Schmollers Jahrbuch 121 (2001), 1 – 26
Duncker & Humblot, Berlin

International Environmental Cooperation in the One Shot Prisoners' Dilemma*

By Alfred Endres and Cornelia Ohl

Abstract

Our analysis simultaneously deals with two types of uncertainty: first, the uncertainty of the behaviour of nature (stochastic or parametric uncertainty) and second, the uncertainty of the behaviour of nations (strategic uncertainty). This *risk-strategic* analysis points out that chances of international coalition formation to protect the global commons depend on the characteristics of the national welfare distributions and the country specific risk attitudes. Focusing on a static two-country-model and a dichotomous choice setting we point out that risk aversion is a prerequisite for transforming a static prisoners' dilemma (according to the order of expected national welfare) into a game of higher cooperation possibility. For different intensities of risk aversion we develop a typology of coopertive behaviour showing that enforcing environmental agreement is not necessarily harder than initiating it. Moreover we investigate how the design of strategies of international risk management (here: emission trading with and without trade restrictions) feeds back to the incentive structure of an international treaty, like the Kyoto protocol. We argue that the traditional judgement criteria of policy assessment in an international setting should be put into a wider context by the criterion of *"cooperative push"*. This criterion reflects the ability of instruments and technologies to initiate and self-enforce international environmental agreements. Thereby it provides the necessary link between local and global concern.

Zusammenfassung

Wir integrieren zwei Typen von Unsicherheit, die in der Literatur zu den internationalen Umweltproblemen bisher getrennt behandelt worden sind, parametrische (stochastische) und strategische Unsicherheit. Die erstgenannte Form bezieht sich auf die den Entscheidungsträgern unbekannten künftigen Zustände der Natur, die zweite darauf, dass für jeden einzelnen Staat das Verhalten der anderen Staaten ungewiss ist. Unsere *risiko-strategische* Analyse ergibt, dass die Chancen, multinationale Koalitionen zum Schutz der globalen Umwelt zu schließen, stark von den län-

* The authors gratefully acknowledge the receipt of grant no. EN 192 / 4 -1 by the German Research Foundation (DFG). They are also indebted to Bianca Rundshagen, University of Hagen, and two anonymous referees for their helpful comments on an earlier draft of this paper.

derspezifischen Risiko-Präferenzen abhängen. Risikoscheu erweist sich als notwendige Bedingung für die Überführung eines statischen Gefangenendilemmas in ein Spiel mit höherer Kooperationswahrscheinlichkeit. Wir zeigen, wie dieser Zusammenhang für ein internationales Risikomanagement genutzt werden kann und wenden die Analyse auf das Kyoto Protokoll an. Bisher hat sich die ökonomische Analyse umweltpolitischer Instrumente auf traditionelle Kriterien wie Effizienz, ökologische Treffsicherheit und dynamische Anreizwirkung konzentriert. Aus unserer Analyse folgt, dass der Katalog im internationalen Kontext um das Kriterium der "Kooperationsfreundlichkeit" (Cooperative Push) ergänzt werden muß. Mit diesem Kriterium gelingt es, die Fähigkeiten von verschiedenen umweltpolitischen Instrumenten, die individuelle Rationalität und die Stabilität von Vereinbarungen zu fördern, zu erfassen.

JEL-Classification: Q20, Q28, C78.

1. Introduction

To limit global environmental risks internationally coordinated action is a must. Global pollutants, defined to show uniform diffusion, are pure public bads (e.g. Siebert 1998). Therefore countries are unable to reach the nationally most preferred pollution (emission) load unilaterally. Consequently nations have to reach an agreement on a globally accepted emission target. However, agreements are difficult to attain and to protect. The provision of global public goods (here: environmental quality) is a voluntary task of sovereign nations which is usually not enforceable by global institutions. Since each country maximises national welfare as a free-rider (problem of self-enforcement; e.g. Barrett 1994, 1997; Heister 1997) globally optimal solutions are generally missing.

Because of national welfare depending on the success of international risk management i.e., on the possibility of cooperation among sovereign nations, economic literature has given close attention to the strategic behaviour of the polluting countries. Up to now the design of mechanisms which initiate and self-enforce international environmental agreements have been in the centre of interest of game theoretic analyses (e.g. Barrett 1992, 1997; Carraro/Siniscalco 1997; Endres 1997; Finus/Rundshagen 1998; Stähler 1998a, b). A contract is said to be self-enforcing when each country maximises national welfare by holding on to the agreement. The property of self-enforcement is of great importance since supranational institutions to monitor international agreements (and if necessary to punish free-riders) are missing.

To reduce complexity, game theoretic literature on this issue usually neglects an important aspect of the problem: National welfare is undetermined because of scientific lacks of knowledge regarding the relationship between environmental pollution and damage. Additional uncertainties arise from the ambiguities of economic damage evaluation. Thus game theoretic analyses only concentrate on strategic uncertainty caused by the interaction of rational players (here: the nations). In contrast to this the interaction of scientific uncertainty and the optimal level of pollution control is the heart of risk theoretic analyses (e.g. Chichilnisky / Heal 1998; Siebert 1998; Welsch 1995; Xepapadeas 1997). However these analyses generally neglect strategic aspects which are as important as aspects of stochastic (parametric) uncertainties, in the international context.

The *risk-strategic* analysis dealt with in this paper tries to bridge the coexisting literature by the integration of both: *strategic uncertainty* caused by the behaviour of nations and *stochastic uncertainty* caused by the behaviour of nature. In this overriding context we discuss the effects of the interaction of stochastic and strategic uncertainties on the possibility of international coalition formation. To keep things simple, the analysis is confined to a two-country-model and a dichotomous policy choice (cooperation versus defection).

We proceed as follows:

The following *section 2* defines the prisoners' dilemma which in the past has been the starting point of many game theoretic analyses. *Section 3* expands this basic model by the integration of stochastic uncertainty. With the help of a "classical" decision criterion, the μ-σ-principle, we link stochastic and strategic uncertainty. *Section 4* introduces the "global alliance of risk". *Section 5* discusses the influence of alternative preference patterns (risk-neutrality, risk aversion and risk-seeking) on the possibility of international cooperation. Here the focus is on how national risk attitudes determine the strategy choices of the involved countries and the stability of an international environmental agreement. Based on the results derived in part 5 of the paper, *section 6* develops a typology of cooperative behaviour in terms of game theory. *Section 7* points out that national choice of environmental technology and policy instruments (as emission taxation or quota restriction) can be designed as catalysts of global environmental cooperation given the risk preferences of the involved countries. Here, the focus is on the consequences of the design of alternative emission trading regimes, as discussed within the Kyoto protocol. *Section 8* summarises the main findings. *Section 9* (*Epilogue*) investigates the consequences of alternative assumptions regarding the global alliance of risk. Particularly, the variability of environmental damages is modelled different from the main body of the paper.

2. The Prisoners' Dilemma

The dependence of national welfare on the behaviour of foreign nations can be described with game theoretic methods. According to the quality of real world interdependencies different types of games are illustrative, like the chicken and assurance game.[1] To characterise the interdependencies of nations suffering from a global pollutant the *prisoners' dilemma* is a relevant game. The structure of the game covers the following incentive scheme: Two nations decide on whether to coordinate environmental measures or not (choice of "cooperation" or "defection"). As argued above, binding agreements are not feasible in a supranational setting i.e., the game is non-cooperative. Therefore nations are not sure which action (cooperation or defection) each other nation takes.[2]

Given the strategy choice of the foreign country we label the maximum welfare level which is feasible for the home country by W_{DC}, W_{CC}, W_{DD} and W_{CD} respectively.

We further presuppose that the incentive structure of the prisoners' dilemma adequately describes the decision problem of the two countries under consideration. Then the following order of a home country's national welfare displays the incentive problems of international environmental agreements:

$$W_{DC} > W_{CC} > W_{DD} > W_{CD} \tag{1}$$

For the foreign country the same order holds.

Ranking (1) states that:

Referring to the status quo (*DD*), welfare of both nations can be (Pareto-) improved if bilateral environmental cooperation (*CC*) takes place ($W_{CC} > W_{DD}$). However if the foreign nation behaves as a free-rider (and the home country cooperates unilaterally) national welfare in the home country is lower than national welfare feasible in the status quo ($W_{CD} < W_{DD}$). This is so because national costs of emission control exceed the damage avoided in the home country. If on the other hand the home country behaves as a free-

[1] For an introduction into game theory see e.g. Luce / Raiffa 1957; for a modern version e.g. Eichberger 1993.

