

Note that we cannot say whether it is the climate irreversibility constraint or the investment irreversibility constraint that represent the larger ex ante costs. This depends on the choice of parameters, and the ‘business as usual’ emissions, $e_0$ and $e_1$. However, if the uncertainty in damage costs increases, we can say for sure that the climate option value increases, but the effect on the investment option value is ambiguous. With reasonable choices of parameters, however, an increase in the uncertainty of abatement costs will increase the investment option value, but tends to decrease the climate option value.

It may be noted, however, that a doubling of the discount factor to account for the investment irreversibility represents the maximum possible implication of this irreversibility, namely for the case where the rate of capital depreciation is zero, which is equivalent to the case analyzed in Kolstad (1996b). Introducing depreciation would reduce this term to somewhere between 1 and 2, depending on the rate of depreciation. If the periods under consideration are 20–25 years or more, the term is probably close to 1, which would be equivalent to the case analyzed by Arrow and Fisher, and Henry discussed in Section 4.2. Note, however, that climate irreversibility is also subject to ‘depreciation’ but the time perspective is substantially longer. The concentrations of $CO_2$, for example, are reduced by 2/3 over a period of 150–200 years, and 10 percent of emissions will remain in the atmosphere for thousands of years.

Thus, within the framework of this model, considering the time perspective of the global warming problem, the introduction of irreversibility in investments is not likely to have large impact. In principle, the requirement for an option value of early abatement to arise is similar to the case of reversible damage costs. However, the investment irreversibility scales down the future climate effects in the same way as an increase in the discount factor would. This implies that the possible cost of choosing a too low initial level of abatement is smaller. The conclusion is nevertheless similar: whether a possible climate option value of early abatement affects policy making depends on the degree of climate irreversibility, but this effect is reduced the more irreversible the investments. This last effect is, however, smaller the longer the time periods we make decisions for. We interpret this as having two option values that point in different directions, and whether the preferred policy option is an ‘early action’ or a ‘wait and see’ strategy depends on the relative size of these option values plus the possible adjustment costs. If the sum of the climate option value plus the adjustment costs due to too low initial abatement exceeds the sum of the investment option value plus the adjustment costs due to too high initial abatement, the net option value is positive, the climate irreversibility effect dominates, and ‘early action’ policies should be preferred. This shows that it is not possible to advocate either an ‘early action’ or ‘wait-and-see’ strategy merely on the basis of the existence of irreversibilities, nor on a direct comparison of option values. The additional expected adjustment costs resulting from deviations from traditional expected utility maximization, which ex ante are always non-negative, must also be considered.

## 4. Decision criteria

The question of irreversibilities in climate change policy has traditionally been addressed in an expected utility framework. In fact, all the papers cited above use this framework, and this literature indicates that climate irreversibilities should not have too much impact on the design of climate policy. It would be interesting to reconsider this result under alternative decision criteria. The discussion of figures 4A, B, and C was not tied to the principle of expected utility maximization. Hence, the figures could be used to consider implications of the use of alternative decision criteria that are based on the economic principle of equalizing marginal costs and benefits, given beliefs about future states of the world. We will in this section indicate briefly how this could be analyzed in the framework of this paper, saving the in-depth analysis for our future work.

The complexity of global warming makes it impossible to completely overlook the consequences of alternative choices. One might therefore ask whether this problem, which exhibits such severe forms of uncertainty, should be analyzed in a framework of ignorance, or at least partial ignorance. Theories of rational behavior under complete ignorance can be found for example in Arrow and Hurwicz (1972). Non-probabilistic criteria build on such a notion of ignorance. Critics of these criteria have put forward that the decision maker must at least have some vague partial information about the true state of nature (Luce and Raiffa 1957). The question remains, however, if this vague partial information is sufficient to assign subjective probabilities to the possible states of the world.

In Bretteville (2004) and Aaheim and Bretteville (2001), we examine the implication of the choice of a number of different decision criteria within a static setting for abatement investments. We found that the choice of criterion to a large degree will influence policy choice and is therefore a political question. It is therefore very interesting to examine how the inclusion of timing of policy, irreversibilities, and the possibility for learning may change the conclusions from a static analysis of climate policy.

The most well-known non-probabilistic criterion is perhaps the *maximin* principle (Rawls 1971). This principle implies maximization of the welfare in the worst possible case, and essentially it allows risk aversion to become infinite. It has been claimed that the conservatism of the maximin principle makes good sense in the context of climate irreversibility. Chevé and Congar (2002), for instance, argue that maximin is consistent with the precautionary principle. However, knowing the actual nature of the worst case is problematic. Another problem with the maximin principle is that the worst case might be a catastrophe of such dimensions that deciding between policy options might have no significant impact on the outcome. A third problem is that the conclusion depends on what you define as the worst case.

Applied to the problem of climate change, the maximin criterion can be interpreted as choosing the level of abatement that maximizes the social welfare in the worst possible state of the world. In the context of this paper, the worst state would mean that the decision in period 0 is based on both high damages and high abatement costs. If there is uncertainty attached to policy effectiveness, ‘early action’ cannot be rationalized as the appropriate maximin strategy because the worst case scenario would be to implement a costly remedial policy that fails to avert severe damages. Bouglet and Vergnaud (2000) analyzes the maximin criterion in a context of irreversibility theory and concludes that it does not necessarily lead to more flexible decisions than expected utility maximization.

Minimax regret (/risk/loss), suggested by Savage (1951) as an improvement on the maximin criterion, aims at minimizing the difference between the best that could happen and what actually does happen. The decision-maker tries to minimize possible regrets for not having, in hindsight, made the superior choice. In this global warming example, it can be interpreted as choosing the strategy that minimizes the difference in benefits between guessing right and wrong. This is easily connected to the discussion in Section 4. We found that the effect of a possible option value of early abatement on the maximum regret is less the more irreversible investments are, but this last effect is smaller the longer the time horizon. The possible mistake of choosing a too low level of abatement in the initial period is thus reduced compared to the case with climate irreversibility only. The preferred policy option depends on the relative size of the two option values and the possible adjustment costs. If the sum of the climate option value and the adjustment costs due to too low initial abatement exceeds the sum of the investment option value and the adjustment costs due to too high initial abatement, the net option value is positive, and the maximum regret is minimized by choosing 'early action' over a 'wait-and-see' strategy.[37]

It is obvious from the discussion above that the probability distribution over states will influence the net option value, and thus the preferred policy. Our model supposes two future scenarios. This can be interpreted as picking two of all the possible scenarios – one in the low cost range and one in the high cost range – or it could be

---

[37] Chevé and Congar (2002), however, claims that the minimax regret criterion is not consistent with the precautionary principle.

interpreted as some sort of a mean or median in the two groups. Out model can thus be used as a framework for analyzing the *Generalized Maximin/Maximax* criterion, also known as *the pessimism-optimism index criterion* of Hurwicz (1951). This criterion states that the level of abatement should be chosen in order to maximize a weighted average of the net benefit in the best and the worst state. The size of the pessimism-optimism index (the weight) should reflect the decision-maker's beliefs about the probabilities of facing different future states of the world. Whether the net option value is positive or negative with this criterion depends on the choice of focus, which states we choose as the worst and the best, and also the choice of weight. If the states in focus are the same as in the discussion above, the conclusions are, of course, unchanged.

Another criterion mentioned in the literature on decision-making under ignorance is *the principle of insufficient reason*, first formulated by Bernoulli in the 17th century. This criterion states that if there is no evidence leading us to believe that one event is more likely to occur than another, the events should be judged equally probable. In our model we have two possible scenarios, A and B, and they should thus each be assigned a probability of 0.5. In our two-period model with uncertainty and learning, this means that when we decide on the initial abatement we treat the two scenarios as equally likely, and when we decide on future abatement we have acquired new information and can adjust our emissions accordingly. If state B really is a low-probability extreme event, it would be weighted too heavily relative to the weight we would assign to it if we had had enough information to apply expected utility maximization. The occurrence of B is fortunate because then we are more likely able to increase our abatement sufficiently to avoid most of the damage. If the future optimal abatement turns out to be higher than the maximum, the climate irreversibility constraint will be encountered. If B does not occur, which is the most likely outcome, we did too much in the initial period and would thus like to increase emissions after learning the true state of nature. If future optimal abatement is less than zero, the investment irreversibility constraint is encountered.

Combinations of probabilistic and non-probabilistic criteria are also possible candidates for decision-making. The *limited degree of confidence* criterion is one example. It implies that we maximize a weighted sum of the expected utility criterion and the maximin criterion. The weight reflects the degree of confidence in the underlying probability distribution. In the case of full confidence, the weight is equal to 1 and the expected utility criterion is used, whereas under complete uncertainty the weight is equal to 0 and the maximin decision rule is applied. Lange (2003) compares expected utility with this criterion in a two-period model. He finds that more weight on the worst case (less weight on EU) may lead to increased first-period emissions and that the irreversibility effect holds if and only if the value of learning is negative.

## 5. Concluding remarks

Because of the irreversibilities related to climate change, a question arises of whether a more environmentally cautious policy (a policy based on beliefs of high future damages), could yield a higher expected benefit than a decision based on expected values. This depends on three factors: first, the size of the climate option value; second, the size of the investment option value; and third, the possible loss from deviating from traditional maximization of expected utility (meaning option values are disregarded). The net option value must thus exceed the expected value of this loss if a cautious initial policy is to be preferred.

The model discussion showed that whether the social loss of an ‘early action’ policy exceeds the social loss of choosing a ‘wait and see’ strategy, in the case of climate irreversibility only (the Arrow and Fisher, and Henry case), depends on the option value of early abatement as well as the possible abatement choice adjustments resulting from initial period mistakes. The option value of early abatement (the climate option value) depends on the probability of encountering the climate irreversibility constraint, which depends on the size of the maximum abatement in the future.

Including irreversibility in abatement investments (the Kolstad case) scales down the effects of climate irreversibility equivalent to the effect of an increase in the discount factor. Ceteris paribus this leads to an allocation of abatement from period 1 to period 0. Total abatement is, however, likely to go down. The possible mistake of choosing a too small abatement level in period 0 is reduced compared to the case with irreversibility only in emissions of greenhouse gases. We found that the effect of a possible climate option value on the policy choice should be less the higher the option value of late investments (the investment option value), but that this effect is smaller the longer the time horizon.

The option values relates only to the costs of encountering the irreversibility constraints. Under a process of learning, additional costs will occur because of adjustments desired as a result of new information. The climate option value is realized only if the desired future abatement, after learning takes place, is higher than what is actually possible. Then the climate irreversibility constraint is binding. Similarly, the investment option value is realized if the desired abatement investment, after learning takes place, is negative. Then the irreversibility constraint on investments is binding. The loss imposed when encountering the irreversibility constraints adds to the costs of the policy adjustments as a result of learning.

Changes in the parameters will affect total abatement, as well as the allocation between periods. If the value of damages increases, total abatement goes up, and future abatement increases relative to present abatement. An increase in expected future abatement unit costs contributes to an allocation of abatement from the future to the present. An increase in the discount rate affects, in this model, only the abatement assigned to period 0 which is reduced. Some of the reduced abatement may be allocated to the future, but the total amount will most likely go down. Changes in uncertainty contribute for the most part to a reallocation of abatement from the present to the future.

Choice of strategy in climate policy is not only a question of comparing costs, but also a choice of decision criterion. The choice of criterion will to a large degree influence policy choice and is therefore a political question. The model in this paper is applicable for analyzing and comparing alternatives to expected utility maximization. For example, if the choice between criteria is subject to a comparison between the required adjustment costs under extreme outcomes, the model can be used to attach relative numerical values to alternative strategies. This is a subject for future research.

## Acknowledgements

We are grateful to Geir Asheim, Scott Barrett, Espen Froyn, Fredric C. Menz, Ottar Mæstad, Jon Vislie, participants at the 12TH Annual Conference of EAERE, Bilbao, June 2003; in particular Ronan Congar, and our colleagues at CICERO for helpful discussions on this material and comments on earlier drafts. We also thank Espen Froyn for excellent Excel macro programming, Lynn Nygaard for editorial assistance, Tone Veiby for drawing Figure 1, Frode Rørvik for computer assistance, and the Research Council of Norway for financial support. Responsibility for errors and opinions are our own.

## APPENDIX 1:

### IMPLICATIONS FOR THE CONSTANTS OF A MEAN-PRESERVING SPREAD OF RISK

The constants on which the optimal solution is based are all subject to uncertainty in damage costs and abatement costs. Each constant will, however, be affected differently depending on which variable is subject to uncertainty. As a consequence, also the value of abatement will change when the policy is adjusted to account for new information, depending on the nature of this information. Below follows a discussion of how a mean-preserving spread of risk will affect the constants defined in Section 4. This affects, in particular, how the abatement is adjusted when new information arrives in period 1.

In general, the constants in Section 4 can be represented by three categories of expected values:

$$EX = \pi x^{A} + (1-\pi)x^{B}, \tag{A1}$$

which applies for all the constants when there is uncertainty only in the damage costs;

$$EX = \pi \frac{1}{y^A} + (1-\pi)\frac{1}{y^B}\,, \tag{A2}$$

which applies for $C_1$, $\Psi_0$ and $\Psi_1$ when there is uncertainty in the damage costs; and

$$EX = \pi \frac{1}{(y^A)^2} + (1-\pi)\frac{1}{(y^B)^2}\,, \tag{A3}$$

which applies for $\Phi_1$ when there is uncertainty in the damage costs.

A mean-preserving spread of risk means that $dX = \pi dx^A + (1\text{-}\pi)dx^B = 0$, or $dx^B = -\pi/(1\text{-}\pi)\ dx^A$. Hence, for (A1), $dEX = 0$ by definition. That is, the expected value is unaffected by a mean-preserving spread of risk in damage costs, if the abatement costs are certain.

For (A2), we have:

$$\begin{aligned} dEX &= \pi \frac{-1}{(y^A)^2} dy^A + (1-\pi)\frac{-1}{(y^B)} dy^B = -\pi \frac{1}{(y^A)^2} dy^A - (1-\pi)\frac{1}{(y^B)}\frac{-\pi}{1-\pi} dy^A \\ &= -\left[\pi \frac{1}{(y^A)^2} dy^A - (1-\pi)\frac{1}{(y^B)^2}\frac{\pi}{1-\pi} dy^A\right] = -\pi dy^A \left[\frac{1}{(y^A)^2} - \frac{1}{(y^B)^2}\right] \end{aligned} \tag{A4}$$

Therefore, $dEX$ is positive or negative depending on whether $y^A$ is higher or lower than $y^B$. If we talk about a spread of risk, $dy^A$ will be negative if $y^A$ is lower than $y^B$, and vice versa. Therefore, $dEX$ will increase with ‘more’ uncertainty in the abatement costs.

For (A3), we have:

$$\begin{aligned} dEX &= \pi \frac{-2y^A}{(y^A)^4} dy^A + (1-\pi)\frac{-2y^B}{(y^B)^4} dy^B = \pi \frac{2y^A}{(y^A)^4} dy^A - (1-\pi)\frac{2y^B}{(y^B)^4}\frac{\pi}{1-\pi} dy^A \\ &= 2\pi dy^A \left[ \frac{1}{(y^A)^3} - \frac{1}{(y^B)^3} \right] \end{aligned} \tag{A5}$$

Since a spread of risk means that $dy^A < 0$ if $y^A$ is lower than $y^B$ and vice versa, $dEX > 0$ when the uncertainty in abatement costs increases.

Table A1 sums up the effects on the expected value of the six model constants from a mean preserving spread of risk in damage costs and abatement costs respectively: 0 denotes unaffected, 1 represents a first-order effect described in equation A2, and 2 the second order effect described in equation A3.

*Table A1 Effects on model constants from a mean preserving spread of risk.*

| Constant | Damage costs | Abatement costs |
|---|---|---|
| $C_0$ | 0 | 0 |
| $C_1$ | 0 | 1 |
| $\Phi_0$ | 0 | 0 |
| $\Phi_1$ | 0 | 2 |
| $\Psi_0$ | 0 | 1 |
| $\Psi_1$ | 0 | 1 |

That is, a spread of risk in the damage costs does not affect the expected value of any of the constants, but a spread of risk in abatement costs does. From the effects on the constants, we can now discuss how the trade-off curves for state A, state B, and the expected outcomes with equal probabilities are affected by a spread of risk. These are summarized in table A2.

*Table A2 Effects of a spread of risk on the trade-off curves*

| Trade-off curve for: | Period 0 | | Period 1 | |
|---|---|---|---|---|
| | Interception with $a_0$-axis | Slope | Interception with $a_0$-axis | Slope |
| *Damages uncertain* | | | | |
| Expected | None | None | None | None |
| State A | Down | Gentler | Down | None |
| State B | Up | Steeper | Up | None |
| *Abatement costs uncertain* | | | | |
| Expected | None | Steeper | Down | Steeper |
| State A | Down | Steeper | ? | Steeper |
| State B | Up | Gentler | ? | Gentler |

The effect on the climate option value from increases in uncertainty depends on the shifts in the trade-off curves for state B. The more uncertain the damage costs are, the higher will the sate B curves for both period 0 and period 1 be, thereby increasing the climate option value. The increase is, to some extent, counteracted by the increased steepness of the state B trade-off curve for period 0, implying that the adjustment costs contribute to a larger part of the total cost of guessing wrong.

If the uncertainty in abatement costs increases, the state B trade-off curves become gentler. In addition, the curve for period 0 shifts upward while the shift in the trade-off curve for period 1 is ambiguous. However, since the stronger shift is in the trade-off curve for period 0, the climate option value decreases.

The cost of the investment option value can be read from the shifts in the interceptions of the state A trade-off curves with the $a_0$-axis. An increase in the uncertainty in damages shifts the interception point of both periods' trade-off curves downwards, leaving the effect on the investment option value open. If the uncertainty in abatement costs increases, however, the interception of the period 0 curve shifts downwards, while the change in the interception point for period 1 is ambiguous. The investment option value is thereby likely to increase.

## APPENDIX 2:

### SENSITIVITY ANALYSIS - A NUMERICAL EXAMPLE

In this numerical example we have, as far as possible, tried to base the choice of parameters on assumptions that seem to be standard in studies of optimal climate policy. The simple structure of the model in this paper nevertheless implies that the results can only be considered as illustrations.

Each period counts 25 years, making coverage of the period 2000 to 2050. We assume that the 'business-as-usual' emissions grow between 1.5 and 2.5 percent per year, proportional to the growth in GDP. What this implies for the exogenous variables $e_0$ and $e_1$ is not easy to say. This is because the model expresses emissions in units of concentrations, but it does not include a natural rate of decay in the concentrations. Hence, using the sum of emissions would, first, exaggerate the contribution to concentrations, and second, disregard the fact that emissions contribute less to concentrations the higher the concentrations are. We have therefore replaced emissions with additions to concentrations in each period from the chosen emissions assumption. The natural decay was set to 1/250, which is an approximation of the $CO_2$ decay rate in the long run.

The unit cost of abatement is based on an assumption that reducing emissions by 10 percent costs 0.75 percent of GDP. This is relatively high compared with other studies. For the damage costs, we assume 5 percent of GDP at a concentrations level of 550 ppmv. Again, this is high. The reason for choosing an upper level for both is that we assume that the policy is to be implemented for the entire world, whereas most cost estimates apply for developed countries, assuming that only these countries abate under the allowance of trading and CDM engagement.

The assumed unit cost of abatement corresponds to a cost lower than 0.5 USD/tC. This is extremely low compared with estimates of marginal costs in other studies (see e.g. IPPC, 2001). The reason is that we assume a constant unit cost of abatement, which means that the cost represents the average, rather than marginal cost, which is usually reported. The damage resulting cost per unit of concentrations is approximately 2 USD/ppmv. Based on the assumption for the damage costs, the constant of the damage cost function was set to 0.0015.

