University of Southampton

Faculty of Social Sciences

Externalities and Agreements in International Rivers

Karima Mohamed Magdy Ahmed Kamal

Submitted for the Degree of Doctor of Philosophy

Department of Economics

July 2003

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCES

Doctor of Philosophy

EXTERNALITIES AND AGREEMENTS IN INTERNATIONAL RIVERS

by Karima Kamal

The thesis investigates the unidirectional impact of an upstream country's activities on the water quantity/quality of an international river in a downstream country, and the effect of this unidirectional externality on reaching self-enforcing water agreements.

The first part of the thesis develops a dynamic game of incomplete information to model strategic behaviour in reaching agreements on sharing the waters of the River Nile, when the actual power of downstream Egypt is unknown to upstream Ethiopia. The main results of the analysis are that beliefs do matter, that being perceived as strong is more important than actually being strong to avoid conflict, and that more accurate information can prevent war although it cannot prevent the conflict.

The second part of the thesis considers a two-player, multi-period dynamic game with complete but imperfect information, to model the effect of asymmetric irreversibility of the actions of two countries sharing a river on the feasibility of reaching a self-enforcing environmental agreement. The main results of the analysis are that full co-operation cannot be achieved if at least one country's actions are irreversible, that higher levels of partial co-operation can be sustained if both countries' actions are irreversible compared to the case where only one country's actions are irreversible, the type of co-operation, whether *gradual* or *immediate* has only an effect if both countries' actions are irreversible.

The final part of the thesis examines two reduced form relationships between various river pollution indicators and economic development indicators for domestic and international rivers. The objective is to determine whether the income-environmental quality relationship is affected by the nature of the river and to determine the effect of ignoring inter-country pollution on the intra-country turning points of the Environmental Kuznets curves. The main results indicate that national income is an important determinant for river pollution, especially for domestic rivers, that upstream effects and especially income related upstream effects, seem to play an important role in river pollution, that all pollutants peak at a lower level of income in domestic rivers than in international rivers, that most pollutants peak a lower level of income when upstream effects are considered.

Table of Contents

Chapter 1 Introduction..1

Chapter 2 Strategic Behaviour in Reaching Agreements on Sharing Waters of International Rivers....................................17

2.1. Introduction..17

2.2. Water in the International World..20

2.2.1. Water as a Natural Resource..20

2.2.2. Water and Water Development Projects..21

2.2.3. Water and National Security in Terms of Food Self Sufficiency..............23

2.2.4. Water and Power Relations..24

2.2.5. Water Disputes and International Law..25

2.2.6. Water Disputes in the Economic Literature..28

2.3. The River Nile Case..30

2.3.1. Ethiopia's Position..33

2.3.2. The Difficult Position of Sudan..33

2.3.3. Egypt's Position..34

2.4. The Model..35

2.4.1. Description of the Model..36

2.4.2. The Parameters, Assumptions and Pure Strategies..............................38

2.4.3. The Extensive Form of the Signalling Game..42

2.4.4. Analysis of the Strategic Form of the Signalling Game..............................43

2.5. Results..52

2.6. Summary and Conclusion..54

Appendix 2.1. The River Nile Case in More Detail..57

Chapter 3 International Environmental Agreements in Water Resources with Unidirectional Externalities....................60

3.1. Introduction..60

3.2. Reversibility Constraints and the Incentive Structure of a Game..........................62

3.2.1. The Classical Prisoners' Dilemma Game (The PD Game)....................63
3.2.2. The Lockwood and Thomas Game (The LT Game)..............................64
3.2.3. The Asymmetric Irreversibility Game (The AI Game)..........................65
3.3. Related Literature..66
3.4. The Model...68
3.4.1. Description of the Model..68
3.4.2. The Actions, Parameters, Assumptions and Pure Strategies....................72
3.5. The Analysis of the Model..77
3.5.1. The Equilibrium Conditions for the Games..78
3.5.1.1. The Equilibrium Conditions for PD Game..............................78
3.5.1.2. The Equilibrium Conditions for the LT Game..........................84
3.5.1.3. The Equilibrium Conditions for the AI Game...........................92
3.5.2. The Necessary Punishment (an Illustrative Example)..........................99
3.5.2.1. Necessary Punishment Across the Games..............................105
3.5.2.2. Necessary Punishment Across Levels of Cooperation..............105
3.5.3. The Interpretation of Credible Punishment in the Games....................108
3.5.3.1. Credible Punishment in the PD Game......................................108
3.5.3.2. Credible Punishment in the LT Game.......................................109
3.5.3.3. Credible Punishment in the AI Game..110
3.5.3.4. Credible Punishment Across the Games..................................111
3.5.3.5. Credible Punishment Across Levels of Cooperation................111
3.5.4. Sustainable Cooperation in the PD, LT and AI Games.........................112
3.6. Results...112
3.7. Summary and Conclusion..116

Chapter 4 Water Quality in Domestic and International Rivers.......119

4.1. Introduction..119
4.2. Literature Review...122
4.2.1. Basic Environmental Kuznets Curve..123
4.2.2. Expanded Environmental Kuznets Curve...128
4.2.3. Alternative Explanations for the Environmental Kuznets Curve.............132

4.3. The Data..138
4.3.1. Water Quality Indicators..138
4.3.1.1. The Oxygen Balance in Rivers and Nitrates..........................138
4.3.1.2. The Pathogenic Contamination..139
4.3.1.3. Heavy Metal Contamination..140
4.3.2. Income Indicators..143
4.3.3. Externality Indicators...144
4.4. Methodology...144
4.5. Results..149
4.6. Summary and Conclusion..155
Appendices...158
Appendix 4.1. Random Effects ML Estimation for Water Quality Indicators..............158
Appendix 4.2. List of International Rivers..173

Chapter 5 Concluding Remarks...175

References...179

List of Tables

2.1. The effect of an increase in parameters on the likelihood of possible outcomes....54

3.1. f^* and f^{**} across the games and across the levels of cooperation.................107

3.2. *Credible* and *necessary* punishment...113

3.3. Equilibrium conditions for the game, levels and types of cooperation..............114

4.1. Statistics summary of the dependent variables..141

4.2. Significance ($prob > \chi^2$) for Income-related Effects of the Oxygen Balance and Nitrates Contamination in Rivers..150

4.3. Significance ($prob > \chi^2$) for Income-related Effects of Faecal Contamination and Heavy Metals Contamination in Rivers..151

4.4. Estimated Critical per Capita Incomes at Peak Pollution Levels.....................154

4.5. Random-effects ML Estimation for Dissolved Oxygen...................................159

4.6. Random-effects ML Estimation for Biological Oxygen Demand.......................161

4.7. Random-effects ML Estimation for Chemical Oxygen Demand........................163

4.8. Random-effects ML Estimation for Nitrates..165

4.9. Random-effects ML Estimation for Faecal Coliform.....................................167

4.10. Random-effects ML Estimation for Cadmium..169

4.11. Random-effects ML Estimation for Lead..171

List of Figures

2.1. The extensive form of the signalling game 38

2.2. The extensive form of the signalling game with payoffs 42

3.1. The payoffs in period t=1 74

3.2. The PD^{∞} game in period $t=1$ and in period $t \geq 2$ 79

3.3. The PD_{∞} game in period t=1 and in period $t \geq 2$ 80

3.4. The PD^2 game in period $t=1$ and in period $t \geq 2$ 81

3.5. The PD_2 game in period $t=1$ and in period $t \geq 2$ 83

3.6. The PD^1 game in period $t=1$ and in period $t \geq 2$ 84

3.7. The LT^{∞} game in period $t=1$ 85

3.8. The LT^{∞} game in period $t \geq 2$ after defection in period $t=1$ 85

3.9. The LT^{∞} game in period $t \geq 2$ without defection in period $t=1$ 85

3.10. The LT_{∞} game in period $t=3$ 87

3.11. The LT_{∞} game in period $t \geq 4$ after defection on period $t=3$ 87

3.12. The LT_{∞} game in period $t \geq 4$ without defection in period $t=3$ 87

3.13. The LT^2 game in period $t=1$ 88

3.14. The LT^2 game in period $t \geq 2$ after defection in period $t=1$ 88

3.15. The LT^2 game in period $t \geq 2$ without defection in period $t=1$ 89

3.16. The LT_2 game in period $t=2$ 90

3.17. The LT_2 game in period $t \geq 3$ after defection in period $t=2$ 90

3.18. The LT_2 game in period $t \geq 3$ without defection on period $t=2$ 90

3.19. The LT^1 game in period $t=1$ 91

3.20. The LT^1 game in period $t \geq 2$ after defection in period $t=1$ 92

3.21. The LT^1 game in period $t \geq 2$ without defection in period $t=1$ 92

3.22. The AI^{∞} game in period $t=1$ 93

3.23. The AI^{∞} game in period $t \geq 2$ after or without defection in period $t=1$ 93

3.24. The AI_{∞} game in period $t=3$ 95

3.25. The AI_{∞} game in period $t \geq 4$ after or without defection in period $t = 3$ 95

3.26. The AI^2 game in period $t = 1$ 96

3.27. The AI^2 game in period $t \geq 2$ after or without defection in period $t = 1$ 96

3.28. The AI_2 game in period $t = 2$ 97

3.29. The AI_2 game in period $t \geq 3$ after or without defection in period $t = 2$ 97

3.30. The AI_2 game in period $t \geq 3$ after or without defection in period $t = 2$ 98

3.31. The AI^1 game in period $t \geq 2$ after or without defection in period $t = 1$ 99

Acknowledgements

I am extremely grateful to my supervisor Robin Mason, whose valuable comments and suggestions have been very much appreciated throughout my studies. I am also very grateful to my supervisor Alan Ingham for his support and encouragement during the first part of my thesis. I would also like to thank my advisor Peter Smith, the Head of Department Grayham Mizon and Jackie Wahba for their advice, support and encouragement.

My thanks also go to the administrative support staff of the Department of Economics and in particular to Gill Crumplin, Caire Caffrey and Fatima Fauzel for always having a smile on their faces and for making the department a pleasant place to work in.

I also wish to thank many of my fellow PhD students and in particular Zheng Wang, Manpreet Sidhu, Grazia Rapisarda, Mokhrazinim Mokhtar, Sugiharso Safuan, Manfred Eckert and Emmanuel Mulenga Pamu for many helpful comments and invaluable interesting discussions about various aspects of life from a multi-cultural perspective.

I am also deeply grateful to the Egyptian government for providing me with a scholarship and financial support throughout my PhD studies at the University of Southampton.

My deepest thanks however go to my husband Tarek and my sons Youssef and Farid for being patient and very supportive to me, especially during those periods when my confidence was low. They have always believed in me and have relentlessly surrounded me with their constant love and care. To them, to my mother and to my recently deceased father this work is dedicated.

Chapter I

Introduction

> *"...There is no economic incentive for cooperation when an upstream country uses an international river to the detriment of the downstream country and that country has no reciprocal power over the upstream country..." LeMarquand (1978, p.151)*

The thesis investigates the unidirectional impact of an upstream country's activities on the water quality/quantity of an international river in a downstream country, and the effect of this unidirectional externality on reaching self-enforcing water agreements. A country that depends heavily on river water for its mere existence, feels tremendously threatened by the possibility of having its water supplies severely constrained by an upstream neighbouring country. Therefore, on the one hand, some pessimists believe that regional water scarcity that is exacerbated by a high dependency of downstream countries on international rivers is likely to spark international tensions and possibly war. On the other hand, many optimists believe that water disputes are very likely to lead to negotiations, non-violent solutions and possibly cooperation. Our objective in chapter 2 and chapter 3 is therefore to demonstrate under which conditions sustainable cooperation can be achieved. Although the quality impact seems less pressing in terms of violent conflicts compared to the quantity aspect, it is equally important to get a comprehensive picture of the overall unidirectional externality effect in international rivers. Chapter 4 is devoted to detect this unidirectional quality impact of pollutants discharged by upstream countries and borne mainly by downstream countries.

Population growth, industrial development and irrigation expansion during the past century has put enormous pressure on both the quantity and the quality of available freshwater resources. Gleick (1998) estimates that freshwater withdrawals have increased sevenfold since 1900[1] and that current and future water scarcity in many arid and semi-arid regions will provide conditions that can lead to regional tensions[2], especially if a water resource has to be shared by two or more countries. In the past the development of water sources in most countries was confined to their national water sources, because there were rarely adequate water agreements between riparians on developing international water resources. Biswas (1999) anticipates that as water scarcity in many countries is becoming more severe, countries are forced to consider the development of their international water sources, which will become the only remaining water sources that can be developed economically. While international fresh water sources include rivers, lakes and aquifers, the focus of the thesis is on rivers only, because the no-impact of the downstream countries' actions on the upstream countries' productivity and welfare is considered the main obstacle to cooperation. Therefore, rivers are a special case in terms of the externality impact compared to other international water sources such as lakes and aquifers, as we discuss later.

Wolf *et al.* (1999) provide the most comprehensive register of international rivers. It counts 261 international river basins covering 45.3% of the earth's land surface distributed among the continents as follows: 60 in Africa, 53 in Asia, 71 in Europe, 39 in North America and 38 in South America. While most of these international river basins are shared by only two countries, 19 river basins are shared by more than five countries. This makes many countries highly dependent on water that originates outside their borders. According to Barrett (1994), only six out of 31 of countries that are highly dependent on water from neighbouring countries are high-income countries, the rest are low or middle-income countries. For example, more than 90% of water sources in Egypt, Hungary, Mauritania, Botswana and Bulgaria originate outside their borders. Over the years stronger riparians have managed to use as much water as possible from a shared water source thereby establishing their historical water rights. Weaker riparians, which have started their

[1] World population has increased from 1,600 million in 1900 to nearly 6,000 million in 1995; world wide irrigated land has increased from 50 million hectares to over 250 million hectares for the same period.
[2] For an excellent chronology of conflict over water in the legends, myths, and history of the ancient Middle East and from 1500 to the present, see Gleick (1998).

development much later, often need a larger share of the international resource to achieve their development goals.

Biswas (1999) considers the biggest challenge facing countries sharing a water source as follows:

> *"...how to develop and manage the various international water sources sustainably and efficiently in full agreement and with cooperation between the co-basins so that the result could be a 'win-win' situation for all the parties concerned..." (p.433)*

However, international agreements have not only to be ratified but most importantly they have to be binding on all sovereign countries involved. Barrett (1994) thinks of this as the main obstacle to international agreements.

> *"...In an intranational dispute, parties which freely commit to an agreement can be made to comply with the terms of the agreement by the courts ... In an international dispute, however, agreements between countries cannot be enforced by a third party. International agreements must be self-enforcing. Self-enforcement is a severe constraint, and may mean that international water resources potentially cannot be managed as efficiently as intranational resources..." (p.6)*

The interdependence among countries sharing a water source differs according to the type of externality[3] each country exhibits. Reciprocal externalities occur when countries use the resource simultaneously and access it without restrictions. In this case each country's action has an impact on the shared resource and each country is affected by the actions of all other countries. These reciprocal externalities can be global, as for example the emission of global pollutants, or regional, as for example the simultaneous use of a lake or an aquifer. According to Barrett (1990), each country has some incentives to eliminate

[3] Whenever natural resources are shared by two or more countries, each country's action has an impact on the productivity or welfare of the other countries. This impact is called externality and varies from no impact to full impact.

negative externalities in the global externality case, even without any cooperation. If one country reduces its emissions unilaterally, it will benefit from the improved environmental quality, unless other countries increase their emission in a way that totally offsets the one country's efforts. The regional reciprocal externality case is somewhat different and often resembles the prisoners' dilemma. According to Mäler (1990), if many countries have simultaneous access to a regional resource, such as a lake, it tends to be overused. Each country continues to use the shared resource as long as it is for its own benefit without regard for the consequences on the other countries. The incentive to cooperate is impeded by the incentive to 'free ride' on the other countries' efforts. Unidirectional externalities occur, when countries access the shared resource consecutively, and not simultaneously, as for example in the case of international rivers. Most often[4] only the upstream country has a positive and/or negative effect on the quantity and/or quality of the downstream country. On the one hand, hydropower production helps to regulate river flows, flood storage provides flood protection downstream, navigational uses and ecological maintenance keep adequate water in the river. On the other hand, irrigation, industrial and municipal diversions remove water from the river system, and wastewater treatment, urban development and agriculture increase river pollution. Under these conditions, upstream countries have no incentive to eliminate negative externalities on the downstream countries, unless they are compensated in one way or the other. The 'free rider' problem mentioned in connection with regional reciprocal externalities, is also relevant for unidirectional externalities whenever there is more than one country that suffers from an upstream country.

What has happened in April 1995 on the Mekong River is a good example of the influence that an upstream country has in an international river. Biswas (1999) has reported that after Cambodia, Lao PDR, Thailand and Vietnam have signed an agreement on cooperation for the sustainable development of the lower Mekong River in 1995, the representatives of the signatory countries decided to take a boat trip on the Mekong River. Unfortunately their boat became stuck because China, the most upstream country of the Mekong River, which

[4]In rare cases, downstream countries can also harm upstream countries. The construction of dams for irrigation, hydropower production or flood control might cause flooding upstream.

is not part of the agreement, was filling up a new major reservoir. Rogers (1993) has indicated that without an agreement[5] a downstream country cannot influence the actions of an upstream country and needs to balance this unidirectional externality with the use of other resources such as economic or military power.

Starting from the above considerations, which describe a standard common resource access problem, this thesis concentrates on negative unidirectional externalities from upstream countries to downstream countries in international rivers. We deal with both the quantity and the quality aspects of the upstream effects. Our purpose is to determine what conditions facilitate self-enforcing water sharing agreements despite the lack of enforcement power on the side of the downstream countries. Is full cooperation possible? Is partial cooperation a satisfactory outcome? We also want to know more about water pollution in international rivers that is caused by pollutants discharged by upstream countries and borne mainly by downstream countries. What effect has ignoring inter-country pollution on intra-country income turning points, i.e., do pollutants in domestic rivers reach their maximum level at a lower level of income than pollutants in international rivers due to the upstream effect? Is the upstream effect income or non-income related? But before turning to the main points in our thesis we provide a brief review of the literature on the potential of reaching water agreements over international rivers[6]

The literature on reaching agreements on sharing waters of international rivers can be largely divided into three broad categories: studies focusing on the market approach in determining the optimal level of the externality; studies applying cooperative game theory; and others applying non-cooperative game theory.