[2] With *C* for cooperation and *D* for defection four activity combinations have to be judged: *DC, CC, DD* and *CD*. The first capital letters point to the action of the country under consideration (the "home / first country"), the second to the action of the other country (the "foreign / second country"). If the foreign country defects, the home country is able to choose cooperation or defection leading to the combinations *CD* and *DD,* respectively. Analogously *CC* and *DC* are obtainable if the foreign country's choice is cooperation.

rider (and the foreign country cooperates unilaterally) the home country's welfare improves compared to the choice of bilateral cooperation ($W_{DC} > W_{CC}$). These coherencies also hold vice versa. Consequently (because of $W_{DD} > W_{CD}$ and $W_{DC} > W_{CC}$), each country does best when choosing to defect. Thus disregarding stochastic uncertainties the choice of defection is a dominant option (the equilibrium choice) and in the end the reason why international cooperation fails (at least in this static setting).

So given that a prisoners' dilemma is at work both nations are expected to hold on to the status quo which is unfavourable for each of them compared to bilateral cooperation ($W_{DD} < W_{CC}$). Following these game theoretic considerations international coalition formation is not feasible. However, some international environmental agreements among sovereign nations have been achieved and enforced, as the Montreal Protocol on CFCs, 1987. Moreover an agreement on the limitation of greenhouse gases has been reached in Kyoto, 1997. These phenomena (and other examples of international coalition formation) can not be explained within the simple model of the static prisoners' dilemma. Extensions of this model are needed to cover complex real world stories. As is well known, switching to dynamic games is a possibility to follow that road (e.g. Pearce 1992; Stähler 1996; Taylor 1987). Another possibility is the introduction of stochastic uncertainties, as shown in the paper at hand.

Analysing the effects of stochastic uncertainties on cooperation in a static setting instead of a dynamic one is theoretically appealing. This is so because the static context is the *worst case scenario* regarding the chances for cooperation. If the introduction of stochastic uncertainties generates incentives to cooperate in this adverse setting, the point regarding the importance of these uncertainties is particularly strong.[3] Moreover, in certain real world applications modelling a static game seems to be particularly appropriate. A dynamic game in general presupposes that the same stage game (e.g., a prisoners' dilemma) is played more than once. However, environmental issues like global warming can not always be characterised in terms of a dynamic setting.

[3] It is well known that the problems of the prisoners' dilemma can be solved in a dynamic game with infinite time horizon and a sufficiently low rate of discount. Apart from the argument in favour of using a static model, given above, it may also be doubted that the dynamic variant is more realistic in this paper's context: E.g., the targets of the Kyoto protocol have to be attained within the years 2008 and 2012 (see Oberthür / Ott 1999). After this period further measures have to be agreed upon in a subsequent protocol. To date, it is not clear which targets and measures for which countries a further protocol will contain nor can we take it for granted that there will be a subsequent agreement at all.

3. On the Integration of Probabilistic Phenomena into the Prisoners' Dilemma Game

Scientific uncertainty as well as uncertainty arising from damage evaluation (Hanley/Spash 1993; Houghton et al. 1996) complicate the forecast of the quality and the degree of manmade and natural environmental change.[4] Due to these lacks of knowledge the advantages of slowing global warming by the reduction of anthropogenic caused greenhouse gases are unknown to the nations. Consequently the assessment of national welfare when limiting environmental risks has to be expressed in terms of a national welfare distribution. In the paper at hand this is specified by the mean μ and the spread of the outcomes σ.[5]

The possibility of *manmade* environmental change was already pointed out by Svante Arrhenius at the beginning of the 20th century for the case of climate variation (Arrhenius 1903). Today this possibility enables international risk management to reduce the level as well as the probability of environmental damage by the limitation of global pollutants (endogenous risk). However anthropogenic effects are overlapped by natural phenomena; in the context of climate change, e.g. the position of earth to sun or the activity of sun spots and their relation to climate variations (exogenous risk). The impact of exogenous and endogenous risk factors on environmental change is therefore hard to distinguish but there is no doubt that human activity is unable to reduce the realisation of exogenous risks. It is also out of dispute that ecological systems, like the climate, are subject to sudden fluctuations. Thus, the possibility remains that even a small effect of an anthropogenic risk may have a significant environmental impact.[6] That is why incentives to reduce manmade risks exist, even if anthropogenically caused risks are not the dominant risk factor. We therefore expand the concept of national welfare maximisation under strategic uncertainty by the integration of national attitudes towards risk.

Risk preferences may imply the desire to minimise fluctuations of national welfare (risk aversion). Alternatively, national preference may be insensitive towards the possible sway of welfare (risk-neutrality). A further possibility is that the variation of welfare is explicitly welcome (risk-seeking). These spread preferences may be recorded by the well established

---

4 E.g., research on the climate system still lacks knowledge of the interaction between oceans and climate or between climate and earth's albedo. (Albedo indicates the intensity relation of incoming radiation from sun to earth and outgoing reflections.)

5 These measures are the basis of many risk economic analyses e.g., in investment- and portfolio theory.

6 Imagine a barrel that is brimful with exogenous risks. Then it is just one drop of endogenous risk which runs the barrel over.

μ-σ-criterion (Huang / Litzenberger 1988; Sinn 1990). In the context of this paper the decision for or against environmental protection depends on the level of expected national welfare σ and the spread of national welfare μ.[7] Both decision parameters implicitly take into account the level of damage as well as of damage probability. The paper at hand determines the preference value (ϕ) of a risky option as:

$$\phi = \mu - \alpha\sigma \,. \tag{2}$$

With the preference value ϕ we are able to transform national welfare W (now expressed in terms of μ and σ) into national risk utility. The level of ϕ is determined by the level of the mean μ which records nationally expected welfare and the standard deviation σ measuring the spread of the national welfare distribution, with σ being weighed by a national risk discount factor (α). Risk-neutrality corresponds to a value of α equal to zero, risk aversion to positive and risk-seeking to negative values. The bigger ϕ the higher is the national welfare measured in terms of risk-utility.

If we replace the national welfare levels under different activity combinations (W_{DC}, W_{CC}, W_{DD} and W_{CD}) considered to be certain in equation (1) above, by their corresponding means the following order results:

$$\mu_{DC} > \mu_{CC} > \mu_{DD} > \mu_{CD} \,. \tag{3}$$

The subscripts of μ show the activity combinations to which the means refer. Again, the first letter points to the action of the home country, the second to that of the foreign country. The order of the means corresponds to the incentive structure of a prisoners' dilemma game: Independent from the foreign action the home country's expected welfare is highest in the case of its defection ($\mu_{DC} > \mu_{CC}$ and $\mu_{DD} > \mu_{CD}$). In this sense we speak of a prisoners' dilemma under stochastic uncertainty.

Choosing D or C, however, it is not only expected welfare that varies but also the spread parameter σ. This is so because global pollutants cross over national boundaries and affect the level of damage as well as damage probability in *each* of the nations. This interaction generates the "global alliance of risk" that is measured by the following order of the standard deviation σ (which aggregates alternative levels of damage as well as damage probability):

---

[7] It should be noted that the method applied here is also relevant in connection with different decision criteria or measures of risk, like the variance or semi-variance of a welfare distribution. For alternative measures of risk e.g. see Brachinger / Weber (1997). Also see e.g. Sinn (1983) for the possibility of transforming a μ-σ-criterion into a risk utility function of the von Neumann-Morgenstern-type. On the approach presented above, also see Endres / Ohl (1998).

$$(4) \qquad \sigma_{DD} > \sigma_{CD} = \sigma_{DC} > \sigma_{CC} > 0\,.$$

Ranking (4) states that:

Starting from the status quo environmental risk can be reduced by each cooperative contribution of the countries ($\sigma_{DD} > \sigma_{CD}$; $\sigma_{DC} > \sigma_{CC}$). Thus we presuppose that the simple principle holds: Each emission unit abated or not emitted in the first place reduces manmade environmental risks. The limitation of anthropogenically caused risk is highest if both countries cooperate simultaneously. Assuming symmetric emission reduction for complexity reduction,[8] each nation unilaterally has the same impact on risk reduction (e.g. climate stabilisation) so that $\sigma_{CD} = \sigma_{DC}$ holds.

4. The Global Alliance of Risk

In the model presented here and (most likely) also in reality, the nations form a *global alliance of risk:* The environmental damages they suffer and the probabilities with which these damages occur depend upon the aggregate effect of all countries' emissions reductions.