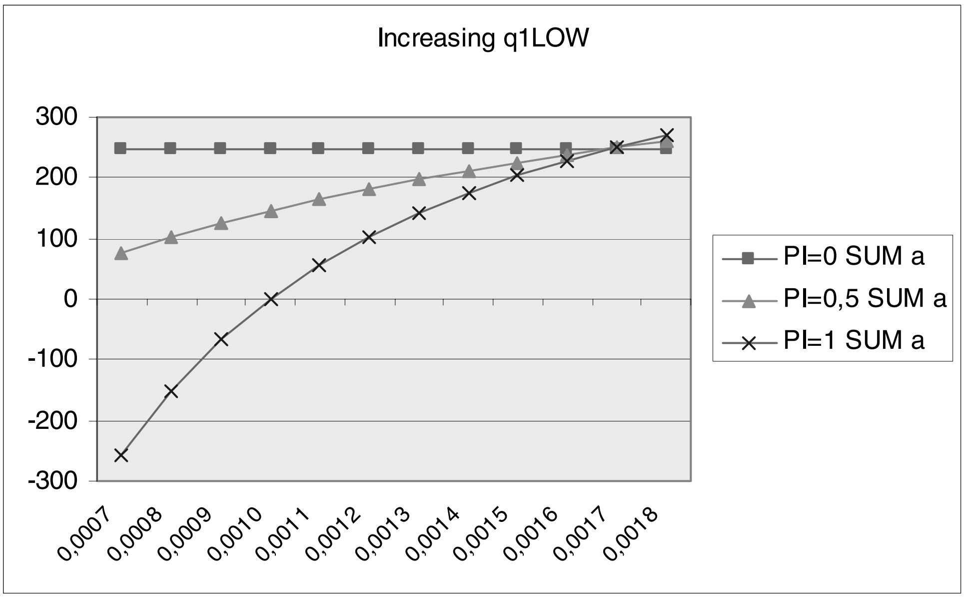
The base case for our numerical example is thus:

| y_hat | e0 | e1 | c0 | c1A | c1B |
|---|---|---|---|---|---|
| 368 | 63 | 116 | 1.15 | 1.1 | 1.3 |

| q0alpha0 | q1alpha1A | q1alpha1B | pi | r |
|---|---|---|---|---|
| 0.001 | 0.0015 | 0.0019 | 0.5 | 1.75 |

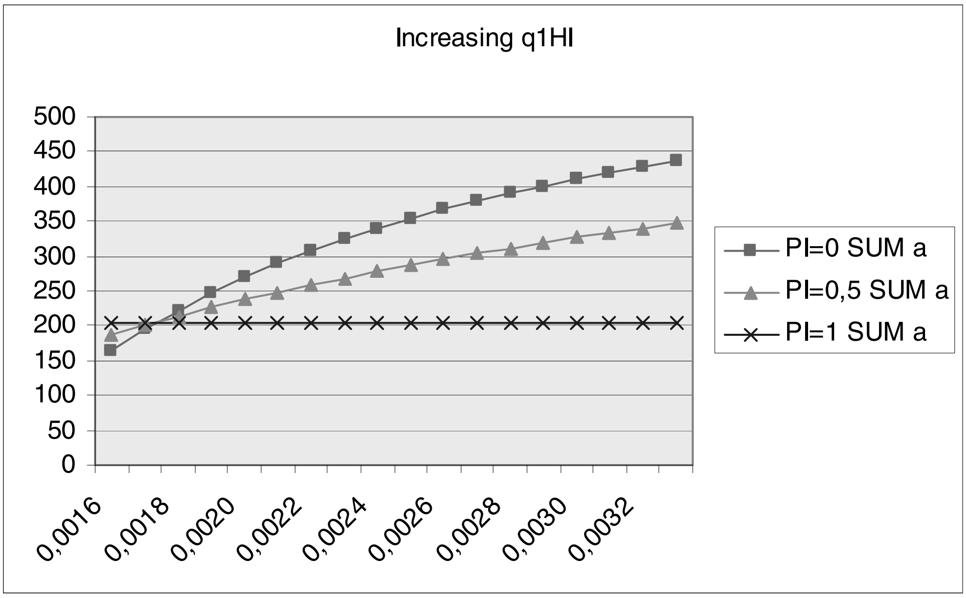
To show how sensitive the abatement decisions are to changes in the exogenous variables, we have made some calculations based on variations in one variable at the time illustrated in the figures below. First we examine the implication of increasing $q_1\alpha_1$ in the damage function. Figure A shows how the sum of the abatement in the two periods change with changes in $q_1\alpha_1$ for state A, which is the low damage scenario, for three different values of π. Recall that π is the probability of state A.

FIGURE A



Similarily, Figure B shows the development in total abatement when $q_1\alpha_1$ for state B increases. State B is the high damage cost scenario.

FIGURE B



The next two figures show the how the allocation of abatement between the two periods, as well as the total amount of abatement, changes with $q_1\alpha_1$ for state A (Figure C) and state B (Figure D) respectively, when the probability of state A is 0.5.

FIGURE C

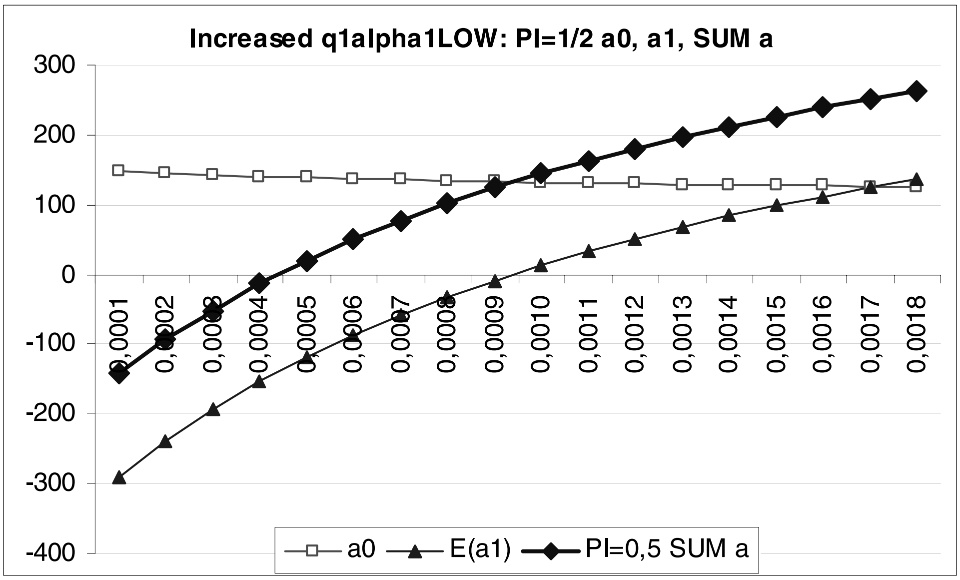
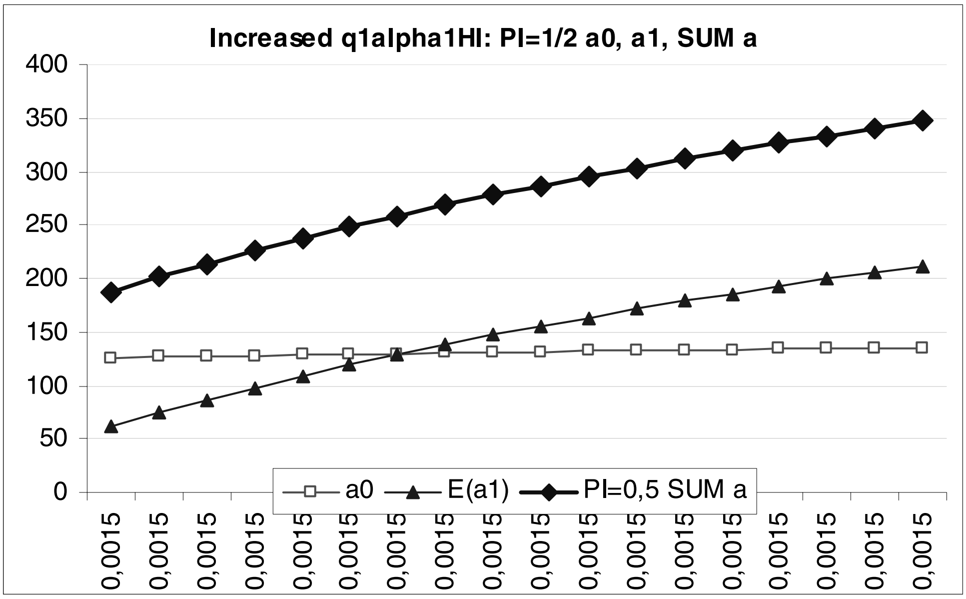


FIGURE D



In the same fashion, we have computed changes in abatement when the abatement cost in period 1 increases, first in the low-cost scenario (A), and second for the high-cost scenario (B). Figures E and F show the development in total abatement in the two cases for three different values of π.

FIGURE E

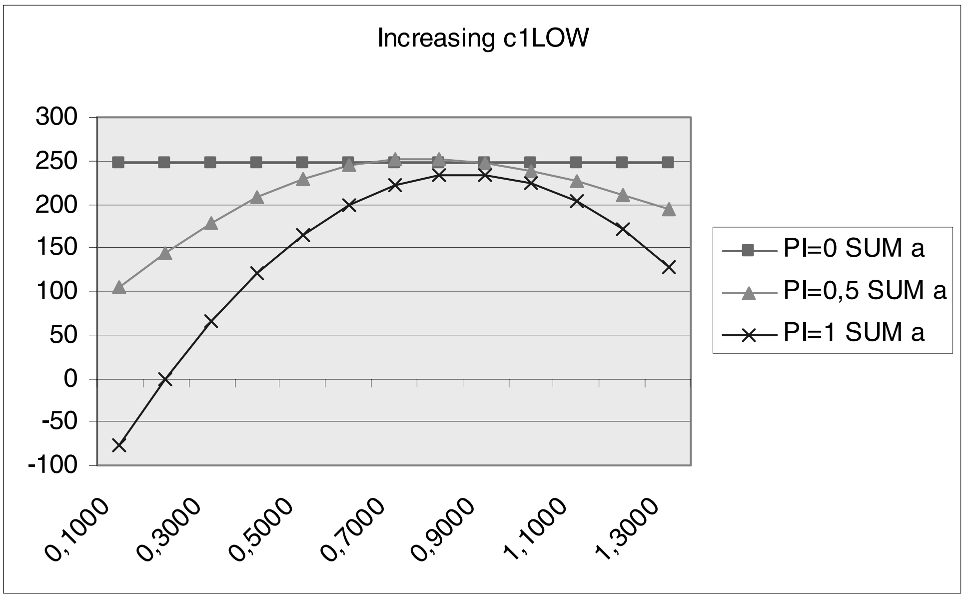
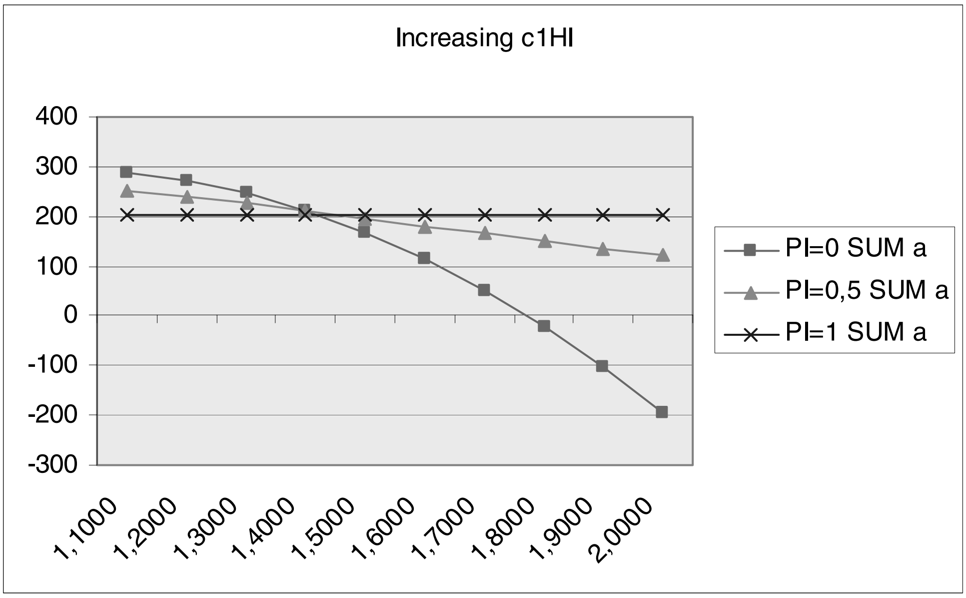


FIGURE F



Figures G and H show the changes in total abatement and how abatement is allocated from period 1 to period 0 when the future abatement costs increase and the probability of state A is 0.5.

FIGURE G

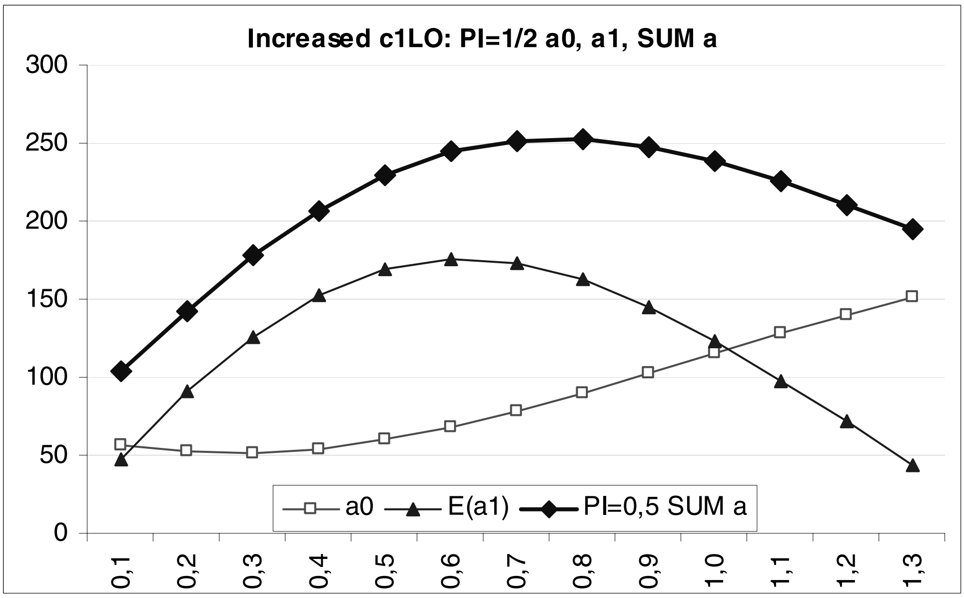
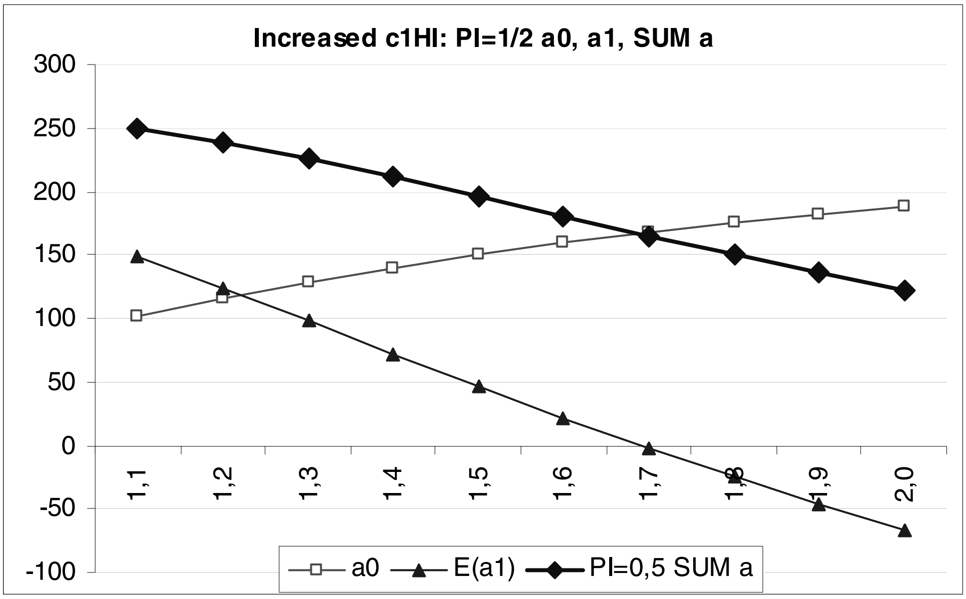


FIGURE H



Lastly we will show the implications on abatement decisions of changes in the discount rate. Figure I shows changes in total abatement for three different values of π.

FIGURE I

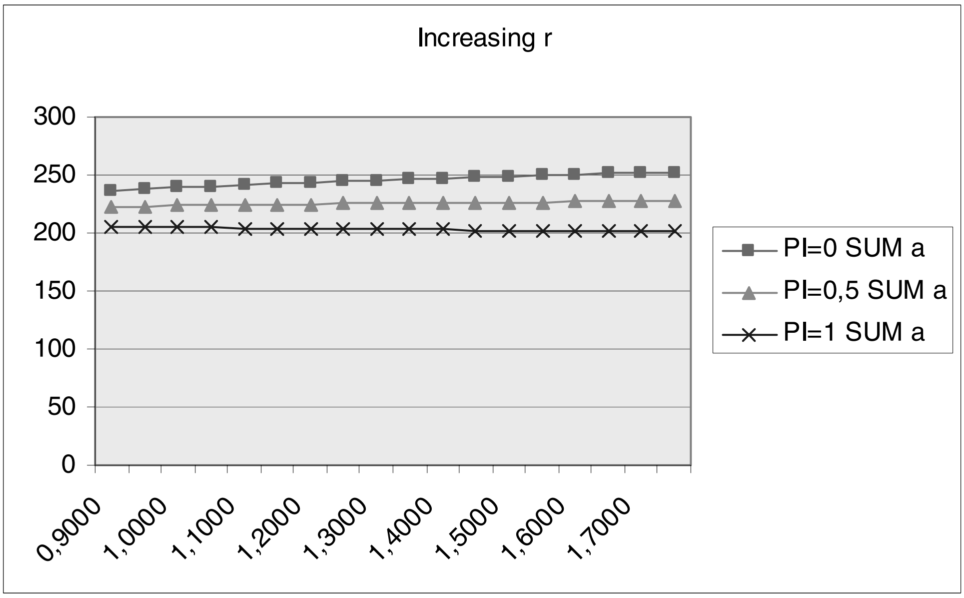
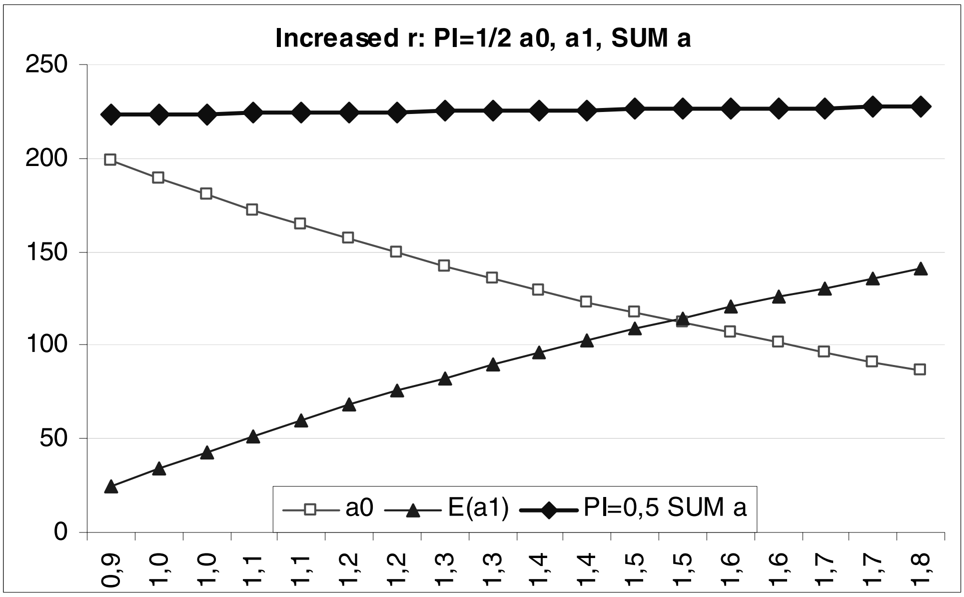


Figure J shows the allocation abatement for $\pi=0.5$. When the discount rate increases abatement is allocated from the present to the future, while the total amount is constant.