Some economic studies that have focused on determining the efficient water quantity to be transferred between two countries sharing a river basin in a bilateral monopoly setting,

[5] A searchable database of more than 400 international, freshwater-related agreements, covering the years 1820 to 2002 are available via http://www.transboundarywaters.orst.edu/

[6] While this literature review concentrates on water agreements, the literature review on river pollution in connection with the Environmental Kuznets curve is reviewed in chapter 4.

consistent with the Coase Theorem. If the cost and benefit functions of both countries are known, the original distribution of property rights is well defined, there are no transaction costs, the externality problem can be handled in isolation from other international relations, and there are no income effects, then the Coase Theorem predicts a unique level of the externality, that is independent of the assignment of property rights. Although the initial property rights do not determine the unique amount of the externality, they determine the set of equilibria, where both countries do at least as well as at their initial property rights. While all equilibria are Pareto efficient and have the same amount of the externality, the ultimate outcome depends on the bargaining power of the two countries. An example provided by Mäler (1990) states that if two countries pollute a shared river, the Coase Theorem guarantees that an agreement will be reached by which the upstream country reduces its pollution discharges to an optimal level. If the upstream country has the initial right to pollute, it will be compensated by the downstream country for reducing its pollution to the optimal level. If the downstream country has the initial right to clean water, it will be compensated by the upstream country for the optimal level of pollution that is still discharged into the river by the upstream country. Because property rights of shared rivers are not well defined, countries are unable to agree on one of the many equilibria. Lekakis (1998) has suggested that a water market, where countries bargain over the price to pay for the quantity of water transferred, offers a solution to the problem, provided that property rights are well defined and the allocation of joint benefits is considered fair by both countries. Frisvold *et al.* (1995) use the Nash Bargaining Solution to determine the potential benefits from water transfer between an upstream seller and a downstream buyer in a bilateral monopoly setting. Their bargaining model includes water quality as well as water quantity. Kilgour, and Dinar (1995) have also determined the efficient water quantity to be transferred, without handling the problem of payments, which were left rather undetermined on the contract curve.

The main criticisms of the market approach are not only the multiple equilibria, but also that some of the underlying assumptions and requirements are difficult to fulfil. According to Mäler (1990), enforcing and monitoring agreements is rather costly, information related to the cost and benefit functions is rarely complete, and assuming independence between the water and other relations is generally unrealistic. Another main obstacle to adopting the market approach in dealing with this kind of water problems, is that property rights of international water sources are not well defined. There usually exists an informal agreement, which is not equally accepted by all riparian countries and is therefore enforced by economic or military superiority by one of the countries. This means that some countries are not willing to accept the market solution at all, which usually means that they have to pay for something they think they are entitled to.

Other economists such as Dinar and Wolf (1994) and Rogers (1969; 1993) have used Cooperative Game Theory to solve the water-sharing problem of international rivers. Although cooperative games can provide unique and efficient solutions in situations where the number of countries is relatively small, its main obstacle is the need for a cooperative infrastructure i.e. some form of mechanism that ensures that agreements are binding and are being enforced[7]. Dinar and Wolf (1994) have used this approach to demonstrate that water trade among potential water users in the Middle East, especially between Egypt and Israel as the main participants, can increase regional welfare substantially. But even based on economic aspects alone, cooperation is very unlikely because of the distributional inequality of the potential gains. Adding political aspects reinforces the previous result and adds to the obstacles of cooperation. The analysis treats economic and political aspects separately and does not provide an integrated model where both aspects are treated simultaneously. Rogers (1969) has used the concepts of cooperative game theory to investigate many different strategies of cooperation between India and former East Pakistan (Bangladesh since 1972) with regard to the lower Ganges and the Brahmaputra Rivers, whose floods are causing a lot of damage to crops, property and human lives in former East Pakistan every year. Later on Rogers (1993) has updated his research on river water disputes and has included Nepal as third country alongside with India and Bangladesh. He points out the importance of political aspects, especially sovereignty, which is considered of utmost importance in choosing the location of any water-related investments. He argues that if utility is not transferable and side-payments are not

[7] Dasgupta (1993) argues that credible threats by one party to force the other party to fulfil the agreement can be included in the notion of a cooperative infrastructure.

acceptable, it is worthwhile to consider planning the river basin as a whole, on condition that the investment resources are used only within the countries of origin. Although this may lead to a substantial lower net benefit, compared to the full cooperative solution, the net benefit may still be significantly higher than the total non-cooperative solution, where each country acts unilaterally, provided that the cooperative game without side-payments has a solution in the first place. Bennett *et al.* (1998) argue that multi-good bargaining is a way to deal with the problem of side-payments, which are usually not only opposed by the country receiving them and the country paying them, but also by the international community, which argues against the 'victim pays' principle in many environmental affairs. Therefore, the importance of national sovereignty, non-transferable utility, and the lack of a cooperative infrastructure at the international level to enforce agreements, are major obstacles to applying cooperative game theory to international environmental problems.

The third category of economic studies has looked into the problem of sharing the waters of international rivers using Non-cooperative Game Theory. In a two-country game with reciprocal externalities it is very likely that the standard Prisoner's Dilemma applies if at least one of the two countries has a strictly dominant strategy not to cooperate. Each country's best response will be not to cooperate as well, leading to a Pareto inferior Nash equilibrium. In some cases where the externalities are unidirectional, the Nash equilibrium can be that one of the countries bears all the cost of a water project, whose benefits accrue mostly to the other country. Barrett (1994) has provided an example for this based on the Columbia River case. Developing the Columbia on the Canadian side would generate a lot of benefits to Canada, and even more benefits to the downstream United States. Although the Nash equilibrium outcome was Pareto efficient, it was nevertheless difficult to accept by Canada, which would bear all the costs of the water project but get a far smaller share of the benefits than the USA. It was therefore reasonable to expect that Canada would very much prefer to get some side payments from the USA, the main beneficiary. But there was no reason for the USA to enter the bargaining process unless there were either other matters linked to the water project, or Canada was able to commit itself to a credible threat in case that the negotiations did not reach an acceptable outcome. The gain from a negotiated outcome was calculated relative to the disagreement point using the Nash bargaining solution assuming once again that utility is transferable. Transforming the disagreement game into a threat game does not alter the net benefits from cooperation, only the distribution of the benefits, as long as such credible threats do not use up too many resources and thereby diminish the potential gains from cooperation. The Nash equilibrium of the pre-negotiated outcome has been for Canada to build a water storage project on its side of the border with the USA receiving a large share of the benefits. Because the USA believed that Canada would develop the Columbia River on its side of the borders anyway, it did not even consider making a side payment to Canada. Only when Canada threatened to abandon this project in favour of another project on another river, did the USA have to choose between a negotiated settlement with side payments to Canada to induce it to go ahead with the initial project, or to build its own facilities on its side of the borders, which would have been less efficient. Because the USA believed that the Canadian threat was credible, Canada could secure itself a larger share of the benefits after the negotiated settlement in the 1961 treaty. The threat game is credible because the upstream country can affect the downstream country both positively and negatively.

Waterbury (1994) has pointed out that in many actual reached agreements, water alone was rarely the only issue. Widening the bargaining process to include other goods or services of interest to both countries, preferably where the externalities or benefits run in the opposite direction to the water project, helps to make the distribution of benefits from cooperation in water related projects less unequal. This would especially attract countries, which are not willing to accept side payments as compensation. Examples include Iraqi oil for Turkish water, Syrian control of Kurdish rebels in partial exchange for Euphrates water, Israeli know-how in water conservation for a part of the saved water, Egypt's help in raising external assistance for all states of the Nile in exchange for a binding understanding of Egypt's needs for water. Bennett *et al.* (1998) show that connecting two games within an infinitely repeated framework, has a better outcome than the sum of the outcomes of the separate games. The authors provide two examples of water sharing problems using hypothetical values for gains and losses. In both cases they achieve a more satisfactory outcome after connecting the water-sharing problem with unsatisfactory Nash equilibria to other externality problems with unsatisfactory Nash equilibria. The success of this approach depends on the hypothetical values chosen for the connected problems, which were nearly equivalent in magnitude though opposite in the effect. Parametrizing the gains and losses, using realistic values for the water problems and the connected other problems, is necessary to shed more light on the usefulness and applicability of this approach.

To conclude this brief review we realize that undefined property rights of international water sources and lack of an international authority to ensure that agreements are binding to all parties, leaves us with the Non-cooperative Game Theory as the most suitable approach for modelling self-enforcing water agreements.

As mentioned before, this thesis concentrates on the negative unidirectional externalities from upstream countries to downstream countries in international rivers. Both the quantity and the quality aspects of the upstream effects will be treated. In chapter 2 and 3 we apply the non-cooperative game theory approach to model the effect of the unidirectional externality, which is specific to international rivers, on the feasibility of reaching self-enforcing environmental agreements. We limit our attention to a two-country setting, where only one upstream country causes the externality and only one downstream country is exposed to this externality, in order to concentrate fully on the externality problem. As mentioned by Mäler (1990), including more than one suffering downstream country adds the incentive to 'free ride' on the efforts of the others, and including more than one externality-causing country adds the problem of how to distribute the abatement cost among them. In chapter 4 we investigate the effect of inter-country water pollution on the alleged income-environmental quality relationship of each country.

The purpose of chapter 2 is to model strategic behaviour in reaching agreements on sharing the waters of the River Nile between Egypt and Ethiopia. The focal point of the model is the incomplete information that one country has about the other countries' military strength. This rather realistic assumption of incomplete information means that the Coase Theorem, in contrast to Non-cooperative game theory, cannot be applied here. After outlining the role that freshwater plays in the international world in general, we discuss the specific case of the River Nile. Ethiopia, the upstream country, where more than 85% of the Nile waters originate has no legal share of the waters of the Nile. Egypt, the downstream country, which was able to secure the biggest share of the Nile waters under the 1959 Nile Agreement, watches vigilantly over any water development projects in Ethiopia that could affect Egypt's historical water rights. As the demand for water in both countries is already close to the limit of supply and is expected to increase in the near future, some researchers such as Guariso and Whittington (1987), Jovanovic (1985), and Whittington and McClelland (1992) have taken great interest in an American study, which was conducted by the US Bureau of Reclamation in 1964. According to this study, reducing the huge evaporation losses in Egypt's water storage reservoir by building a new water storage reservoir in Ethiopia, where evaporation losses are much less, will guarantee a larger water share for Ethiopia without affecting Egypt's share. The main obstacle to this project is Egypt's refusal to give up the control over the River Nile to Ethiopia, which can use the large storage facilities upstream to block the Nile flow for political reasons.

Drawing on the idea of the American study and using a similar framework as Güner (1998), we apply a simple signalling game, i.e., a dynamic game of incomplete information, to demonstrate that cooperation between Egypt and Ethiopia over the development of the River Nile is possible. Our model captures the interdependence of Egypt's and Ethiopia's strategies with regard to the available information on military power. It turns out that the beliefs, which Ethiopia has about Egypt's military strength, is crucial for the ultimate outcome. It not only affects Ethiopia's decision whether to cooperate or not to cooperate, but more importantly it also affects Egypt's decision in the first place whether to cooperate with Ethiopia or not. This is in contrast to the results obtained by Güner (1998) from his signalling game between Turkey and Syria. In the specific case of the Euphrates River, beliefs do not matter in the strategic behaviour of Turkey, because Turkey does not agree to any water treaty whatever Syria's beliefs about Turkey's strength are. Other important results of our analysis are, that conflict between Egypt and Ethiopia can be avoided if Ethiopia believes Egypt to be strong enough. Being perceived as strong is even more important for Egypt when its vulnerability is increased, as is expected to happen after an agreement is reached. An interesting result is that although more accurate information can prevent war between the two countries, it does not help to put the conflict between them to an end. While many economic studies on the River Nile conflict have concentrated on the significant gains of cooperation, which should motivate downstream Egypt to accept upstream development in Ethiopia, our study contributes to the economic literature by providing an analysis that takes strategic considerations and beliefs explicitly into account. Another contribution of our model is to the applied literature. A comparable signalling game has been applied to the Euphrates River, with the result that beliefs do not matter and do not affect the behaviour of Turkey, the upstream country.

While the signalling game in chapter 2 deals with reaching an agreement over the River Nile, our dynamic game in chapter 3 is about reaching agreements over the development of international rivers in general. This time the driving force that motivates an agreement is the prospect of increasing cooperation in the future. Lockwood and Thomas (2002) have used an infinite dynamic game to prove that the irreversibility of the players' actions[8] prohibits the players from reaching sustainable full cooperation. If two countries, for example, agree to reduce their fishing in a shared lake to prevent overfishing, they can do so by reducing the fishing trips for each boat, or by destroying some of their fishing boats. While the first option is reversible by simply increasing the fishing trips again, the second option is irreversible and new fishing boats need to be bought if one country intends to defect on the agreement. We transform the analysis into a finite framework and extend it to consider the unidirectional externality in international rivers, which we express as an asymmetry in the irreversibility of the countries' actions. While the upstream country's actions are assumed to be reversible at any time, the downstream country's actions are assumed to be irreversible. The asymmetric irreversible case is still rather unexplored despite its important implications for International Environmental Agreements (IEAs) for countries sharing a river.

Our purpose is to demonstrate that different levels of cooperation can be supported by different irreversibility constraints of the players' actions. Therefore, we will compare the

[8] Players can only increase their level of cooperation or keep it constant, but cannot reduce it.

sustainable levels of cooperation for three different reversibility constraints on the players' actions in a two-player setting: symmetric reversibility (as in the repeated Prisoners' Dilemma case), symmetric irreversibility, as in the model analysed by Lockwood and Thomas (2002), and finally the asymmetric irreversibility which is the main theme in chapter 3.

The main results of the analysis are: (1) full co-operation can only be sustained if both players' actions are reversible; (2) higher levels of partial co-operation can be sustained under symmetric irreversibility of the players' actions than under asymmetric irreversibility of the players' actions; (3) while *gradual* co-operation improves the situation for the symmetric irreversibility case compared to *immediate* co-operation, the type of co-operation has no effect on the other cases (4) a higher discount (i.e., a lower discount rate) factor is conducive to co-operation in all three base cases, for all levels and for types of co-operation. Our results confirm the general intuition, that despite obvious gains from cooperation, asymmetric irreversibility of the players' actions makes reaching even a partial agreement between countries sharing a water body with unidirectional externalities, such as a river, more difficult than reaching an agreement between countries sharing a water body with reciprocal externalities, such as a lake. Interpreting the unidirectional externality in the context of irreversible actions has not been covered in the literature before as far as I know and has proved to be one way of showing how international water agreements might be achieved. The approach to adapt an infinite dynamic game into a finite dynamic game allows tractable analysis of a very complicated, but highly relevant extension of the Lockwood and Thomas (2002) model.

In chapter 4 we concentrate solely on the quality aspect of the unidirectional externality of international rivers. Many researchers such as Cole *et al.* (1997), Grossman and Krueger (1995), Holtz-Eakin and Selden (1995), Mason and Swanson (2003), Roberts and Grimes (1997), Selden and Song (1994), Shafik (1994), Stern *et al.* (1996) have taken part in the ongoing debate about the inverted U-shaped relationship between environmental deterioration and income, usually called the 'Environmental Kuznets curve' or simply

EKC[9]. They do not only find empirical evidence for the existence of the EKC for at least some pollutants, but do also confirm an observed pattern for air quality indicators. It has been noticed that air pollutants with direct or local impacts have turning points at lower levels of income than air pollutants with more indirect, regional or global impacts. While this different behaviour of local and global air pollutants has been widely studied, the same has not been done for water pollutants. Dependent on the nature of the river, the water pollutant is local if the river is a domestic river and is regional if the river is an international river. The question that emerges is: how to detect the unidirectional externality effect with pollutants that are discharged by upstream countries and borne mainly by the downstream countries? To answer this question we first need to investigate whether the alleged income-environmental quality relationship, known as 'Environmental Kuznets curve', is affected by the nature of the river, i.e., whether the river is domestic or international. This gives an indication of a possible upstream effect in international rivers. Therefore, we examine a basic reduced form relationship, based on the approach used by Grossman and Krueger (1995) between various water quality indicators and economic development indicators for domestic and international rivers. Our second purpose is to determine whether this externality effect is income related or not. Conte Grand (1999) has introduced the externality effect of international lakes and rivers, via the existence or non-existence of water treaties. Sigman (2002) has presented the externality effect of international rivers through dummy variables, which specify whether a water quality monitoring station is an upstream, a downstream or a border station. In both studies an externality effect has been detected but is not related to income. As there is little doubt in the economic literature that income plays an important role in the income-environmental quality relationship, we expand the previous reduced form relationship and include the per capita income of the upstream country as a new explanatory variable. Then we estimate the critical levels of income at which each water quality indicator is expected to reach its maximum or minimum to facilitate the comparison among the different types of pollutants and between the two regression models.

[9] The EKC hypothesis states that environmental degradation increases first with income at low levels of income, then decreases later with income at high levels of income.

The main results of the analysis are: (1) national income is an important determinant for river pollution, especially for domestic rivers, (2) the Upstream-downstream model is better suited for international rivers than the Basic model, because it takes upstream effects explicitly into account, (3) upstream effects seem to play an important role in river pollution, (4) income related upstream effects seem to play a more important role in river pollution than non-income related upstream effects, (5) all water pollutants peak at lower level of income in domestic than in international rivers, and (6) most of the water quality indicators in international rivers peak at a lower level of income when upstream effects are considered. The distinction between local water pollutants and regional water pollutants through the nature of the river has not been covered in the literature before as far as I know. A second contribution is the detection of a possible income related externality effect from upstream countries to downstream countries in an attempt to separate income related and non-income related externality effects. A third contribution is that we confirm the observed pattern that local air pollutants tend to peak at lower income levels than regional or global pollutants, for water pollutants. This has not been done for water quality indicators as far as I know.

Overall, the thesis highlights the specific difficulties caused by the unidirectional externalities in international rivers. It covers both the quantity aspects of the upstream effects, which are the primary concern in water scarce developing countries[10], and the quality aspects of the upstream effect, which are a major concern in industrialized countries[11]. Both aspects need to be taken into account in recommending any policy interventions, whether on national level or on a more regional level through international environmental agreements. Chapter 2 and chapter 3 demonstrate that the solution to externality problems faced by river sharing countries depends among other things on the choice of the theoretical model. The signalling game in chapter 2 shows that the continuous credible threat of war is one way to achieve continuous full co-operation in a

---

[10] Developing countries are mostly agricultural economies and agricultural activities are known to consume up to twenty times more water than industrial activities world wide.
[11] The water discharged back to the river system after being used for industrial activities is usually heated and polluted.

water scarce region. A very different model in chapter 3 illustrates that the prospect of increasing co-operation gradually over time is another way to achieve limited continuous partial co-operation. Chapter 4 discusses whether improving water quality in rivers is a by-product of economic growth once a certain level of per capita income has been achieved or whether it can only happen through regional agreements among river sharing countries.