There are (at least!) two important issues to be dealt with when modelling a global alliance of risk:

- In traditional environmental economic analysis it is assumed that environmental damage continuously varies with the level of emissions. The expected damage is taken to be a well behaved function of emissions, "preferably" with positive first and second order derivatives.[9] If we take the global environmental problem dealt with in this paper to be of this conventional kind, it is completely appropriate to model it with the von Neumann-Morgenstern approach. This is what we do in this paper, using the μ-σ-criterion. An alternative view would be to describe global environmental problems to be "catastrophic": With a very low probability they might lead to the extreme form of damage, extinguishing life on our planet. It is sometimes argued that for global warming, a certain threshold of emission concentrations might exist, the transgression of which would lead to global catastrophe. Just like in the case of nuclear power plants, this type of risk is not adequately modelled using the expected utility framework.[10]

---

[8] In case of cooperation each cooperating country is expected to avoid the same amount of global pollutants. Of course, the model is open for a generalisation.

[9] See, e.g., Endres (2000), Kolstad (2000), or any other environmental economics textbook.

[10] See, e.g., Chichilnisky (2000).

Of course, whether global environmental problems are of the conventional or of the catastrophic type is an empirical matter. It is possible that the answer will depend upon the type of global environmental problem we are dealing with. In the context of global warming there is considerable scientific dispute upon whether a critical threshold exists (Roughgarden / Schneider (1999), Nordhaus (1994)). If it does exist, it is more likely to be relevant in the long run (Manne / Richels (1999)), than in the short or intermediate run (Pizer (1999)).[11] In the negotiations for the Kyoto protocol, a critical threshold of CO_2- concentrations in the atmosphere did not play any role for the definition of the emission reduction targets specified in the agreement.

All in all, even though it may not be appropriate for each specific global problem in any context, expected utility theory, as used in this paper, is probably helpful to explain and to evaluate the behaviour of countries in the international environmental arena.

- Of course, using the μ-σ-criterion does not say how the means and standard deviations of environmental damages vary with emissions.

  There will be little dispute with our assumption (3), supposing that the structure of our problem is an expected prisoners' dilemma.[12] Our assumption on the standard deviations is likely to be more controversial.[13]

  We have already given a brief rationale for our presumption that the degree of uncertainty decreases with the degree of cooperation when we introduced (4) in chapter 3, above. Let us elaborate: Emissions are closely correlated with economic variables, like the GDP. Future development of GDP is uncertain. In case the countries do not commit to stabilize their emissions (i.e., they defect, *DD* in (4)), the uncertainty of future emission levels will be a consequence of the GDP-uncertainty. This is different in the case of both countries agreeing to certain target levels of emissions (*CC* in (4)). In this case the effect of the sway in the GDP on the sway of emissions will be counterbalanced by environmental policy.[14] Thus, in the

---

11 Consequently, WBGU (2000) characterizes global warming by two different types of risk: One conventional (intermediate) and one catastrophic (long run).

12 The prisoners' dilemma is the economist's traditional interpretation of global (and other) environmental problems (see Brockmann / Stronzik / Bergmann (1999) for a recent application to global warming). Still, there is no natural law that environmental problems are of the prisoners' dilemma type, indeed. After all, this is an empirical matter like, alas, so many issues dealt with in this paper. E.g., Lipnowski / Maital (1983) argue that many environmental problems are described better by the Chicken Game.

13 Indeed, both referees were concerned about this assumption.

14 Stabilization of national emissions to meet the standards of an international agreement is complete if the countries use a transferable discharge permit system to regulate emissions. If fully enforced, this system operates with complete "ecological accuracy".

case of bilateral cooperation the variance of emissions is plausible to be smaller than in the case of bilateral defection. If one country cooperates and the other one defects (*CD* in (4)), the variance is plausible to be in-between.

Of course, we would be happy to present empirical evidence instead of speculations in discussing (4). Unfortunately, this kind of evidence on the relationship between the variability of damages and the level of emissions is fragmentary at best.[15] The data base for an empirically supported formulation of (4) is much too weak and spotty. Moreover, in areas of application where we have some evidence, it seems to point into different directions: A case where the standard deviation seems to have increased with the level of emissions (global temperature) is rain. In the process of global warming not only did the average amount of rain per incidence increase, but also the periods of drought on the one hand and of "flash floods" on the other hand. However, the variability of daily temperature seems to constitute a counter example: Even though average temperature increased, its daily amplitude decreased.

What should we conclude for economic analysis from the fact that sciences do not give a clear cut answer about how to specify (4), and perhaps never will?

To our opinion, economic analysis should produce *conditional* answers: If the inequality signs in (4) point one way, what is the consequence for the nations' propensity to cooperate? If they point the other way, how does the propensity change?

This is the route we have taken in the present paper. In its main body we carry our analysis *given* the standard deviation decreases with increasing levels of cooperation (i.e. with decreasing levels of emissions), as assumed in the formulation of (4) given in chapter 3, above. In the epilogue (Ch. 9) we investigate what it does to our results if the unequality signs in (4) are reversed.

5. Risk Preferences and Strategy Choice

If nations apply the μ-σ-principle introduced above and if they seek to maximise national risk utility, cooperation is preferred to defection if the former choice yields a higher ϕ than the latter. Dependent on the action chosen by the foreign country we have to distinguish two cases:

[15] These remarks draw upon our personal correspondence with Hartmut Graßl (MPI and University of Hamburg). We would like to thank Professor Graßl for his patience in explaining scientific complications to some economists.

a) The *foreign* country defects

An incentive to cooperate exists in the *home* country if:

(5) $$\phi(CD) = \mu_{CD} - \alpha\sigma_{CD} > \phi(DD) = \mu_{DD} - \alpha\sigma_{DD} \Rightarrow$$

(I) $$\alpha > (\mu_{DD} - \mu_{CD})/(\sigma_{DD} - \sigma_{CD}) \equiv \alpha_{\mathrm{I}}^{\min} > 0$$

b) The *foreign* country cooperates

The *home* country is expected to cooperate, too, if:

(6) $$\phi(CC) = \mu_{CC} - \alpha\sigma_{CC} > \phi(DC) = \mu_{DC} - \alpha\sigma_{DC} \Rightarrow$$

(II) $$\alpha > (\mu_{DC} - \mu_{CC})/(\sigma_{DC} - \sigma_{CC}) \equiv \alpha_{\mathrm{II}}^{\min} > 0$$

Given the strategy choice of the foreign country the minimum degree of risk aversion which supports cooperation in the home country is labelled $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$, respectively. Both thresholds are determined by the stochastic terms of the model. Inequality (I) refers to an incentive to cooperate unilaterally (condition of incentive compatibility: $\alpha > \alpha_{\mathrm{I}}^{\min}$). Inequality (II) refers to an incentive for bilateral cooperation (condition of self-enforcement: $\alpha > \alpha_{\mathrm{II}}^{\min}$). The conditions (I) and (II) show that internationally coordinated measures on the one hand depend on the level of the parameters μ and σ (the 'objective' risk assessment) and on the other hand on the intensity of national risk preferences (the 'subjective' risk assessment) expressed by the value of α.

Given the order of μ and σ as of (3) and (4), cooperative behaviour is only to be expected if α takes a positive value ($\mu_{DD} > \mu_{CD}$ and $\sigma_{DD} > \sigma_{CD}$ as well as $\mu_{DC} > \mu_{CC}$ and $\sigma_{DC} > \sigma_{CC}$). Thus, in equilibrium the choice of cooperation is only to be expected if the countries under consideration are risk averse ($\alpha > 0$). Neither risk-neutrality ($\alpha = 0$) nor risk-seeking ($\alpha < 0$) strengthens the possibility of international cooperation, here. Consequently country-specific risk preferences are able to explain why even nations showing identical profiles of national welfare distributions do not necessarily act in the same way.

Although risk aversion is a prerequisite for international cooperation, given (3) and (4) hold, the supposition of risk aversion is not sufficient to improve environmental quality. Only if the intensity of risk aversion exceeds one of the thresholds, incentives for cooperative environmental protection arise. In contrast to this, traditional risk economic analyses, generally excluding strategic interdependencies, lead to the result that environmental quality goes up if risk aversion is at work (e.g. Kreuzberg 1994; Siebert

1998). So integrating strategic and parametric uncertainties, as done in the paper at hand, really makes a difference.