FIGURE J



## References:

Aaheim, H. A. (2004), ‘The Implication of Irreversibilities and Composites of Abatement Costs for Climate Policy’. MIMEO, CICERO.

Aaheim, H. A. and C. Bretteville (2001), ‘Decision-making frameworks for climate policy under uncertainty’. Working Paper 2001-02. CICERO, Oslo, Norway.

Arrow, K.J. and A.C. Fisher (1974), ‘Environmental Preservation, Uncertainty, and Irreversibility’. *Quarterly Journal of Economics* **88** (2), 312–319.

Arrow, K.J. and L. Hurwicz (1972), ‘An Optimality Criterion for Decision-making under Ignorance’, in Uncertainy and Expectations in Economics, C.F. Carter and J.L. Ford (eds.), Augustus M Kelly Publishers, New Jersey.

Aslaksen, I., K.A. Brekke, T.A. Johnsen, and H.A. Aaheim (1990), ‘Petroleum resources and the management of national wealth’ in Recent Modeling

Approaches in Applied Energy Economics, O. Bjerkholt, Ø.Oslen, and J. Vislie (eds.), Chapman and Hall, London.

Bouglet, T. and J.-C.Vergnaud (2000), ‘A non-Bayesian approach to decision-making with irreversibilities’, MIMEO, EUREQUA-University of Paris

Brekke, K.A. and Lystad (1999). Optimal CO2 emissions with irreversible, but depreciating abatement capital. MIMEO, Statistics Norway.

Bretteville, C. (2004), ‘Decision criteria, scientific uncertainty, and the global warming controversy’. Forthcoming, *Mitigation and Adaptation Strategies for Global Change.*

Chevé, M. and R. Congar (2002), ‘Managing environmental risk under scientific uncertainty and controversy’, in Proceedings of the International Conference on Risk and Uncertainty in Environmental and Resource Economics, June 2002, van Ierland, Weikard and Wesseler (eds.), Wageningen University, The Netherlands

Fisher, A.C. (2001), ‘Uncertainty, Irreversibility, and the Timing of Climate Policy’. Prepared for the conference on the ‘Timing of Climate Change Policies’, Pew Center on Global Climate Change, October 2001.

Fisher, A.C. and U. Narain (2003), ‘Global Warming, Endogenous Risk, and Irreversibility.’ *Environmental & Resource Economics* **25**, 395-416.

Hanemann, W.M. (1989), ‘Information and the Concept of Option Value’, Journal of Environmental Economics and Management **16**, 23-37.

Hanley, N., J.F. Shogren and B. White (1997). Envionmental Economics in Theory and Practice, OxfordUniversity Press, Oxford.

Heal, G. and B. Kriström (2002), ‘Uncertainty and Climate Change’, *Environmental & Resource Economics* **22**, 3-39.

Henry, C. (1974): “Option Value in the Economics of Irreplaceable Assets”, *The Review of Economic Studies* **41**, Issue Symposium on the Economics of Exhaustible Resources, 89–104.

IPCC(1995): *Climate Change 1995. Economic and Social Dimension of Climate Change. Contribution from Working Group III to the Second Assessment Report of the Intergovernmental Panel of Climate Change*. Cambridge University Press. Cambridge

IPPC (2001): *Climate Change 2001. Mitigation. Contribution from Working Group III to the Third Assessment Report of the Intergovernmental Panel of Climate Change*. Cambridge University Press. Cambridge

Kolstad, C.D. (1996a), ‘Fundamental Irreversibilities in Stock Externalities’, *Journal of Public Economics* **60**, 221-233.

Kolstad, C.D. (1996b), ‘Learning and stock effects in environmental regulation: The case of greenhouse gas emissions’, Journal of Environmental Economics and Management **31**, 1-18.

Lange, A. (2003), ‘Climate Change and the Irreversibility Effect – Combining Expected Utility and MaxiMin’, *Environmental & Resource Economic*s **25**, 417-434.

Lorentsen, L., K. Roland and A. Aaheim (1985): “Cost Structure and Profitability of North Sea Oil and Gas Fields”, in O. Bjerkholt and E. Offerdal (ed.): *Macroeconomic Prospects for a Small Oil Exporting Country*, Martinus Nijhoff Publishers. Dordrecht, The Netherlands.

Luce, R.D. and H. Raiffa (1957), Games and Decisions – Introduction and Critical Survey, Dover Publications Inc., New York

Pindyck, R.S. (2002), 'Optimal timing problems in environmental economics', *Journal of Economic Dynamics and Control*, **26,** 1677-1697

Rawls (1971), *A Theory of Justice*. Cambridge, MA.: Harvard University Press.

Savage (1951), 'The Theory of Statistical Decision'. *Journal of the American Statistical Association* **46,** 55–67.

Statistics Norway (2003), Statistics of Norway: Oil and Gas Activity 2003.

Torvanger, A. (1997), 'Uncertain Climate Change in an Intergenerational Planning Model'. *Environmental & Resource Economics* **9**: 103-124.

Ulph, A. and D. Ulph (1995): 'Global Warming, Irreversibility and Learning', *Discussion Paper* No. 9601, Department of Economics, University of Southampton.

# Essay no. 3:

# Regional versus Global Cooperation for Climate Control

CAMILLA BRETTEVILLE FROYN,[38] JON HOVI,[39] and FREDRIK C. MENZ[40]

## Abstract

This paper considers whether the climate change problem is better dealt with through regional cooperation than through a global treaty. Previous research suggests that, at best, a global treaty will achieve very little. At worst, it will fail to enter into force. Using a simple dynamic game-theoretic model, with weakly renegotiation proof equilibrium as solution concept, we demonstrate that two agreements can sustain a larger number of cooperating parties than a single global treaty, even when the cost of reducing emissions is the same. The model provides upper and lower bounds on the number of parties under each type of regime. It is shown that a regime with two agreements can Pareto dominate a regime based on a single global treaty. Thus, should the Kyoto Protocol not enter into force, cooperation based on regional agreements can be a good alternative. If Kyoto *does* enter into force, regional cooperation might be an option for future commitment periods.

**Key words:** Climate change, international environmental agreements, regional cooperation, the Kyoto Protocol, non-cooperative game theory, public goods, weak renegotiation proofness

---

[38] CICERO E-mail: camilla.bretteville@cicero.uio.no

[39] CICERO and Department of Political Science, University of Oslo E-mail: jon.hovi@stv.uio.no

[40] CICERO and Faculty of Economics and Finance, School of Business, Clarkson University, Potsdam, NY 13699, USA E-mail: menzf@clarkson.edu

## 1. Introduction

No international regime on climate change will be fully effective without the involvement of countries that have not yet committed to reduce emissions levels under the Kyoto framework, such as China, India, other developing countries, the United States, and Australia. In this paper we argue that individual countries might have a greater incentive to join a regional agreement than a global one. Regional arrangements for climate control are thus more likely to enter into force than a single global treaty. Furthermore, we show that an arrangement with more than one agreement can also sustain a larger number of cooperating parties. A global agreement like the Kyoto Protocol (was intended to be), requiring agreement among a large number of countries, is therefore not necessarily the best way to effectively control emissions of greenhouse gases.

There is speculation about possible U.S. interest in a regional climate arrangement that would include NAFTA members (Canada, Mexico, and the United States) and (perhaps) Australia. While such an arrangement is complicated because Canada is still within the Kyoto agreement and Mexico is exempt as a developing country, a western hemispheric continental agreement could be facilitated by the North American Commission on Environmental Cooperation. This existing institutional framework could oversee and monitor implementation of a North American emissions trading region should one be established. It has also been suggested that the United States cannot afford “climate isolationism” and should join with China or other developing countries in reducing greenhouse gas emissions. In either case, there are potential gains to the United States from extending emissions reduction possibilities beyond its borders, particularly in the form of an extended scope for emissions trading.

Regional cooperation on climate change is already taking place in the European Union (EU). In March 2000, the European Commission launched the European Climate Change Programme (ECCP) to coordinate the EU effort to reduce greenhouse gas emissions. This initiative has already resulted in ratification of Kyoto by the EU Commission in May 2002 and has implementation of an emissions trading scheme within the European Union scheduled for 2005. The EU-wide emissions trading system is intended not only to ensure that the EU achieves the eight percent cut in emissions by 2008–2012 to which it is committed under the Kyoto Protocol, but also to reduce member nation abatement costs and lessen competitive impacts of achieving compliance. While the EU directive governs only trades within the region, it provides that the EU may enter into separate agreements with non-EU countries to extend the scope of emissions trading beyond the Union.

Given the possibility of at least two regional climate policy arrangements and the likelihood of more in the future as developing countries commit to reduce greenhouse gas emissions, the next phase of the global accord for climate control might devolve into a series of more limited regional alliances, with global emissions trading among and within the regions. Previous research suggests that, at best, a global treaty will achieve very little and at worst, might fail to enter into force (e.g., Barrett 1999, 2003; Carraro 1998). It is therefore interesting to consider whether multiple regional treaties might be more successful than a single global treaty.

Using a simple dynamic game-theoretic model, we show that a regime based on two agreements can sustain a larger number of cooperating parties than a global treaty. This is true even when the cost of reducing emissions is the same for both types of regime. The model provides upper and lower bounds on the number of parties under each regime. It is shown that a system with two agreements can Pareto-dominate a regime based on a global treaty. We conclude that, should the Kyoto Protocol not enter into force, cooperation based on regional agreements could be a good alternative. And even if Kyoto *does* enter into force, regional cooperation might still be an option for future commitment periods.

The paper is organized as follows. The next section provides some background and reviews the existing literature on the conditions for a global climate treaty. The third section compares these conditions to the conditions for regional climate agreements in a simple dynamic model. The fourth section briefly discusses the possibility that a regime based on regional agreements could actually be chosen in future negotiations over a new design for the international climate regime. The paper closes with some concluding remarks.

## 2. Background and Literature Review

Pure public goods provide non-rival and non-excludable benefits for a given (local, national, or international) set of agents. Climate change mitigation is a pure public good on a global scale. Each country's emissions of greenhouse gases add to the world's atmospheric stock and each country's efforts will have only a minor impact on total global emissions. All parties thus have a strong incentive to free ride on other countries' efforts to reduce greenhouse gas emissions. A good is said to be impurely public to the extent that benefits are either partially non-rival or partially non-excludable. If the benefits can be excluded at an affordable cost, then pseudo-market arrangements in the form of clubs can collect tolls or membership fees to finance the good, reducing incentives for free-rider behavior (Buchanan 1965; Mueller 1989). In the absence of an exclusion mechanism, a real concern exists as to how countries will confront the incentive to free ride on the actions of others. In the extreme, no nation may act, but a more likely scenario is suboptimal provision by a few well-to-do countries (Olson 1965; Sandler 1998).

There is a considerable body of literature addressing under-provision of global international pollution control[41]. One focus of this literature is the conditions leading to the formation of multilateral agreements or coalitions. A fundamental assumption is that international agreements must be self-enforcing since there is no supranational authority to enforce compliance (Barrett 1994). The main problem analyzed by these models is that a country can – at least temporarily – achieve net gains by free-riding, or by joining a smaller but more stable coalition. There are two types of incentive for free riding: the incentive for a country to not sign the agreement and thus benefit from the signatories' abatement efforts; and the incentive for a signatory to violate commitments made in an agreement (non-compliance). Because of these free-rider incentives, there will generally be suboptimal equilibrium coalition structures in global pollution control (Finus and Rundshagen 2001).

To be effective, an international treaty needs to specify a strategy that can deter free riding and enforce compliance. Moreover, it must be in the best interest of each party to act in accordance with this strategy. A strategy is credible if no country is worse off accepting the agreement (*individual rationality*) and no sub-coalition of two or more countries can achieve a higher joint payoff by concluding a partial agreement (*collective rationality*). Furthermore, collective rationality implies that an equilibrium agreement must be renegotiation proof. In particular, it must be in the (collective) best

---

[41] Notable examples are Barrett (1991, 1994), Carraro and Siniscalco (1992, 1993), Hoel (1992), and Tulkens (1979).

interest of other countries to insist that a non-compliant country be punished before cooperation can be resumed (Barrett 1997; Finus 2001).

International environmental agreement models differ with respect to the specification of the utility functions of governments and the stability concept they employ. However, they can roughly be divided into two groups – dynamic game models and reduced-stage game models (Finus and Rundshagen 2001). The dynamic game models typically assume an infinitely repeated game where governments agree on a contract in the first stage that has to be enforced in subsequent stages by using credible threats (e.g., Barrett 1994; 1997; 1999).[42] Main results include that the allocation of abatement burdens crucially affects the success of international environmental agreements, that a grand coalition is unlikely to form, and that a sub-coalition may achieve more than the grand coalition. The reason for these results is that the larger the number of parties to an agreement (k), the greater the harm suffered by the k-1 other countries when they impose the punishment needed to deter a unilateral deviation, and consequently the less credible the threat.

Reduced-stage game models depict coalition formation as a two-stage game. In the first stage, countries decide on coalition formation. In the second stage, they choose abatement levels and how the gains from cooperation will be distributed (e.g., Chandler and Tulkens 1992; Carraro and Siniscalco 1993; Hoel 1992)). Some of these models define equilibria with both internal stability, meaning that no signatory has an incentive to leave the coalition, and external stability, meaning that no non-signatory wants to accede to the agreement. A key result is that the number of signatories generally falls

---

[42] Note that it is generally easier to sustain cooperation in infinitely repeated games than in finitely repeated games.

short of the grand coalition. Often, the equilibrium coalition is rather small. A second result is that the coalition typically achieves results far from the social optimum.[43]

Other reduced stage models (Bloch 1997; Carraro 1998; 1999; 2000; Carraro and Siniscalco 1998; and Siniscalco 1999) have addressed the possibility of giving countries the freedom to negotiate more than one climate agreement, much in the same way as the international trade regime allows countries to be parties to both regional and global agreements. A two-stage coalition game is used to show that when more than one coalition is possible, the equilibrium coalition structure that endogenously emerges from the negotiation process is characterized by several coalitions. It has also been shown that social welfare can be higher with multiple agreements than with a single global accord. Below we analyze the question of global versus regional agreements in an infinitely repeated game framework. Our results support the conclusions reached by Carraro and others using a reduced-stage game framework.

The model presented below follows Barrett (1999), but while Barrett (1999) uses the assumption of strong collectively rationality,[44] the equilibrium concept implied here is weak renegotiation proofness[45] (Farrell and Maskin 1989). Barrett (2002), also using weak renegotiation proofness, shows that a single treaty can be broadened to incorporate all countries (a consensus treaty), but at the cost of limiting the per-country level of cooperation. Allowing the depth of cooperation to vary, he demonstrates that countries might prefer a ‘broad but shallow’ treaty over one that is ‘narrow but deep’.

---

[43] Other models apply the concept of the *core* to determine the equilibrium coalition structure. A key result of these models is that by choosing a cleverly designed transfer scheme, the grand coalition establishing the social optimal emission vector *can* under special circumstances be an equilibrium, depending on the number of countries and on the degree of farsightedness. See Finus and Rundshagen (2001) for a review of these models. For overviews of the coalition literature, see Bloch (1997) and Finus (2001).

[44] The notion of strong collective rationality is explained in section 4.
[45] Weakly renegotiation proofness is defined formally in Section 3.2.

Thus, compared to Barrett (1999), Barrett (2002) changes both the equilibrium concept *and* the model. By contrast, we show in this paper that overall participation can be increased even *without* changing the model, using weakly renegotiation proof equilibrium as solution concept.

Sandler and Sargent (1995) discuss a number of coordination games where potential ratifiers are uncertain about the actions of others. They argue that if countries are distrustful of one another, a mixed strategy makes sense. Each country must then anticipate the likelihood that other countries will cooperate. Distrust is relevant whenever some parties have, or might have, an incentive to free-ride. There is little in the Kyoto agreement itself that ensures that ratifying countries will actually fulfill their obligations. However, the Marrakesh Accords provide details for a compliance mechanism for the climate regime. As argued by Barrett (2002; 2003) there are a number of problems with this mechanism.[46] It is therefore far from clear that Kyoto's enforcement mechanism will be able to ensure compliance. Hence, external means of enforcement are potentially relevant as an alternative – or a supplement – to the provisions of the Marrakesh Accords.[47] Regional agreements may contribute to this, because countries in the same region tend to be highly integrated. A high level of interdependence implies that a host of options are available (via issue linkages) for providing responses to non-compliance in any one particular issue area.[48] In addition,

---

[46] Barrett's main objections to the compliance mechanism are that (i) the punishment might be forever delayed; (ii) the anticipation of being punished is likely to induce countries to hold out for a generous allowance for the second period; (iii) there are no provisions for enforcement of failure by a non-compliant country to accept the punishment; (iv) a country that is being punished might choose to withdraw from the Kyoto Protocol; and (v) the compliance mechanism is not legally binding, and can be made so only through an amendment that must be ratified by the member countries.

[47] See Hovi (forthcoming) and Stokke (forthcoming) for discussions of the potential relevance and effectiveness of external enforcement in relation to the international climate regime.

[48] There is a whole body of literature suggesting that a solution to offset countries' free-riding incentives is the linkage of environmental negotiations to other economic issues (issue linkage). The idea is to link an issue with excludable benefits (a club good) to the public good provision. It has for example been

countries that are in close geographic proximity also tend to be culturally close, to have similar economic and political systems, and therefore to have similar preferences.[49] All of this might lower the costs of reaching agreement in the first place. Countries may thus not only be more likely to comply with a regional agreement, but also more inclined to join a regional agreement in the first place.

To sum up, a regime based on regional agreements might reduce uncertainty, ensure intra-treaty homogeneity, enhance trust, reduce negotiation costs, and hence encourage compliance and participation. All of these features represent a potential rationale for regional agreements, and could make a regime based on regional agreements an attractive option in the continuing negotiations on the international climate regime. However, we show below that a regime based on multiple agreements can increase participation even when the above advantages to a regional regime are disregarded. In fact, multiple agreements can be attractive even when countries are completely homogenous, so that no special cultural, economic, or political ties exist between particular subsets of countries.

## 3. A Simple Dynamic Model of Regional Climate Agreements

Consider a situation where the world consists of two regions (A and B) that are identical in all relevant respects. Assume that there are two options for an international regime to reduce emissions of greenhouse gases. Option 1 is a global agreement on emissions reductions and option 2 is a regime consisting of two separate agreements,

---

suggested to link the climate change negotiations with negotiations on trade liberalization (i.e. Barrett 1995; 2003) or research and development cooperation (i.e. Barrett 2003; Carraro and Siniscalco 1995; 1997).

[49] Of course, this is not always the case. In the Middle East, for instance, some neighbouring countries are likely to have very different preferences.

one for each region.[50] To isolate the effect of regime type, we assume that in both types of agreement, emission reductions are uniform *within* as well as *across* regions. Regardless of which option is chosen, the (periodic) cost of compliance is $c$.

In every period of the game, each country must choose to cooperate (i.e., reduce emissions) or to defect (not reduce emissions). Following Barrett (1997; 1999), in the case of a *global agreement* a country's periodic payoff if it chooses *cooperate* (C) is

$$\prod(C) = dk - c\,,$$

where d is a constant (d>0), and $k$ is the total number of countries that cooperate. For a country that chooses *defect* (D), the periodic payoff is similarly

$$\prod(D) = bk\,,$$

where $b$ is another constant ($b$>0). As a special case, we may have $b=d$. Similarly, with two *regional agreements*, the payoffs are:

$$P(C) = d(k_A + k_B) - c$$

$$P(D) = b(k_A + k_B)$$

---

[50] One commentator objected to this formulation saying that, since geography is not explicitly included, the real question addressed by the model is if two agreements are better than one. We agree. However, with two agreements, one would need a criterion to decide what countries to include in which agreement. Clearly, one possible criterion might be region, although other criteria are certainly also conceivable.

where $k_A$ and $k_B$ are the number of cooperating countries in regions A and B, respectively.[51] Like Barrett, we assume that $c > d \geq b > 0$.[52] We also assume that $b(N-1) > dN - c > 0$, meaning that full participation is *not* a Nash equilibrium of the stage game, but Pareto dominates the situation where no country participates. These assumptions are sufficient to show that *defect* is dominant in the stage game, and that the resulting Nash equilibrium (that all countries play defect), is inefficient. This means that the stage game is a special version of the N-person Prisoners' Dilemma.