Chapter 2

Strategic Behaviour in Reaching Agreements on Sharing Waters of International Rivers: The River Nile Case

2.1. Introduction

The purpose of this chapter is to model strategic behaviour in reaching an agreement on sharing the waters of the River Nile between Egypt and Ethiopia. Most studies on the River Nile conflict concentrate on the significant gains of cooperation, which should motivate downstream Egypt to accept upstream water development projects. These studies do not provide an answer to the greatest obstacle to cooperation, namely the loss of Egypt's control over its water resources, if Ethiopia is to build its own over-year storage on the Blue Nile. Our study contributes to the economic literature by providing an analysis, which takes strategic considerations and beliefs explicitly into account. Our model demonstrates that perceived credible threats are crucial in reaching self-enforcing water agreements.

Egypt and Ethiopia are the most important riparians on the River Nile. Egypt, the downstream country, consumes nearly 75% of the estimated total flow to which it contributes nothing. Ethiopia, the upstream country, contributes over 86% of the estimated total flow and its consumption is so minimal that it does not affect the total flow. According to the 1959 Nile Agreement the estimated total flow of the River Nile is 84 BCM (billion cubic metres). Egypt is allocated 55.5 BCM, Sudan is allocated 18.5 BCM and an estimated 10 BCM is written off for evaporation losses. This means that Ethiopia, the main water provider has no legal share of the waters of the River Nile.

The demand for water in both Egypt and Ethiopia is already close to the limit of supply and is expected to increase in the near future. Current trends of population growth and the ambitious economic development plans both countries wish to implement, will need far more water than is currently available. But even current supply is not certain and might decrease if the anticipated climate change leads to an increase in evaporation losses caused by higher average global temperature. Gleick (1992) considers the increase in evaporation losses the greatest certain threat from the climate change in connection to fresh water resources in water scarce regions. Yet current and future evaporation losses from the River Nile can be substantially reduced, as discussed below. If the climate change does not occur, reducing the current evaporation losses has still a lot of gains because more water becomes available.

In the absence of alternative water resources in Egypt and the exhaustion of alternative water resources in Ethiopia, the River Nile is likely to become a source of conflict between the two countries. Ethiopia is fighting for a legal share of a resource, which originates in its territories, and Egypt is fighting to maintain its legal share under the 1959 Nile Agreement. Only an increase in the total flow of the River Nile would allow some share for Ethiopia without decreasing Egypt's share.

One possible way of increasing the total flow of the River Nile, suggested by the US Bureau of Reclamation, is to change the location of the main water reservoir of the River Nile, which is currently in southern Egypt, to another location in Ethiopia, where evaporation losses are much less. Researchers such as Guariso and Whittington (1987) and Whittington and Haynes (1985) who approve of this theoretical solution are aware that Egypt not only opposes any reductions in its current legal share of the 1959 Agreement, but also refuses rigidly to give up any control over its water supplies. Any water development projects upstream in Ethiopia with large storage facilities triggers Egypt's fears that Ethiopia may use the project to block the Nile flow for political reasons. So far, Egypt has managed to block any international funding[12] for projects in Ethiopia that have potential to harm Egypt. But is Egypt's fear justified? Is it in any interest of Ethiopia to use

[12] The World Bank does not fund any water development projects unless the riparian countries agree among themselves to avoid that water development projects cause uncompensated harm to any of the affected counties.

its favourable upstream position to harm Egypt? Are international rivers always a source of conflict or can they promote cooperation?

Gleick (1992) points out that Egypt has threatened various times to use military force to protect its water supplies. This is because Egypt is afraid that once Ethiopia has large-scale storage facilities that it will use these against Egypt as a political weapon. Ethiopia will only have an incentive to do so if Egypt does not respond with war. Are these threats credible? Will Egypt always respond with war if Ethiopia actually takes more water without Egypt's prior consent? Shapland (1997) reports that Egypt has fought a brief border war with Libya in July 1977 for giving support to Ethiopia against Egypt's interests. In September1988, the International Court of Justice has settled a border dispute between Egypt and Israel over the Taba area. This is an indication that war is not the only possible reaction for Egypt in a conflict situation. If Egypt is relatively strong it is more likely to resort to war; otherwise it is more likely to seek some sort of international mediation. Although Egypt is very well aware of how it will behave in any conflict situation, Ethiopia might not be so sure about Egypt's actual response. Ethiopia needs information about Egypt's strength to determine whether it is worthwhile to defect on an agreement or to take unilateral actions or not.

This chapter demonstrates that there can be cooperation over the River Nile, despite present conflict. The beliefs that Ethiopia has about Egypt's strength are crucial for the ultimate outcome. It not only affects Ethiopia's decision whether to cooperate of to defect, but more importantly it affects Egypt's decision in the first place of whether to cooperate with Ethiopia or not. This interdependence of Egypt's and Ethiopia's strategies with regard to the available information on military power is captured in a dynamic game of incomplete information.

This chapter has two main purposes. First, to demonstrate that beliefs do matter and that the decisive factor, which determines the ultimate outcome for the River Nile case, is not the actual military power of the threatened downstream country Egypt, but how this military power is perceived by upstream Ethiopia. This is done by modelling strategic behaviour in reaching an agreement between Egypt and Ethiopia on sharing the River Nile, when the actual military power of Egypt is unknown to Ethiopia. Secondly, to establish whether an increased interdependency between Egypt and Ethiopia, as a result of an agreement, will necessarily increase the potential of conflict between the two countries and therefore the possibility of war.

The structure of the chapter is as follows. In section 2 we look into the role that water plays in the international world. In section 3 the specific case of the River Nile is explained. In section 4 we develop a dynamic model of one-sided incomplete information, which captures how perceived power determines whether agreements are reached and whether they are sustainable, applied to the River Nile. Section 5 reports the main results and section 6 concludes.

2.2. Water in the International World

To appreciate the role that water plays in the international world we start by looking at water as a natural resource. We then identify the effects that the location of major water development projects might have on the total water availability and on the relations among riparians. Then, we discuss the illusion of reaching food self-sufficiency in already water-scarce regions. After discussing the main power attributes that affect riparian relations, we look at the role of international law in dealing with water related problems among sovereign countries. Finally, we close this section with two examples in the economic literature that have dealt with the problems of sharing international rivers.

2.2.1. Water as a Natural Resource

There are four major features of water, as a natural resource, which contribute to increased tensions among water sharing countries. Frey (1993) describes water as extremely important, relatively scarce, unevenly distributed and has to be shared. Among all natural resources, water is considered the most important one, because it enters almost all economic activities. Its importance increases dramatically if fresh water is also relatively scarce as for example in the Middle East and North Africa. Because fresh water is very unevenly distributed, not only within the same country, but also among different countries it has a political significance in determining the economic and political strength of a country. This political significance becomes even more apparent, if the water crosses national borders and has therefore to be shared among sovereign countries. World-wide, more than two hundred rivers are shared by two or more sovereign countries. Gleick (1998) mentions that if countries sharing a freshwater resource are relatively close in economic and military strength, they are more likely to engage in negotiations with cooperative outcomes. If on the other hand countries sharing a freshwater resource are relatively unequal in economic and military strength, the stronger country is more likely to take unilateral and inequitable decisions, as the weaker country will rarely start a military conflict.

2.2.2. Water and Water Development Projects

The location of major water development projects is important in three aspects. First, water supply systems are preferred targets in any military confrontation. Secondly, water storage systems are perceived as a political tools by downstream countries. Thirdly, the location of water development projects affects the total amount of water availability.

Water development projects are usually among the first military targets. This is because systems of water storage, delivery, flood control, power grids and pipelines are very difficult to defend even within a country's borders and are almost indefensible if they are outside the country's borders. Gleick (1992) reports that although hydroelectric dams were repeatedly bombed during World War II and the Korean War, the importance of water supply systems as military targets is much greater in water scarce regions such as the Middle East. During the Persian Gulf War in 1991, dams, desalination plants and other water supply systems were under attack from both sides.

Although most fresh water resources are renewable, the fact that any water resource is usually controlled by one country gives this country the opportunity to use it as a tool to achieve its political aims. Therefore, even the perception that water can be restricted or used as a political tool by the upstream country against the downstream countries can lead to serious tensions among them and to expensive precautionary measures in downstream countries. Gleick (1998) provides some illustrative examples of possible threats that were not carried out. In 1986 North Korea announced its intention to build a major hydroelectric dam on the Han River upstream Seoul, the capital of South Korea. South Korea started immediately to build a series of levees and check dams to defend itself against a possible sudden release of the reservoir's waters, which would destroy most of Seoul. During the first Gulf War in 1991 the possibility of using the Ataturk Dam in Turkey to shut off the flow of the Euphrates River to Iraq was discussed but never implemented. Although Turkey states that it would never restrict water to downstream countries as a means of political pressure, both Syria and Iraq still have their concerns.

Although the total flow (or runoff) of any river depends on natural rainfall and evaporation, the location of water supply systems affects the amount of water that is captured for agricultural, industrial, hydroelectric and municipal uses. In the past water development projects were built assuming that climate conditions are constant, and were based on historical data of existing flows and the size and frequency of expected floods and droughts. But historical data may be an unreliable guide to the future, especially as all water development projects are designed to last at least 50 years and often more than a hundred years. Gleick (1998) reports that if the global average temperature is expected to rise, as a result of the anticipated climate change, evaporation will increase as well. He mentions some studies of river basins in the western United States of America, which indicate that an increase in temperature between 2 and 4 ºC could result in a decrease in river runoff of up to 20%, provided that there is no change in precipitation. This in turn would have tremendous effects on all other water-related activities. According to some estimates a 10% decrease in the average flow of the Colorado River would result in a 30% decrease in reservoir storage, a 30% decrease in hydroelectricity generation, and a violation of water quality standards in the lower river in almost all years.

Although many countries are very successful in defending their existing water shares from intended reductions by other riparians, they are less equipped to deal with unintended reductions of the river flow resulting from climate change. Current and expected evaporation losses should be taken explicitly into account whenever existing agreements are not sustainable any more or new agreements are under negotiation. Evaporation losses have high opportunity costs in some water scarce regions of the world already, and might become increasingly difficult to afford in the future

2.2.3. Water and National Security in Terms of Food Self Sufficiency

Pearce and Turner (1990) have reckoned that while irrigation absorbs roughly 70% of the available fresh water world-wide, it absorbs up to 90% in developing countries, especially if they are located in arid water scarce regions. Hence, a small saving in water use from irrigation makes a big difference and will free enough water for increasing industrial and domestic demand. But in many developing countries self-sufficiency in food production is usually an important aim and sometimes a necessity that leads to protection of the agricultural sector, regardless of its inefficient use of water. Some developing countries are unable to substitute for water by importing food. They face shortages of foreign exchange and are forced to produce more food domestically even if it is cheaper on the world market. Shapland (1995) has looked at those countries in the Middle East, which have expanded their agricultural sector to achieve food self-sufficiency. For Egypt, Syria and Iraq, the main downstream countries in the Middle East this vision has remained an illusion. Allan (1997) believes that as long as there is enough water globally any water deficit in a specific region is not as serious as it seems. The Middle East was the first major region in the world to run out of water during the 1970s. Since then it has become increasingly dependent on global water by importing water intensive commodities such as wheat. Water security for the Middle Eastern countries could more likely be achieved by economic systems via trade in 'virtual water' i.e. water embedded in key water intensive commodities, and not by hydrological and water engineering systems. Therefore, any reductions in the allocation of water for agriculture would merely reinforce the current trend of importing water embedded in grains rather than creating a totally unfamiliar situation.

2.2.4. Water and Power Relations

There are three attributes of power that play an important role in defining the ultimate power relations among countries sharing a river: first, each country's riparian position; secondly, each country's actual military power; and thirdly, each country's relation with other powerful countries with vital interests in region.

The riparian position refers to being either upstream, midstream or downstream relative to the other countries in the river basin. A country can be either upstream, if it is the first in line, or downstream, if it is the last in line, or midstream, if it is anywhere between the upstream and downstream country. Access to rivers is not open and can be restricted by upstream countries, because rivers are accessed consecutively and not simultaneously. Therefore the impact of each country's action on the productivity or welfare of the other countries in the river basin, flows only in one direction, downstream, and is called a unidirectional externality. This unidirectional externality of the river flow, where an upstream country affects the quantity and/or quality of a downstream country's water can be negative as well as positive. Base load hydropower production helps regulate rivers; flood storage provides flood protection downstream; navigational uses and ecological maintenance keep adequate water in the river. Irrigation, industrial and municipal diversions remove water from the river system. Wastewater treatment, urban development and agriculture increase the pollution of the river. Rogers (1993) believes that the upstream country is always in a naturally advantaged position in any possible conflict, and without an agreement the downstream country has no influence on the actions of the upstream country. A downstream or mid-stream country has therefore to balance this unidirectional externality by the use of other resources such as economic or military power.

Power, in the context of river disputes, refers, according to Frey (1993), to military power, whether defensive or offensive. It determines the country's ability to defend its water resources against others or helps it to achieve its water goals by the use of force against others. Many involuntary or imposed river-sharing agreements happened in the past when a dominant power in the basin imposed its solution on the others. While the Ottoman Empire decided on the Euphrates River, the British Empire decided on the Nile River. Now current powers such as midstream-become-upstream Israel in the Jordan basin, upstream Turkey in the Euphrates basin and downstream Egypt in the Nile basin impose their interests. Therefore, although the riparian position is very important, it can be neutralised or even overcome by military and/or economic power.

The third attribute of power, which can either reinforce or overcome a country's military and/or economic power, is the riparian's relation to other powerful countries, which have vital interests in the region. As all armed conflicts will have world-wide side effects on other countries, each riparian's power might be enhanced or undermined by international intervention in the conflict. The First Gulf War in 1991 shows clearly how vital oil interests in a small helpless country, Kuwait, were protected by the international community. A change in regime, a change in the backing superpowers or the collapse of some superpowers can have tremendous effects on whether a country is perceived as strong or weak, regardless of its actual military strength.

2.2.5. Water Disputes and International Law

The role of international law in dealing with conflicts over international water resources is rather limited. One apparent reason is the lack of a unique legal water doctrine, but according to Frey (1993) even if there were such a consensus it would be rather difficult to enforce, because of the lack of effective institutions for adjudication and enforcement. Secondly, many international water treaties have been signed by colonial powers on behalf of their administered territories, and Shapland (1997) has found that in most cases the newly independent states reject those pre-independence treaties. Thirdly, international organisations such as the International Court of Justice (ICJ) can only assist in water disputes if the water sharing countries consent to its jurisdiction. Still, the legal principles on water resources that are recognised by international law, have a large impact on public opinion. Serious violations of the more widely accepted legal principles managing the use of water resources has damaging effects on the image of the offending country.

Since the early 1950s most regions of the world face controversies over water resources crossing political borders. Although there exist several legal principles recognised by international law dealing with transboundary water resources, they often present opposing interests. Depending on their riparian position, upstream or downstream, countries want to adopt the principles that favour their own particular interests. Upstream countries support the *geography-based principle of total sovereignty*, which gives them unrestricted rights over the water that originates in their territory. In contrast, downstream countries support the *history-based legal principle of acquired rights and total territorial integrity*, which guarantees them an unaltered flow of the waters that enters their territory in terms of quantity and quality.

Waterbury (1994) describes the international efforts to reach a principle that would make the opposing interests more compatible. Many international bodies such as the Institute of International Law (IDI), the International Law Commission (ILC), the International Law Association (ILA) spent decades debating and commenting on a water doctrine. No such doctrine has emerged so far. In an effort to reconcile the conflicting interests of the riparians over the development of their transboundary water resources *the principle of limited territorial sovereignty* has gained wide acceptance, especially among countries unable to impose their preferred principle and by most third-part countries. The principle of limited territorial sovereignty aims to achieve *'reasonable and equitable use'* for all riparians without causing *'appreciable harm'* to any of them. The Helsinki rules announced by the International Law Association (ILA) in 1966 lays out 11 principles in an attempt to transform these rather difficult to define terms into more practical and applicable terms. But the principles of *'reasonable and equitable use'* and *'appreciable harm'* seem also to be conflicting in various instances. Wouters (1997) reports that while upstream countries stress the principle of *'reasonable equitable use'* of the shared resource, downstream countries insist on the principle of *'no harm'*. This leads to favouring existing uses of downstream countries and justifies their objection to any water use in upstream countries that might affect them adversely. Therefore, in 1997, in a Committee formed by the UN General Assembly, countries have voted that reconciling the conflicting interests of all riparians will be best achieved under the principles of *'equitable utilisation'*, which aims to mitigate harm and discuss compensation. All countries that have voted against this new convention, China, France, Turkey and Tanzania, are upstream countries.

Because there is no legal doctrine for solving international river disputes, the role of international law is more restricted to existing international agreements over water issues. There are over 280 international water treaties. Almost two-thirds of these treaties have been signed by countries in Europe and North America. Shapland (1997) maintains that many other water treaties have caused problems, because they were signed by colonial powers on behalf of their administered territories. After the withdrawal of the colonial powers most countries have adopted the 'Nyerere doctrine' named after the former president of Tanzania. He argues that former colonies should not be bound automatically by treaties signed by colonial powers, because they had no role in negotiating those treaties.

Wouters (1997) emphasizes that international organisations such as the International Court of Justice can only assist in water disputes if the riparians consent to its jurisdiction. In 1997, the first water-related dispute for over fifty years was heard by the International Court of Justice between Hungary and the Slovak Republic over the Danube River. It is very unlikely that countries will solve their water disputes by arbitration if at least one of the riparians thinks that it could do better by using its own power, especially if there are reasons to expect an unfavourable outcome. Therefore, arbitration will only be considered if direct military confrontation is not possible. Frey (1993) believes that this could be because war is too costly or because of the recent more active role of the UN Security Council in discouraging military conflicts.

In the absence of an international agreement and the unwillingness of countries to seek the judgement of the International Court of Justice, international law still has a big impact on public opinion. Serious violations of the more widely accepted legal principles dealing with transboundary water resources might have damaging effects on the image of the offending country. The International Water Tribunal is a private judicial entity supported by 85 European environmental organisations. Frey (1993) believes that the publicity that surrounded its verdicts in 19 pollution cases has led to corrective actions.

2.2.6. Water Disputes in the Economic Literature

Frey (1993) uses a Political Accounting System (PAS) approach to predict the conflict potential among countries sharing a river. His approach is based on three factors, the importance of water to each riparian, the relative military power of each riparian, and each country's riparian position. He assigns weights ranging from 1 (weak) to 5 (strong) to each factor. The total ranking for each riparian is determined by summing up the weights of each riparian's factors. Comparison of the total rankings for each riparian implies that the more uniform the rankings the higher the conflict potential. Frey (1993) applies this approach to the three most important Middle Eastern river basins, the Euphrates, the Jordan and the Nile. His results suggest, that the Euphrates has the least conflict potential of all three basins, because Turkey's upstream position is reinforced by its military power. On the other hand, the Nile basin has the greatest conflict potential, because downstream Egypt is powerful enough to overcome its unfavourable riparian position and to defend its interests against the Sudan and Ethiopia. The main drawbacks of this approach are: first, that it depends on the subjective and ordinal measurement of the factors included and may therefore lack accuracy and consistency; and secondly, that it rates river basins according to their conflict potential by comparing the overall power of the basin countries, assuming that each factor has a similar effect on the overall power of each country.