Moreover conditions (I) and (II) reveal that the claims on the minimum degree of risk aversion necessary to foster cooperation adapt to country-specific differences in the levels of the stochastic parameters, μ and σ. This will e.g. be the case if differences in national objective risk assessment arise (the levels of μ and σ then might nationally vary because of different expertise or in case of asymmetric nations because of differences in national damage expectation). In consequence the thresholds in the home country may differ from that of the foreign country. This suggests (also see part 7 of the analysis) that an agreement among countries differing in culture and socio-economic characteristics is not necessarily harder to attain and to protect than an agreement among homogenous countries.[16]

Consequently acquiring knowledge of national risk attitudes (their nature *and* strength) is as important as gathering information on scientific uncertainties (here: on μ and σ) and their relation to the strategic (emission) behaviour of nations. Knowledge of each is a prerequisite first, to build up expectations regarding the cooperative behaviour of foreign countries and second, to predict the success of international risk management.

Having this knowledge, the condition of incentive compatibility (I) and the condition of self-enforcement (II), that determine the thresholds $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$, provide a claim profile with which national risk preferences can be distinguished according to their cooperative push.

6. A Typology of Cooperative Behaviour

The *risk-strategic* analysis presented above demonstrates that even in a static setting certain risk preferences are able to push international cooperation within the unfavourable expectation structure of a prisoners' dilemma game.[17] The cooperative behaviour of nations (here: for the case of risk aversion) can be illustrated by different types of games depending on the re-

---

16 Consider for example two nations (1, 2) differing in national reduction costs (e.g. assume: $\mu_{CD}^1 > \mu_{CD}^2$ and $\mu_{CC}^1 > \mu_{CC}^2$ with the superscript indicating the country to which the means refer). Then the thresholds in country 1 ($\alpha_{\mathrm{I}}^{1\,\min}$; $\alpha_{\mathrm{II}}^{1\,\min}$) are lower than the thresholds in country 2 ($\alpha_{\mathrm{I}}^{2\,\min}$; $\alpha_{\mathrm{II}}^{2\,\min}$) i.e., $\alpha_{\mathrm{I}}^{1\,\min} < \alpha_{\mathrm{I}}^{2\,\min}$ and $\alpha_{\mathrm{II}}^{1\,\min} < \alpha_{\mathrm{II}}^{2\,\min}$. In this setting bilateral cooperation may still prevail. First, if $\alpha^1 = \alpha^2 \geq \alpha_{\mathrm{II}}^{2\,\min}$ and second, if $\alpha^1 \neq \alpha^2$ but $\alpha^1 \geq \alpha_{\mathrm{II}}^{1\,\min}$ as well as $\alpha^2 \geq \alpha_{\mathrm{II}}^{2\,\min}$ holds.

17 Hence problems of the prisoners' dilemma type, like the case of global warming, do not necessarily have to be redefined (according to their expected cost-utility-structure) as games of a higher cooperation possibility (as e.g. done by Edenhofer 1996 or supposed by Ostrom 1999) to explain that international measures of risk management are chosen by independent countries. Nor is it necessary to put them into a dynamic setting for this purpose.

lation between the thresholds $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$, and the national risk attitude α. Since it can not be generally stated whether the threshold of incentive compatibility ($\alpha_{\mathrm{I}}^{\min}$) is lower (equal or higher) than the border of self-enforcement ($\alpha_{\mathrm{II}}^{\min}$),[18] there are three cases to distinguish: $\alpha_{\mathrm{I}}^{\min} = \alpha_{\mathrm{II}}^{\min}$ (case 1), $\alpha_{\mathrm{I}}^{\min} < \alpha_{\mathrm{II}}^{\min}$ (case 2) and $\alpha_{\mathrm{I}}^{\min} > \alpha_{\mathrm{II}}^{\min}$ (case 3).

Case 1: $\alpha_{\mathrm{I}}^{\min} = \alpha_{\mathrm{II}}^{\min}$

In the special case where $\alpha_{\mathrm{I}}^{\min}$ equals $\alpha_{\mathrm{II}}^{\min}$, each intensity of risk aversion of a home country lower than the thresholds ($\alpha < \alpha_{\mathrm{I}}^{\min}; \alpha_{\mathrm{I}}^{\min}$) leads to a higher degree of national risk utility (ϕ) if this country defects irrespective of the foreign strategy choice ($\phi_{DD} > \phi_{CD}$ and $\phi_{DC} > \phi_{CC}$). For each intensity of risk aversion higher than the thresholds ($\alpha > \alpha_{\mathrm{I}}^{\min}; \alpha_{\mathrm{I}}^{\min}$) the opposite holds: The preference value of a cooperating home country is always higher than its preference value in case of defecting ($\phi_{DD} < \phi_{CD}$ and $\phi_{DC} < \phi_{CC}$).

With this, case 1 delivers two general results holding in each of the three cases:

If actual risk aversion overleaps both, the threshold of incentive compatibility ($\alpha_{\mathrm{I}}^{\min}$) and the threshold of self-enforcement ($\alpha_{\mathrm{II}}^{\min}$), cooperation is in the self-interest of each nation irrespective of what the other country does. In terms of game theory, the thresholds $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$ separate between a game of the prisoners' dilemma type (which is played when $\alpha < \alpha_{\mathrm{I}}^{\min}, \alpha_{\mathrm{II}}^{\min}$) and a no conflict game (which is played when $\alpha > \alpha_{\mathrm{I}}^{\min}, \alpha_{\mathrm{II}}^{\min}$). Thus if (I) and (II) are simultaneously met in each of the nations the global alliance of risk is able to completely absorb international incentives to free ride. Below the thresholds environmental cooperation does not constitute an equilibrium choice and therefore should not be expected.

We now drop the restrictive assumption of the special case where $\alpha_{\mathrm{I}}^{\min} = \alpha_{\mathrm{II}}^{\min}$ holds and allow the two thresholds to deviate from each other. Consider the case $\alpha_{\mathrm{I}}^{\min} < \alpha_{\mathrm{II}}^{\min}$.

Case 2: $\alpha_{\mathrm{I}}^{\min} < \alpha_{\mathrm{II}}^{\min}$

In addition to what has been said for case 1 we now have to consider a further game in the case where the country-specific risk attitude α lies in-between the thresholds ($\alpha_{\mathrm{I}}^{\min} < \alpha < \alpha_{\mathrm{II}}^{\min}$).

If we consider case 2 and it is only condition (I) that is met ($\alpha_{\mathrm{I}}^{\min} < \alpha < \alpha_{\mathrm{II}}^{\min}$) an incentive to cooperate unilaterally exists. However, since the condition of self-enforcement is not fulfilled ($\alpha < \alpha_{\mathrm{II}}^{\min}$) this incentive is unstable as soon as the other country chooses to cooperate. In equili-

18 This crucially depends on the relation between the differences of the means and the standard deviations as given in (I) and (II). For more details on this issue see Endres / Ohl (2000a) in the context of the quadratic μ-σ-principle.

brium therefore "only" unilateral cooperation is induced. In addition to the corner solutions: prisoners' dilemma- and no conflict game, we thus have a game of chicken if $\alpha_{\mathrm{I}}^{\min} < \alpha < \alpha_{\mathrm{II}}^{\min}$ holds.

In the chicken dilemma the question is who the cooperating party is. Since each country is better off if the other country acts cooperatively no country might wish to cooperate first. This may lead to a wait-and-see-strategy where each country speculates on maximising national risk welfare as a free-rider (by the choice of defection). However sooner or later unilateral cooperation takes place (an unilateral incentive to cooperate exists; $\alpha > \alpha_{\mathrm{I}}^{\min}$). To forecast which of the countries turns out to be the "chicken", further assumptions have to be made. E.g. in the case of asymmetric nations (regarding their intensity of risk aversion), it is very suggestive that the country with the higher degree of risk aversion is pushed into cooperation first. The reason is that in case of holding on to bilateral defection a highly risk averse nation incurs a higher loss of national risk welfare than its less risk averse counterpart.[19]

Case 3: $\alpha_{\mathrm{II}}^{\min} < \alpha_{\mathrm{I}}^{\min}$

If we have $\alpha_{\mathrm{II}}^{\min} < \alpha_{\mathrm{I}}^{\min}$ and it is condition (II) that is exclusively fulfilled ($\alpha_{\mathrm{II}}^{\min} < \alpha < \alpha_{\mathrm{I}}^{\min}$) incentives to cooperate unilaterally are missing. Nevertheless, bilateral cooperation could be supported by the promise of reciprocal (conditional) cooperation. Signalling cooperation trustworthily is, here, an important factor to actually transform the prisoners' dilemma into a game of stag hunt (a variant of the assurance game) which is played in-between the thresholds.[20]

The stag hunt points to a situation where nations prefer either the status quo or a bilateral cooperation. However, since bilateral cooperation is preferred to bilateral defection each country prefers to contribute to a cooperative risk management (which is here a focal point in the sense of Schelling 1960) as long as it expects the other country to cooperate, too.