Following Barrett, we assume that compliance in the global agreement is sustained by way of the "Getting Even" strategy.[53] This strategy specifies that a participating country plays c*ooperate* except if another participating country has been the sole deviator from Getting Even in the previous period, in which case *defect* is played. Non-participating countries play *defect* after any history.[54] It is well known that a global agreement of this type admits only a very limited number of parties (e.g., Barrett 1999).

In the case of two regional agreements, compliance is likewise sustained on the basis of a close relative to Getting Even. This strategy, which we refer to as "*Regional* Getting Even," specifies that a participating country plays cooperate except if another participating country *in its own region* has been the sole deviator from Regional Getting Even in the previous period, in which case defect is played. As in the case of a global

---

[51] We also did the analysis with cross regional heterogeneity in b and d, that is $b_A \neq b_B$ and $d_A \neq d_B$, but this did not change the results significantly.

[52] See Barrett (1997, pp.11-13).

[53] There are other strategies that might potentially sustain cooperation as a renegotiation proof equilibrium. Different strategies might potentially produce different results.

[54] This formulation of the Getting Even strategy differs slightly from Barrett's. In Barrett's formulation, a participating country is presented with the prospect of a future reward even if it does *not* take part in the punishment of a deviating country. Barrett's formulation requires that an additional condition be checked, which leads to non-trivial calculations. However, this additional condition does not seem to be an essential part of the strategy, and does not impinge on the main results presented in this paper. Hence, we prefer the slightly simpler formulation of the strategy given in the text.

treaty, non-participating countries play defect after any history. It may be noted that if all countries use this strategy, then a defection by a participating country in region A will be punished by other parties in region A, but *not* by parties to the agreement for region B, and vice versa.[55]

In order to be a weakly renegotiation-proof equilibrium (in the sense of Farrell and Maskin 1989), a strategy profile must satisfy two requirements. The first is that no player is able to gain by a one-period deviation in some state. This requirement is necessary and sufficient for the strategy profile to be a subgame perfect equilibrium. The second is that not all players gain if they deviate collectively from implementing the threatened punishment, given that a defection has taken place. The latter requirement ensures that the subgame perfect equilibrium is weakly renegotiation proof.[56] In the following, we compare a global agreement to a regime based on two agreements with respect to each of these two requirements. We assume throughout that all countries discount future payoffs using a common discount factor $\delta$ $(0 < \delta < 1)$.

### *3.1 The subgame perfection requirement*

Consider first the case of a global agreement. In this case there are two states to check. The first state is that there is no single deviator in the previous round. Getting Even then prescribes that all participating countries cooperate. As long as all participating countries follow this prescription, country *j* collects a payoff of $dk_g^* - c$ in each period,

---

[55] We show below that this is part of the equilibrium behaviour.

[56] Various concepts of renegotiation-proof equilibrium exist in the literature (see Bergin and MacLeod 1993 for a discussion). When we use that term in this article we always refer to the concept of weakly renegotiation-proof equilibrium, formalized for 2×2 games by Farrell and Maskin (1989), and extended to N-person games by Barrett (1994) and by Finus and Rundshagen (1998). Recall that every weakly renegotiation-proof equilibrium is a subgame perfect equilibrium (but not vice versa).

where $k_g^*$ is the number of parties in the global agreement. By contrast, suppose country *j* defects in period *t* and reverts to Getting Even at time $t+1$. It then gets a payoff of $b\left(k_g^*-1\right)$ in period t. In period t+1 it is punished by the other players and receive $d-c$. And at time t+2 cooperation is restored, meaning that country *j* obtains $dk_g^*-c$ in that and in any future period. It is individually rational to stick to Getting Even unless *j* gets a strictly larger payoff in periods *t* and $t+1$ by deviating. Formally, this condition is satisfied if $(1+\delta)\left[dk_g^*-c\right]\geq b\left[k_g^*-1\right]+\delta[d-c]$. Solving for $k_g^*$ gives:

$$(1)\ k_g^* \geq \tfrac{c-b+\delta d}{d-b+\delta d},$$

which may be rewritten as

$$(2)\ k_g^* \geq 1+\tfrac{c-d}{d-b+\delta d}.$$

If condition (2) is fulfilled, then the strategy profile where all countries play Getting Even is a Nash equilibrium. It may be noted that if $d=b$, then the right-hand side of condition (2) approaches $\tfrac{c}{d}$ as $\delta$ approaches 1. Thus, for cooperation to be a Nash equilibrium when $b=d$ and $\delta$ is close to 1, the number of cooperators must simply be large enough to secure each cooperating party a payoff above the zero payoff it would receive with no cooperation.[57]

---

[57] The intuition behind this result is that in an infinitely repeated game with no discounting all payoffs above the no cooperation baseline are sustainable in equilibrium (cf. the folk theorem).

The second state to check is the one where a single deviation *has* taken place in the previous period. If the country that deviated at time t 'accepts' the punishment and reverts to Getting Even at time t+1, it receives $d-c$ in period t+1 and $dk_g^*-c$ in any future period. By contrast, if it defects for one more period before it reverts to Getting Even, it receives a payoff of zero at time t+1 and $d-c$ at time t+2. Thus, it is individually rational to revert to Getting Even (at time t+1), provided that

$$(3)\ d-c+\delta\left[dk_g^*-c\right]\geq 0+\delta(d-c)\,.$$

Solving for $k_g^*$ gives

$$(4)\ k_g^*\geq \tfrac{c-d+\delta d}{\delta d}\,,$$

which may be rewritten as

$$(5)\ k_g^*\geq 1+\tfrac{c-d}{\delta d}\,.$$

Note that $1+\tfrac{c-d}{\delta d}\geq 1+\tfrac{c-d}{d-b+\delta d}$ as long as $d\geq b$ and $c\geq d$. Both of the latter conditions hold by assumption in our model. Hence, (5) provides the binding requirement for subgame perfection. This condition places a lower bound on the number of parties needed to make an agreement sustainable.[58]

---

[58] This is one possible explanation why most multilateral treaties (including the Kyoto protocol) provide a clause stating that their entry into force requires ratification by a certain number of countries.

Next, consider a regime with two regional agreements. In this case, there are three states to be considered to ensure subgame perfection. The first is the state where no single violation has taken place in any of the regions in the previous period. If country *j* in region A deviates in period *t* and reverts to Getting Even at time $t+1$, then in period t it collects a payoff of $b(k_A^* - 1 + k_B^*)$, where $k_S^*$ is the number of parties in the agreement for region S (S=A, B). In period *t*+1, the deviating party pays penance by being the only country in region A to comply, whereas cooperation continues as before in region B. Country *j* thus obtains a payoff of $d(1+k_B^*) - c$. At time $t+2$ compliance is restored, and so *j* receives $d(k_A^* + k_B^*) - c_r$ in that and in any future period. Had country *j* not deviated, it would have received $\mathrm{d}(k_A^* + k_B^*) - c$ in every period. So, it is individually rational to stick to Regional Getting Even, provided that:

$$(6) \qquad (1+\delta)\left[d(k_A^* + k_B^*) - c\right] \geq b(k_A^* - 1 + k_B^*) + \delta\left[d(1+k_B^*) - c\right]$$

By symmetry, we know that the number of parties in the two agreements must be identical. Setting $k_A^* = k_B^* = k_S^*$ and solving for $k_S^*$ gives:

$$(7) \qquad k_S^* \geq \tfrac{c-b+\delta d}{\delta d + 2(d-b)} \text{ (S=A,B),}$$

which may be rewritten as

$$(8) \qquad k_S^* \geq 1 + \tfrac{c-d-(d-b)}{\delta d + 2(d-b)} \text{ (S=A,B)}$$

Condition (8) ensures that the strategy profile where all countries use Regional Getting Even is a Nash equilibrium.[59]

The second state to check is the one where a single deviation has occurred in only one of the two regions in the previous period. Suppose the deviation takes place in region A. By reverting to Regional Getting Even at time t+1 the deviating country obtains a payoff of $d(1+k_B^*)-c$ in that period, and $d(k_A^*+k_B^*)-c$ in period t+2. By contrast, if it defects for one more period, and then reverts to Regional Getting Even, it receives $bk_B^*$ at time t+1 and $d(1+k_B^*)-c$ at time t+2. In either case, cooperation is restored from period t+3 onwards. Hence, it is individually rational to revert to Regional Getting Even (at time t+1), provided that

$$(9)\ d(1+k_B^*)-c+\delta\left[d(k_A^*+k_B^*)-c\right]\geq bk_B^*+\delta\left[d(1+k_B^*)-c\right]$$

Setting $k_A^*=k_B^*=k_S^*$ and solving for $k_S^*$ gives:

$$(10)\ k_S^*\geq \tfrac{\delta d+c-d}{\delta d+d-b},$$

which may be rewritten as

$$(11)\ k_S^*\geq 1+\tfrac{c-d-(d-b)}{\delta d+d-b}$$

---

[59] It may be noted that in the case where $d=b$, the right-hand side of condition (5) approaches $\frac{c}{d}$ when $\delta$ approaches 1.

For this state, we also need to check that after a violation in one region, it is individually rational for the parties in the *other* region to continue to comply. If a defection in region A occurs in period *t*, and all countries use Regional Getting Even from period t+1, then country *r* in region B receives a payoff of $d \cdot 1 + dk_B^* - c_r$ in period t + 1 and $d(k_A^* + k_B^*) - c_r$ in period t + 2. By contrast, if *r* deviates (defects) in period *t*+1 and reverts to Regional Getting Even in period *t*+2, it receives $b \cdot 1 + b(k_B^* - 1)$ and $d \cdot (1 + k_A^*) - c_r$ in those two periods. In either case, cooperation will be reestablished in both regions in period t + 3. Thus, country r has no incentive to deviate from Regional Getting Even if:

$$(12) \qquad k_S^* \geq 1 + \tfrac{c - d - (d - b)}{\delta d + d - b}.$$

Condition (12), which turns out to be identical to condition (11), says that if the number of participating countries in each region is sufficiently large, then it is not worth suffering a (temporary) break-down of cooperation in one's own region in order to punish a defection in the other region.

Finally, we must check the state where there is a sole deviator in *both* regions in the previous round. If all countries, including the deviator in the other region, play Regional Getting Even from time t+1 onwards, a deviating country obtains $2d - c$ at time t+1 and $dk_A^* + dk_B^* - c$ at time t+2. By contrast, if it defects for one more period (and then reverts to Regional Getting Even), it obtains *b* at time t+1 and $d + dk_B^* - c$ at time t+2. In either case, cooperation is restored in both regions from time t+3 onwards. Hence, it is individually rational to revert to Getting Even (at time t+1) if

$$(13)\;\; 2d - c + \delta\left[dk_A^* + dk_B^* - c\right] \geq b + \delta\left[d(1 + dk_B^* - c\right].$$

Setting $k_A^* = k_B^* = k_S^*$ and solving for $k_S^*$ gives

$$(14)\;\; k_S^* \geq \tfrac{\delta d + c + b - 2d}{\delta d},$$

which may be rewritten as

$$(15)\;\; k_S^* \geq 1 + \tfrac{c - d - (d - b)}{\delta d}$$

It is easy to verify that for $d = b$, conditions (8), (11), (12) and (15) all becomes identical. For $d > b$, condition (15) is stricter than (11), which in turn is identical to (12) and stricter than (8). Hence, (15) is the effective lower bound on the number of parties in a regime based on two agreements, and represents the binding condition for the regime based on two agreements to be a subgame perfect equilibrium.

To illustrate the difference between a global and two regional agreements, assume that $\delta = 0.95$, $c = 2$, and $d = b = 1$. We then find that a global regime requires at least three participating countries.[60] With two agreements, by contrast, the minimum number of parties is three in each region. For a second illustration, assume that the cost of compliance is eight rather than two. Then the minimum number of signatories in a global regime is nine, as opposed to nine in each agreement with a regional regime. In either case, the minimum number of parties in the regional regime is twice that of the global regime. If $b < d$, then the difference between the two types of regime can be

---

[60] Recall that this number must be an integer.

smaller. For example, assume that $\delta = 0.95$, $c = 8$, $b$ = 1/2, and $d$ = 1. Then the minimum number of parties with regional agreements becomes sixteen (eight in each region), as opposed to nine in the global agreement.

### *3.2 The renegotiation proofness requirement*

Somewhat loosely formulated, an equilibrium is renegotiation proof if the players as a group cannot find a better alternative (Bergin and MacLeod 1993). More precisely, a subgame perfect equilibrium in an infinitely repeated game is weakly renegotiation-proof if there does not exist two continuation equilibria such that all players are better off in one continuation equilibrium than in the other (Farrell & Maskin 1989:330-31). Assume that country $j$ defects in period $t$. For the *global agreement* to be weakly renegotiation proof, the other parties must be better off by carrying out the punishment than by restarting cooperation at once (i.e., without punishment).  If the punishment is carried out in accordance with Getting Even, then in period $t$ + 1 each of the other players receives a payoff of $b$. By contrast, if cooperation is restarted immediately, then each country receives $dk_g^* - c$ in that period. In either case, each player receives $dk_g^* - c$ in every future period, so all we need to consider is period $t$ + 1. Thus, renegotiation proofness requires that

$$(16) \qquad dk_g^* - c \leq b ,$$

which reduces to

$$(17) \qquad k_g^* \le \tfrac{b+c}{d}.$$

Thus, if the number of participating countries is too large, the payoff from renegotiation (i.e., resuming cooperation immediately) exceeds the payoff from carrying out the punishment.

While the subgame perfection requirement places a *lower* bound on the number of parties in the agreement, the renegotiation proofness requirement places an *upper* bound on this number. To illustrate, suppose that $d = b = 1$ and that $c = 2$. Then renegotiation proofness requires that the number of signatories in the global agreement is at most three. Similarly, if the cost of compliance is eight rather than two, the maximum number of signatories is nine. In other words, the maximum number of parties *increases* with the cost of cooperation. This result may be explained as follows. After a defection it must be better for other parties to punish than to renegotiate (i.e., to restart cooperation without punishment). Given that cooperation is sustained by the Getting Even strategy, this is equivalent to a requirement that choosing *defect* when only one country cooperates must be at least as good as the outcome where all participating countries choose *cooperate*. This condition defines the maximum number of parties. Increasing the cost of cooperation is equivalent to a constant, negative shift in the payoff function for cooperation, meaning that cooperation (and hence renegotiation) becomes less attractive. Thus, the maximum number of parties becomes larger.

Next, consider *regional agreements*. Assume that country $j$ is a party to the agreement in region A, and that this country defects in period $t$. If the punishment is carried out in accordance with Regional Getting Even, then each of the other countries in region A receives $b(1+k_B^*)$ in that period. By contrast, if cooperation is restarted at once, each of them receives $\mathrm{d}(k_A^* + k_B^*) - c$ in period t+1. Renegotiation proofness therefore requires that

$$(18) \qquad \mathrm{d}(k_A^* + k_B^*) - c \le b(1 + k_B^*)$$

With identical regions, the condition reduces to:

$$(19) \qquad k_S^* \le \tfrac{b+c}{2d-b} \text{ (S=A, B).}$$

It follows that a regime with two separate agreements allows more countries to cooperate in weakly renegotiation proof equilibrium than a regime based on a single global agreement. For example, if $c = 2$, and $d = b = 1$, then with two agreements a total of 3 + 3 = 6 countries can cooperate in renegotiation proof equilibrium. Similarly, if $c = 8$, then the two regional agreements admit a total of 9 + 9 = 18 countries. In either case, the regional regime doubles the maximum number of cooperating countries, compared to a single global agreement.

In combination with the subgame perfection requirement, the renegotiation proofness requirement leaves a very narrow set of possibilities regarding the number of parties in an international regime. In fact, the number of parties is sometimes fully determined. For example, with $\delta = 0.95$, $c = 8$ and $d = b = 1$, a global agreement requires exactly nine parties. By contrast, a regime based on two regional agreements must have exactly eighteen parties (nine in each region).

### *3.4 Discussion*

The model analyzed above has a number of interesting implications. First, it provides upper and lower bounds on the number of parties in each type of agreement. As we have seen, the two boundaries might even fully determine this number. The participating countries' identities are, however, *not* generally determined.[61]

Second, a regime based on regional agreements admits a larger number of parties (for the two regional agreements taken together) than the global regime. At the same time, however, the required number of parties in the regional regime is also larger than the number required by a global treaty.

Third, while the model analyzed here includes only two regions, one might reasonably ask why one should stop there. Why not design a regime with an even larger number of agreements? This idea can certainly not be dismissed on the basis of the above model. We have seen, however, that there is a bottom line to the number of parties in each regional agreement. Hence, with a given pool of countries to select from (not all of which are likely to join any kind of climate regime, at least not at the present time) there is an upper limit to the number of agreements that can be sustained in equilibrium.

Fourth, the model suggests that a regime based on regional agreements has certain advantages in terms of efficiency. Note first that a global treaty is Pareto inefficient if

$$(20)\ bk_g^* < dN - c$$

---

[61] The parties' identities are determined only if the number of participating parties in the two agreements add (exactly) up to the total number of countries, i.e., $k_A^* + k_B^* = N$. If $k_A^* + k_B^* < N$, the selection of participating countries might represent a substantial co-ordination problem.

i.e., if the total pool of countries is sufficiently large compared to the number of parties in the global agreement. Next, compare the following two scenarios:

- A global regime where $k_g^*$ countries cooperate.
- A regime with two regional agreements where $k_A^* \leq k_g^*$ countries cooperate in one agreement, and $k_B^* \leq k_g^*$ other countries cooperate in another agreement, such that $k_A^* + k_B^* > k_g^*$.

With linear payoff functions, it follows immediately that countries included as parties under *both* regimes prefer the regional regime. The same is true for countries that are *non-parties* in both cases. The reason is simply that, from the point of view of these countries, the number of *other* countries that cooperate is larger in the regional than in the global regime. The question is therefore whether the regional agreements regime would be acceptable to the members of the additional agreement, i.e., to countries which are allowed to free ride in the first scenario, but required to cooperate in the second. With a global agreement, each of these countries receives a (periodic) payoff of $bk_g^*$. With two regional agreements, each of them obtains $d(k_A^* + k_B^*) - c$. Hence, if $d(k_A^* + k_B^*) - c \geq bk_g^*$, i.e, if

$$(21) \quad k_A^* + k_B^* \geq \tfrac{bk_g^{'} + c}{d},$$

each of these countries is at least as well off being a party to one of the two regional agreements as it is free riding on a global agreement. It follows that, whenever condition (21) is fulfilled, a regime based on regional agreements Pareto dominates a global agreement. To see that this condition is always fulfilled in the regional agreements equilibrium, recall that it is feasible for any participating country in the additional agreement to deviate by always playing defect. If country *j* starts defecting in period t, other countries in the same region will also defect from period t+1 on, thereby effectively terminating the additional agreement. Defecting from a regional agreement in this way would give the deviator a higher payoff than it would obtain with the global agreement. However, it follows from the subgame perfection requirement that such a deviation is not profitable. Hence, whenever the subgame perfection requirement is fulfilled, two regional agreements Pareto dominate a single, global agreement.