The model we develop in this chapter demonstrates more precisely when a potential conflict situation of opposing interests can be avoided, when it might escalate into an actual conflict situation, and what the likely outcome of this conflict situation will be.

Güner (1998) uses a signalling game (a dynamic two-person game of one-sided incomplete information) to assess the role of beliefs in understanding strategic choices by Turkey and Syria with respect to a water treaty over the Euphrates river. Turkey, the upstream country, is willing to negotiate a water treaty with Syria, the midstream country, if Syria ceases to support Kurdish rebels, who wish for an independent Kurdish state in eastern Anatolia in Turkey. Turkey moves first by deciding whether to agree to a water treaty or not. Then, Syria moves by deciding whether to support the Kurdish rebels or not. It is assumed that Syria is uncertain about Turkey's preferences with respect to the mutual conflict. Syria does not know what will happen if it continues to support the Kurdish rebels. Only a strong Turkey would retaliate if Syria continues to support the Kurdish rebels after a water treaty is signed. The element of incomplete information is Syria's uncertainty of Turkey's type, whether Turkey is *strong* or *weak.* If Syria is certain that Turkey is strong, it will stop any support for the Kurdish rebels, otherwise it will continue to support them. The analysis identifies three pooling equilibrium outcomes. The first one represents the current status quo of mutual conflict: Syria supports the Kurdish rebels and Turkey does not agree to a water treaty with Syria. The two other equilibrium outcomes represent a unilateral concession by Syria: Syria does not support the Kurdish rebels and Turkey does not agree to a water treaty with Syria. Only Syria's best response changes according to different Syrian beliefs about Turkey's strength. Turkey's best response is not affected at all by Syria's beliefs. In this specific case of the Euphrates River beliefs do not matter in the strategic behaviour in reaching a water treaty, because Turkey does not agree to any water treaty whatever Syria's beliefs about Turkey's strength are. The reason for this outcome is that the already vulnerable downstream country has incomplete information about the naturally advantaged upstream country. This is not always the case as will be demonstrated in our study by exploring the case of the River Nile. Our model shows that beliefs do matter in the strategic behaviour of both Ethiopia and Egypt.

Our study combines elements of both Frey's (1993) and Güner's (1998) studies to model strategic behaviour of downstream Egypt and upstream Ethiopia in reaching an agreement over the River Nile. We use a comparable signalling game to Güner (1998) with a change that makes all the difference to the outcome of the game. In our model, it is not the vulnerable downstream country, but the naturally advantaged upstream country that has incomplete information about the strength of the vulnerable downstream country. Although most of the factors involved in our model are similar to Frey's factors, there is an important difference. Frey's factors are accumulated to arrive at the overall power of the country. The factors in our model are not accumulative, i.e., they exert their individual effects on payoffs, timing of actions, or type of country, which seems more realistic. The importance of water to each country is embodied in the payoffs to each country. The riparian position is captured partly in the relative power and partly in the timing of the actions of each country. Only the upstream country can take unilateral actions or defect on an agreement that will harm the downstream country. The relative power of the countries is captured in the type of the downstream country. The downstream country can be either *weak* or *strong* relative to the upstream country. While trying to keep our model as simple as possible and as detailed as necessary, it is able to demonstrate that beliefs do matter in the strategic behaviour of both Egypt and Ethiopia, with a variety of equilibrium outcomes according to whether these different beliefs are confirmed or not.

2.3. The River Nile Case

The River Nile[13] one of the world's longest rivers, flows through ten northern African countries. These are Egypt, Sudan, Ethiopia, Eritrea, Zaire, Uganda, Kenya, Tanzania, Rwanda and Burundi. It has two main tributaries, the Blue Nile and White Nile, which converge near Khartoum, the capital of Sudan, before flowing to Egypt and emptying into the Mediterranean. For all six countries on the White Nile upstream of Sudan, and for

Eritrea on the Blue Nile, irrigation is not essential for cultivation because rainfall is much higher than in northern Sudan and Egypt. Our study is concerned with the Blue Nile, which originates in Ethiopia, converges with the White Nile in Sudan, flows as the main Nile through Egypt, where it empties into the Mediterranean. This part of the River Nile is crucial for three countries, Egypt, Sudan and recently Ethiopia in descending order. Ethiopia is upstream, Sudan midstream and Egypt downstream. Shapland (1997) believes that even without taking the possibility of any reductions in water supply resulting from climatic change into account, potential future demand will be much higher than existing supply. This is because all three countries have a fast growing population, which is expected to reach nearly 285 million in 2025. If all three basin countries are to achieve at least some of their ambitious and much needed development plans some sort of co-operation that increases the total water supply will be necessary.

During the colonial period several agreements were signed among the colonial powers, Britain, Italy, France and Belgium, on behalf of their administered territories in the Nile basin. These agreements covered almost the whole basin of the River Nile and could theoretically serve as a starting point for cooperation among the now independent basin countries. But after the withdrawal of the colonial powers most of these agreements were not accepted by the newly independent countries, except by Egypt whose interests were protected by those pre-independence agreements. In 1959, a new bilateral agreement, the Agreement for the Full Utilisation of the Nile Waters, was negotiated and ratified by Egypt and Sudan. The 1959 Nile Agreement, which made the building of the Aswan High Dam possible, allocates the total flow of the Nile between Egypt and Sudan after accounting for evaporation and seepage losses from Lake Nasser, the Aswan High Dam reservoir. The approach adopted by Egypt and Sudan in the Agreement, that other riparians have to seek the approval of Egypt and Sudan if they want to make use of the Nile waters, does not find acceptance by the upstream countries. Shapland (1997) reports that since 1956, while the 1959 Nile Agreement was still under negotiation, Ethiopia has stressed repeatedly its rights to use the waters of the Nile to the benefit of its people. Gleick (1998) states, that after Ethiopia declared its intentions in 1978 to construct dams on the headwaters of the Blue

---

[13] For more details on the River Nile see Appendix 2.1.

Nile, Egypt immediately threatened to respond with war and has repeatedly declared the vital importance of the Nile waters to Egypt whenever the issue has appeared. Therefore, any water development projects upstream, especially in Ethiopia, can lead to serious tensions over water in the region if no progress is made in reaching some sort of agreement which includes Ethiopia.

Whittington *et al.* (1995) think that cooperation over the use of the Nile is not strictly a zero-sum game, even if Ethiopia is given a share for itself. But this is conditional on increasing the long-term yield of the river, which theoretically can be achieved by reducing the huge evaporation losses from Lake Nasser, the Aswan High Dam reservoir in Egypt and from the Sudd swamps in Sudan. Evaporation losses of dam reservoirs are side-effects with international implications, because evaporation uses up water that could be used elsewhere in the basin, if the reservoir was in a cooler or more humid location. In 1964, the US Bureau of Reclamation carried out an economic study about the irrigation and hydropower potential for the Blue Nile. This study, which is cited by Guariso and Whittington (1987), Jovanovic (1985), and Whittington and McClelland (1992) suggests to shift part of the over-year storage from Lake Nasser in southern Egypt, which has one of the highest evaporation rates in the world, to another more appropriate location further upstream on the Blue Nile in Ethiopia, where evaporation losses are 50% of those in Sudan and Egypt. This will lead to water savings of 4 to 5 BCM (billion cubic metres) per year. Thus, by operating Lake Nasser at a lower level to reduce evaporation losses, some water will be freed for Ethiopia's much needed irrigation plans. This basic idea of constructing a Blue Nile reservoir in Ethiopia, where evaporation losses are much less than in Egypt, is the main motivation for building our model. The key question is whether Egypt will ever delegate the security of its water supplies to another country. Allan (1999), Guariso and Whittington (1987), Jovanovic (1985), Swain (1997), and Whittington and McClelland (1992), who are acquainted with this problem, are optimistic and do not consider Egypt's refusal an insurmountable problem, if appropriate incentives and assurances from Ethiopia are able to diffuse Egypt's fear of not being in full control of its water supplies.

2.3.1. Ethiopia's Position

The Blue Nile originates in Ethiopia, which contributes around 86% to the total flow of the Nile without having any legal share of it. Ethiopia has eight rivers flowing across its borders and in the past, development efforts were concentrated on its eastern watershed. Waterbury (1987) expects that increasing population will inevitably shift Ethiopia's attention to the western watershed, which includes the Blue Nile, the Sobat (Baro) River, which flows into the White Nile, and the Atbara River, which flows into the Main Nile. Jovanovic (1985) has estimated that the available flow to Egypt and Sudan would be reduced by up to 23% if Ethiopia irrigated land inside the Nile catchment only, and by up to 39% if irrigation is extended to irrigable land outside the Nile Basin. If on the other hand, Ethiopia implements all the proposed projects in the previously mentioned American study, Guariso and Whittington (1987) reckon that the total annual flow of the Blue Nile into Sudan will be reduced by 8.5% only. Swain (1997) emphasizes that Ethiopia maintains the view, that in absence of any binding agreements, it is the nation's sovereign right to develop all the resources within its boundaries. Ethiopia bases its case on the legal principle of 'equitable use', which guarantees each riparian a 'fair', but not necessarily equal share of the international river. Ethiopia also recognises the complementary legal principle of not causing 'appreciable harm' to other riparians, but in a much wider sense. According to Shapland (1997), Ethiopia argues that harm is not only inflicted by upstream countries on downstream countries, but that the opposite is equally true. Egypt's desert reclamation programme is an example of the latter case, because Egypt pre-empts water that Ethiopia wants to use as its equitable share.

2.3.2. The Difficult Position of Sudan

Sudan is midstream and until now not using its all of its water share according to the 1959 Nile Agreement, which it has signed with Egypt. But this unused water share will soon be insufficient to meet all of Sudan's future water needs. Waterbury (1987) reports that the ongoing water development projects on the White Nile to increase the amount of the total flow, especially in swamps of southern Sudan, had to be stopped since 1983 because of civil unrest in Sudan. If Sudan restricts its attention to the development of the White Nile it would have to bear alone the social and ecological disruption in the development regions, while the gains would be rather modest and far more important for Egypt than for Sudan. On the other hand, Sudan would gain a lot by cooperating with Ethiopia on the development of the Blue Nile. A large water reservoir in Ethiopia would even out the flow for Sudan and would reduce a lot its problems in operating its Roseires reservoir. But under the 1959 Nile Agreement Egypt and Sudan have to take a common position against any other riparian. Historically, Sudan has always been an important ally for Egypt, but its present political instability makes it difficult to rely on as an ally for either Egypt or Ethiopia.

2.3.3. Egypt's Position

Egypt is downstream to the Nile and very vulnerable, because it has no other significant fresh water resources. It is by far the largest consumer of the Nile waters while it contributes nothing to its flow. Egypt currently uses its whole water share according to the 1959 Nile Agreement and even has to borrow some water from Sudan's share. With a population growing at the rate of one million people every nine months, higher consumption in upstream countries will have serious effects not only on future development plans but on existing economic activities as well.

Shapland (1997) believes that there were two main achievements of the 1959 Nile Agreement between Egypt and Sudan, which enabled Egypt to build the Aswan High Dam. First, the gain of total control of the Nile; and secondly, the guarantee of a fixed water share respected by Sudan. The choice for the present location of the Aswan High Dam was therefore a strategic rather than a technical or an economic choice. The construction of a single massive dam within Egypt's borders would guarantee Egypt the absolute control over the water, which flows into Lake Nasser, one of the largest reservoirs in the world. Enough water could be saved in years of high flow to use in years of low flow, thereby avoiding any possible conflict with upstream riparians. Egypt has made it repeatedly clear, that it will not agree to any solutions that reduce its water share according to the 1959 Nile Agreement or that reduce its control its over its water supplies. It bases its case on the legal principle of 'acquired rights' and more recently on the legal principle of avoiding 'appreciable harm'. Egypt's interpretation of 'harm' is somewhat different form the Ethiopian interpretation, as it refers only to upstream works that cause reductions of the flow of the Blue Nile. Allan (1999) considers Egypt's announcement of the southern New Valley project in 1997, a clear message to Ethiopia that any water savings will not be shared with upstream riparians. This ambitious desert reclamation programme demonstrates that any savings made from using water more efficiently will be easily absorbed in the newly reclaimed land. Shapland (1997) therefore reckons that the capital necessary for the reclamation programmes are not only an investment in Egypt's productive capacity, but also a means to reinforce Egypt's claim to its existing water share and to improve its negotiating position in any future negotiations.

2.4. The Model

The model takes its basic idea from the previously mentioned study carried out by the US Bureau of Reclamation in 1964. Therefore, we only tackle the problem of reaching an agreement between Egypt and Ethiopia over the utilisation of the Blue Nile by reducing the evaporation losses. The construction of the Blue Nile Reservoir in Ethiopia, where evaporation losses are much lower than in Egypt, frees enough water that will be allocated to Ethiopia. Ethiopia bears all the costs of constructing the reservoir, in addition to compensating Egypt for giving up the control of its water supplies and any other losses associated with shifting the over-year storage from Egypt to Ethiopia. We assume that the amount of the compensation that Ethiopia has to pay to Egypt is exogenous and can be determined separately by a theory of bargaining. Now Ethiopia has the facilities to defect on the agreement and to take even more water, especially during droughts, at no additional costs. Egypt is afraid that if Ethiopia has the ability to withhold water from Egypt, especially during droughts, than it will use water as a strategic weapon against Egypt. As explained earlier, even the perception that access to fresh water can be used as a political weapon leads to serious tensions. Although Ethiopia always has an incentive to defect on the agreement during droughts, it will only do so if it thinks that Egypt will not respond with war. It is more likely that Egypt responds with war to any defection of the agreement or to any unilateral change of the status quo by Ethiopia without Egypt's prior consent, if Egypt is strong rather than weak. Therefore, Ethiopia has to consider carefully whether a defection from the agreement or a unilateral change of the status quo is worthwhile or not. Although Egypt has threatened many times to go to war, if anybody adversely affects its water quota, such a threat is only credible if Egypt is powerful enough relative to the offending country. Shapland (1997) mentions that Egypt has fought a brief border war with Libya in 1977, when Libya threatened Egypt's interests by supporting Ethiopia over its claims on the Nile. Kemp and Ben-Eliezer (2000) describe how Egypt managed to solve a border dispute with Israel over the Taba area peacefully through the International Court of Justice in September 1988. Therefore, responding with war in a conflict situation seems only reasonable for Egypt if it is relatively strong, and seeking some sort of international mediation or arbitration in a conflict situation seems reasonable for Egypt if it is relatively weak. Some power attributes of a country are more easily detected than others; and some change rather quickly due to changes in the political regime or the country's relation with other powerful countries. Therefore, we assume that Ethiopia is uncertain about Egypt's actual power. The objective of the model is to understand how Ethiopia's uncertainty about Egypt's military power affects the strategic choices of both Egypt and Ethiopia with regard to their water problem. We assume that both Egypt and Ethiopia are risk neutral and therefore aim to maximize their expected payoff from any strategy they adopt.

2.4.1. Description of the Model

This conflict situation is captured in a simple signalling game: a dynamic Bayesian game of incomplete information, which involves two players, a Sender (S) and a Receiver (R).

Applied to the River Nile case, the two players involved are Egypt (Eg) and Ethiopia (Eth). Egypt, the Sender, has information that is not available to Ethiopia; Ethiopia, the Receiver, has no information to which Egypt does not have access. Egypt knows whether it is *'strong'* or *'weak'*, but Ethiopia does not know that; it has only a probability distribution over the two types of Egypt. First, Nature chooses the type (t_i) of Egypt whether *'strong'* or *'weak'* according to a probability distribution known to both Egypt and Ethiopia.

$$T = \{strong, weak\}$$
$$prob\{strong\} = r; prob\{weak\} = 1 - r$$

Egypt cannot choose its own type (t_i), it only knows what Nature has chosen, whether it is *'strong'* or *'weak'*. Egypt has the first move and sends a message (m_j). It decides whether to *'agree'* to the construction of the Blue Nile Reservoirs or to *'not agree'* to it.

$$M = \{agree, not\ \ agree\}$$

Ethiopia observes Egypt's message, and updates its beliefs about Egypt's type accordingly using Bayes' rule. $\mu(t_i|m_j)$ is Ethiopia's belief about Egypt's type t_i conditional on Egypt's message m_j. The posterior belief that Egypt is *strong* is denoted by *p*, if an agreement has been reached and is denoted by *q*, if no agreement has been reached.

$$\mu(t_1|m_j) = p\ \ or\ q \qquad\qquad \text{where } t_1 = strong$$

Then Ethiopia chooses its action (a_k), whether to *'cooperate'* or to *'defect'*.

$$A = \{cooperate, defect\}$$

If Ethiopia decides to *'cooperate'* the game ends. If Ethiopia decides to *'defect'*, the game continues and Egypt chooses its reaction (r_l) whether to go to *'war'* or to go to *'arbitration'*. Egypt makes its final move (reaction) depending on its true type, whether *'strong'* or *'weak'*.

$$R = \{war, arbitration\}$$

The payoffs for the Sender, Egypt (Eg) and for the Receiver, Ethiopia (Eth) are as follows:

$$U_{Eg}(t_i, m_j, a_k, r_l)$$

$$U_{Eth}(t_i, m_j, a_k, r_l)$$

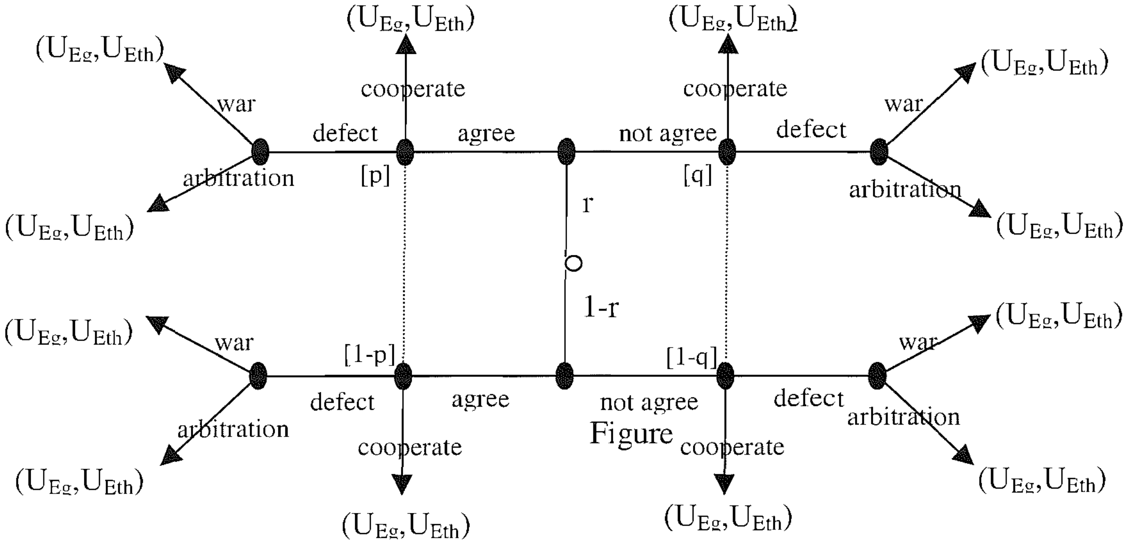

Figure 2.1. The extensive form of the signalling game

2.4.2. The Parameters, Assumptions and Pure Strategies

The Parameters

B_1 benefits for Ethiopia from an agreement with Egypt

B_2 benefits for Ethiopia from taking unilateral actions

C_1 compensation paid by Ethiopia to Egypt for an agreement

C_2 costs of facilities for Ethiopia to use more water

C_3 costs for Ethiopia to use more water than agreed upon after an agreement is reached