Moreover, since the stability border $\alpha_{\mathrm{II}}^{\min}$ is lower than $\alpha_{\mathrm{I}}^{\min}$, actual values of α that are characterised as incentive compatible also fulfil the criterion of self-enforcement. Thus, given case 3 we are able to conclude that if α exceeds the border of incentive compatibility (I), nations do not play the chicken game as in case 2, but the no conflict game instead ($\alpha_{\mathrm{II}}^{\min} < \alpha_{\mathrm{I}}^{\min} < \alpha$).

---

19 We will show in section 7., below, that considering different types of policy instruments to perform environmental cooperation may also be able to solve the chicken dilemma by transforming this game into one with a higher cooperation possibility, like the no conflict game or the game of stag hunt.

20 The importance of reciprocal trust to induce bi- or multilateral cooperation under anarchy (in a non-cooperative game) has also been pointed out e.g. by Axelrod 1984 in a dynamic setting.

It has been shown that depending on the actual intensity of risk aversion (α) and its relation to the threshold values different types of games evolve. Table 1 summarises the different types of cooperative behaviour:

Table 1

A typology of cooperative behaviour

Type of game	Incentive structure	Demands on the risk attitude (α)
prisoners' dilemma	$\phi(DC) > \phi(CC) > \phi(DD) > \phi(CD)$	$\alpha < \alpha_{\mathrm{I}}^{\min}; \alpha_{\mathrm{II}}^{\min}$
chicken	$\phi(DC) > \phi(CC) > \phi(CD) > \phi(DD)$	$\alpha_{\mathrm{I}}^{\min} < \alpha < \alpha_{\mathrm{II}}^{\min}$
stag hunt	$\phi(CC) > \phi(DC) > \phi(DD) > \phi(CD)$	$\alpha_{\mathrm{I}}^{\min} > \alpha > \alpha_{\mathrm{II}}^{\min}$
no conflict	$\phi(CC) > \phi(DC) > \phi(CD) > \phi(DD)$	$\alpha_{\mathrm{I}}^{\min}; \alpha_{\mathrm{II}}^{\min} < \alpha$

The analysis presented above points out that country-specific risk preferences α as well as the stochastic parameters, μ and σ, defining the thresholds $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$, codetermine the incentive for and the stability of international environmental agreements among sovereign nations. The relation between the thresholds and the actual intensity of risk aversion determines whether an incentive to cooperate exists and if so whether this incentive is stable. Our analysis suggests that enforcing an agreement is not necessarily harder than initiating it. E.g. the possibility of $\alpha_{\mathrm{II}}^{\min} < \alpha_{\mathrm{I}}^{\min}$ rises ceteris paribus the higher reduction costs in case of unilateral cooperation are.[21] Thus the relation between the thresholds can be shifted by strategies of *risk management*. This is especially useful in cases where $\alpha < \alpha_{\mathrm{I}}^{\min}; \alpha_{\mathrm{II}}^{\min}$ holds, so that the prisoners' dilemma remains unsolved. Moreover, in the context of a global pollutant also partial cooperation (as in the game of chicken) is unsatisfactory. The reason is that defection of one country causes the threat of undermining unilateral emission reductions by emission increases in the defecting nation (e.g. Hoel 1991). Consequently enlarging cooperation by policy measures (decreasing the value of $\alpha_{\mathrm{II}}^{\min}$) in the context of global environmental risk is as important as initiating it. Some possibilities to do that are discussed in the subsequent chapter.

21 In this case μ_{CD} is small compared to μ_{DD}. Consequently the border of incentive compatibility, $\alpha_{\mathrm{I}}^{\min}$, is relatively high. For the effects on the threshold values in case of different policy instruments also see Endres / Ohl (2000b).

7. Strategies of Risk Management to Shift the Borders of Incentive-Compatibility and Self-Enforcement

The borders of incentive compatibility ($\alpha_{\mathrm{I}}^{\min}$) and self-enforcement ($\alpha_{\mathrm{II}}^{\min}$) are determined by the stochastic parameters μ and σ. Since both values are not only determined by national damage assessment but also by the choice of environmental technology or policy instrument (as measures of actually performing international cooperation) global risk management is able to move both boundaries. To show this, we clarify the relation between the thresholds ($\alpha_{\mathrm{I}}^{\min}$; $\alpha_{\mathrm{II}}^{\min}$), μ and σ, as defined in equations (I) and (II), in the following μ-σ-diagram:

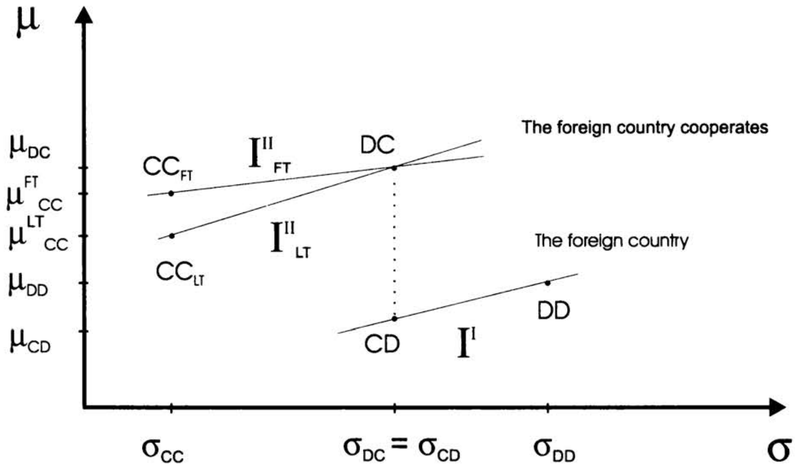

Figure 1: Designing an international environmental treaty (case 1)

Equations (I) and (II) define the slope of a straight line. Thus, in figure 1 above, the slopes of the straight lines represent the values of $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$, respectively, as given by (I) and (II). Each curvature is the location of different μ-σ-combinations leading to the same demands on the national risk attitude α given the strategy choice of the foreign country. Consequently, we can speak of I^{I} and I^{II} as of Iso-requirement-curves of cooperation.[22] Each μ-σ-tuple specifies the impact of national damage assessment, technology choice and / or policy instrument applied on the national welfare distribution. The status quo for example is displayed by *DD* which refers to a business-as-usual-scenario where no measures of international risk management are used.

22 The subscripts *LT* and *FT*, used in figure 1, will be introduced later in this section.

In case of one country or both countries choosing to cooperate each cooperating nation has to decide which environmental instrument or technology to use to perform international cooperation. Generally different types of instruments and technologies are available. Thus the values of the means as well as the values of the standard deviation which specify the μ-σ-tuples *CC* and *CD* depend on politically determined factors. Therefore usually a range of μ-σ-combinations has to be judged. For example comparing emission taxation and quota restriction, the former is (at least in theory) more efficient than the latter. Moreover quotas usually are more accurate in reaching a target (in terms of the level of emission abatement) than taxes. These instrument-specific pros and cons will influence the levels of μ and σ in different ways.[23] Consequently the slopes of the above curves (the relation of the tuples *CD* and *DD* as well as *DC* and *CC*) and hence the possibility of international coalition formation generally depend on measures of global risk management. That is why decision makers are able to shift the demands on the actual α (the slope of the curves) by the selection and the design of environmental policy instruments and technologies.

Consider for example an international quota regime as discussed within the Kyoto protocol to limit greenhouse gases (e.g. Heister et al. 1991). To date it is unclear which part of the assigned amounts[24] will be licensed for trade. Especially the US state that they will not accept any restrictions on trade. On the other hand the European countries claim for trade restrictions performed by a buyers or a sellers cap on the assigned amounts.[25] With the help of figure 1 above and figure 2 below we are able to highlight that the design of an international trade regime (that defines the endogenous risk pattern, here: of global environmental change) may be crucial for the enforcement of international agreements like the Kyoto protocol.

We assume that independent of the applied trade regime the order of the means as well as the order of the standard deviations as given by (3) and (4) hold. Thus the standard deviation is lowest in case of bilateral cooperation

23 For example $\mu_{\text{Tax}} > \mu_{\text{Quota}}$ and $\sigma_{\text{Tax}} > \sigma_{\text{Quota}}$ may hold.

24 The assigned amounts are determined by the country-specific emission targets of the protocol. The US, Japan and the European countries for example agreed to limit their individual amount of six major greenhouse gases (CO_2, CH_4, N_2O, HFCs, PFCs, SF_6) to 7%, 6% and 8%, respectively, below the level of 1990 (base year) emissions during the five-year budget period 2008–2012.