For an analytical demonstration, let the additional agreement be in region A. Multiplying (6) by $1-\delta$ and rearranging terms gives

$$(22)\ \ d(k_A^* + k_B^*) - c \geq (1-\delta)b\left[k_A^* - 1 + k_B^*\right] + (1-\delta)\delta\left[d(1+k_B^*) - c\right] + \delta^2\left[d(k_A^* + k_B^*) - c\right] .$$

By rearranging (9) we get

$$(23)\ \ (1-\delta)\left[d(1+k_A^*) - c\right] + \delta\left[d(k_A^* + k_B^*) - c\right] \geq bk_B^* .$$

Combining inequalities (22) and (23) gives

$$(24)\ \ d(k_A^* + k_B^*) - c \geq (1-\delta)b\left[k_A^* - 1 + k_B^*\right] + \delta b k_B^* ,$$

which may be rewritten as

$$(25)\ \ d(k_A^* + k_B^*) - c \geq bk_B^* + (1-\delta)b\left\lfloor k_A^* - 1 \right\rfloor.$$

Since $k_A^* > 1$, condition (21) follows with *strict* inequality. Hence, even the participating countries in region A strictly gains by forming an additional agreement.[62]

Fifth, the above model also identifies an important condition for a regime based on regional agreements to be sustainable. Recall that a regime based on regional agreements not only admits, but also requires, a larger number of signatories than a global agreement. This means that only if a sufficiently large number of countries are willing to accept international commitments to reduce emissions of greenhouse gases will a regime based on regional agreements be an attractive option. If this number is too small to sustain two regional agreements, then a global agreement might be the only viable option.

Finally, an important driving force in Barrett's model is the way in which the agreement is enforced. The reason why the maximum number of signatories is so limited is that a single violation triggers punishment by all other parties. Our model provides less pessimistic results because it allows each agreement to be enforced regionally. The enforcement mechanism is similar to Barrett's in the sense that all parties to the agreement for region A punish a violation of the agreement *for region A*. However, since cooperation in region B is allowed to continue undisturbed by a violation in region A, the total number of punishers becomes smaller than in Barrett's model. It is this feature that allows the maximum number of signatories to become larger with a regime based on regional agreements.

---

[62] We are indebted to Geir Asheim for this proof. Note that the analytical argument follows the verbal one.

## 4. Regional Agreements, Enforcement, and Participation

In the previous section we have shown that two agreements can accommodate a larger number of parties than what is possible in the scenario with a single global agreement studied by Barrett (1999). It is important to note that we are using a weaker concept of collective rationality than Barrett (1999). In our model, we rely on the concept of weakly renegotiation-proof equilibrium, which requires that participating countries cannot do better collectively than to impose the prescribed punishment if a single deviation has occurred in the previous period. By contrast, Barrett (1999) adds the requirement that if these countries can choose to impose other punishments than the one specified by the treaty, they can not gain collectively by doing so (strong collective rationality).

The use of weakly renegotiation proof equilibrium also opens other avenues for broadening the number of parties, in addition to the multiple agreements approach highlighted in this paper. While the informal arguments presented in Section 2 suggest that a regional approach might be an attractive option, other alternative enforcement regimes would also admit a larger number of signatories than Barrett's design for a global treaty. For example, the maximum number of parties in a *global* agreement might be increased by designing an enforcement scheme whereby a single defection triggers punishment only by a subset of the other parties.[63] Indeed, even a global agreement could – in principle – be enforced regionally, and so the effects analyzed here could also be achieved within the setting of a global agreement.[64] However, we find it rather

---

[63] In fact, in the appendix to Barrett (1999) this type of regime is examined, but rejected on the ground that it fails to be strongly collectively rational.

[64] Another option could be to have both a global *and* a set of regional supplementary agreements.

implausible that a system of regional enforcement would be adopted except as part of a regime based on regional agreements.

This brief discussion reflects the fact that the above model admits a multitude of equilibria (involving various degrees of cooperation). Is it likely – or at least possible – that the regional agreements equilibrium would actually be chosen in future negotiations over a new design for the international climate regime? Obviously, we can not provide a full discussion of this question here. However, an attractive feature of the two agreements equilibrium is that the punishment is carried out equally by all parties in the same region. Thus, the right to reap the benefits connected with punishment would seem to be fairly distributed (provided that defections are evenly distributed between regions). Granted, a fair distribution of the right to punish might also be possible – at least in principle – in equilibria based on a global agreement where only a subset of the participating countries punishes a deviation. With homogenous countries, however, this assumes that over time, the right to punish is allocated to different subsets of countries, so that in the long run, all countries are allowed to reap the benefits of punishing equally. To the extent that deviations are rare phenomena (which they should be, according to the theory), and countries discount future benefits, the fairness problem might well prove to be real with this kind of enforcement mechanism. Of course, with heterogeneous countries, other difficulties (such as different notions of fairness) are likely to come in addition.

One might also ask why a defection in one region ought to be enforced by other members of the same agreement. An obvious alternative is that a sole deviation in one region is punished by all countries in the *other* region. Again, the distribution of the rewards of punishment would seem fair (assuming that defections are evenly distributed over regions). However, with heterogeneous countries, punishment by other countries in the same region arguably comes closer to the notion of having a defector judged by a 'jury of its peers'. If so, it might be easier for a defecting country to accept punishment by countries in its own region than to accept punishment by countries that are more distant in terms of both geography and culture.[65]

## 5. Concluding Remarks

This paper has used a simple dynamic game-theoretic model to analyze how a regime based on two agreements would affect the prospects for cooperation to reduce emissions of greenhouse gases. A main finding is that two treaties are able to encompass a larger number of parties than a single global treaty, assuming that the cost of reducing greenhouse gases is the same with both types of regime. We have also shown that a system with two agreements can Pareto-dominate a regime based on a global treaty. These results support the findings of previous research by Carraro et al., using a different type of model. [66]

The model analyzed in this paper follows Barrett (1999), but while Barrett (1999) uses the assumption of strong collectively rationality, the equilibrium concept implied here is weakly renegotiation proofness (in the sense of Farrell and Maskin 1989). We have emphasized that the use of weakly renegotiation proof equilibrium also opens other avenues for broadening the number of parties besides the multiple agreements approach highlighted in this paper. However, informal arguments presented in Section 2 suggest that a regime based on regional agreements can reduce uncertainty,

---

[65] Interestingly, empirical studies on economic sanctions suggest that sanctions imposed by friends and allies tend to be more effective than sanctions imposed by more distant countries (Hufbauer, Schott and Elliott 1990).

[66] For example, Carraro (1998, 1999, 2000), Carraro and Siniscalco (1998), and Siniscalco (1999).

ensure intra-treaty homogeneity, enhance trust, and reduce negotiation costs. These factors further strengthen the main conclusions of the formal analysis in section 3. Thus, should the Kyoto Protocol not enter into force, cooperation based on regional agreements can be a good alternative. If Kyoto *does* enter into force, regional cooperation might be an option for future commitment periods.

## Acknowledgements

We would like to thank Geir Asheim for detailed and extremely helpful comments to several previous versions of this paper. We are also indebted to Scott Barrett, Leif Helland, Ottar Mæstad, Kevin Siqueira, Jon Vislie, our colleagues at CICERO, in particular Odd Godal and Asbjørn Torvanger, and two anonymous referees for helpful discussions and comments on earlier drafts. Finally we thank the Research Council of Norway for financial support, Lynn Nygaard for editorial assistance, and Frode Rørvik for computer support. The usual caveats apply.

## References

Barrett, S. (1991), ‘The Paradox of International Environmental Agreement’, MIMEO, London Business School, London.

Barrett, S. (1994), ‘Self-enforcing International Environmental Agreements’, *Oxford Economic Papers* 46, 878-94.

Barrett, S. (1995), ‘Trade restrictions in International Environmental Agreements’, London Business School, London.

Barrett, S. (1997), ‘A Theory of International Cooperation’, MIMEO, London Business

School, London.

Barrett, S. (1999), ‘A Theory of Full International Cooperation’, *Journal of Theoretical Politics* **11,** 519-541.

Barrett, S. (2002), ‘Consensus Treaties’, *Journal of Institutional and Theoretical Economics* **158,** 529-547.

Barrett, S. (2003), Environment & Statecraft. The Strategy of Environmental Treaty-Making, Oxford University Press, New York.

Bergin, J. and B. MacLeod (1993), ‘Efficiency and Renegotiation in Repeated Games’, *Journal of Economic Theory* **61**, 42-73.

Bloch, F. (1997), ‘Non-cooperative Models of Coalition Formation in Games with Spillovers’, Chapter 10 in C. Carraro and D. Siniscalco (Eds.), New Directions in the Economic Theory of the Environment, Cambridge University Press, Cambridge.

Buchanan, J.M. (1965), ‘An Economic Theory of Clubs’, *Economica* **32,** 1-14.

Carraro, C. (1998), ‘Beyond Kyoto, A Game Theoretic Perspective’, in the Proceedings of the OECD Workshop on Climate Change and Economic Modeling. Background Analysis for the Kyoto Protocol, Paris, September 17-18.

Carraro, C. (1999), ‘The Structure of International Agreements on Climate Change’, in C. Carraro (Ed.), International Environmental Agreements on Climate Change, Kluwer Academic Pub, Dordrecht.

Carraro, C. (2000), ‘The Economics of Coalition Formation’, in J. Gupta and M. Grubb (Eds.) Climate Change and European Leadership, pp. 135-156, Kluwer Academic Publishers.

Carraro, C. and D. Siniscalco (1992), ‘The International Dimension of the Environmental Policy’, *European Economic Review* **26,** 379-387.

Carraro, C. and D. Siniscalco (1993), ‘Strategies for the International Protection of the Environment’, *Journal of Public Economics* **52,** 309-328.

Carraro, C. and D. Siniscalco (1995), Policy Coordination for Sustainability: Commitments, Transfers, and Linked Negotiations, in I. Godin and A. Winters (Eds.), The Economics of Sustainable Development, Cambridge University Press, Cambridge.

Carraro, C. and D. Siniscalco (1997), R&D Cooperation and the Stability of International Environmental Agreements, in C. Carraro (Ed.), International Environmental Agreements: Strategic Policy Issues, E. Elgar, Cheltenham.

Carraro, C. and D. Siniscalco (1998), ‘International environmental agreements: Incentives and political economy’, *European Economic Review* **42,** 561-572.

Chander, P. and H. Tulkens (1992), ‘Theoretical Foundations of Negotiations and Cost Sharing in Transfrontier Pollution Problems’, *European Economic Review* **36,** 388-398.

Farrell, J. and E. Maskin (1989), “Renegotiation in Repeated Games.” *Games and Economic Behavior* 1:327-60.

Finus, M. (2001), Game Theory and International Environmental Cooperation, E. Elgar, Cheltenham.

Finus, M. and B. Rundshagen (1998), “Renegotiation-proof Equilibria in a Global Emission Game when Players Are Impatient.” *Environmental and Resource Economics* **12**(3), 275-306.

Finus, M. and B. Rundshagen (2001), 'Endogenous Coalition Formation in Global Pollution Conrol', FEEM Working Paper No. 43.

Hoel, M. (1992), 'International Environmental Conventions: The Case of Uniform Reductions of Emissions', *Environmental and Resource Economics* **2,** 141-159.

Hovi, J. (forthcoming), 'The Pros and Cons of External Enforcement', in O.S. Stokke, J. Hovi, and G. Ulfstein (eds.), International Compliance: Implementing the Climate Regime. London: Earthscan.

Hufbauer, G.C., J.J. Schott, and K.A. Elliott (1990), Economic Sanctions Reconsidered: History and Current Policy, 2d ed., Washington, DC: Institute for International Economics.

Mueller, D.C. (1989), Public Choice II, Cambridge University Press, Cambridge, UK.

Olson, M. (1965), The Logic of Collective Action, Harvard University Press, Cambridge, MA.

Sandler, T. (1998), 'Global and Regional Public Goods: A Prognosis for Collective Action', *Fiscal Studies* **19,** 221-47.

Sandler, T. and K. Sargent (1995), 'Management of Transnational Commons: Coordination, Publicness, and Treaty Formation', *Land Economics* **71,** 145-162.

Siniscalco, D. (1999), 'Global Externalities Sovereign States', Paper presented at the 1st Annual Bank Conference on Development Economics in Europe, Paris, June, 1999.

Stokke, O.S. (forthcoming), 'Trade Measures, WTO, and Climate Compliance: The Interplay of International Regimes', O.S. Stokke, J. Hovi, and G. Ulfstein (eds.), International Compliance: Implementing the Climate Regime. London: Earthscan.

Tulkens, H. (1979), ‘An Economic Model of International Negotiations Relating to Transfrontier Pollution’, in K. Krippendorff (Ed.), Communication and Control in Society, pp. 199-212, Gordon & Breach, New York.

Ulfstein, G. and J. Werksman (2004), ‘Hard Enforcement and Due Process’, forthcoming in O.S. Stokke, J. Hovi and G. Ulfstein (eds.), International Compliance: Implementing the Climate Regime.

This essay has been accepted for publication in
*International Environmental Agreements: Politics, Law, and Economics*

# Essay no. 4:
# Sectoral Opposition to Carbon Taxes in the EU - a Myopic Economic Approach

CAMILLA BRETTEVILLE FROYN[67] and H. ASBJØRN AAHEIM[68]
CICERO - Center for International Climate and Environmental Research, Oslo

## Abstract:

Over ten years ago the EU Commission proposed a directive on carbon taxes, but faced so much domestic resistance that agreement was not reached until last year – and after it had been considerably watered down. The aim of this paper is to look into economic reasons for the political infeasibility of extensive carbon taxes. Since opposition is believed to arise prior to the policy implementation, the cost estimates have a *myopic* character compared with market estimates, in the sense that sectors are presumed to take into account their own substitution opportunities, but disregard changes in other sectors as well as the macroeconomic welfare gains from a tax regime. With this myopic approach, we estimate and compare costs of emissions cuts across sectors and across countries in the EU, showing how different sectors might have anticipated the impacts from an expected carbon tax. This focus illustrates that what seems to be cost-effective and to the best for the region on paper turned out too controversial to be politically feasible.

Three main conclusions are drawn. First, common measures across countries are generally not attractive since a particular measure may be advantageous to one country in order to keep the national cost of climate policy down, but may in another country

---

[67] camilla.bretteville@cicero.uio.no
[68] aaheim@cicero.uio.no

spur opposition that could be avoided if exceptions were allowed. Second, the electricity sector plays a key role in climate policy in every EU country and in the potential opposition to any policy measure. Conditions in this sector vary greatly among countries, and a long-term strategy to reduce conflicts of interest would be to introduce a common electricity market. Third, before a common electricity market can be established, further differentiation of national emission targets could represent an important opportunity for mitigating conflicts of interests across countries.

**Key words:** climate policy, EU, interest conflicts, opposition, political infeasibility, regional carbon tax

## 1. Introduction

The influence of interest groups is a common theme in the study of environmental politics. In an open society, environmentalist groups and their opponents, generally the corporations and businesses that resist the costs involved in complying with environmental regulation are presumed to spend resources on trying to influence policy-makers. The policy-makers need the votes, the money, the moral approbation, and the publicity these groups might provide in exchange for policy stances that gain approval, and avoid disapproval. Environmental policy is thus a function of the different pressures emanating from these (and other) interest groups, and hence seldom fully reflects the interests of any one of them (Barkdull and Harris 2002).

Conflicts of interest within the European Union (EU) countries have represented a major obstacle for the achievement of a common climate policy. Organized interests have resisted increased costs and have used arguments like that of competitiveness with success. National positions in the policy negotiations are the result of political solutions to internal conflicts resulting from different preferences among political actors, perceptions of fairness, and anticipated economic effects of policy measures.

The EU Commission’s carbon tax directive proposal, launched in 1992, was met with intense lobbying. Why was the opposition to taxes so great? One obvious explanation is the anticipated cost increases; another is that most economic actors do not take into account the macroeconomic welfare gains from an efficient policy regime. The aim of this paper is to look into economic reasons for the political infeasibility of extensive carbon taxes. We suggest that interest groups do not take all market effects of the policy into account when they present their case to policymakers because they have incentives to exaggerate the negative effects of regulation. By using a myopic approach the paper illustrates how differently economic sectors might have anticipated the impacts from a common carbon tax. The differences in these estimated cost anticipations are presumed to explain, at least partly, the opposition to carbon taxes in the EU. The study is based on estimating and comparing costs of emissions cuts in six countries; France, Germany, Italy, the Netherlands, Spain, and the United Kingdom[69]. Since opposition is believed to arise prior to the policy implementation, the cost estimates have a myopic character, compared with market estimates, in the sense that sectors are presumed to take into account their own substitution opportunities, but disregard changes in other sectors as well as the macroeconomic welfare gains from a tax regime. We argue that this narrow focus helps explain why what seems to be cost-effective and to the best for the region turned out to be too controversial to be politically feasible.

---

[69] The analysis is based on Aaheim and Bretteville (1999, 2000).

The following sections present some background information (section 2), explain the methodology (section 3), summarize and compare carbon tax costs in six EU countries (section 4 and 5), and draw conclusions (section 6).

## 2. Background

Three main regulatory options have been considered most relevant for environmental regulation: first, traditional command-and-control regulation, where proportional targets for individual sources are defined; second, a uniform emissions tax on all sources; and third, auctioned as well as grandfathered permit markets, where polluters are given their initial distribution of permits free of charge, typically according to historical emissions[70]. The government has – according to public choice theory – an incentive to prefer environmental taxes because they contribute to revenues, although the primary concern of elected officials is presumed to be getting re-elected. Climate policy affects all economic sectors, but the various policy options will influence sectors differently. Interest groups will therefore try to influence the policy choice.

In public choice theory it is assumed that the direction in which these interest groups will try to push the policy choice will depend on the distribution of costs and benefits from regulation (Mueller 1989; Svendsen 1998). With a tax solution, every emitter must pay for all emitted units, which increases production costs substantially and thus lowers the political acceptability of the instrument. Buchanan and Tullock (1975) show that existing firms will prefer traditional command-and-control measure to taxes because the former create a barrier to entry and are more open to negotiations and lobbying influence. Grandfathered permit markets provide an even stronger barrier to

---

[70] See for instance Baumol and Oates (1988) for a comprehensive treatment of these options.

entry because the initial permits are free of charge, suggesting that this option will be even more attractive to existing firms. Permit trading also leads to reduced abatement costs and more flexibility than is possible with both command-and-control regulation and emission taxes (Svendsen 1998).

Taxes have been prescribed as a corrective for externalities at least since Pigou (1920), and it was only natural that they should be considered as policy tools in the EU, both for efficiency and revenue reasons (Howe 1994). Thus, the EU Commission launched a carbon tax directive proposal in 1992, but this attempt was opposed by several member states. For instance, the United Kingdom was against common EU taxes in principle, wishing to deal with the $CO_2$ problem on their own. Spain, Portugal and Greece refused to accept increased production costs until they reach an economic level similar to that of the more wealthy northern EU countries (Svendsen 1998). In March 2003, however, the EU finally reached political agreement on a proposed common framework for energy taxation. The framework is seen as a key element in the EU’s climate policy, but its influence will be felt only over time because minimum EU tax rates have been watered down considerably compared to the Commission’s initial proposal (Environmental News Service, August 26, 2003).

The historical discrepancies between the declared targets and implemented actions in EU climate policy have been explained by the different decision layers which influence lobbying behavior. Historically, interests hostile to climate policy have had more influence and been able to prevent the use of stringent instruments that might lead to real emissions reductions (Michaelowa 1998). In discussing the goals of different interest groups and their influence on EU institutions and the climate policy process, Michaelowa (2000) suggests that a strengthening of pro-abatement interests compared to those of emitters lead to more demanding national climate policies in the EU. His analysis of selected sectors shows that at the EU level the importance of emissions-intensive sectors (e.g., fossil fuel based energy, metal producers, and chemical industry) is declining, while sectors that profit from climate policy (e.g., producers of renewable energy, nuclear energy, and hydro electricity) gain importance. As the influence of emissions-intensive sectors diminishes, stronger packages of climate policy instruments are likely at the national level (Michaelowa 2000). This may be an explanation for why the EU finally managed to reach agreement on a common energy tax.