D_1 damage to Egypt if Ethiopia defects after an agreement

D_2 damage to Egypt if Ethiopia acts unilaterally without an agreement

W_H high war costs

W_L low war costs

A arbitration costs

r probability that Egypt is strong

p Ethiopia's beliefs that Egypt is strong if an agreement is reached

q Ethiopia's beliefs that Egypt is strong if no agreement is reached

The Assumptions

ASS(1): $B_1 > C_1 + C_2$

ASS(2): $0 \leq C_3 < C_2$

ASS(3): $D_1 - C_1 > D_2$

ASS(4): $W_L < A < W_H$

ASS(5): $C_1 - D_1 - W_L \leq 0$

ASS(6): $C_1 - D_1 - A_L \leq 0$

ASS(7): $B_2 < W_H$

ASS(1) means that the benefit from an agreement between Egypt and Ethiopia must be greater than all the other costs associated with implementing the agreement, otherwise there is no incentive for reaching the agreement in the first place. ASS(2) means that after an agreement between Egypt and Ethiopia is reached, Ethiopia has all the facilities to take more water than agreed upon without any additional construction cost, which basically means that defecting after the agreement is much easier for Ethiopia than taking unilateral action without an agreement. Nevertheless there are other, less tangible costs associated with taking more water than agreed upon. These costs range from 0 to less than C_2. If $C_3 \geq C_2$ an agreement does not increase Egypt's vulnerability to upstream abstractions from the River Nile. ASS(3) means that the damage to Egypt if Ethiopia defects after an agreement is reached is greater than the damage it suffers if Ethiopia takes unilateral actions without an agreement, even if compensation is deducted. This assumption can be relaxed, so that the compensation is big enough to make a defection after the agreement less damaging than unilateral actions without an agreement. ASS(4) means that arbitration costs are greater than low war costs, but smaller than high war costs. If Egypt is *strong* it incurs only low war costs but inflicts high war costs on Ethiopia, whereas if Egypt is *weak* it incurs high war costs and inflicts only low war costs on Ethiopia. This assumption is essential, otherwise defection or unilateral actions will always lead to war or never lead to war, which contradicts reality. ASS(5) and ASS(6) mean that Egypt is always worse off, relative to the status quo position, if Ethiopia defects after an agreement, regardless of Egypt's response. So if the status quo is positive instead of zero, ASS(5) and ASS(6) can be relaxed. ASS(7) means that the benefits from defecting are smaller than the high war costs, otherwise it is always worthwhile to defect. Whereas ASS(1), ASS(2), ASS(4) and ASS(7) are crucial, ASS(3), ASS(5) and ASS(6) can be relaxed as explained above.

The Pure Strategies Available to Both Egypt and Ethiopia

In a three stage signalling game the Sender sends a message dependent on the Sender's type. The Receiver acts upon receiving the message according to his or her beliefs. At the final stage, the Sender reacts to the Receiver's action according to his or her true type. Both the Sender and the Receiver have a variety of strategies available to choose from. A player's strategy is a complete plan of action i.e., it specifies a feasible action in every contingency in which the player might have to act.

In our signalling game, a pure strategy for Egypt is a function $m(t_i)$ that specifies which message is sent for each type that Nature might choose, and a function $r(a_k, t_i)$ that specifies which reaction is chosen for each type of Egypt and for each action chosen by Ethiopia. A pure strategy for Ethiopia is a function $a(m_j)$ that specifies which action is chosen for each message that could have been sent by Egypt, according to Ethiopia's updated beliefs about Egypt's type. A strategy is considered to be a *pure pooling strategy* if all types of Egypt send the same message, and is considered to be a *pure separating strategy* if all types of Egypt send different messages.

Egypt has a lot more pure strategies available than Ethiopia (48 compared to 4) because it has to act twice. A pure strategy for Egypt consists of six entries: the first entry if Egypt is strong, the second entry if Egypt is weak; the third entry if Egypt is strong and agrees and Ethiopia defects, the fourth entry if Egypt is strong and does not agree and Ethiopia defects, the fifth entry if Egypt is weak and agrees and Ethiopia defects, and the sixth entry if Egypt is weak and does not agree and Ethiopia defects. In entry one and two Egypt decides, in the first stage of the game, whether to agree or not to agree to the construction of the Blue Nile reservoirs in Ethiopia. In entry three to six, in the final stage of the game, Egypt reacts to Ethiopia's defection. Some examples of pure strategies for Egypt are:

(1) *(agree, agree; war, war, war, war)* i.e. agree whether strong or weak, then go to war if Ethiopia defects whether strong or weak and whether there is an agreement or not.

(2) *(agree, agree; war, war, arbitration, arbitration)* i.e. agree whether strong or weak, then go to war if strong and Ethiopia defects whether there is an agreement or not, and go to arbitration if weak and Ethiopia defects whether there is an agreement or not.

(3) *(not agree, not agree; war, war, arbitration, arbitration)* i.e. don't agree whether strong or weak, then go to war if strong and Ethiopia defects whether there is an agreement or not, and go to arbitration if weak and Ethiopia defects whether there is an agreement or not.

(4) *(not agree, agree; arbitration, arbitration, arbitration, arbitration)* i.e. don't agree if strong, agree if weak, then go to arbitration if Ethiopia defects whether strong or weak and whether there is an agreement or not.

A pure strategy for Ethiopia will determine what Ethiopia will do in the second stage of the game after it receives the message from Egypt and updates its beliefs accordingly. Ethiopia has only four pure strategies available, each strategy has two entries (the first entry if Egypt agrees to the construction of the Blue Nile reservoirs, the second if it does not agree to it). Ethiopia's pure strategies, according to its updated beliefs, are:

(1) *(cooperate, cooperate)* i.e. co-operate whether Egypt agrees or does not agree

(2) *(cooperate, defect)* i.e. co-operate if Egypt agrees, defect if Egypt does not agree

(3) *(defect, cooperate)* i.e. defect if Egypt agrees, co-operate if Egypt does not agree

(4) *(defect, defect)* i.e. defect whether Egypt agrees or does not agree

2.4.3. The Extensive Form of the Signalling Game

Figure 2.2 shows the extensive form of a dynamic model of one-sided incomplete information with payoffs applied to case of the River Nile.

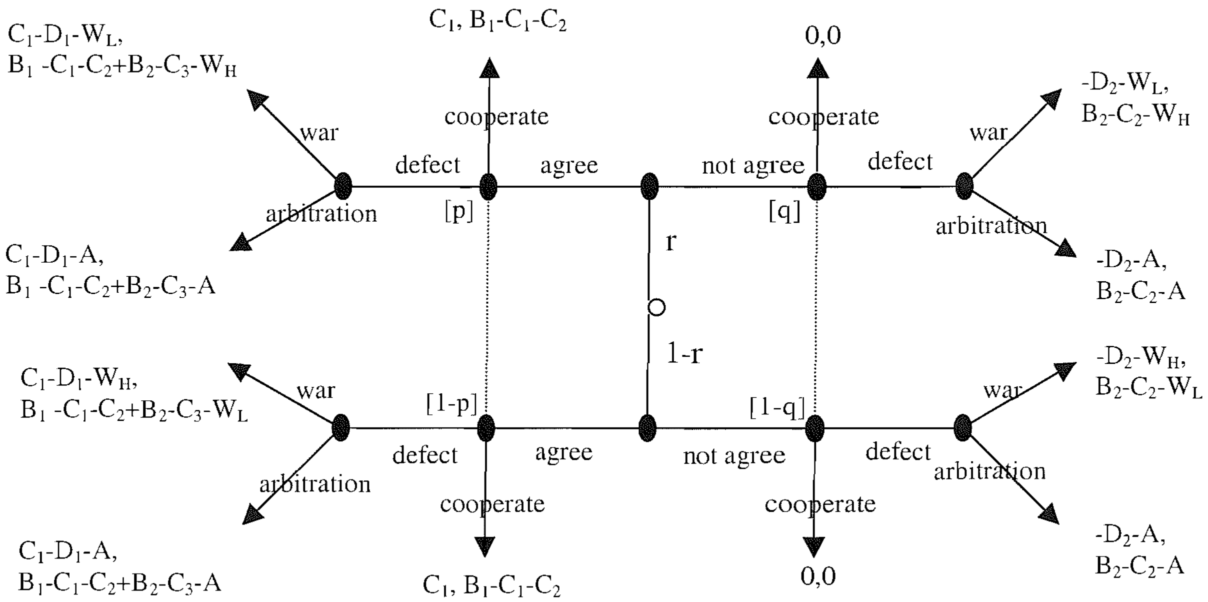

Figure 2.2. The extensive form of the signalling game with payoffs

2.4.5. Analysis of the Strategic Form of the Bayesian Game

In a Bayesian game, as explained in Gibbons (1992), the players' beliefs are made explicit and are as important as strategies in the definition of equilibrium. A perfect Bayesian equilibrium consists of a strategy for each player and a belief for each player at each information set at which the player has to move. In equilibrium, players do not only choose credible strategies but they must also hold reasonable beliefs both on the equilibrium path and off the equilibrium path. At any given information set the player's action is based partly on the player's belief at this information set and partly on the players' subsequent strategies. This belief in turn depends on the players' actions higher up in the game tree. But these actions higher up in the game tree are partly based on the players' subsequent strategies including the action at the original information set. This circularity implies that working backwards through the game tree is not enough to establish a perfect Bayesian equilibrium.

Formally, a pure-strategy perfect Bayesian equilibrium in a signalling game consists of a pair of strategies $m^*(t_i)$ and $a^*(m_j)$ and a belief $\mu(t_i|m_j)$ which satisfy the following requirements:

a) After observing any message m_j from M, the Receiver must have a belief about which type could have sent this message. Because the Receiver's choice is at a non-singleton information set, his or her beliefs follow a probability distribution $\mu(t_i|m_j)$, where $\mu(t_i|m_j) \geq 0$ for each t_i in T and $\sum_{t_i \in T} \mu(t_i|m_j) = 1$. The Receiver's belief at his or her information set corresponding to m_j must follow from Bayes' rule and the Sender's strategy: $$\mu(t_i|m_j) = \frac{p(t_i)}{\sum_{t_i \in T} p(t_i)}$$

b) For each m_j in M, the Receiver's action $a^*(m_j)$ must maximize the Receiver's expected utility, given his or her belief $\mu(t_i|m_j)$ about which type could have sent the message. Therefore $a^*(m_j)$ solves

$$\max_{a_k \in K} \sum_{t_i \in T} \mu(t_i|m_j) U_R(t_i, m_j, a_k, r_l)$$

c) For each t_i in T, the Sender's message $m^*(t_i)$ must maximize the Sender's utility, given the Receiver's strategy $a^*(m_j)$ and the Sender's reaction to the Receiver's strategy $r^*(t_i, a_k)$. Therefore, $m^*(t_i)$ solves

$$\max_{m_j \in M} U_S(t_i, m_j, a^*(m_j), r^*(t_i, a_k))$$

and $r^*(t_i, a_k)$ solves $$\max_{r_l \in R} U_S(t_i, m^*(t_i), a^*(m_j), r(t_i, a_k))$$

If the Sender's equilibrium strategy is pooling, the equilibrium will be a *pooling pure-strategy perfect Bayesian equilibrium.* If on the other hand, the equilibrium strategy of the Sender is separating, the equilibrium will be a *separating pure-strategy perfect Bayesian equilibrium.* Fudenberg and Tirole (1991) explain that in a pooling equilibrium, the Receiver does not update his beliefs when he or she observes the equilibrium message i.e. his or her posterior beliefs are his or her prior beliefs. In a separating equilibrium, the Receiver has complete information when he or she observes the equilibrium message and knows exactly at which node he or she is.

In the last stage of this particular signalling game, Egypt's decision nodes are singletons, because Egypt has perfect and complete information. Egypt knows its own type and has observed all previous moves before it chooses its next move. Both Egypt and Ethiopia know the payoffs from all feasible combinations of moves in this last stage of the game. To maximize Egypt's utility from its last move

$$r^*(t_i, a_k) \quad \text{solves} \quad \max_{r_l \in R} U_S(t_i, m_j, a_k, r_l) \quad \text{by backward induction.}$$

Given ASS(1) to ASS(7) and assuming rational behaviour by Egypt, there is a unique solution for each feasible combination of previous moves and types. For this particular game, Egypt plays *war, war, arbitration, arbitration* if it reaches the upper left node, the upper right node, the lower left node and lower right node in the third stage of the game respectively. If we denote this backward induction outcome from the third stage of the game by $R(t_i, m_j, a_k)$, then Ethiopia's maximization problem becomes

$$a^*(m_j) \quad \text{solves} \quad \max_{a_k \in K} \sum_{t_i \in T} \mu(t_i | m_j) U_R(t_i, m_j, a_k, R(t_i, m_j, a_k))$$

and Egypt's maximization problem becomes

$$m^*(t_i) \quad \text{solves} \quad \max_{m_j \in M} U_S(t_i, m_j, a^*(m_j), R(t_i, m_j, a_k))$$

This reduces the possible pure equilibrium strategies available to Egypt to only four pure strategies. These are:

(1) (agree, agree; war, war, arbitration, arbitration) i.e. agree whether strong or weak
(2) (agree, not agree; war, war, arbitration, arbitration) i.e. agree if strong, disagree if weak
(3) (not agree, agree; war, war, arbitration, arbitration) i.e. disagree if strong, agree if weak
(4) (not agree, not agree; war, war, arbitration, arbitration) i.e. disagree whether strong or weak

To determine the pure-strategy perfect equilibria of this signalling game, the possible pure-strategy perfect equilibrium candidates are analysed. First, the four possible pure-strategies of the Sender (Egypt) are proposed: (1) pooling on 'agree'; (2) pooling on 'not agree'; (3) separating with *strong* playing 'agree'; and (4) separating with *strong* playing 'not agree'. Then, the optimal response of the Receiver (Ethiopia) to each proposed strategy is determined. Finally, to determine whether the Sender's proposed strategy is an equilibrium strategy, the Receiver's response to an unexpected move by the Sender has to be specified, to make sure that the Sender's strategy is optimal and cannot be improved upon. Whether a particular pair of strategies form a pure-strategy perfect equilibrium depends on the relative values of the payoff parameters and beliefs.

The four possible pure-strategies are as follows:

I. Pooling on 'agree'

If there is an equilibrium in which Egypt plays (agree, agree), Ethiopia's information set corresponding to 'agree' is on the equilibrium path and Ethiopia's beliefs (p, 1-p) are determined by Bayes' rule. In the case of a pooling strategy of the Sender the posterior belief about the probability distribution equals the prior probability distribution i.e. p = r.

Given this belief Ethiopia's best response to 'agree' is:

co-operate if $r \geq \frac{B_2 - C_3 - A}{W_H - A}$

defect if $r < \frac{B_2 - C_3 - A}{W_H - A}$

To determine whether both types of Egypt would optimally choose 'agree', Ethiopia's best responses to Egypt's off-equilibrium actions have to be determined. As there are no restrictions on Ethiopia's off-equilibrium beliefs, q can be anything between 0 and 1.

If Egypt plays 'not agree' Ethiopia's best response would be:

co-operate if $q \geq \frac{B_2 - C_2 - A}{W_H - A}$

defect if $q < \frac{B_2 - C_2 - A}{W_H - A}$

If ASS(1) to ASS(7) hold, there will be two pooling pure-strategy perfect Bayesian equilibria where Egypt plays 'agree':

(1) [(agree, agree), (co-operate, co-operate), p = r, q] is a perfect Bayesian equilibrium if

$r \geq \frac{B_2 - C_3 - A}{W_H - A}$ and $q \geq \frac{B_2 - C_2 - A}{W_H - A}$

(2) [(agree, agree), (co-operate, defect), p = r, q] is a perfect Bayesian equilibrium if

$r \geq \frac{B_2 - C_3 - A}{W_H - A}$ and $q < \frac{B_2 - C_2 - A}{W_H - A}$

The outcomes of both (1) and (2) mean, that sustainable co-operation is reached in equilibrium, and none of the players has an incentive to deviate from its equilibrium strategy. This sustainable co-operation is most likely to result if the prior subjective probability held by Ethiopia that Egypt is strong is high enough, regardless of Ethiopia's off-equilibrium beliefs.

II. Pooling on 'not agree'

If there is an equilibrium in which Egypt plays (not agree, not agree), Ethiopia's information set corresponding to 'not agree' is on the equilibrium path and Ethiopia's beliefs (q, 1-q) are determined by Bayes' rule. In the case of a pooling strategy of the sender the posterior belief about the probability distribution equals the prior probability distribution i.e. q = r.

Given this belief Ethiopia's best response to 'not agree' is:

co-operate if $r \geq \frac{B_2 - C_2 - A}{W_H - A}$

defect if $r < \frac{B_2 - C_2 - A}{W_H - A}$

To determine whether both types of Egypt would optimally choose 'not agree', Ethiopia's best responses to Egypt's off-equilibrium actions have to be determined. As there are no restrictions on Ethiopia's off-equilibrium beliefs, p can be anything between 0 and 1.