25 For further details on this issue see e.g. Michaelowa / Koch 1999. Also notice that in climate negotiations two group of countries can be distinguished: First, the "umbrella group" with the US as the leading country prefering modest emission controls and second, the "EU-bubble" that includes countries supporting tight emission and trade controls (e.g. the European countries and the small islands group called the AOSIS). This enables us to focus on a two-country-model (with one country reflecting a group of countries) as the most simple way to approximate the interest schemes of the different countries envolved in the process of climate negotiations.

and it increases in case of one country or both countries defecting, as is displayed at the abscissas of figures 1 and 2 ($\sigma_{CC} < \sigma_{CD} = \sigma_{DC} < \sigma_{DD}$). The order of the means is that of a prisoners' dilemma as it is shown at the ordinates of the figures ($\mu_{DC} > \mu_{CC} > \mu_{DD} > \mu_{CD}$).[26]

We first focus on the simple case (case 1, figure 1, above) where the design of the trade regime exclusively causes a difference in expected national welfare. Because of differences in national reduction costs each country's welfare improves when implementing an unlimited trade regime compared to a regime with constraints on trade. (The transformation of differences in national reduction costs into national welfare gains is not bounded by trade restrictions.) We thus have:

$$\mu_{CC}^{FT} > \mu_{CC}^{LT} \tag{7}$$

with μ_{CC}^{FT} the level of expected welfare in case of free trade (*FT*) and μ_{CC}^{LT} the level of expected welfare in case of limited trade (*LT*). The policy-specific differences in the levels of the means as given by (7) affect the slope of the *CC-DC*-curve and hence the level of the threshold $\alpha_{\mathrm{II}}^{\min}$. In case of free trade the threshold is lower than in case of limited trade i.e., the *LT*-curve is steeper than the *FT*-curve ceteris paribus.[27] Thus an unlimited trade regime may lead to the case where $\alpha_{\mathrm{I}}^{\min} > \alpha_{\mathrm{II}}^{\min}$ holds whereas trade restrictions may induce the case $\alpha_{\mathrm{I}}^{\min} < \alpha_{\mathrm{II}}^{\min}$ so that $\alpha_{\mathrm{II}FT}^{\min} < \alpha_{\mathrm{I}}^{\min} < \alpha_{\mathrm{II}LT}^{\min}$ results (see figure 1, above). Consequently, if the actual intensity of risk aversion is close to $\alpha_{\mathrm{I}}^{\min}$ ($\alpha_{\mathrm{I}}^{\min} \cong \alpha$), the selection of a quota regime supporting free trade is able to induce stable bilateral cooperation whereas a regime with barriers of trade fails to do this job ($\alpha_{\mathrm{II}FT}^{\min} < \alpha_{\mathrm{I}}^{\min} \cong \alpha < \alpha_{\mathrm{II}LT}^{\min}$). Thus, ceteris paribus, free trade has a higher cooperative push than trade restrictions: In the *FT*-case a no conflict game evolves whereas in the *LT*-case a chicken game is played. Allowing free trade therefore might enhance the chances of the protocol being ratified by the parties. This confirms the statement (e.g. made in Brockmann / Stronzik / Bergmann 1999) that only the flexibility within the mechanisms of the Kyoto protocol is able to induce countries like the US to take over legally binding emission targets. Moreover, if the US is less risk averse than the EU ($\alpha^{\mathrm{US}} < \alpha^{\mathrm{EU}}$) figure 4 suggests that a reason for ratifying a protocol supporting trade restrictions by the EU but not by the US could be that $\alpha_{\mathrm{II}FT}^{\min} < \alpha^{\mathrm{US}} < \alpha_{\mathrm{II}LT}^{\min} < \alpha^{\mathrm{EU}}$ holds. So varying degrees of risk aversion

---

26 For a classification of the global warming issue as a prisoners' dilemma game also see Brockmann / Stronzik / Bergmann 1999.

27 Notice that the *CD-DD*-curve is not altered by the design of the trade regime. In a two-country-modell trade restrictions only matter in the case of bilateral cooperation. In the case of unilateral emission reduction trading of the amounts avoided among countries does not happen anyway (we neglect the possibility of trading among each group of countries for complexity reduction).

are one explanation for differences in the acceptance of the mechanism design of treaties.

We now focus on the case (case 2, figure 2, below) where not only the means but also the standard deviations depend on the choice of the trade regime. Regarding the implementation of the Kyoto protocol trade restrictions are expected to have some impact on the global level of greenhouse gases. The reason is that the assigned amounts mainly of the Russian Federation and the Ukraine (as given by the zero reduction target) cause the problem of "hot air". The term "hot air" stands for emission rights that are allowed by the Kyoto protocol but will not be called on because of declining economic (and hence emission) growth in some of the eastern European countries. In the case of trade restrictions it should become impossible to sell or buy the whole amount of "hot air". Therefore the level of global emissions can be affected by the design of the trade regime.[28] Regarding these interdependencies it is not only the means that vary but also the standard deviations. In figure 2 we assume that free trade still induces the higher level of expected welfare (i.e., efficiency gains from free trade outweigh savings from damage reduction that origin from trade restrictions freezing the amount of tradable "hot air"). That is $\mu_{CC}^{FT} > \mu_{CC}^{LT}$ as given by (7) holds. However, additional to case 1 (see figure 1, above), we further assume that the expected difference in the global emission level causes a difference in the standard deviations (σ_{CC}).[29] In accordance with the global alliance of risk (as introduced in chapter 3) we assume that

$$\sigma_{CC}^{FT} > \sigma_{CC}^{LT} \tag{8}$$

holds. Figure 2 illustrates this case.[30]

If the actual α of each nation equals $\alpha_{\mathrm{I}}^{\min} (\alpha \cong \alpha_{\mathrm{I}}^{\min})$ and $\alpha_{\mathrm{II}FT}^{\min} > \alpha_{\mathrm{I}}^{\min} > \alpha_{\mathrm{II}LT}^{\min}$ holds, now, a no conflict game evolves in the *LT*-case and a chicken game is played in the *FT*-case.[31] The reason is that in the case of free trade the dis-

[28] There may be other elements of the design of the trade regime affecting the level of global emissions. However, to highlight the role of hot air-trading we focus on this issue, here.

[29] Herold 1998 e.g. calculates that "hot air" could affect the global emission level of greenhouse gases by an amount of 516 – 650 mio. tons (only CO_2) during the first budget period (2008 – 2012). For the problem of "hot air" also see Michaelowa / Koch 1999.

[30] If the *FT*-curve is actually steeper than the *LT*-curve, as it is assumed in figure 2 below, can not be answered within this theoretical paper. Here, focusing on a trade regime like it is discussed in connection with the Kyoto protocol is only to illustrate the method applied and does not claim for any empirical validity even if the calculations of Herold 1998 (see footnote 28, above) seem to point in this direction.

[31] Thus, as argued in section 6, above, the choice of instruments is also able to improve the situation by transforming the chicken into a game of a higher cooperation possibility.

advantages of the higher sway of the outcomes may overcompensate the advantages of the higher mean. If this effect enforces the agreement critically depends on the intensity of the countries' actual risk aversion. If the actual α falls short of the thresholds ($\alpha < \alpha^{\min}_{IIFT}$; $\alpha^{\min}_{IILT}$) a ratification of the protocol is not to be expected, irrespective of trade being restricted or not. However, if α exceeds the threshold $\alpha^{\min}_{IILT}$ but not the threshold $\alpha^{\min}_{IIFT}$ ($\alpha^{\min}_{IIFT} > \alpha > \alpha^{\min}_{IILT}$ the spread effect dominates and trade restrictions will be characterised by a higher cooperative push than free trade.