The usual economic approach assumes that individual opinions on economic policies are based on their calculations on the post-policy market equilibrium. In this study, however, *anticipated* costs of a carbon tax are calculated, disregarding the effect on market prices of the reduction in energy demand and structural changes that would follow in general equilibrium models. The reason for this choice is that we wish to look into economic reasons for the political infeasibility of carbon taxes. We assume that the anticipation of costs of a carbon tax would create different levels of opposition in different sectors and countries in the EU, and that sectors where the costs of climate policies are low will be less reluctant to accept such policies. Furthermore, opposition is presumed greater when the costs are unevenly distributed because this gives rise to organized lobbying. Since economically strong, well organized, groups have substantial political power in Western democracies (Bang 2004), opposition is believed to arise prior to an implementation of a disadvantageous policy. Interest groups are assumed not to anticipate – at least not officially – all possible market effects of a tax *ex post* and the cost estimates on which the analysis is based have a *myopic* character, compared with market estimates. Our myopic approach assume that actors consider their own opportunities for switching to less emissions-intensive technologies, but they do not take all market effects into account when they present their case to policymakers because they have incentives to exaggerate the negative effects of regulation.

## 3. Measures and Measurements

Assume that if the government announces a carbon tax, sectors will anticipate the associated added cost by considering two factors. One is the carbon content in their energy demand, which is the share of fossil fuels in the sector's total energy demand, including the indirect carbon content in their electricity demand. Thus, if a large share of the electricity supply is generated by coal-fired power plants, a sector with a high demand for electricity also contributes to high emissions. A tax on carbon emissions thereby affects production costs in this sector more than if electricity were generated from less carbon-intensive energy sources.

The second factor determining the anticipated cost of a carbon tax is the possibilities for switching to less emissions-intensive technologies or activities. For a specific firm or individual, these substitution possibilities depend first and foremost on available technologies and individual flexibility when it comes to adaptation to a change in relative prices. On a national level, the substitution possibilities also depend on the economic structure of the economy. Countries with large energy-intensive industries are likely to be more vulnerable than service-oriented economies, to the extent that transport constitutes a large part of this sector. The opportunities to switch to less emission-intensive technologies are greater if most of the energy in the service sector is used for heating purposes.

The country study consists of two elements. One is a formal analysis of the anticipated, added cost of production following a change in the relative prices of the input arising from a carbon tax and the resulting substitution between input factors. To give additional insight we add informal discussions, where the substructures of the main sectors, as well as expected technology advancements and opportunities in the years to come regardless of a carbon tax are taken into account.

The formal analysis is based on a standard, macroeconomic demand scheme, which has been used frequently to study the cost of emissions control in general equilibrium models (see e.g., Edmonds and Barns 1992, Oliveira-Martins *et al.* 1992, Peck and Teisberg 1994, or Jorgenson and Wilcoxen 1994). We divide the economy of each country into two sectors, the energy sector, and production of other commodities and services, which produce the output $x_i$ ($i$ = 1,2). Both sectors use labor, $n_i$, and energy, $u_i$, as input, so the production function is $x_i = f^i(n_i, u_i)$. Strictly speaking, this means that input of real capital and other intermediate input are considered constant. However, since our interest is to compare possibilities for substitution when fossil fuels are being taxed, labor may be interpreted in a broader context, giving some representation of 'other input'. We assume CES technology:

$$x_i = A_i\left[\alpha_i n_i^{\rho} + (1-\alpha_i)u_i^{\rho}\right]^{\frac{1}{\rho}} \qquad (i = 1,2)$$

and we derive the demand for $n_i$ and $u_i$ as functions of output and factor prices of labor, $w$, and energy, $q$ under the assumption of profit maximization in competitive markets:

$$n_i = n^i(w,q,x_i) \qquad n_i = \frac{x_i}{A_i}\left[\left(\frac{w}{\alpha_i}\right)^{\frac{1}{\rho-1}} + \left(\frac{q^\rho}{(1-\alpha_i)}\right)^{\frac{1}{\rho-1}}\right]^{\frac{-1}{\rho}}\left(\frac{w}{\alpha_i}\right)^{\frac{1}{\rho-1}} \quad (i = 1,2)$$

$$u_i = u^i(w,q,x_i) \qquad u_i = \frac{x_i}{A_i}\left[\left(\frac{w}{\alpha_i}\right)^{\frac{1}{\rho-1}} + \left(\frac{q^\rho}{(1-\alpha_i)}\right)^{\frac{1}{\rho-1}}\right]^{\frac{-1}{\rho}}\left(\frac{q}{1-\alpha_i}\right)^{\frac{1}{\rho-1}} \quad (i = 1,2)$$

Energy can be considered an aggregate of fossil fuels, $f$ and electricity $e$ ($\delta$ and $\mu$ are constants):

$$u_i = \left(\delta_i f^{\mu_i} + (1-\delta_i)e^{\mu_i}\right)^{\frac{1}{\mu_i}}$$

The demand functions for fossils and electricity are derived in a similar fashion as those for labor end energy. Fossil fuels are an aggregate of coal, oil and gas. The energy input in the electricity sector consists only of fossil fuels. In addition, electricity may be provided by other energy sources, such as nuclear power or hydropower. These sources are not explicitly included in the demand scheme, since they are considered to be unaffected by a carbon tax. In countries where a large part of the electricity supply comes from sources other than fossil fuels (e.g., France), the price of electricity would be relatively insensitive to a carbon tax and the immediate effect on costs of a carbon tax would be relatively small. On the other hand, the emissions intensity of the overall economy is low because of the large amount of non-fossil fueled based power[71]. As a result, switching to less emissions-intensive power sources is more costly than in economies with coal-based electricity.

---

[71] Low emissions intensity is used throughout as equivalent to not emissions intensive; and high emissions intensity as equivalent to emissions intensive.

Each firm takes market prices as exogenous. A change in prices will lead to substitution between input factors, resulting in a new composition of oil, coal and gas in the fossil fuel aggregate. In the electricity sector, fossil fuels may be replaced or replace other input, represented here by labor. The composition of fossil fuels and electricity will change in the commodity and service sector, resulting in a change in the relative input of energy and labor. The rates at which these changes will take place depend on the increase in prices. Thus, if the price of oil coal and gas increase by a vector $\Delta p$ that corresponds to a $CO_2$ charge, we can calculate the resulting increase in the price aggregate, first for the output from the electricity sector. This is then used to calculate the changes in the energy input for the commodity and service sectors, and finally the cost increase per unit of production for the commodity and service sectors. We call this increase in the production cost for *the cost-push indicator*.

Households demand goods from the two sectors. Hence consumption is an aggregate of energy and other goods and services, and the consumer's utility can be expressed ($\beta$ and $\gamma$ are constants, and $p$ and $q$ are prices):

$$U = \left[\beta x^{\gamma} + (1-\beta)u^{\gamma}\right]^{\frac{1}{\gamma}}$$

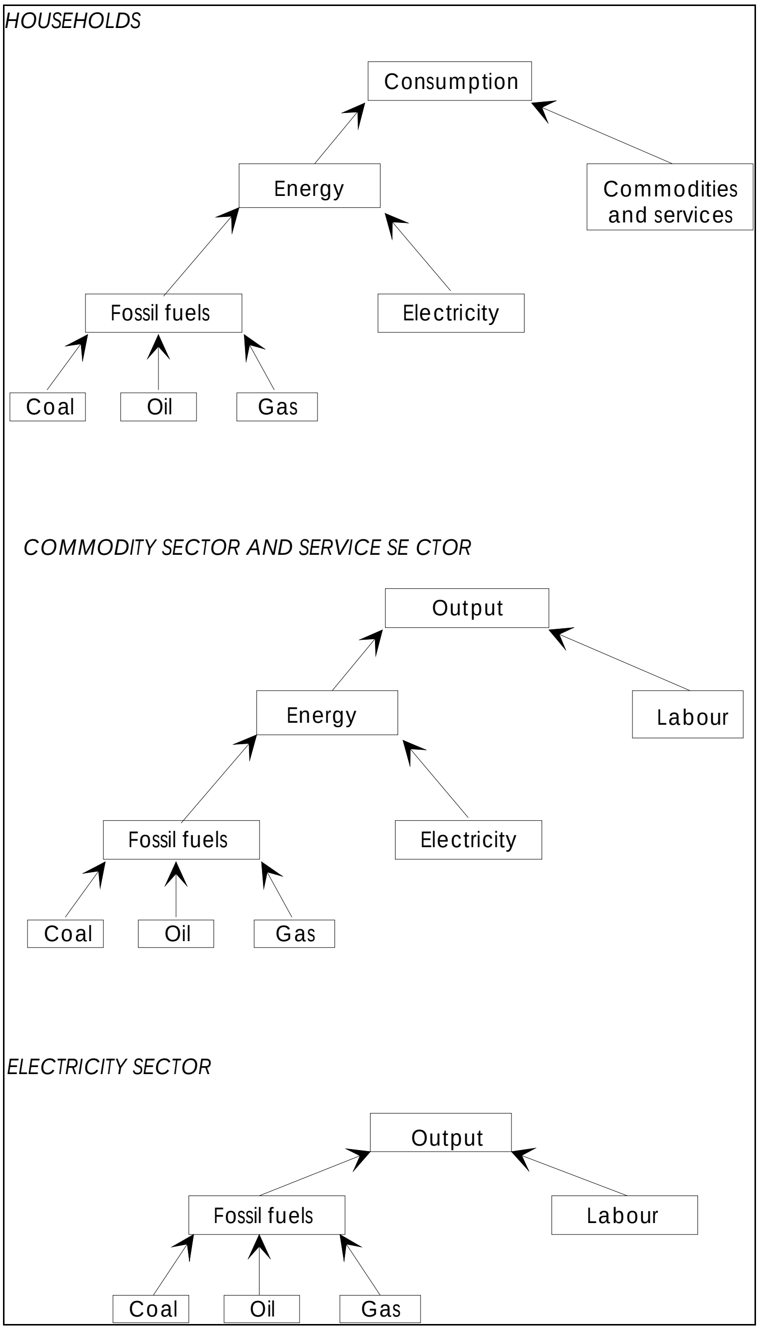
Assuming utility maximization, the demand functions are:

$$x = \frac{1}{\beta^{\frac{1}{1-\gamma}} p^{\frac{\gamma}{\gamma-1}} + (1-\beta)^{\frac{1}{1-\gamma}} q^{\frac{\gamma}{\gamma-1}}} \left(\frac{p}{\beta}\right)^{\frac{1}{1-\gamma}}$$

$$u = \frac{1}{\beta^{\frac{1}{1-\gamma}} p^{\frac{\gamma}{\gamma-1}} + (1-\beta)^{\frac{1}{1-\gamma}} q^{\frac{\gamma}{\gamma-1}}} \left(\frac{q}{1-\beta}\right)^{\frac{1}{1-\gamma}}$$

The cost-push indicator for households is calculated in the same manner as for production sectors by first calculating the added cost of fossil fuels resulting from the

Figure 1: Demand scheme for fossil fuels



carbon tax after substitution has taken place, then finding the cost of the energy aggregate, where the electricity price is also subject to the tax, and finally calculating the increase in total expenditures. The second order effects on prices due to changes in demand and supply are thus not taken into account in this type of analysis, that is, prices are not subject to equilibrium constraints. An illustration of the aggregates is given in Figure 1.

The characteristics of each of the six countries studied in this paper can be captured by calibration of the demand functions. The magnitude of the cost-push indicators gives an indication of whether the opposition to carbon taxes in European countries has an economic explanation. However, the cost-push indicator depends to a large extent on the elasticity of substitution[72] in the domestic industries. On the aggregated level used in this study, this elasticity for the same aggregate in one industry may vary from country to country, but there were no reliable estimates available. Therefore, we chose the same elasticity for the same aggregate in all countries, as shown in Table 1. The constant terms and the distribution parameters were calibrated specifically for each country. To compensate for the lack of estimates for the substitution parameters, we also perform an informal analysis of the sectors in each country economy, where we focus particularly on the composition of the commodity and service sectors, and on the opportunities in the electricity sector.

High elasticity means that a given change in relative prices leads to a large change in the composition of input factors, and conversely that low elasticity leads to a small change. Hence, the manufacturing sectors and the household sectors exhibit a relatively flexible use of input. The service sectors are assumed to be relatively inflexible with respect to the use of energy because transport constitutes a large share of the service sector. The electricity sectors are also assumed to be inflexible at the macro level[73].

---

[72] The elasticity of substitution shows the relative change in the composition of input factors due to a relative change in the factor prices.

[73] The elasticities are thus chosen by assumption to the best of our ability.

Table 1. Elasticities of substitution by sector. All countries

| Aggregates and sub-aggregates<br>Fossil fuels | Energy | Output | Sector |
|---|---|---|---|
| 0.625 | - | 0.588 | Electricity |
| 0.714 | 0.909 | 0.800 | Industry |
| 0.500 | 0.444 | 0.980 | Services |
| 0.909 | 0.909 | 0.645 | Households[1] |

[1] 'Output' in the household sector is total consumption.

The demand functions were calibrated with data from different sources: Economic macroeconomic figures were taken from OECD (1996). Energy data in physical terms were provided from IEA (1995), and energy prices from IEA (1993) and Boug and Brubakk (1996).

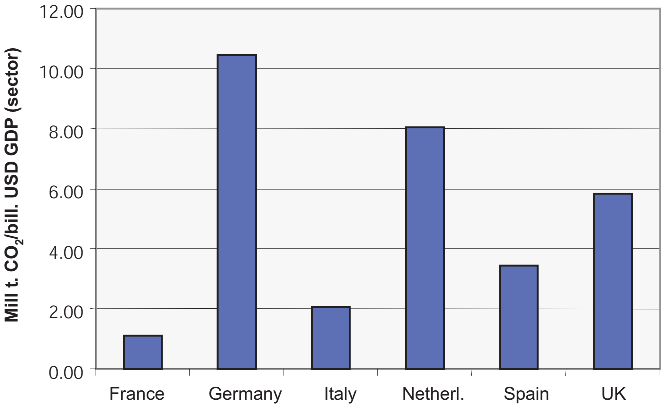
The purpose of the demand schemes (Figure 1) is to find how sensitive the demand for fossil fuels is to a carbon tax. This is found by imposing a supposed carbon tax on the use of fossil fuels in order to obtain a 10 percent reduction in $CO_2$ emissions for each country. Thus, the tax would be the same for all sectors within a country, but would vary from one country to the other. The abatement effort would thus be cost effectively distributed within the countries, something that means that the costs would be unevenly distributed. To calculate the expected costs of abatement for the different sectors, we use two indicators; one related to production costs, and one related to abatement costs. The first is the *cost-push indicator* (defined above) that indicates the increase in the unit cost of production for a sector subject to the carbon tax. The indicator can thus be interpreted as the elasticity of production costs with respect to national carbon cuts. The second is the *anticipated unit cost of abatement,* which is the cost-push per unit of $CO_2$ emissions reductions in the local sector or country. The anticipated unit cost of abatement is thus an indicator for the amount of emissions reductions you get per dollar. Both indicators will vary across sectors and across countries, while they would be the same in an equilibrium model. Thus, if the policy was implemented, the sectors with low anticipated costs might in fact have to abate more than indicated by the calculations, and the sectors with high anticipated costs might have to reduce less than indicated. Alternatively, the indicators reveal in which sectors a myopic approach exaggerates the final costs, and in which sectors it understates them.

## 4. Differences in energy structure

The six countries studied in this paper; France, Germany, Italy, the Netherlands, Spain, and the United Kingdom (UK), account for nearly 85 percent of the EU's total population and more than 85 percent of the region's $CO_2$ emissions. The energy structure varies substantially between the countries. The Netherlands has by far the most energy intensive economy, due to its relatively high consumption of natural gas. This proportionally large share of gas means that its emissions intensity in terms of $CO_2$ per billion USD of gross domestic product (GDP) is the same as for Germany and the UK. These three countries' level of $CO_2$ per billion USD of GDP is approximately twice that of France, which has the lowest emissions intensity, while the level for Italy and Spain lies between these extremes.

The national emissions intensities are closely related to those of the electricity sectors. Figure 2 shows emissions intensities in the electricity sector by country for 1994. German power plants emit 10 million tons of $CO_2$ per billion USD of GDP. The corresponding figures for the Netherlands and the UK are lower, but significantly higher than for the three other countries. For France, the explanation is the substantial contribution of nuclear power, which amounts to about 75 percent of the total supply of electricity. For Italy and Spain, the reasons are less obvious. Italy has banned nuclear power, and the electricity sector is heavily based on imported oil. As a consequence, electricity is expensive and energy intensity is low, both because the high price encourages effectiveness and because it leads to a high contribution to the GDP. Spain's electricity sector is mainly coal based, but nuclear power and hydropower also contribute noticeably. This increases the economic value of power generation in Spain, and implies that the emissions intensity of the electricity sector is low, that is, the production of electricity per unit input of fossil fuels is high. In contrast, the German, the Dutch, and the British electricity sectors can be characterized as emissions intensive.

*Figure 2. Emission intensities in the electricity sector by country (1994)*



There are fewer differences in energy use patterns in other sectors, with the exception of the Netherlands, which in general exhibits high energy intensity. The manufacturing sectors divide their energy use between coal, oil, gas, and electricity, but gas contributes to a higher share in Italy, the Netherlands, and the UK than in the other three countries. The German and the French manufacturing sectors use more coal than manufacturing sectors in the other countries. The service sector's energy use is dominated by oil in all countries, which is due to transport. In the household sectors, one finds the same three ‘gas users’ as for the manufacturing sectors, British and German households are more dependent on coal than the households in other countries, while electricity is the main source of energy in French households. Table 2 shows the main indicators for all six countries.[74]

The distribution of cost-push among sectors exhibits approximately the same pattern across countries. Recall that in this model the national cost-push is perceived as the anticipated increase in the cost of the national product of a carbon tax sufficient for a 10 percent reduction in emissions below business as usual. The cost-push is two to three times higher in the electricity sectors than in the manufacturing sectors, while it is one and a half to two times higher in the manufacturing sectors than in the service sectors. The unit cost of emissions cuts is highest in the manufacturing sectors, and usually very low in the electricity sectors. However, the levels of the cost-push indicators are different and shown in Table 2 and Figure 3.

To concentrate on the effects from the myopic approach, new options or technological ‘jumps’ are not taken into account in the demand scheme presented in Section 3. However, large potentials have been documented in several countries in many studies. These may represent technological jumps with respect to emissions cuts (Kram and Hill 1996, COHERENCE 1991). According to Krause and Koomey (1996), the potential for new technologies with substantially lower emissions of $CO_2$ seems to be most promising in the electricity sector and the manufacturing sector. This implies that the possibilities differ between countries, since the potential is large in most coal-fired plants and in heavy manufacturing industry. Economies with a greater share of

---

[74] A more thorough study of country characteristics is given in Aaheim and Bretteville (1999).