If Egypt plays 'agree' Ethiopia's best response would be:

co-operate if $p \geq \frac{B_2 - C_3 - A}{W_H - A}$

defect if $p < \frac{B_2 - C_3 - A}{W_H - A}$

If ASS(1) to ASS(7) hold, there will be two pooling pure-strategy perfect Bayesian equilibria where Egypt plays 'not agree':

(3) [(not agree, not agree), (defect, co-operate), q = r, p] is a perfect Bayesian equilibrium

if $\quad r \geq \frac{B_2 - C_2 - A}{W_H - A} \quad$ and $\quad p < \frac{B_2 - C_3 - A}{W_H - A}$

(4) [(not agree, not agree), (defect, defect), q = r, p] is a perfect Bayesian equilibrium

if $\quad r < \frac{B_2 - C_2 - A}{W_H - A} \quad$ and $\quad p < \frac{B_2 - C_3 - A}{W_H - A}$

Outcome (3) means a unilateral concession of Ethiopia. Egypt is not willing to reach an agreement with Ethiopia on the construction of the Blue Nile reservoirs, and Ethiopia will not take any unilateral actions. This outcome is the current status quo of the situation, which Egypt tries to hold to as long as possible. It is most likely to occur if the prior subjective probability held by Ethiopia that Egypt is *strong* is high enough, but Ethiopia's posterior off-equilibrium beliefs that Egypt is *strong* are relative low.

Outcome (4) means mutual conflict between Egypt and Ethiopia, which can escalate into war. Egypt is not willing to reach an agreement with Ethiopia on the construction of the Blue Nile reservoirs, but Ethiopia takes unilateral action to get some of the water it desperately needs. If Egypt is *strong*, war will be the ultimate outcome, if Egypt is *weak,* the conflict will be solved by arbitration. This outcome is most likely to occur if the prior subjective probability held by Ethiopia that Egypt is *strong* is low enough and Ethiopia's posterior off-equilibrium beliefs that Egypt is *strong* are low enough as well.

III. Separating with Egypt playing (agree, not agree)

If there is an equilibrium with Egypt playing (agree, not agree), both of Ethiopia's information sets are on the equilibrium path and Ethiopia's beliefs are determined by Bayes' rule and Egypt's strategy. In this case p = 1 and q = 0.

Given this belief Ethiopia's best response to 'agree' is

co-operate if $B_2 \leq W_H + C_3$

defect if $B_2 > W_H + C_3$

and its best response to 'not agree' is

co-operate if $B_2 \leq C_2 + A$

defect if $B_2 > C_2 + A$

If ASS(1) - ASS(7) hold, there will be no separating pure-strategy perfect Bayesian equilibria where Egypt plays (agree, not agree).

IV. Separating with Egypt playing (not agree, agree)

If there is an equilibrium with Egypt playing (not agree, agree), both of Ethiopia's information sets are on the equilibrium path and Ethiopia's beliefs are determined by Bayes' rule and Egypt's strategy. In this case p = 0 and q = 1.

Given this belief Ethiopia's best response to 'agree' is

co-operate if $B_2 \leq A + C_3$

defect if $B_2 > A + C_3$

and its best response to 'not agree' is

co-operate if $B_2 \leq C_2 + W_H$

defect if $B_2 > C_2 + W_H$

If ASS(1) - ASS(7) hold, there will be no separating pure-strategy perfect Bayesian equilibria where Egypt plays (not agree, agree).

Although the amount of compensation C_1 that Ethiopia pays to Egypt does not affect the prior or posterior beliefs that Ethiopia holds about Egypt's strength, it can make the net damages of defection after an agreement $D_1 - C_1$ greater [ASS(3)] or smaller [ASS(3`)] than the damages of defection without an agreement $D_2$. So, if ASS(3) is relaxed and replaced by ASS(3`) so that $D_1 - C_1 \leq D_2$, there will still be no separating pure-strategy perfect Bayesian equilibria, but the following new pooling pure-strategy perfect Bayesian equilibrium

(5) [(agree, agree), (defect, defect), p = r, q] is a pure-strategy perfect Bayesian equilibrium

if $r < \dfrac{B_2 - C_3 - A}{W_H - A}$ and $q < \dfrac{B_2 - C_2 - A}{W_H - A}$.

will replace the former pooling pure-strategy perfect Bayesian equilibrium

(4) [(not agree, not agree), (defect, defect), q = r, p] is a perfect Bayesian equilibrium

if $r < \dfrac{B_2 - C_2 - A}{W_H - A}$ and $p < \dfrac{B_2 - C_3 - A}{W_H - A}$

Outcome (5) means mutual conflict between Egypt and Ethiopia that might escalate into war. Egypt has reached an agreement with Ethiopia on the construction of the Blue Nile reservoirs, but Ethiopia defects on the agreement and takes more water than agreed upon. If Egypt is *strong*, war will be the ultimate outcome, if Egypt is *weak*, the conflict will be solved by arbitration. This outcome is most likely to occur if the prior subjective probability held by Ethiopia that Egypt is *strong* is low enough and Ethiopia's posterior off-equilibrium beliefs that Egypt is *strong* are low enough as well.

2.5. Results

This signalling game has four possible pooling equilibrium outcomes. These are:

(a) bilateral co-operation as in outcome (1) and (2)

(b) imposed co-operation with unilateral concessions by Ethiopia as in outcome (3)

(c) mutual conflict leading to arbitration as in outcomes (4) and (5) if Egypt is *weak*

(d) mutual conflict leading to war as in outcomes (4) and (5) if Egypt is *strong*.

Outcome (3) describes the observed status quo i.e., Egypt does not agree to the agreement over the Blue Nile and Ethiopia does not take any unilateral actions that would harm Egypt. The other equilibrium outcomes show that beliefs matter because Ethiopia's beliefs not only affected Ethiopia's best responses but affects Egypt's best responses as well.

Conflict between Egypt and Ethiopia can be avoided altogether if Egypt is perceived as *strong* by Ethiopia, whether it is actually strong or weak. Mutual conflict leading to war will most likely occur if a *strong* Egypt is wrongly perceived as weak by Ethiopia, whereas mutual conflict leading to arbitration is most likely to occur if a *weak* Egypt is correctly perceived as weak by Ethiopia. War is a possible outcome, whether the agreement is actually reached or not. If ASS(3)[14] holds, war is only likely if no agreement is reached, whereas if ASS(3)[15] is relaxed , war is only likely if an agreement is reached.

The amount of compensation C_1 determines whether ASS(3) holds, i.e. whether outcome (4) or outcome (5) is possible. Outcomes (1), (2) and (3), are not affected by the amount of compensation as long as ASS(1) holds.

An agreement increases the potential for conflict (given ASS(3) does not hold). Without an agreement Ethiopia will defect if its prior subjective probability that Egypt is *strong* is less

[14] The damage to Egypt if Ethiopia defects after an agreement is reached is greater than the damage it would suffer if Ethiopia takes unilateral actions without an agreement, even if compensation is deducted.

[15] The damage to Egypt if Ethiopia defects after an agreement is reached is smaller than the damage it would suffer if Ethiopia takes unilateral actions without an agreement, if compensation is deducted.

than $\frac{B_2 - C_2 - A}{W_H - A}$. After an agreement Ethiopia will defect if its prior subjective probability that Egypt is *strong* is less than $\frac{B_2 - C_3 - A}{W_H - A}$, which is greater by the difference of C_2 and C_3 given ASS(2). Therefore, there might be some truth in the notion that increasing the interdependency between or among sovereign countries via agreements increases the possibility of tensions between or among them.

There are no separating pure–strategy perfect Bayesian equilibria in this signalling game. In a separating equilibrium the Receiver has complete information about the Sender's type upon receiving the message and exploits this knowledge to his or her own benefit. Because ASS(3) cannot hold and be relaxed at the same time, one type would always have an incentive to deviate from the proposed equilibrium strategy.

Besides the prior subjective probability distribution about Egypt's strength r, Ethiopia's posterior beliefs about Egypt's strength p or q, and the amount or the compensation paid by Ethiopia to Egypt C_1, there are other parameters which affect the likely outcome of the game. These are the benefits of Ethiopia from defecting B_2, which are likely to increase with economic development in Ethiopia, the costs of facilities for Ethiopia to use more water C_2 which might be affected by the willingness of international institutions to finance water development projects in Ethiopia, the costs to Ethiopia of defecting on the agreement C_3, arbitration costs A, and high war costs W_H.

The following table gives an indication of the effects that an increase in C_2, C_3, B_2, and W_H have on the four potential outcomes of the game (via their effect on the threshold values of p and q which will determine Ethiopia's actions $\frac{B_2 - C_2 - A}{W_H - A}$ and $\frac{B_2 - C_3 - A}{W_H - A}$).
Arbitration costs is not included in the table, because it affects both nominator and denominator at the same time.

Table 2.1. The effect of an increase in parameters on the likelihood of possible outcomes

Parameters	Outcomes of the game			
	Bilateral cooperation	unilateral concession	conflict without agreement	conflict with agreement
C_2 cost of facilities to use more water	no effect	positive	negative	no effect
C_3 cost of facilities to use more water after agreement	positive	no effect	no effect	negative
B_2 benefits of taking unilateral action	negative	negative	positive	positive
W_H high war costs	positive	positive	negative	negative

2.6. Summary and Conclusion

We have attempted in this chapter to model strategic behaviour on reaching an agreement on sharing the waters of the River Nile between Egypt and Ethiopia. Our study contributes to the economic literature by providing an analysis, which takes strategic considerations and beliefs explicitly into account. Most other studies on the River Nile conflict have concentrated on the significant gains of cooperation, which should motivate downstream Egypt to accept upstream water development in Ethiopia. Results of a comparable signalling game applied to the River Euphrates indicate that beliefs do not matter in the strategic choices of Turkey, the key player. The main results on our River Nile model are: first, beliefs do matter and Ethiopia's beliefs about Egypt affect the behaviour of both Ethiopia and Egypt; secondly, conflict between Egypt and Ethiopia can be avoided if Ethiopia perceives Egypt as strong enough; thirdly, more accurate information about Egypt cannot prevent conflict between Ethiopia and Egypt, but can prevent war between them;

and finally being perceived as strong is even more important for Egypt when its vulnerability is increased as is expected to happen after an agreement is reached.

Beliefs do matter in this signalling game and being perceived as *strong* is more important than actually being *strong* to avoid conflict. War threats from downstream countries in this regard might not mean that a country is intending to carry out its threats, but to create or to maintain the image of a strong country that is able to defend its vital interests by military force if necessary. War is the least preferred outcome for both countries and would only happen if Ethiopia wrongly perceived Egypt as *weak*. This means that a lot can be gained in terms of avoiding wars, if the countries involved have more accurate information about each others actual military power.

More accurate information can avoid war between Ethiopia and Egypt, but can it avoid conflict? If Ethiopia correctly perceives Egypt as *weak* enough and defects on the agreement or takes unilateral actions, conflict will be inevitable. The increase or decrease of the potential of the possible outcomes of the game depends on the parameters, which determine Ethiopia's actions. An increase in the construction costs C_2 decreases the potential for conflict in terms of Ethiopia taking unilateral actions, but increases the potential for unilateral concessions by Ethiopia. An increase in the cost of defecting on an international agreement C_3 decreases the potential for conflict in terms of Ethiopia defecting on an international agreement and increases the potential for bilateral co-operation between Egypt and Ethiopia.

Egypt should not be tempted into an agreement by a big compensation $D_1 - C_1 \leq D_2$ unless C_3 is reasonably high. This is because the increased potential for conflict after an agreement is reached arises from ASS(2) that defecting after an agreement is less costly than defecting without an agreement. Even if there are no additional construction costs for defecting after an agreement, as stated in the model, there might be other, less tangible costs for the defecting country. The closer these costs are to the cost of defecting without an agreement the less will the potential for conflict increase after an agreements is reached. Shapland (1997) believes that Ethiopia would indeed pay a heavy cost in terms of its relations with the western world and the Arab world if it were to use water as a weapon against Egypt. So, even if the world community cannot directly force sovereign countries to honour their agreements, it has an important role in encouraging sustainable agreements by making defecting as costly as possible.

Appendix 2.1. : The River Nile in More Detail

The River Nile case was chosen for several reasons. First, it is one of the three problem rivers in the Middle East, where tensions are likely to occur in the near future. Secondly, it differs from the two other problem rivers in the Middle East, the Jordan and the Euphrates, in that the downstream country is the most developed and powerful country if compared to the upstream or midstream countries. To gain a better understanding of the unique problem of the River Nile, we start with a short description of the River Nile, identify then the main problems facing the basin countries, and discuss finally some of the several plans which can increase the available water to the basin countries. The information in this appendix depends on Smith and Al-Rawahy (1990), Swain (1997), Waterbury (1987), Whittington and McClelland (1992), and Whittington *et al.* (1995).

The Nile is the longest international river system in the world and flows through ten countries: Egypt, the Sudan, Ethiopia, Eritrea, Zaire, Uganda, Kenya, Tanzania, Rwanda and Burundi. Its two main tributaries, the Blue Nile and the White Nile converge near Khartoum, the capital of the Sudan, before it flows to Egypt and empties into the Mediterranean. The White Nile has its main source in the Equatorial Lakes (14%), mainly Lake Victoria and Lake Albert. Lake Victoria is shared by Uganda, Kenya and Tanzania and is fed by the Kagera River, which lies in Rwanda and Burundi. Lake Albert, which is fed by the Semliki River, is shared between Zaire and Uganda. An equivalent important source of the White Nile is the Sobat River (14%) which originates in Ethiopia. The Blue Nile has its main source in Lake Tana (59%) in the Ethiopian Highlands and the Atbara River which originates also in Ethiopia and contributes 13% to the Nile flow and empties into the main Nile in Sudan. Only some minor and very fluctuating flows originate in Eritrea. This makes Ethiopia by far the largest provider of the waters of the Nile (86%) compared to the contribution of the Equatorial Lakes region, which is merely 14%.

The White Nile and the Blue Nile do not only differ in their contribution to the Main Nile but in another important aspect. The White Nile, which is relatively clear of sediments, has a steady monthly flow, which was not very much affected by the recent droughts on East and Central Africa. The Blue Nile, which is heavily laden with sediments, has a very unsteady seasonal flow, which is reduced to a trickle in some months and many of its tributaries dry up entirely. But because the water also varies greatly between years the Blue Nile is much more affected by periodic droughts than the White Nile.

The ten sovereign countries, which share the River Nile are not equally dependent on its waters. Rwanda, Burundi and Zaire are indifferent about any water utilization in the Nile Basin and the Equatorial Lakes. Kenya, Tanzania and Uganda, which share the Equatorial Lakes, are only concerned about what they may be asked to give up if the Equatorial Lakes are used for storage of water to increase the downstream discharge of the Nile. Eritrea has only some minor flows, which would not affect any water utilization scheme in the Blue Nile. Although the source of the Blue Nile, Lake Tana, is highly inaccessible and the Blue Nile flows unimpeded through Ethiopia, its utilization becomes more and more important to achieve economic growth. This leaves only Egypt and the Sudan for which the utilization of the Nile is a matter of economic existence. Therefore only Egypt, Sudan and Ethiopia are the main beneficiaries or cost bearers of any utilization of the Nile Waters and the current amount of water available is by no means enough for the ambitious economic development plans for all three nations.

Although the total flow of the Nile is dependent on natural rainfall, water development projects can affect the amount of the water that can be captured for agricultural, industrial, hydroelectric and municipal uses. Most of the water development projects that were built in the past were for flood control, low flow augmentation and hydropower production, such as the Jebel Aulia Barrage in the Sudan, the Owen Falls Dam in Uganda and the Assuan High Dam in Egypt. Of the ongoing projects that aim to increase the amount of utilizable water of the Nile, especially in the Sudd swamps in southern Sudan, is the Jonglei Canal project, which was stopped in 1983 because of civil unrest. Although several plans have emerged, which could increase the amount of utilizable water of the Nile and thereby make the water scarcity problem less severe, the different countries involved have different preferences towards them.

1. The Completion of the Jonglei Canals I and II through the Sudd swamps in southern Sudan will drain water from the swamps, thereby increasing the downstream yield of the river by approximately 3.8 BCM and 3.2 BCM respectively.

2. An increase of storage at Lake Albert and a reduction of storage at Assuan High Dam would reduce evaporation losses.

3. The elimination of the Jebel Aulia Reservoir on the White Nile, where annual evaporation losses are currently 2.8 BCM per year would increase the available water.

4. The construction of the Blue Nile Reservoirs in Ethiopia, (based on a study by the US Bureau of Reclamation in 1964), which shift the over-year storage from the Assuan High Dam in southern Egypt to Ethiopia, where evaporation losses are 50% of those in Sudan and Egypt, could lead to water savings of 4 to 5 BCM per year, because of the lower evaporation rates and lower surface-to-volume ratios in the Canyon of the Blue Nile Reservoir.

Chapter 3

International Environmental Agreements in Water Resources with Unidirectional Externalities

3.1. Introduction

International Environmental Agreements (IEAs) are agreements among sovereign countries. There is no international authority that can guarantee the enforcement of an agreement or that can punish non-compliance of a sovereign country. Therefore, IEAs have to be voluntary and self-enforcing i.e., they have to form a non-cooperative equilibrium.

Self-enforcing IEAs might become even more difficult to sustain when the externality causing the inefficient outcome is unidirectional, as when countries share a river. Upstream countries affect downstream countries but not vice versa[16]. IEAs with unidirectional externalities not only face an asymmetry in payoffs, but also an asymmetry in enforcement power due to an asymmetry in the irreversibility of the players' actions[17]. Mäler (1990) has indicated that in the case of international rivers, the upstream country that causes the externality has no incentive to take the effect of its action on downstream countries into consideration and will not engage in any cooperation unless it is compensated in one way or the other. While Finus (2000) has discussed transfers as a measure to balance such asymmetries, others like Cesar and De Zeeuw (1996) and Spagnolo (1996) have investigated issue linkage. Although Finus (2000) considers transfers to be successful in balancing asymmetric payoffs, he regards them as being only limited effective in balancing the enforcement power among the players. Monetary transfers can be very effective in punishing a recipient country if it defects, but can increase the free-rider incentive of the

---

[16] In rare cases, downstream countries can also harm upstream countries. The construction of dams for irrigation, hydropower production or flood control might cause flooding upstream.

[17] An irreversible action means that a player can only increase his or her level of cooperation or keep it constant, but cannot reduce it.

donor country even more, and are therefore unsuitable to sustain an IEA on the side of the donors. In-kind transfers are less suited than monetary transfers to punish recipient countries because they might have included irreversible investments. They also cannot be used to punish donor countries for the same reasons as monetary transfers. It has been suggested that linking one issue that leads to an asymmetric welfare distribution to another issue that leads to a more or less reversed asymmetric welfare distribution might facilitate agreements. Although issue linkage can help to balance asymmetries in welfare distribution, this is not always the case. Results from Finus (2000) and Spagnolo (1996) imply that if issues are substitutes, linking them could facilitate cooperation, because punishment would become harder and defection would become less valuable. But if issues are complements, linking them has the opposite effect, because defections would become more valuable and punishment would become less costly.