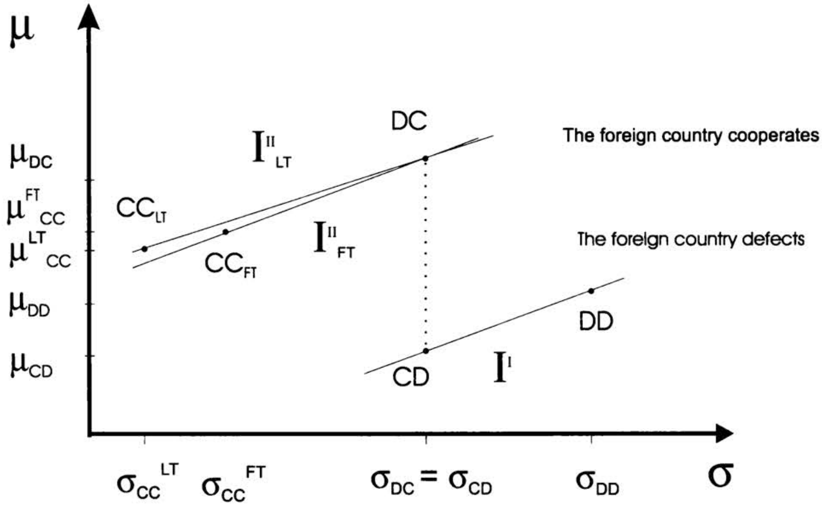

Figure 2: Designing an international environmental treaty (case 2)

In case of the Kyoto protocol it is suggestive that the US are not only less risk averse than the EU ($\alpha^{EU} > \alpha^{US}$) but in addition that their assessment of the risks and costs of global warming differs from that of the EU as well (the stochastic parameters of the model differ, too). Consequently the threshold of self-enforcement could be lower in case of free trade for the US as displayed in figure 1. The risk of "hot air" trading is not expected to alter the spread of the welfare distribution by the US. If, on the other hand, the European countries expect "hot air" trading to affect the spread of the welfare distribution, the threshold of self-enforcement could be lower when trade is restricted as is shown in figure 2. In this case we have: $\alpha^{\min}_{IIFT} < \alpha^{US} < \alpha^{\min}_{IILT}$ and $\alpha^{\min}_{IILT} < \alpha^{\min}_{IIFT} < \alpha^{EU}$. Improving the possibility of cooperation then requires to set the rules of the agreement in accordance to the incentive structure (the prerequisites of cooperation) of the country showing the lowest intensity of risk aversion. The reason is that to induce a country with low risk aversion to cooperate it is necessary to push the threshold in this

DOI https://doi.org/10.3790/schm.121.1.1 | Generated on 2025-08-11 07:39:50

country down as far as possible. Here: Only allowing free trade leads to bilateral cooperation ($\alpha^{EU} > \alpha^{\min}_{IIFT}$ and $\alpha^{US} > \alpha^{\min}_{IIFT}$). The price of enlarging the agreement consequently is that less risk averse countries are able to "dictate" the rules of the game (the mechanism design of the protocol).

Thus, to sum up, the cooperative power of an environmental policy instrument critically depends upon the interaction of the expected stochastic and strategic uncertainties. To make these interdependencies (i.e., the interaction between the choice of instruments / technologies and the possibility of international cooperation) work as a catalyst for stable international coalition formation we recommend to introduce a further judgement criteria for international policy assessment: *"the cooperative push"*. This criterion reflects the ability of instruments (and technologies) to initiate and self-enforce international cooperation to protect the global commons. *"The cooperative push"* provides the necessary (and still missing) link between local and global concerns.[32]

8. Conclusion

Our analysis simultaneously deals with two types of uncertainty: first, with the uncertainty of the behaviour of nature (stochastic uncertainty) and second, with the uncertainty of the behaviour of nations (strategic uncertainty). This *risk-strategic* analysis points out that both, research on national risk preferences as well as scientific uncertainties is needed to predict whether international environmental cooperation takes (should take) place or not. In a static setting we focused on two countries suffering from global environmental risks. We studied how the chances for international cooperation depend on the characteristics of national welfares.[33] To do that, we first developed a profile (in terms of the criteria-specific thresholds $\alpha^{\min}_{I}$ and $\alpha^{\min}_{II}$) to be able to distinguish national risk attitudes according to their cooperative power.

The thresholds are determined by the objective risk pattern (μ and σ). Dependent on the activity chosen by the nations (cooperation versus defection) we assumed the order of expected welfare μ in each home country to follow the incentive structure of a prisoners' dilemma game. Regarding the global

32 Using the analysis of alternative designs of trade mechanisms within the Kyoto protocol as an example, limited trade has the higher cooperative push compared to free trade given the conditions of case 2. The opposite holds if the assessment of risk has to be reflected by case 1 of the analysis above.

33 To model international cooperation we chose a dichotomous choice setting with cooperation versus defection. National welfares have been characterised by the means μ, the standard deviations σ as well as the national attitudes towards risk α.

alliance of risk we assumed σ to decrease with each cooperative contribution. Given these assumptions both thresholds are positive. Thus, only risk aversion is able to foster international environmental cooperation.

If national risk preferences fall short of both thresholds ($\alpha < \alpha_{\mathrm{I}}^{\min}; \alpha_{\mathrm{II}}^{\min}$) internationally coordinated measures are not to be expected. Only, if nations are risk averse and if their actual intensity of national risk aversion exceeds one of the thresholds, a possibility of international cooperation is induced. The step over $\alpha_{\mathrm{I}}^{\min}$ ensures that incentives to cooperate unilaterally exist (condition of incentive compatibility: $\alpha > \alpha_{\mathrm{I}}^{\min}$). Stepping over $\alpha_{\mathrm{II}}^{\min}$ confirms that this incentive is stable (condition of self-enforcement: $\alpha > \alpha_{\mathrm{II}}^{\min}$). In cases where unilateral cooperation is not initiated ($\alpha < \alpha_{\mathrm{I}}^{\min}$) but the condition of self-enforcement is fulfilled ($\alpha > \alpha_{\mathrm{II}}^{\min}$), taking over unilateral actions by one country is able to initiate further cooperative measures in a second country. Conditional cooperation may happen.

Thus the integration of both, stochastic as well as strategic uncertainty also helps to identify the advantages and disadvantages of a "first mover"-policy in international risk management.

Moreover, the *risk-strategic* analysis points out that the traditional criteria of policy assessment (like efficiency and ecological accuracy that initially have been developed for local concerns) should be put into a wider context if used in an international setting:

The thresholds ($\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$) are determined by the stochastic terms of the model (μ and σ) that reflect the shape of the national welfare distributions. In the case of *manmade* environmental risks, μ and σ depend on national / international policy measures that alter the level of damage as well as of damage probability (endogenous risks). Consequently, if cooperative behaviour is not supported by actual risk preferences environmental policy is able to stimulate international cooperation by a shift of the endogenous parameters, here: μ and σ. Varying these terms (the national welfare distribution) by the choice of environmental policy (i.e., by a skilful selection of instruments and technologies) the thresholds, $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$, can be pushed down. Hence, if environmental instruments and technologies can be ordered according to their impact on the stochastic parameters (here: μ and σ), the instrument / technology leading to the lowest demands on the actual intensity of risk attitude able to foster international cooperation (here: the minimum level of $\alpha_{\mathrm{I}}^{\min}$ and $\alpha_{\mathrm{II}}^{\min}$) could be selected. With this choice the chances for effective global risk management are at their best.

It was further shown that in the case of asymmetric nations (according to their intensity of risk aversion) that assess risk differently (the thresholds differ among the nations) the enlargement of environmental cooperation may require to set the rules of the game so as to push the threshold of the country with the lowest risk aversion as far down as possible. Hence, the desire of maximum coalition formation could strengthen the position of countries, here, with low risk aversion in the negotiation process. So they may be able to dictate the rules of the game. These results suggest to supplement the traditional criteria of policy assessment by a further criterion: *"the cooperative push"*. This criterion highlights the fact that national policy measures can be designed to stimulate self-enforcing international coalition formation to reduce global environmental risks. Implementing *"the cooperative push"* may lead to new impulses for the concept of sustainable development claimed by the UN Framework Convention agreed on in Rio de Janeiro 1992 and in many subsequent international declarations.

9. Epilogue

According to chapter 4, it is impossible to establish a general rule of how the standard deviation of environmental damages (σ) varies with the emission behaviour of nations. Therefore, even though we gave some rationale for the ordering assumed in (4), above, it seems to be wise to be prepared for things being different. So let us assume that contrary to what has been said above, the global alliance of risk is characterized by

$$(4') \qquad 0 < \sigma_{DD} < \sigma_{CD} = \sigma_{DC} < \sigma_{CC} \, .$$

Ranking (4′) states that the variability of damages increases with each cooperative contribution. Consequently (see section 5, above), (I) and (II) turn into

$$(\mathrm{I}') \qquad \alpha < (\mu_{DD} - \mu_{CD})/(\sigma_{DD} - \sigma_{CD}) \equiv \alpha_{\mathrm{I}}^{\min} < 0$$

$$(\mathrm{II}') \qquad \alpha < (\mu_{DC} - \mu_{CD})/(\sigma_{DC} - \sigma_{CC}) \equiv \alpha_{\mathrm{II}}^{\min} < 0 \, .$$

Thus, given the reversed order of σ as assumed by (4′) it is only the risk seeking nations which can be expected to cooperate.