*Table 2. Main indicators. All countries.*

**GDP** (billion USD) **and consumption (1994)**

| Country | Total | Commodities | Services | Electricity | Household consumption |
|---|---|---|---|---|---|
| France | 1235 | 382 | 827 | 25 | 976 |
| Germany | 1753 | 666 | 1052 | 36 | 1262 |
| Italy | 1127 | 338 | 753 | 56 | 902 |
| The Netherlands | 305 | 32 | 207 | 6 | 225 |
| Spain | 511 | 174 | 317 | 20 | 409 |
| The United Kingdom | 1009 | 303 | 676 | 30 | 668 |

**Energy intensity** (PJ/GDP)

| Country | Total | Commodities | Services | Electricity | Households |
|---|---|---|---|---|---|
| France | 5.26 | 4.80 | 4.20 | 11.58 | 0.91 |
| Germany | 7.86 | 4.88 | 3.89 | 112.26 | 1.99 |
| Italy | 5.86 | 5.18 | 2.85 | 17.84 | 1.32 |
| The Netherlands | 19.1 | 8.49 | 4.77 | 107.7 | 1.91 |
| Spain | 6.95 | 5.57 | 4.73 | 36.11 | 0.85 |
| The United Kingdom | 8.53 | 6.12 | 4.58 | 64.72 | 1.96 |

**Emissions intensity** (mill. t CO2/GDP)

| Country | Total | Commodities | Services | Electricity | Households |
|---|---|---|---|---|---|
| France | 0.31 | 0.27 | 0.27 | 1.09 | 0.03 |
| Germany | 0.55 | 0.28 | 0.27 | 10.44 | 0.10 |
| Italy | 0.36 | 0.28 | 0.19 | 2.06 | 0.07 |
| The Netherlands | 0.58 | 0.47 | 0.31 | 8.03 | 0.09 |
| Spain | 0.47 | 0.31 | 0.32 | 3.42 | 0.04 |
| The United Kingdom | 0.57 | 0.36 | 0.30 | 5.85 | 0.10 |

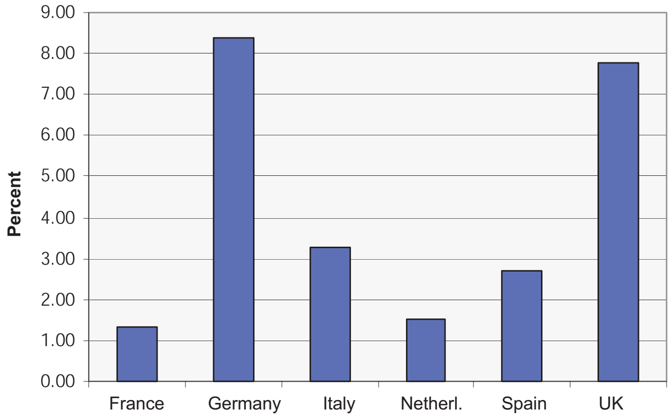
**Cost-push indicator**

| Country | Total | Commodities | Services | Electricity | Households[75] |
|---|---|---|---|---|---|
| France | 1.33 | 1.59 | 1.14 | 3.88 | 4.07 |
| Germany | 8.37 | 12.92 | 5.02 | 22.21 | 22.03 |
| Italy | 3.28 | 4.11 | 2.32 | 11.87 | 12.48 |
| The Netherlands | 1.51 | 1.99 | 1.18 | 5.22 | 5.32 |
| Spain | 2.72 | 3.52 | 1.59 | 13.51 | 12.99 |
| The United Kingdom | 7.79 | 10.93 | 5.79 | 22.83 | 22.74 |

---

[75] The cost-push indicators for households are calculated as the cost-push for total energy use and cannot be directly compared to indicators for other sectors.

*Figure 3. Cost-push indicators by country: Percent increase in the cost of the national product for a carbon tax sufficient for a 10 percent reduction in emissions below reference level.*



heavy manufacturing industry and greater reliance on coal-fired electric utilities should have greater opportunity to cut emissions and thereby the costs of emissions reductions.

In addition, structural differences within sectors may explain why opinions about climate policy may vary across nations. For example, substitution possibilities within the service sector depend primarily on $CO_2$ emissions from transportation sources. At present, the transportation sector depends heavily on oil, and there are currently few practical alternative fuels. Therefore, it is probably more difficult to reduce emissions in the service sector if transportation sources comprise the largest share of energy use in that sector. It should be noted, however, that although the technical potential for energy reductions in commercial buildings seems to be large, energy consumption in these buildings is usually relatively insensitive to prices (see e.g. Flavin and Lensen 1994). In the manufacturing sector, which includes agriculture, additional opportunities may be available if the contribution to the total greenhouse gases emissions from gases other than $CO_2$ is large.

In the following the estimated cost-push indicators and anticipated unit costs of abatement constitute the basis for the assessment of sector and country opposition. Since these estimates are based on very general assumptions about sector technology, the possible conflicts of interest are also discussed in light of certain characteristics of sectors and countries not captured by the macro functions.

## 5. Anticipated cost differences of a carbon tax

The indicators displayed in Table 2 portray important differences among countries and among sectors within countries. These differences show how different economic sectors might have anticipated the impacts from an expected carbon tax. Figure 4 presents the anticipated unit costs of $CO_2$ abatement by sector, while Figure 5 shows the estimated change in demand for energy commodities by sector as a result of the carbon tax.

### *5.1 Distribution of costs and changes in demand*

The anticipated costs of emissions cuts in **France** are dominated by the low emissions intensity in the electricity sector, which is explained by the fact that approximately 75 percent of the electricity is produced by nuclear power plants. A 10 percent reduction in emissions thus implies a relatively small amount in terms of tons of carbon reductions, and this is why the cost-push is low. The reductions are, however, relatively concentrated. One third of the emission cuts are due to a reduction in the demand for oil in the service sector. Most likely this will turn out to be quite costly because transport constitutes a considerably higher share of the service sector in France than in the other countries. In fact, the $CO_2$ emissions from transport comprise 35 percent of fuel-related emissions in France, while in the other countries the share is between 15 and 25 percent. The transport sectors in the EU are growing as a result of the increasing specialization of products that has been taking place since the start of the common market. The share of these products that are transported by road is also rising.

*Figure 4 Anticipated unit cost of abatement by sector. USD/kg $CO_2$*

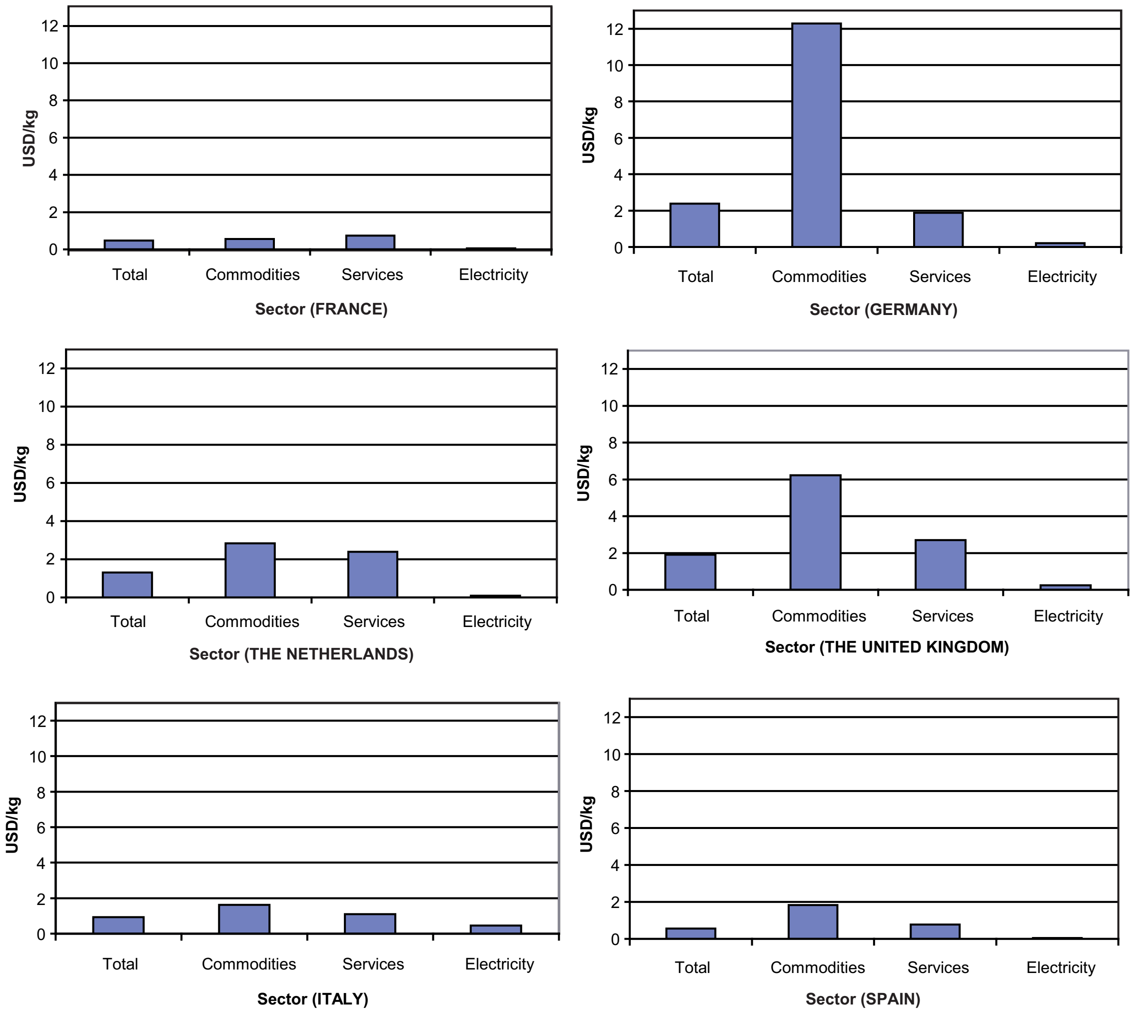
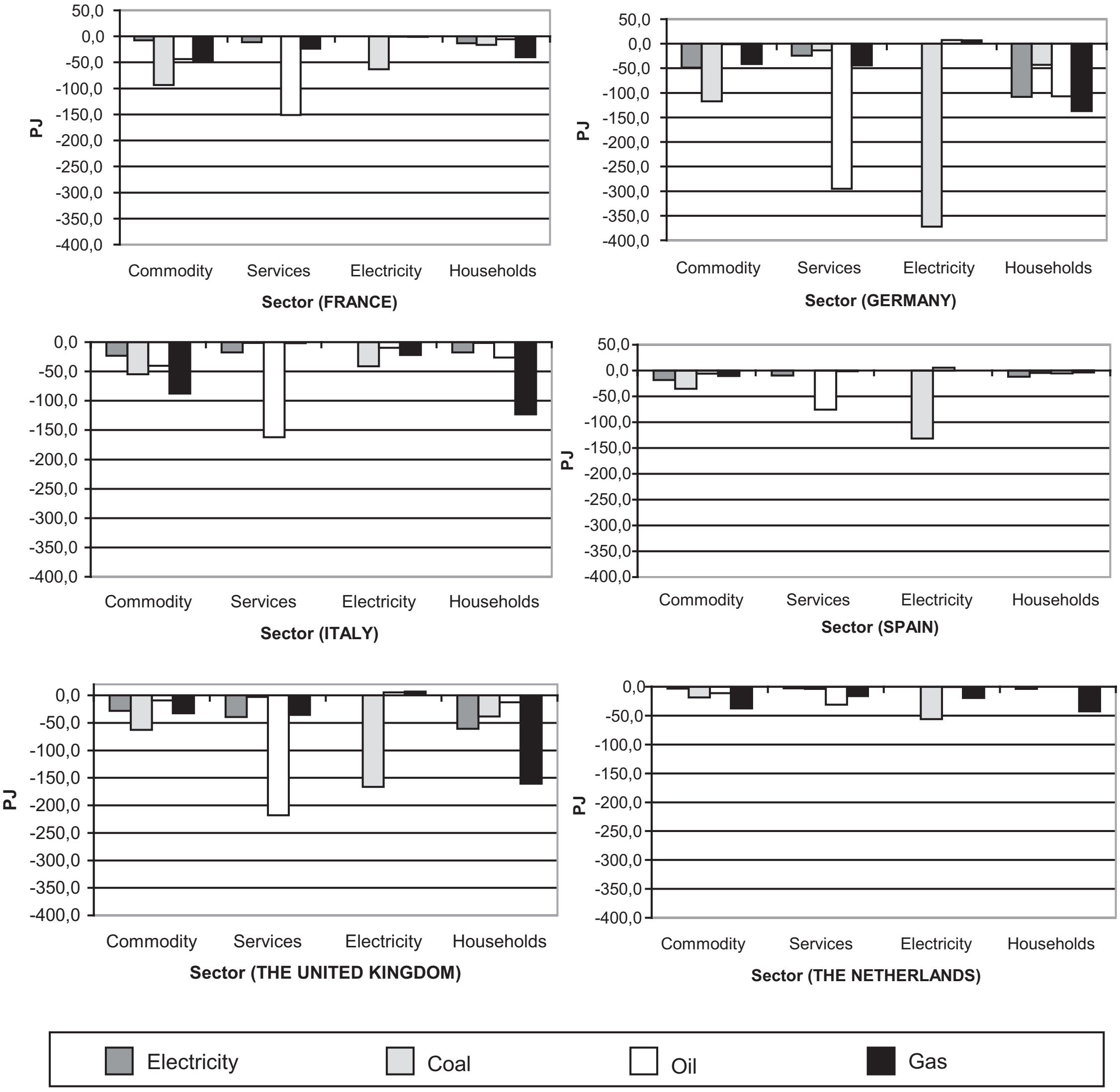


*Figure 5 Change in the demand for energy commodities in PJ (petta joule) by sector as a result of a carbon tax.*



The second largest potential contributor to emissions cuts in France is the reduction in the demand for coal in the manufacturing sector. This sector is expected to cut 94 PJ of coal consumption, which is high in comparison to other countries. The structure of energy use in France makes it relatively cheap to reduce emissions also in the manufacturing sector (Figure 4). However, the possibilities of finding alternative technologies are strictly limited in the electricity sector and the service sector. This may put stronger pressure on the manufacturing sector, where some opportunities may be available.

**Germany** has the highest cost-push of all six countries due to the high levels of coal consumption. This is particularly evident in the electricity sector (see Table 2). The cost-push for electricity is, in fact, higher than the cost-push for direct use of fossil fuels in all sectors. This means that all the sectors will increase their direct use of fossil fuels relative to their use of electricity as a result of a carbon tax. Germany's manufacturing sector faces the highest cost-push of all the countries' manufacturing sectors, partly because the emissions cuts are large, and partly because the anticipated unit costs in the sector are the highest.

It has shown to be very hard to phase out the use of coal in Germany for three main reasons: their big coal reserves, the strong labor unions, and the political power of the German coal-based energy producers (Bang 2004). Moreover, the German manufacturing sector has a lucrative voluntary agreement that makes them resist both a tax regime and emissions trading. Traditionally strong sectors, like the German manufacturing sector, have based their production on coal-based electricity. They have argued that it is hard for them to cut emissions without cutting production, and hence decreasing employment. Germany also has a strong labor movement that would lobby against a carbon tax[76].

Germany's own reserves of fossil energy sources – oil, coal, and natural gas – in 1991 amounted to less than 1 percent of world reserves. Furthermore, the country's oil and natural gas production was not enough to cover its own requirements. At the time, oil met a total of 41 percent of Germany's primary energy requirements, and gas a total

---

[76] See e.g. Bang (2004) for a detailed analysis of German environmental policy.

of 18 percent. Hence, some 99 percent of Germany's oil requirements and 96 percent of its gas requirements had to be met through imports. The government has stated that Germany needs a balanced energy mix that includes hard and lignite coal, oil, natural gas, nuclear power and renewable energies to secure a reliable energy supply. Coal met about one third of Germany's energy needs in the 1990s, and has been an important part of the country's electricity production, particularly since it is the most abundant domestic fossil fuel source. The energy-mix trend shows that the use of hard and lignite coal as energy sources declined from 1990 to 1996, whereas the use of oil, gas and nuclear increased. But still in 1997 about half the electricity production was based on coal, while about a third was based on nuclear power (Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, 1997).

Also in Germany, the service sector would share a large share of the total emissions cuts. The transport sectors contribute a relatively low share of the emissions, slightly above 15 percent. The household sector contributes to the estimated emissions reductions by reducing the demand for all energy factors. The apparent possibilities for reducing emissions in the electricity sector and enhancing energy effectiveness in a large part of the manufacturing sector may reduce the costs of cuts in Germany substantially. Enhancing energy effectiveness through command-and-control regulation has in fact been a main goal in German climate policy.

**Italy** is characterized by large natural gas consumption, and the electricity sector is thus not very carbon-intensive, which contributes to a rather atypical pattern in the estimated emissions reductions. It also helps explain why the cost-push in Italy is the third highest among the countries. Compared with the other countries, the unit cost of emissions cuts in the electricity sector is higher relative to other sectors. The electricity sector thus contributes to a remarkably low share of total reduction – only 15 percent, which is even lower than for the electricity sector in France. Oil in the service sector contributes to more than 25 percent of the total estimated reductions in Italy. This abatement will be costly since transport represents a substantial part of the service sector. Moreover, the household sector contributes to more than 20 percent of total estimated reductions, mainly due to a reduction in the demand for gas.

Generally, and particularly for the service and the household sectors, the anticipated costs for Italy are high. The competitive industries will thus have an extra strong incentive to lobby against taxes to obtain less costly measures to protect their international competitiveness. Against the background of these results, it is likely that Italy will be reluctant to support an aggressive carbon tax policy at the EU level.

The energy consumption pattern in **the Netherlands** is connected to the availability of gas from their own reserves. Apart from being the country with the highest per capita gas consumption, it also has low energy prices, not only for gas, but also for coal and oil (see Birkelund et al. 1993). As a result, energy intensity in the Netherlands is very high, despite the large contribution from gas. This indicates that the gas fired power plants in the Netherlands are ineffective, and that abatement thus is relatively cheap, something which explains why emissions cuts in the Netherlands have a lower cost-push than any other country in this study. The implication is that Dutch power plants have a large potential for low-cost energy and emissions cuts. Potential of this magnitude is not assumed in other studies (see e.g. Kram 1998, and Capros 1998). One explanation for this difference may be that the present study applies price data for 1991, when the energy prices were low in the Netherlands, according to IEA (1992).

The large potential for energy and emissions cuts also applies for the manufacturing sector, although the difference between the anticipated unit costs is larger than usual between the manufacturing sector and the electricity sector. Consequently, the potential is considered to be largest in the electricity sector.

The energy intensity can be explained by the low energy prices, particularly for gas, that prevailed for many years. The Dutch gas reserves are, however, diminishing and prices have been increasing. The adaptation to higher prices will lead to an increase in energy efficiency, even without emissions targets for $CO_2$. The flexibility embedded in an inefficient energy system and the expected increase in basic energy prices make it unlikely that carbon cuts will lead to high costs in the Netherlands. Low costs give the authorities an opportunity to choose among a larger set of measures than in any of the other countries.

Like Germany and the UK, **Spain** is a heavy coal user. In Spain, however, a larger share of coal consumption is found in the electricity sector. In the manufacturing sector, oil and gas constitute approximately 85 percent of fossil energy use, while oil dominates in the service and households sectors. The cost-push effect in Spain is, however, much lower than in Germany and the UK, and this may partly be explained by the low carbon intensity of the electricity sector. Hence, the opportunities for low cost emissions cuts in this sector through the introduction of new technologies seem to be considerably less than in the other countries.

The lack of low cost opportunities for switching to less emissions-intensive production in the electricity sector is important for Spain because more than 40 percent of the estimated emissions reductions are found in this sector. The reduction in demand for oil in the service sector, a slight reduction in coal consumption in the manufacturing sector, and the cuts in the electricity sector put together constitute nearly 80 percent of the reductions in Spain. In other words, the bulk of the reductions would have to be concentrated within a few economic activities and this may spur opposition. In addition, the electricity sector in Spain is more concerned with expansion to meet the expected future growth in the demand for electricity. A carbon tax is thus controversial also in Spain.

**The United Kingdom** shares many similarities with Germany. While emissions and energy intensities are high, there are probably many opportunities for reducing them by introducing new and more effective technologies. There are, however, important differences between the two countries. In particular, the cost-push in the manufacturing sector is about half as high in the UK. On the other hand, the potential of an efficient restructuring of the manufacturing sector in the UK is not the same as in parts of the German manufacturing sector. The British households contribute to a substantial part of the estimated cuts by reductions in the demand for gas.

There is a clear potential for reducing the emissions of $CO_2$ in the electricity sector. The manufacturing sector in particular may benefit from this by saving considerably in terms of indirect cuts in reducing the demand for electricity since the entire cost-push for the manufacturing industry in the UK can be explained by the increase in the price of electricity. Consequently, the manufacturing sector in the UK has a competitive advantage in the export markets by improving their terms of trade, implying that this sector should in fact be in favor of a common carbon tax.