This chapter is concerned with facilitating agreements over developing international rivers, when there are no suitable issue linkages available. We investigate the effect of another incentive, namely future benefits, on the possibility of reaching self-enforcing river sharing agreements. Our study on cooperation in water resources with unidirectional externalities draws on the work of Lockwood and Thomas (2002). They have used an infinite dynamic game to prove that the irreversibility of the players' actions prohibits the players from reaching sustainable full cooperation. We extend the analysis to consider the unidirectional externality in international rivers, which we express as an asymmetry in the irreversibility of the countries' actions. While many of the upstream country's actions or decisions are reversible at any time, the downstream country's actions or decisions are mostly irreversible. Let us consider, for example, that an upstream country builds a water storage system with the financial help of the downstream country in exchange for delivering a certain water quantity at specific times to its downstream neighbour. In this case, once the dam is built, the downstream country cannot reverse its action or decision. The upstream country, on the other hand can reverse its actions or decision simply by withholding water, delivering too much water, and/or by delivering the water at unsuitable times. After transforming the model of Lockwood and Thomas (2002) into a finite framework to make it more tractable, we demonstrate that the reversibility constraints on the players' actions determine *how* an agreement changes the incentive structure of the game and the ability of the countries to carry out threats. Once the future gains from maintaining an increasing level of mutual cooperation are insufficient to offset the incentive to deviate, the corresponding level of cooperation cannot be sustained in equilibrium. The objective then is to demonstrate that different levels of cooperation can be supported by different irreversibility constraints of the players' actions. We investigate how this will affect the feasibility of reaching self-enforcing IEAs for countries sharing a river. We therefore compare the sustainable level of cooperation for three different reversibility constraints on the players' actions in a two-player setting: symmetric reversibility, symmetric irreversibility and asymmetric irreversibility of the players' actions. Our results confirm that despite obvious gains from cooperation, reaching even a partial agreement is more difficult for countries with asymmetric irreversibility constraints on their actions, which result from the unidirectional externality from the upstream country to the downstream country, than for countries with symmetric irreversibility constraints on their actions, which result from reciprocal externalities between them. Interpreting the unidirectional externality in the context of irreversible actions has not been covered in the literature before as far as I know. Our analysis demonstrates the importance of this issue for sustainable international water agreements. The approach to adapt an infinite dynamic game into a finite dynamic game allows tractable analysis of a very complicated, but highly relevant extension of the Lockwood and Thomas (2002) model.

3.2. Reversibility Constraints and the Incentive Structure of a Game

Because credible threats and promises about future behaviour can influence current behaviour, partial or full cooperation can only be achieved, if defection by any player can be credibly punished by the other players. Partial cooperation means that some player chooses an action that leads to payoffs that are higher than in the stage-game Nash equilibrium. Full cooperation means that all players choose actions that maximize the joint payoffs of the players. The ability to carry out threats depends crucially on the irreversibility constraints on the players' actions and whether the intended cooperation is immediate or gradual. Immediate cooperation means that the intended level of cooperation is reached in one period i.e., there is no cooperation at all in the previous period. Gradual cooperation means that the intended level of cooperation is not reached in one period i.e., a lower level of cooperation has already been reached in the previous period. Three base cases can be distinguished: the PD game (symmetric reversibility actions), the LT game (symmetric irreversibility actions) and the AI game (asymmetric irreversibility actions). Although all three games have the basic structure of the Prisoners' Dilemma in each period, they differ in their ability to punish deviation, as will be shown in our model.

3.2.1. The Classical Prisoners' Dilemma Game (The PD Game)

According to Finus (2001), the Prisoners' Dilemma game is the most common way to present the difficulties of reaching self-enforcing IEA in the economic literature, whenever there are negative externalities involved. Suppose two countries share a water body with reciprocal externalities, such as a lake, where each country's action has an impact on the shared resource and therefore also on the other country. This happens if the two countries have unrestricted access to the lake and use it simultaneously. To prevent over-exploitation of the lake, for example, the countries might want to enter an agreement that restricts fishing trips. Although both countries will gain from mutual cooperation, each country gains even more if it free rides on the cooperation of the other country. If the game is played only once, then a Pareto-inferior equilibrium of no cooperation would be the unique outcome of the game, because the dominant strategy of each country is not to cooperate. If, on the other hand, the game is repeated infinitely[18], or for an uncertain time period, some strategic considerations of the countries could lead to more cooperative results. If, for example, both countries agree on a 'trigger strategy', where each country threatens to reverse to the stage-game Nash equilibrium, temporarily or permanently, if the other

[18] See Finus (2001) and Fudenberg (1991) for a detailed explanation of the Folk theorems.

country defects on the agreement, then above some critical discount factor, an efficient level of cooperation can be attained exactly and immediately. Full cooperative equilibria are supported by punishment phases. The punishment is credible, because the players' actions are reversible. The question is: what sort of cooperation can be sustained in equilibrium when the players' actions are irreversible? The answer comes from the second game, as presented by Lockwood and Thomas (2002) which we call 'the LT game' throughout this chapter.

3.2.2. The Lockwood and Thomas Game (The LT Game)

Lockwood and Thomas (2002) refer in their paper to games that have a Prisoners' Dilemma structure in every period, but actions are irreversible i.e., players can only increase their level of cooperation or keep it constant, but never reduce it, i.e.,

$$c_{i,t} \geq c_{i,t-1}, i = 1,2, t = 1,2,..., \quad \text{where } c_{i,t} \text{ is } i\text{'s action in period } t.$$

These irreversibility constraints imply that the game is dynamic rather than repeated, as the game changes over time, but the PD structure is preserved in all periods. The irreversibility constraint has two opposing effects. On the one hand it helps cooperation because deviation becomes less profitable; on the other hand it makes sustaining cooperation more difficult because punishment becomes less severe. Suppose again that two countries share a lake as in the previous example. This time, measures to prevent over-exploitation of the lake could include destroying fishing boats. Clearly, this action is rather difficult or expensive to reverse as punishment for deviation. Another example could be relocating industries away from the lake. Reversing the relocation as punishment for defection is not only expensive but also time consuming. Therefore, the only plausible punishment in this case would be to threaten to withhold benefits, which would accrue from increasing the level of mutual cooperation in the future. Lockwood and Thomas (2002) have established that although full cooperation cannot be sustained, gradual partial cooperation is still achievable, even if the players' actions are irreversible. But what sort of cooperation can be sustained if one player's actions are reversible while the other player's actions are irreversible? This is the main theme of this chapter and leads us to the third game which we call 'the AI game' (the Asymmetric Irreversibility game) throughout this chapter, because the players have asymmetric irreversibility constraints on their actions.

3.2.3. The Asymmetric Irreversibility Game (The AI Game)

This game refers to dynamic games, that have a Prisoners' Dilemma structure in every period, but where the actions of one player are irreversible, while the actions of the other player are reversible. Suppose two countries share a water body with unidirectional externalities, such as a river, which can only be accessed consecutively and not simultaneously, as was the case in the two previous games. In this case the access to the shared resource can be restricted by the upstream country, so that only the upstream country has an impact on the downstream country and not vice versa. Major development projects, such as dams, are usually under the control of the country where the dam is located. Cooperation by a downstream country can be in form of approving upstream water developments. International institutions do not assist in financing major water development projects, unless all riparian countries agree on the project. This is to guarantee the completion of the projects, which could otherwise be under constant threat from disapproving riparians. Therefore, an explicit approval from downstream countries is vital for upstream countries, especially if they depend on international financial aid to build their water projects. In return for their cooperation (to agree to upstream development projects) downstream countries can, for example, receive some international financial aid for water related or water unrelated projects, any other form of compensation, and the promise to get the water quantity and quality agreed upon. Countries sharing a river differ tremendously in their abilities to punish deviators. If the water development project is located upstream and the downstream country goes back on its agreement, it can be severely punished by the upstream country, which can release too much or too little water. On the other hand, if the upstream country defects on its agreement, the downstream country can only withhold future cooperation to punish the upstream country for not complying with the agreement. This implies that although the incentive to defect is as high for the upstream country in the AI game as in the PD game, and much higher than in the LT game, the ability of the downstream country to punish defection is the least in the AI game.

3.3. Related Literature

Schelling (1960) was among the first to realize the impact of irreversibility on cooperation.

> *"...What makes many agreements enforceable is only the recognition of future opportunities for agreement that will be eliminated if mutual trust is not created and maintained, and whose value outweigh the monetary gain from cheating in the present instance ... even if the future will bring no recurrence, it may still be possible to create equivalence of continuity by dividing the bargaining issue into consecutive parts..." (p.45)*

Admati and Perry (1991) and Marx and Matthews (1997) are some of those who have considered dynamic voluntary contribution games in their research.

Admati and Perry (1991) discuss public projects whose benefits accrue only after their total completion. As there are no enforceable contracts, contributions to the project are voluntary. Costs are sunk once they are made (irreversibility aspect). Each partner prefers to contribute nothing and to free ride on the contribution of the others. Do the dynamics of the game alleviate or aggravate this free-rider inefficiency compared to the static game? Admati and Perry (1991) use a discrete time two-person dynamic model where the players take turns to make contributions to the public project. At each stage the player with the turn decides how much to contribute. The closer the projects comes to its completion, one of the players will be willing to complete the project without the other. As both players realize this, each has a strong incentive to deviate in anticipation that the project will be completed by the other player. The punishment of free riding is to forego the completion benefit and to lose all previous contributions. Both deviator and punisher are worse off than the status quo. So if a player intends to deviate or suspects that any other player might deviate, he or she will prefer not to contribute at all and the project will not be built. The punishment increases over time after each sunk contribution costs. The analysis shows that if the cost of the project are not too high, the project will be completed in equilibrium. Otherwise, the project will not be completed in equilibrium even if it is socially desirable.

Marx and Matthews (1997) consider public projects, which generate a flow of public benefits. The players contribute simultaneously to a public project over a finite or infinite number of periods and observe only the aggregate of all players' past contributions. Each player receives benefits along the way that are linear in the sum of cumulative contributions, and an additional benefit when the project is completed. Marx and Matthews (1997) show that if the number of periods is large enough and the players are patient enough, or if the period length is short enough, then the project will eventually or asymptotically be completed. Although future contributors increase the incentive for current contributors to lower their contributions, the potential future contributors can punish this free-riding by lowering their own contributions in the future. Thus, the dynamics of the game can alleviate the free-riding inefficiency.

The paper of Lockwood and Thomas (2002) is closest to our study. They show that whether an efficient level of cooperation is achieved by infinitely repeating the game depends, among other things, on the reversibility of the players' actions each period. They demonstrate how irreversibility leads to gradual cooperation and that full cooperation is never reached in finite time. The shape of the payoff function determines at what level of cooperation the efficient symmetric equilibrium path converges. They show that if payoffs are smooth functions of actions, the efficient symmetric equilibrium path will converge to a level strictly below the full cooperation level, no matter how patient the players are. But if payoffs are linear up to some joint cooperation level and constant or decreasing thereafter, then above some critical discount factor full cooperation can be attained asymptotically. They extend the study to allow a small amount of reversibility i.e. the ability to reduce cooperation gradually over a period of time. Their results imply that reversibility improves on the irreversibility in terms of payoffs from the efficient equilibrium path and is therefore desirable. The main limitation of their study is that the actions of both players are equally irreversible. But this might not always be the case, as we show in our study.

Our study is an extension to the *symmetric* irreversibility of the players' actions case of Lockwood and Thomas (2002). We include *asymmetric* irreversibility of the players' action within a simplified, and therefore more tractable finite game. This enables us to compare the difficulties in reaching agreements on sharing water bodies with reciprocal externalities (symmetric irreversibility) to the difficulties in reaching agreements on sharing water bodies with unidirectional externalities (asymmetric irreversibility).

3.4. The Model

After describing the model in section 3.4.1., we specify the actions, parameters, assumptions and pure strategies in section 3.4.2..

3.4.1. Description of the Model

Lockwood and Thomas (2002) consider an infinite dynamic game in which each of two players decides on a level of cooperation, which can only be increased or kept constant. They show the effect of the irreversibility of the players' actions on the maximum level of cooperation that can be sustained in equilibrium. They characterize the efficient symmetric equilibrium path over time and determine to what limit it will converge eventually.

We attempt to show the effect of three different constraints on the reversibility of the players' actions, on the levels and types of cooperation that can be sustained in equilibrium. Our model is adapted from Lockwood and Thomas (2002) into a finite framework. Whereas the gains from following a cooperative strategy are the same for all games, the gains from defection and the resulting punishment differs a great deal across the games. This can be captured in a finite framework as well as in an infinite framework. For the purpose of this study, namely, to analyze the effect of different irreversibility constraints on the players' actions on the feasibility of IEAs, it is sufficient and more convenient to use a finite framework.

The games we analyze in our study have a classical Prisoners' Dilemma structure in every period i.e., the only subgame-perfect equilibrium that can be sustained in equilibrium, within a finite framework, is the Pareto-inferior Nash equilibrium of the stage game. To support a more preferred outcome as a subgame-perfect equilibrium, at least one other Nash equilibrium is needed in the stage game. A standard approach recommended by Gibbons (1992) and Finus (2001) is to extend the action space of the original PD game to add another Nash equilibrium to the original stage game.

To understand how punishment works in our analysis, we differentiate between three simultaneous move repeated (or dynamic) games, which all have a Prisoners' Dilemma structure in every period[19] as follows:

1. The repeated (or dynamic) finite PD game with a unique Nash equilibrium: If the stage game G has a unique Nash equilibrium (as in Figure A) then, for any *finite* period T, the repeated game G(T) has a unique subgame-perfect outcome: the Nash equilibrium of G is played in every stage t.

2. The repeated (or dynamic) infinite PD game with a unique Nash equilibrium: If the stage game G has a unique Nash equilibrium (as in Figure A) and is repeated *infinitely* then, there may be subgame-perfect outcomes of the infinitely repeated game where the

[19] For a more detailed explanation of this concept see Gibbons (1992) and Finus (2001).

Nash equilibrium of the stage game G is not an outcome in any of the stages t, provided that the discount factor δ is sufficiently high and both players adopt a 'trigger strategy'. Adopting a trigger strategy means that each player cooperates until someone defects, which triggers a switch to non-cooperation forever. In this example, playing C_i in every period is optimal if and only if

$$\frac{4}{1-\delta} \geq 5 + \frac{\delta}{1-\delta} \qquad \text{or} \qquad \delta \geq 4/5$$

	D	C
D	**1,1**	5,0
C	0,5	4,4

Figure A

	D	C	P
D	**1,1**	5,0	-2,-2
C	0,5	4,4	-2,-2
P	-2,-2	-2,-2	**-1,-1**

Figure B

3. The repeated (or dynamic) finite PD game with multiple Nash equilibria: If the stage game G has multiple Nash equilibria (as in Figure B), then, there may be subgame-perfect outcomes of the repeated game G(T) in which for any $t < T$ the outcome of stage t is not a Nash equilibrium of G. In Figure B the classic PD game has been extended to include a punishment action (P, P). We assume that the game is played twice without discounting. If the players agree to play (C,C) in the first period and (D,D) in the second period, the payoff stream is 4+1=5. If a player deviates in the first period he or she gets 5+(-1)=4. Since 5>4, free-riding can be deterred.

The idea of the two-stage repeated game is similar to the idea of the infinite repeated game: if the players cooperate today they get a high payoff tomorrow; otherwise they get a low payoff tomorrow. The difference between the infinitely repeated game and the two-stage repeated game is that the low-payoff equilibrium is artificially added to the stage game and represents the punishment for not cooperating in the first period. If we have more than two periods, then the low-payoff equilibrium can be interpreted as the forgone benefits of not cooperating in the future periods.

While Gibbons (1992) chooses to add a 'reward action' with a higher payoff Nash equilibrium than the original Nash equilibrium, Finus (2001) chooses to add a 'punishment action' with a lower payoff Nash equilibrium than the original Nash equilibrium. Both procedures allow results to be achieved in finite dynamic games, which originally had only one Pareto-inferior NE in every period, that are comparable to the results of infinite dynamic games with the same Pareto-inferior NE in every period. An important limitation of this approach (the finite approximation to infinite games) needs to be mentioned here: while the '*folk theorems* for infinitely repeated games allows in the limit of extreme patience virtually any payoff to be an equilibrium outcome (Fudenberg and Tirole (1991), this is not the case for finitely repeated games, even if they have more than one Nash equilibrium in every period. Nevertheless, the simplicity of this approach allows a tractable analysis of a very complicated, but highly relevant extension of the LT model. While reducing all future punishment actions to a single artificial action might be open to criticism, we consider it the most suitable approach from the expositional point of view.

We use Finus' (2001) version of this approach to analyse the effect of different irreversibility constraints on the players' actions on the feasibility of IEAs. The punishment action P represents future punishment in a static game. Consider, for example, that a major water development project in an upstream country could maximize the joint benefits of two countries sharing a river. The downstream country would assist in financing the project in exchange for a promised water quality/quantity or hydropower. The downstream country is afraid to help funding this upstream water project in case the upstream country defects on its promise to deliver the promised water quality/quantity and/or hydropower once the water project is completed. Cooperation is rather unlikely to happen under these circumstances. If, on the other hand, there are plenty of opportunities for cooperation between the two countries (a number of smaller water development project, for example), cooperation could happen gradually over a longer period of time. Starting with the first project, the downstream country can test the upstream country on its compliance with the agreement. Any following project will only start if the upstream country does not defect on any previous agreement. If the upstream country defects in any period on any of the previous agreements, then cooperation between the two countries will stop forever from the following period onwards. The upstream country definitely gains by defecting on any previous agreements, but it loses all future benefits that would happen if cooperation would not only continue over time but would also increase whenever a new project is added. Defection in any period is followed by punishment in the following period. The payoffs from the punishment action (P,P) are therefore the foregone benefits from not cooperating in the future.

The adapted model is a finite dynamic game in which players' payoffs are common knowledge and moves within a period are simultaneous. For simplicity, there are only four levels of cooperation: no cooperation, a low level of cooperation, a high level of cooperation, and full cooperation. This is the minimum number of cooperation levels necessary to achieve the objective of this study i.e., to show how different irreversibility constraints support different levels of cooperation, but any larger number of cooperation levels will do. As our study tackles three different irreversibility constraints, at least three levels besides the non-cooperation level are needed. These irreversibility constraints are:

(1) the symmetric reversibility of the players' actions as in the classic PD game

(2) the symmetric irreversibility of the players' actions as in the LT game

(3) the asymmetric irreversibility of the players' actions as in the AI game

3.4.2. The Actions, Parameters, Assumptions and Pure Strategies

The Actions

The following capital letters are used for the actions that are available to the two players:

P future punishment in a static game

C_0 no cooperation

C_1 low level of cooperation

C_2 high level of cooperation

C_∞ full cooperation

The Parameters

The small Latin letters stand for the payoffs of the players' actions. The small Greek letter δ stands for the common discount factor. It is assumed that (P, P) is an available NE in every stage game. Assumptions 1 to 11 guarantee for each player, that P is always a best response to P, and that P is not a best response to C_0, C_1, C_2, or C_∞.