Analogously to what has been said in the main body of the paper, it may not be possible to shift the countries' risk preference parameter (α) to meet (I′), (II′). However, also in analogy to what has been said for the case of risk averse countries, above, the critical thresholds ($\alpha_{\mathrm{I}}^{\min}$, $\alpha_{\mathrm{II}}^{\min}$) may be influenced by international risk management in order to make cooperation more likely.

Taking these results for a modified global alliance of risk (4') together with the results of the paper's main body assuming the initial global alliance of risk (4), the following conclusions prevail:

- No matter what form the global alliance of risk takes, (4) or (4′), it is the involved counrties' risk preferences that decide whether the road to cooperation will be taken in a static expected prisoners' dilemma game.
- No matter what form the global alliance of risk takes, it will not be enough to trigger cooperation that countries possess the favourable type of risk preference. It is an additional prerequisite that this preference is strong enough to overcome a certain threshold. There is one threshold for unilateral and one for bilateral cooperation.
- The form of the global alliance of risk decides which kind of risk preference is favourable for cooperation: If variability decreases with emission reduction, risk aversion turns out to be a necessary condition. If, on the other hand, variability increases with emission reduction, risk aversion leads to defection.

Literature

Arrhenius, S. A. (1903): Lehrbuch der kosmischen Physik, Leipzig.

Axelrod, R. (1984): The Evolution of Cooperation, New York.

Barrett, S. (1992): International Environmental Agreements as Games in: Pethig, R. (Ed.): Conflicts and Cooperation in Managing Environmental Resources, Berlin a.o., 11–37.

– (1994): Self-Enforcing International Environmental Agreements, Oxford Economic Papers, 46, 878–894.

– (1997): Towards a Theory of International Environmental Cooperation, in: Carraro, C. / Siniscalco, D. (Ed.), New Directions in the Economic Theory of the Environment, 239–280.

Brachinger, H.W. / *Weber,* M. (1997): Risk as a Primitive: A Survey of Measures of Perceived Risk, OR Spectrum, 4, 235–250.

Brockmann, K.L. / *Stronzik,* M. / *Bergmann,* H. (1999): Emissionsrechtehandel – eine neue Perspektive für die deutsche Klimapolitik nach Kioto, Heidelberg.

Carraro, C. / *Siniscalco,* D. (1997): The International Protection of the Environment: Voluntary Agreements among Sovereign Countries, in: Dasgupta, P. / Mäler, K.-G. / Vercelli, A. (Ed.), The Economics of Transnational Commons, Oxford, 192–209.

Chichilnisky, G. (2000): An Axiomatic Approach to Choice under Uncertainty with Catastrophic Risks, Resource and Energy Economics, 22, 221–231.

Chichilnisky, G. / *Heal,* G. (1998): Global Environmental Risks, in: Chichilnisky, G. / *Heal,* G. / Vercelli, A. (Ed.), Sustainability: Dynamics and Uncertainty, Dordrecht, 23–46.

Edenhofer, O. (1996): Das Management globaler Allmenden, in: Diekmann, A. / Jaeger, C.C. (Ed.), Umweltsoziologie, Sonderheft 36, 390–419.

Eichberger, J. (1993): Game Theory for Economists, San Diego.

Endres, A. (1997): Negotiating a Climate Convention: The Role of Prices and Quantities, International Review of Law and Economics, 17, 147 – 156.

– (2000): Umweltökonomie, Stuttgart.

Endres, A. / *Ohl,* C. (1998): Die Begrenzung globaler Umweltrisiken – eine risiko-strategische Betrachtung, Zeitschrift für Umweltpolitik und Umweltrecht, 4, 511 – 517.

Endres, A. / *Ohl,* C. (2000a): Das Kooperationsverhalten von Staaten bei der Begrenzung globaler Umweltrisiken: Zur Integration stochastischer und strategischer Unsicherheitsaspekte, Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 136(2), 181 – 206.

– (2000b): Taxes versus Quotas: New Insights into an old Affair to Limit Global Environmental Risks, Environmental Economics and Policy Studies, forthcoming.

Finus, M. / *Rundshagen,* B. (1998): Toward a Positive Theory of Coalition Formation and Endogenous Instrumental Choice in Global Pollution Control, Public Choice, 96, 145 – 186.

Hanley, N. / *Spash* C.L. (1993): Cost Benefit Analysis and the Environment, Aldershot.

Heister, J. (1997): Der internationale CO_2-Vertrag: Strategien zur Stabilisierung multilateraler Kooperation zwischen souveränen Staaten, Tübingen.

Heister, J. et al. (1991): Umweltpolitik mit handelbaren Emissionsrechten: Möglichkeiten zur Verringerung der Kohlendioxid- und Stickoxidemissionen, Tübingen.

Herold, A. (1998): Hot air and the Kyoto Protocol, in: WWF Forum on Environment and Development (Ed.), Emission Trading in International Climate Protection, Proceedings, Workshop, June 25 – 26, 1998, Bonn Bad Godesberg, 50 – 52.

Hoel, M. (1991): Global Environmental Problems: The Effects of Unilateral Actions Taken by One Country, Journal of Environmental Economics and Management, 20, 55 – 70.

Houghton, J.T. et al. (1996): Climate Change 1995: The Science of Climate Change, Cambridge, UK.

Huang, C. / *Litzenberger,* R. H. (1988): Foundations for Financial Economics, Amsterdam, New York.

Kolstad, Ch. D. (2000): Environmental Economics, New York, Oxford.

Kreuzberg, P. (1994): CO_2-Emissionen unter Risikoaversion: Vermeidungskosten als Prämie für "Greenhouse Insurance", Wissenschaft und Studium, 9, 463 – 466.

Lipnowski, I. / Maital, S. (1983): Voluntary Provision of a Pure Public Good as a Game of Chicken, The Journal of Public Economics, 20, 381–386.

Luce, R. D. / *Raiffa,* H. (1957): Games and Decisions, New York.

Manne, A. / *Richels,* R. (1999): On Stabilizing CO_2 Concentrations – Cost-Effective Emission Reduction Strategies, in: Carraro, C. (Ed.), International Environmental Agreements on Climate Change, Dordrecht.

Michaelowa, A. / *Koch,* T. (1999): Critical Issues in Current Climate Policy: "Hot Air", Multi-Level Emission Trading Registries and Changes in Emission Commitments Due to International Conflicts, HWWA-Report 194.

Nordhaus, W.D. (1994): Expert Opinion on Climate Change, American Scientist, 82, 45–51.

Oberthür, S. / *Ott,* H.E. (1999): The Kyoto Protocol, Berlin a. o.

Ostrom, E. (1999): Die Verfassung der Allmende: jenseits von Staat und Markt, Tübingen.

Pearce, D. G. (1992): Repeated Games: Cooperation and Rationality, in: Laffont, J.-J. (Ed.), Advances in Economic Theory, Cambridge, 132–174.

Pizer, W.A. (1999): Choosing Price or Quantity Controls for Greenhouse Gases, Climate Issue Brief No. 17, www.rff.org/issue_briefs/pdf_files/ccbrf17.pdf.

Roughgarden, T. / *Schneider,* S. (1999): Climate Change Policy: Quantifying Uncertainties for Damages and Optimal Carbon Taxes – Towards More Comprehensive Calculations, Energy Policy, 27(7), 415–429.

Schelling, T. C. (1960): The Strategy of Conflict, Cambridge.

Siebert, H. (1998): Economics of the Environment: Theory and Policy, Berlin a. o.

Sinn, H.-W. (1983): Economic Decisions under Uncertainty, Amsterdam a. o.

– (1990): Expected Utility, μ-σ Preferences, and Linear Distribution Classes: A Further Result, Journal of Risk and Uncertainty, 3, 277–281.

Stähler, F. (1996): Reflections on Multilateral Environmental Agreements, in: Xepapadeas, A. (Ed.), Economic Policy for the Environment and Natural Resources: Techniques for the Management and Control of Pollution Cheltenham, 174–196.

– (1998a): Economic Games and Strategic Behaviour: Theory and Application, Cheltenham.

Stähler, F. (1998b): On the Economics of International Environmental Agreements, in: Michaelis, P. / Stähler, F. (Ed.), Recent Policy Issues in Environmental and Ressource Economics, Heidelberg, 155–173.

Taylor, M. (1987): The Possibility of Cooperation, Cambridge.

WBGU (2000): World in Transition: Strategies for Managing Global Environmental Risks, Annual Report 1998, Berlin a. o.

Welsch, H. (1995): Greenhouse Gas Abatement under Ambiguity, Energy Economics, 17, 91–100.

Xepapadeas, A. (1997): Advanced Principles in Environmental Policy, Cheltenham.