### *5.2 Conflicts of economic interest and political feasibility*

One conclusion to be drawn from this is that differences in the anticipated effects of a carbon tax in the electricity and manufacturing sectors are the key to understanding the infeasibility of a carbon tax as the main climate policy instrument in the EU. The analysis above indicates were one would expect to find the greatest opposition. For instance would a significant share of total reductions have to take place in the electricity and manufacturing sectors because they contribute to large shares of national emissions. These sectors represent large firms with high emissions, often with significant political power. Consequently, they may have considerable influence over the government's policy choices, including their choice of policy measures. Because all fiscal decisions under the European convention require approval by consensus, European businesses would have to persuade only one state to block the tax proposal.[77] Thus, as long as there was opposition by key sectors within each country, approval was not very likely. This explains why EU energy tax rates have been watered down compared to the Commission's initial proposal. The harmonized minimum tax rates should reduce distortions of competition between EU states and between energy products. Furthermore, this system provides a basis for the EU to collectively raise energy prices over time, thus increasing incentives for more efficient usage. However, due to the difficulty of reaching agreement, the law is full of derogations and transition periods for various countries and sectors (Environmental News Service, March 21, 2003).

The main difference between the electricity and manufacturing sectors, however, is that while the manufacturing sectors face a high anticipated unit cost, the unit cost in

---

[77] See i.e. Skjærseth and Skodvin 2003, p.174.

the electricity sectors are very low. Hence, the incentives to search for new technologies as a result of the emissions targets are quite different. However, the electricity sectors in Germany, the Netherlands and the UK are presently undergoing a substantial restructuring, which leads to lower emissions regardless of climate policy targets. As a consequence, one ought to study the cost-push in the manufacturing sectors in the light of these changes.

As pointed out earlier, the differences between sectors across countries, measured by the cost-push indicators, follow a rather similar pattern to the total cost-push for the countries shown in Figure 3. How important the differences in the cost-push are, however, depend on the degree of international competition for the given sector. It is difficult to distinguish clearly between internationally competing activities and non-competing activities, especially within the EU. All of the various sectors are exposed to competition from other countries to one degree or another. In general, manufactured goods are often used to illustrate changes in the competitiveness of an economy. Manufacturing products also compete to a considerable extent with industries in countries outside of the EU. The manufacturing sector may therefore be considered more vulnerable to a loss of competitiveness than other sectors.

One way to make discrimination between sectors within a country legitimate is to argue that such discrimination is necessary to avoid a decline in competitiveness. Our analysis shows that the three northern countries have more opportunities because the electricity system is very emissions intensive, and they can deliver their electricity at low prices, partly due to subsidies. It is difficult to point at specific measures taken within countries that could contribute to mitigate the differences in the cost-push across the sectors of different countries. However, a development of a common European market for electricity would help mitigate these differences to a large degree.

Looking specifically at economic conflicts of interest seems to strengthen the impression that there is a need for differentiation between the EU countries when it comes to the commitments related to the Kyoto Protocol. The EU went further than any other of the large signatories to Kyoto and proposed a 15 percent reduction in greenhouse gas emissions from their 1990 levels by 2010. A prerequisite for this proposal was exactly that the EU managed to agree on an internal differentiation of targets. The differentiation was based on the so-called Triptych approach (Phylipsen *et al.* 1999). This approach differentiates emission allowances according to standard of living, existing fuel mix, and the competitiveness of internationally oriented activities. While the Triptych approach in many respects is based on similar criteria as those discussed above, potential sectoral conflicts of interest are not taken explicitly into account. National differences in energy prices, which have an important impact on the results of this study, are also not accounted for.

In the Kyoto protocol, the EU ended up with 8 percent reduction and the targets for the six countries are (percent change from 1990 to the period 2008-2012): France 0.0 percent; Germany -21.0 percent; Italy -6.5 percent; the Netherlands -6.0 percent; Spain 15.0 percent; and the UK -12.5 percent. Whether the consideration of the possible conflicts of interest, emphasized in this study, would change the distribution of targets for the EU countries is difficult to determine because it is not entirely clear what the assumed baseline (business as usual) emissions were for 2010. For example, the standard-of-living category was used as an argument to give additional emission allowances to the so-called Cohesion-countries, which include Spain, but it is difficult to say exactly how much this means in terms of the emissions targets. The Triptych approach may thus moderate some of the internal conflicts of interest, but according to this study there are a number of potential conflicts left. It is reason to believe, however, that the flexibility mechanisms in the Kyoto protocol; the Clean Development Mechanism (CDM) and Joint Implementation (JI), as well as emissions trading within and outside the EU might lighten the burden for the countries substantially.

It seems that the Netherlands has achieved a good agreement, considering the low estimated costs of emission cuts in the analysis above. One reason why the Netherlands may have received a relatively moderate target is the low price-level of gas in 1990, which was not explicitly taken into account in the Triptych approach.

It is difficult to explain why Germany should reduce their emissions nearly twice as much as the UK. The German manufacturing sector is not pleased with the German emissions reduction target. One explanation that has been given is that GHG emissions reductions were an immediate effect of the unification of the eastern and the western federal states. The experiences from the drastic reductions in energy consumption as a result of unification indicated to German politicians that energy savings were both possible and economically sound. The belief at the time was that further energy savings were readily achievable, and that such savings would outweigh the heavy costs of unification in terms of infrastructure investments, economic subsidies, economic modernization, and so forth (Bang 2004). The net transfers of public funds from western to eastern federal states, however, amounted to DM 106 billion in 1991, DM 115 billion in 1992, DM 129 billion in 1993, and DM 125 billion in 1994. Hence, the GHG emissions reductions in the early 1990s were not entirely ‘a free lunch' for Germany (Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, 1997).

It appears that Italy's target would be very difficult to achieve without great conflicts of interest. Again, the prices of energy may explain why it is difficult for Italy to reduce emissions without opposition, while at the same time, the Triptych approach, which emphasizes the energy mix, prescribes cuts at approximately the same level as for the Netherlands. On the other hand, the most severe interest conflicts in Spain may be overcome by the allowance of a 15 percent increase in its emissions. The fact that energy prices, which is assumed to be high in Spain in this analysis, is not explicitly considered in the Triptych approach, indicates that the differentiation according to the different levels of the standard of living may have been overemphasized when assessing targets for individual EU countries.

Apart from the fact that differentiation is important to avoid conflicts, one should also be aware that if it is considered unfair or inappropriate by one of more parties, new and more serious interest conflicts might occur as a result. Differentiation of targets within the EU under the Kyoto protocol will probably contribute to mitigate interest conflicts, although some factors could probably be emphasized more in order to avoid them. In particular, the different price levels for energy in various countries do not seem to have been taken sufficiently into account. More competition in the electricity market may reduce the importance of this aspect, but, on the other hand, make the 'winners' and 'losers' of the present differentiation more visible.

The three northern countries have more opportunities because the electricity system is very emission intensive, and they can deliver their electricity at low prices, partly due to subsidies. The change in competitiveness due to the supposed tax analyzed here is thereby both unfair and inefficient. It is difficult to point at specific measures taken within countries that could contribute to mitigate the differences in the cost-push across the sectors of different countries. However, a development of a common European market for electricity would level out price variations and thereby contribute to mitigate these differences. This is clearly an issue where a co-ordination of the EU is advantageous, and probably also ahead of many other regions. Provided that the EU succeeds in establishing a common electricity market, the possibility to take directional leadership in climate polity would be improved, and in addition they would probably take directional leadership in energy policy.

Other measures are often preferred to taxes even though taxes may lead to a cost-effective result. This may be due to practical difficulties in establishing a tax regime, or simply because a tax regime itself may generate opposition if it is considered unfair or unnecessary. The analysis shows that the opposition to carbon taxes in the EU in the nineties may, at least partly, have resulted from anticipation of the costs of such a tax. The basic assumption is that the differences in the anticipated costs would create opposition because unevenly distributed costs makes lobby group organizing easier. However, sectors where the costs of climate policies are low are presumed to be less opposed to a tax policy. The possibility of using alternatives to taxes is perceived to be limited, particularly in the case of $CO_2$ emissions, because there is a large number of relatively small, widely dispersed sources (individuals, households, small firms), which leads to high administrative costs.  The costs of administering direct measures for a smaller number of large enterprises are typically less. However, the possibilities of ineffective control, due to for example asymmetric information and uncertainty about future costs, are greater.

There are many possible ways to confuse the connection between emissions targets and the costs of achieving them, e.g. by means of an appropriate tax. A common argument is that the industry needs a stable framework under which the authorities want them to operate, partly because of the long-term nature of capital-based production. They would prefer to know the future quantity they are allowed to emit rather than to have to adapt to an adjustable charge. In principle, permits may be just as effective as taxes, but permits would then have to be differentiated among activities. Assessing the appropriate differentiation is extremely difficult and the result is, in fact, just as costly to the individuals as taxes, except for those who are allocated free permits. Emissions permits are therefore assumed to be appropriate only for large enterprises.

A combined system of both emissions permits and taxes could be advantageous for those who obtain permits at the expense of those subject to taxes. One reason is that the uncertainty around the national cost of achieving the target is put upon those who pay the tax. Second, the most significant options to reduce greenhouse gas emissions are often found in huge but ineffective industries. A cost-effective strategy then implies that these industries reduce emissions by much more than others. The combined system may then represent a considerable opportunity for lightening the climate policy burden for the big enterprises, since they will gain from any target milder than the cost-effective one, while the authorities may easily regard any permit stronger than the average as a ‘heavy burden’.

Indeed, a mixed system is exactly what the EU has agreed upon when in June 2003 the European Parliament approved the world's first international emissions trading system. The initiative enables the establishment of a multibillion EU-wide market for rights to emit carbon dioxide. Companies will be allowed to buy and sell the difference between their actual emissions and their limit set by member countries. Some 10,000 establishments including steelworks, power generators, oil and gas refineries, paper mills, glass factories, and cement installations will be affected starting in 2005. The parliament also adopted amendments that could bring aluminum and chemical producers into the program later. Governments are expected to distribute most pollution permits for free, but the law allows them to auction up to 5 percent of the allocated permits until 2007, and up to 10 percent after that (Environmental News Network, July 3, 2003) [78].

## 6. Concluding Remarks

More than ten years after the EU Commission originally proposed a directive on carbon taxes they reached agreement on a proposed framework for energy taxation. The 1992 proposal faced so much domestic resistance that agreement was not reached until the directive had been considerably watered down. The lobbying succeeded in having the tax altered significantly, with an energy component added, energy-intensive sectors exempted, and the entire tax package made conditional on other OECD countries undertaking precisely the same tax.

In addition, the Commission, which fought against an emissions trading system at the international level, arguing that the flexibility mechanisms of the Kyoto Protocol weakened the incentive to undertake domestic emissions reductions, nevertheless relented in 2003 and agreed to adopt the world’s first regional emissions trading system. The weakening of the carbon tax compared to the robustness of the emissions trading system thus looks like a de facto shift in policy.

---

[78] See e.g. Bang et al. (2004) for a brief summary of the EU’s shift towards emissions trading.

It has been argued that the transition from resistance to policy innovation in EU climate policy is due to considerable pressure in the climate negotiations from the US. However, internal factors, like the political infeasibility of an extensive carbon tax regime, have had their share of the explanation. The analysis above suggests that opposition due to conflicts of interest emerging from market actor's anticipation of the costs of a carbon tax may explain why carbon taxes have been politically difficult to implement. The mixed system (that is, the combination of carbon taxes and emissions trading) eventually adopted by the EU suggests that implementation of a regional carbon tax needs to be accompanied by other strategies in order to be politically feasible. Both the energy taxation framework and the permit market bill, however, allow exceptions.

The aim of this paper was to study economic reasons for the political infeasibility of extensive carbon taxes. More specifically, we have pointed out reasons for why the opposition to carbon taxes was so great. The estimates and comparisons of costs of expected emissions cuts across sectors and across countries in the EU illustrated how different economic sectors might have anticipated the impacts from an expected carbon tax. This focus shows exactly how what seems to be cost-effective and to the best for EU as a region on paper may, because of the myopic vision of the affected sectors, turn out too controversial to be politically feasible.

The decision on the directive for emissions trading makes the EU a pioneer in the large-scale application of market based instruments in climate policy. This development would have been unthinkable a few years earlier (Butzengeiger et al. 2003). It would therefore be very interesting to look further into reasons for why a system of emissions trading was more politically feasible than carbon taxes. One obvious explanation is the decision making procedures in the EU. While the carbon tax requires consensus because it is a fiscal measure, the EU directive on a framework for GHG emissions trading is a burden-sharing measure and as such requires only a qualified majority (Skjærseth and Skodvin 2003). When a majority rule is applied, the opposition must persuade a sufficiently large minority to block the proposal. It is no longer sufficient to convince only one country since the member countries have no veto right. A goal for further work would thus be to analyze other, non-institutional, reasons for the difference in opposition to the two systems.

From the analysis it is clear that the selection of measures is vital, and that different countries have different needs when it comes to policy design. While certain measures may be acceptable to some countries, they are likely to generate opposition in others. The conflicts of interest are largely the result of dependency on fossil fuels, which is higher in the three northern countries and lower in the three southern EU countries in this study. In general, the greater the interest conflicts, the greater are the opportunities to find ways to reduce emissions because of the relative low efficiency in these countries' electricity production.

Three main conclusions can be drawn. First, common measures across countries are generally less attractive since a particular measure may be advantageous to one country in order to keep the national cost of climate policy down, but may in another country spur opposition that could be avoided if exceptions were allowed. Second, the electricity sector plays a key role in climate policy in every EU country and in the potential opposition to any policy measure. Conditions in the electricity sector vary greatly among countries, and a long-term strategy to reduce opposition due to internal as well as external conflicts of interest would, therefore, be to introduce a common electricity market. Third, before a common electricity market can be established, further differentiation of national emissions targets could represent an important opportunity for mitigating conflicts of interests across countries.

## ACKNOWLEDGEMENTS

We are grateful to Geir Asheim, Guri Bang, Scott Barrett, Espen Froyn, Fred Menz, Ottar Mæstad, Jon Vislie and two anonymous referees of INEA for helpful discussions on this material and comments on earlier drafts. We also thank Lynn Nygaard for excellent editorial assistance, Tone Veiby for figure and table editing, Frode Rørvik for computer assistance, and the Research Council of Norway for financial support. Responsibility for errors and opinions are our own.

## REFERENCES

Aaheim H.A. and C. Bretteville (1999), ‘Sources of conflicts in climate policy within the EU – An economic analysis’. CICERO Report 1999:3.

Aaheim H.A. and C. Bretteville (2000), ‘Emission reductions in EU countries – Sources of conflicts’. Chapter 12 in J. Gupta and M. Grubb (eds.), Climate Change and Eropean Leadership, 135-156. Kluwer Academic Publishers, the Netherlands.

Bang, G. (2004), Sources of influence in climate change policy making: A comparative analysis of Norway, Germany, and the United States. PhD thesis, Dep. Of Political Science, the University of Oslo.

Bang, G., H.-C. Lee, and J. Vevatne (2004), ‘Meeting Kyoto Commitments: EU Influence on Climate Policy in Norway and Germany’. Chapter proposal to a book on global warming and European foreign policy, CICERO.

Barkdull, J. and P.G. Harris (2002): ‘Environmental Change and Foreign Policy: A survey of Theory’, *Global Environemntal Politics* 2:2, May 2002, pp. 63-89.

Baumol, W.J. and W.E. Oates (1988). The theory of environmental economics. Second edition. Cambridge University Press

Birkelund, H., E. Gjelsvik, and M. Aaserud (1993). Carbon/Energy Taxes and the Energy Market in Western Europe. Discussion Paper no. 81, Statistics Norway, Oslo.

Boug, P. and L. Brubakk (1996): ‘Impacts of Economic Integration on Energy Demand and $CO_2$ emissions in Western Europe’, Documents 96/15, Statistics Norway

Buchanan, J.M. and G. Tullock (1975). “Polluters’ Profits anf Political Response: Direct Control versus Taxes”, *The American Economic Review*, Vol. 65, Issue I, March 1975, 139-147.

Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (1997), Second National Communication to the UNFCCC by the Federal Republic of Germany. Bonn.

Butzengeiger, S., A. Michaelowa, and S. Bode (2003): Europe- a Pioneer in Greenhouse Gas Emissions Trading. History of Rule Development and Major Design Elements’, *Intereconomics*, July/August 2003, pp. 219-228.

Capros, P. (1998): “Economic and Energy System Implications of European $CO_2$ Mitigation Strategy: Synthesis of Results from Model Based Analysis”, in *Economic Modelling of Climate Change*, OECD Workshop REPORT 17-18 Sept. 1999, OECD, Paris

COHERENCE (1991): *Cost-effectiveness Analysis of $CO_2$-reduction Options; Synthesis Report.* JOULE CEC DGXII Report.

Edmonds, J.A., and D.W. Barns (1992): ‘Factors Affecting the Long Term Cost of Global Fossil Fuel $CO_2$ emissions Reducions’, *International Journal of Global Energy Issues*, [**4**] (3), 140-166.

Environmental News Network, July 3 (2003), ‘E.U. parliament approves greenhouse gas trading system’, http://www.enn.com/news/2003-07-03/s_6158.asp

Environmental News Service, August 26 (2003), ‘Europe Reaches Agreement on Energy Tax Framework’, http://ens-news.com/ens/mar2003

Flavin, C. and N. Lensen (1994): *Power Surge. Guide to the Coming Energy Revolution*. World Watch Institute, W.W. Norton & Company. New York, London.

Howe, C.W. (1994), ‘Taxes versus Tradable Permits: A Review in the Light of the U.S. and European Experience’, *Environmental and Resource Economics* **4**, 151-169

International Energy Agency (1992). Energy Policies and Programmes of Member Countries. IEA, Paris.

International Energy Agency (1993): Energy Prices and Taxes 1991. OECD/IEA, Paris

International Energy Agency (1995): Energy Balances for 1993/1994. OECD/IEA, Paris

Jorgenson and Wilcoxen (1994). The Economic Effects of a Carbon Tax in Climate Change: Policy Instruments and their Implications. Proceedings of Tsakuba Workshop of IPPC Working Group III, Tsukuba.

Kram, T. (1998): ‘The Energy Technology Systems Analysis Programme: History, the ETSAP Kyoto Statement and Post-Kyoto Analysis’ in *Economic Modeling of Climate Change*, OECD Workshop REPORT 17-18 Sept. 1999, OECD, Paris

Kram T. and D. Hill (1996): “A Multinational Model for $CO_2$ Reduction”, *Energy Policy* **24** (1), 39-51.

Krause and Koomey (1996 ), ‘Electricity Sourcebook B: Low-Carbon Power Supply Options in Western Europe’, Part 3 in Energy Policy in the Greenhouse Volume Two, Prepared for the Dutch Ministry of Housing, Physical Planning and Environment, International Project for Sustainable Energy Paths, El Cerrito, California.

Michaelowa, A. (1998), ‘Impact of interest groups on EU climate policy’, *European Environment,* **8** (5) 152-160.

Michaelowa, A. (2000), ‘The strength of different economic interests in shaping EU climate policy’, *Energy and Environment,* **11** (3) 277-292.

Mueller, D.C. (1989), Public Choice II, Cambridge University Press.

OECD (1996), *OECD Economies at glance*, OECD/OCED, Paris

Oliveira-Martins, J., Burniaux, J.-M., J.P. Martin, and G. Nicoletti (1992): “The Cost of Reducing $CO_2$ Emissions: A Comparison of Carbon Tax Curves with GREEN. *OECD Economic Working Papers*, no. 118, Paris.

Peck, S.C. and T.J. Teisberg (1994): “International $CO_2$ Emissions control. An Analysis Using CETA”, *Energy Policy* **[23]** (4/5), 297-308.

Phylipsen, G.J.M., J.W. Bode, K. Blok, H. Merkus and B. Metz (1998):”A Triptych sectors approach to burden differentiation; GHG emissions in the European bubble”, *Energy Policy* **[26]** (12), 929-943.

Pigou, A.C. (1920), The Economics of Welfare, MacMillian and Co., London

Skjærseth J.B. and T. Skodvin (2003). Climate change and the oil industry: Common problem, varying strategies. Manchester University Press.

Svendsen, G.T. (1998). Public Choice and Environmental Regulation, Edward Elgar Publishing