The per-period payoff for each player $\pi(C_{1i}, C_{2i})$[20] depends on his or her own level of cooperation and the level of cooperation of the other player. It is assumed that π is strictly decreasing in that player's level of cooperation and strictly increasing in the other player's level of cooperation (until some upper level of full cooperation is reached). This ensures the Prisoners' Dilemma structure in every period. The payoffs over the finite horizon are discounted by the common discount factor δ. The payoffs from the each cooperation level are assumed to be symmetric to restrict attention to the enforceability problem. The punishment payoffs on the other hand can be symmetric as well as asymmetric, to show the effect of the irreversibility constraints on the ability to punish defectors. Bold payoffs show the Nash equilibria of the stage game. The payoffs $\pi(P, z)$ and $\pi(z, P)$ where z is anything but P are constructed in such a way that (P,P) is an available Nash equilibrium in every period. To simplify and to keep notation to a minimum, $\pi(P, z)$ and $\pi(z, P)$ equal x,y. Neither the x's nor the y's have to be equal; the only constraint they must fulfil is to be smaller than the minimum of all other payoffs.

[20] The subscripts denote both the players (1 and 2) and the level of cooperation (high or low).

	C_∞	C_2	C_1	C_0	P
C_∞	a, a	l, m	p, q	r, s	x, y
C_2	m, l	b, b	j, k	n, o	x, y
C_1	q, p	k, j	c, c	h, i	x, y
C_0	s, r	o, n	i, h	**d, d**	x,y
P	x,y	x, y	x, y	x, y	**e, f**

Figure. 3.1.: Payoffs in period $t \leq T$

The Assumptions

The assumptions guarantee that the modified Prisoners' Dilemma structure is preserved in each stage game.

(1)	$a > b > c > d > e, f$	payoffs from mutual coopration or mutual punishment
(2)	$s > q > m > a$	payoffs to one player if the other plays C_∞
(3)	$o > k > b > l$	payoffs to one player if the other plays C_2
(4)	$i > c > j > p$	payoffs to one player if the other plays C_1
(5)	$d > h > n > r$	payoffs to one player if the other plays C_0
(6)	$s > o > i > d$	payoffs to the other player if one player plays C_0
(7)	$q > k > c > h$	payoffs to the other player if one player plays C_1
(8)	$m > b > j > n$	payoffs to the other player if one player plays C_2
(9)	$a > l > p > r$	payoffs to the other player if one player plays C_∞
(10)	$0 \leq \delta \leq 1$	the discount factor

(11) $x, y < min\ [a,b,c,d,e,f,g,h,i,j,k,l,m,n,o,p,q,r,s]$ to guarantee that (e,f) is NE

The Pure Strategy of the Players

A strategy is a complete plan of action that specifies a feasible action for each player in every contingency in which the player might have to act. The players move simultaneously in each period, and the outcome of each period is observed before the next period begins. This is important because the payoffs for the whole finitely repeated game are simply the sum of payoffs from all stages of the game. Therefore, at the beginning of each period, the player knows the complete history of the game so far, which is the sequence of all action combinations played up to this period. With the history of the game in mind, the player chooses at the beginning of each period between playing cooperatively (S) or playing non-cooperatively (D), thereby triggering a punishment that starts from the following period. The superscript on S^t means a strategy leading to *immediate* cooperation and the subscript on S_t means a strategy leading to *gradual* cooperation. The ability to defect differs according to the irreversibility constraints and changes through the different levels and types of cooperation. The superscript on D^t means deviating from an *immediate* equilibrium strategy and the subscript D_t means deviating from a *gradual* equilibrium strategy. For example, in the PD game each player can defect by reducing his or her level of cooperation up to the non-cooperation level whatever level of cooperation has been reached so far, and whether the cooperation has been *immediate* or *gradual*. In the LT game, on the other hand, each player can only defect by keeping his or her level of cooperation constant. Therefore, not only the level of cooperation, but also whether the cooperation is *immediate* or *gradual* determines what level of cooperation can be supported in equilibrium. Finally, the AI game is a combination of the PD game and the LT game. While one player can defect in any period by reducing his or her level of cooperation up to the non-cooperative level (whether the cooperation has been *immediate* or *gradual)*, the other player can only defect by keeping his or her level of cooperation constant, as long as full cooperation has not been reached yet.

We need to remember, that the t in C_t denotes the level of cooperation, while the t in S_t indicates the time period. As this is a finite game, T denotes the last period.

Each player plays the following general strategy for *immediate* cooperation:

For $t < T$, S^t={ play C_t if both players have played C_t in every previous period, otherwise play P}

For $t = T$, S^t={ play the stage game Nash equilibrium[21] with the highest joint payoff if both players have played C_t in every previous period, otherwise play P}

$S = (S^t)_{t=0,1,...,T}$.

Each player plays the following general strategy for *gradual* cooperation:

For $t < T$, S_t={ play C_t if both players have played C_τ in every previous period $\tau \le t-1$, until the intended level of cooperation is reached, thereafter play C_t if both players have played C_t in every previous period after reaching the intended level of cooperation, otherwise play P}

For $t = T$, S_t={ play the stage game Nash equilibrium with the highest joint payoffs if both players have played C_t since reaching the intended level of cooperation, otherwise play P}

[21] In a Nash equilibrium (NE) each player's strategy must be a best response to the other players' strategies, so that no player chooses a strictly dominated strategy.

$S = (S_t)_{t=0,1,\ldots,T}$.

3.5. The Analysis of the Model

The analysis of the model proceeds in four steps, as follows:

First, we determine the equilibrium conditions that support *immediate* and *gradual* cooperation for each of the three nonzero levels of cooperation in the three base cases: the PD game, the LT game and the AI game.

Secondly, we determine the *necessary punishment* that supports *immediate* and *gradual* cooperation for each of the three nonzero levels of cooperation in the three base cases: the PD game, the LT game and the AI game for the specific case where T=4.

Thirdly, we determine the *credible punishment* according to the chosen definition. The interpretation used in this analysis is *the forgone future benefits* i.e., the discounted present value of the payoffs from the Nash equilibrium of the stage game minus the discounted present value of the payoffs from continuing increasing cooperation. This interpretation allows the available punishment to differ according to the irreversibility constraints of the players' actions and according to the types and levels of cooperation, whether full or partial, *immediate* or *gradual.*

In the end, we determine the types and levels of cooperation that can be sustained in equilibrium for each of the three base cases by comparing the *necessary punishment* from the second step with the *credible punishment* from the third step.

3.5.1. The Equilibrium Conditions for the Games

A necessary and sufficient condition for any equilibrium outcome is that unilateral deviation is not profitable i.e., the payoffs from defecting must be smaller than the cooperation payoffs. We consider only the most profitable defections. Therefore, we assume that defection takes place only from the intended level of cooperation that we want to support and not from any previous levels. We assume also that the punishment P is an available NE in every period of the game. Its value depends on the interpretation of the punishment.

It does not matter whether the PD game or the LT game is solved for the row-player or the column-player, because the actions and constraints of both players are symmetric. But once the AI game is solved for the player with the reversible actions, there is no need to solve it for the player with the irreversible action, as we need only to consider the case with the most profitable defection. Throughout this chapter we assume in the AI game that the actions of the column-player are reversible and that the actions of the row-player are irreversible. To simplify notation and to be able to make comparisons among the three different games, we solve all the games for the column-player.

3.5.1.1. The Equilibrium Conditions for the PD Game

To support S^{∞} as an equilibrium strategy for the PD game i.e., a strategy leading *immediately* to full cooperation in all periods but the last one, defection must not be profitable in any period to any of the players.

> $S^{\infty(PD)} \Rightarrow$ play C_{∞} in period $t=1$; then play C_{∞} in period t as long as (C_{∞}, C_{∞}) has been played in period $t-1$ and $t<T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_{∞}, C_{∞}) has been played in period $T-1$, otherwise play P.

The outcome of each period is observed before the next period begins. At the beginning of each period each player decides whether to play C_∞ or whether to defect by playing $C < C_\infty$, knowing that (P, P) will be played from the following period. Each player plays C_∞ as long as his or her payoffs from the equilibrium $\pi(S^\infty)$ are bigger than his or her payoffs from defecting $\pi(D^\infty)$. Therefore, *f* must satisfy $\pi(S^\infty) > \pi(D^\infty)$, where

$$\pi(S^\infty) = \sum_{t=1}^{T-1} \delta^{t-1} a + \delta^{T-1} d \qquad (1)$$

$$\pi(D^\infty) = s + \sum_{t=2}^{T} \delta^{t-1} f \qquad (2)$$

	C_∞	C_0	P
C_∞	a, a	r, s	x, y
C_0	r, s	**d, d**	x,y
P	x, y	x, y	**e, f**

Figure 3.2.:The PD^∞ game in period $t = 1$ and in period $t \geq 2$

To support S_∞ as an equilibrium strategy for the PD game i.e., a strategy leading *gradually* to full cooperation except for the last period, defection must not be profitable to any of the players.

$S_{\infty(PD)} \Rightarrow$ play C_1 in period $t = 1$; then play C_2 in period $t = 2$ if (C_1, C_1) has been played in $t = 1$, otherwise play P; then play C_∞ in period $t = 3$ if (C_2, C_2) has been played in $t = 2$, otherwise play P; then play (C_∞, C_∞) in period $t > 3$ as long as (C_∞, C_∞) has been played in the previous period and $t < T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_∞, C_∞) has been played in period $T-1$, otherwise play P.

At the beginning of the first three periods, each player decides whether to increase the level of cooperation by playing $C_t > C_{t-1}$ for the first three periods and C_∞ thereafter except for the last period, or whether to defect by playing $C_t < C_{t-1}$ in the third period (we are only considering the most profitable defection), knowing that (P,P) will be played from the following period onwards. Each player plays $C_t > C_{t-1}$ for the first three periods and C_∞ thereafter, except for the last period, as long as his or her payoffs from the equilibrium $\pi(S_\infty)$ are bigger than his or her payoffs from defecting $\pi(D_\infty)$. Therefore, f must satisfy $\pi(S_\infty) > \pi(D_\infty)$, where

$$\pi(S_\infty) = c + \delta b + \sum_{t=3}^{T-1} \delta^{t-1} a + \delta^{T-1} d \qquad (3)$$

$$\pi(D_\infty) = c + \delta b + \delta^2 s + \sum_{t=4}^{T} \delta^{t-1} f \qquad (4)$$

	C_∞	C_2	C_1	C_0	P
C_∞	a, a	l, m	p, q	r, s	x, y
C_2	m, l	b, b	j, k	n, o	x, y
C_1	q, p	k, j	c, c	h, i	x, y
C_0	s, r	o, n	i, h	**d, d**	x, y
P	x, y	x, y	x, y	x, y	**e, f**

Figure 3.3.: The PD_∞ game in period $t = 1$ and in period $t \geq 2$

To support S^2 as an equilibrium strategy for the PD game i.e., a strategy leading *immediately* to a high level of cooperation in all periods but the last one, defection must not be profitable to any of the players.

> $S^{2(PD)} \Rightarrow$ play C_2 in period $t=1$; then play C_2 in period t as long as (C_2, C_2) has been played in period $t-1$ and $t<T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_2, C_2) has been played in period $T-1$, otherwise play P.

At the beginning of each period each player decides whether to play C_2 or whether to defect by playing $C<C_2$, knowing that (P,P) will be played from the following period. Each player plays C_2 as long as his or her payoffs from the equilibrium $\pi(S^2)$ are bigger than his or her payoffs from defecting $\pi(D^2)$. Therefore, f must satisfy $\pi(S^2)>\pi(D^2)$ where,

$$\pi(S^2)=\sum_{t=1}^{T-1}\delta^{t-1}b+\delta^{T-1}d \qquad (5)$$

$$\pi(D^2)=o+\sum_{t=2}^{T}\delta^{t-1}f \qquad (6)$$

	C_2	C_0	P
C_2	b, b	n, o	x, y
C_0	o, n	**d, d**	x, y
P	x, y	x, y	**e, f**

Figure 3.4.: The PD^2 game in period $t=1$ and in period $t \geq 2$

To support S_2 as an equilibrium strategy for the PD game i.e., a strategy leading *gradually* to a high level of cooperation in all periods but the last one, defection must not be profitable to any of the players.

> $S_{2(PD)} \Rightarrow$ play C_1 in period $t = 1$; then play C_2 in period $t = 2$ if (C_1, C_1) has been played in $t = 1$, otherwise play P; then play C_2 in period $t > 2$ as long as (C_2, C_2) has been played in the previous period and $t < T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_2, C_2) has been played in period $T-1$, otherwise play P.

At the beginning of the first two periods, each player decides whether to increase the level of cooperation by playing $C_t > C_{t-1}$ for the first two periods and C_2 thereafter except for the last period, or whether to defect by playing $C_t < C_{t-1}$ in the second period (we are only considering the most profitable defection), knowing that (P, P) will be played from the following period onwards. Each player plays $C_t > C_{t-1}$ for the first two periods and C_2 thereafter except for the last period, as long as his or her payoffs from the equilibrium $\pi(S_2)$ are bigger than his or her payoffs from defecting $\pi(D_2)$. Therefore, f must satisfy $\pi(S_2) > \pi(D_2)$, where

$$\pi(S_2) = c + \sum_{t=2}^{T-1} \delta^{t-1} b + \delta^{T-1} d \qquad (7)$$

$$\pi(D_2) = c + \delta o + \sum_{t=3}^{T} \delta^{t-1} f \qquad (8)$$

	C_2	C_1	C_0	P
C_2	b, b	j, k	n, o	x, y
C_1	k, j	c, c	h, i	x, y
C_0	o, n	i, h	**d, d**	x, y
P	x, y	x, y	x, y	**e, f**

Figure 3.5: The PD_2 game in period $t = 1$ and in period $t \geq 2$

To support S^1 as an equilibrium strategy for the PD game i.e., a strategy leading *immediately* to a low level of cooperation in all periods but the last one, defection must not be profitable to any of the players. As the low level of cooperation is by assumption the first level of cooperation, the equilibrium strategy leading to this first level of cooperation *gradually* is equivalent to the equilibrium strategy leading to the first level of cooperation *immediately* i.e., $S_{1(PD)} = S^{1(PD)}$. Therefore, there is no need to describe the equilibrium strategy leading to *gradual* cooperation separately.

$S^{1(PD)} \Rightarrow$ play C_1 in period $t = 1$; then play C_1 in period $t > 1$ as long as (C_1, C_1) has been played in period $t-1$ and $t < T$, otherwise play P from period *t* onwards; then play C_0 in the last period *T* if (C_1, C_1) has been played in period $T-1$, otherwise play P.

At the beginning of each period, each player decides whether to play C_1 or whether to defect by playing $C < C_1$, knowing that (P, P) will be played from the following period. Each player plays C_1 as long as his or her payoffs from the equilibrium $\pi(S^1)$ are bigger than his or her payoffs from defecting $\pi(D^1)$. Therefore, *f* must satisfy $\pi(S^1) > \pi(D^1)$, where

$$\pi(S^1) = \sum_{t=1}^{T-1} \delta^{t-1} c + \delta^{T-1} d \qquad (9)$$

$$\pi(D^1) = i + \sum_{t=2}^{T} \delta^{t-1} f \qquad (10)$$

	C_1	C_0	P
C_1	c, c	h, i	x, y
C_0	i, h	**d, d**	x, y
P	x, y	x, y	**e, f**

Figure 3.6.: The PD^1 game in period $t = 1$ and in period $t \geq 2$

3.5.1.2. The Equilibrium Conditions for the LT Game

To support S^∞ as an equilibrium strategy for the LT game i.e., a strategy leading *immediately* to full cooperation for all periods, defection in the *first* period must not be profitable to any of the players.

> $S^{\infty(LT)} \Rightarrow$ play C_∞ in period $t=1$; then play C_∞ as long as (C_∞, C_∞) has been played in period $t-1$ and $t < T$, otherwise play P from period t onwards; then play C_∞ in the last period T if (C_∞, C_∞) has been played in period $T-1$, otherwise play P.

At the beginning of the *first* period, each player decides whether to play C_∞ or whether to defect by playing $C < C_\infty$, knowing that (P, P) will be played from the following period. Each player plays C_∞, if his or her payoffs from the equilibrium $\pi(S^\infty)$ are bigger than his or her payoffs from defecting $\pi(D^\infty)$. Therefore, *f* must satisfy $\pi(S^\infty) > \pi(D^\infty)$, where

$$\pi\left(S^{\infty}\right)=\sum_{t=1}^{T}\delta^{t-1}a \tag{11}$$

$$\pi\left(D^{\infty}\right)=s+\sum_{t=2}^{T}\delta^{t-1}f \tag{12}$$

	C_{∞}	C_0	P
C_{∞}	a, a	r, s	x, y
C_0	r, s	**d, d**	x, y
P	x, y	x, y	**e, f**

Figure 3.7.: The LT^{∞} game in period $t=1$

	C_{∞}	C_0	P
C_{∞}	a, a	r, s	x, y
P	x, y	x, y	**e, f**

Figure 3.8.: The LT^{∞} game in period $t \geq 2$ after defection in period $t=1$

	C_{∞}	P
C_{∞}	a, a	x, y
P	x, y	**e, f**

Figure 3.9.: LT^{∞} game in period $t \geq 2$ without defection in period $t=1$

To support S_∞ as an equilibrium strategy for the LT game i.e., a strategy leading *gradually* to full cooperation, defection in the *third* period must not be profitable to any of the players.

> $S_{\infty(LT)} \Rightarrow$ play C_1 in period $t = 1$; then play C_2 in period $t = 2$ if (C_1, C_1) has been played at $t = 1$, otherwise play P; then play C_∞ in period $t = 3$ if (C_2, C_2) has been played in $t = 2$, otherwise play P; then play C_∞ in period $t > 3$ as long as (C_∞, C_∞) has been played in the previous period and $t < T$, otherwise play P from period *t* onwards; then play C_∞ in the last period *T* if (C_∞, C_∞) has been played in period $T-1$, otherwise play P.

At the beginning of the first three periods, each player decides whether to increase the level of cooperation by playing $C_t > C_{t-1}$ for the first three periods and C_∞ thereafter, or whether to defect by playing $C_t = C_{t-1}$ in the third period (we are only considering the most profitable defection), knowing that (P, P) will be played from the following period onwards. Each player plays $C_t > C_{t-1}$ for the first three periods and C_∞ thereafter as long as his or her payoffs from the equilibrium $\pi(S_\infty)$ are bigger than his or her payoffs from defecting $\pi(D_\infty)$. Therefore, *f* must satisfy $\pi(S_\infty) > \pi(D_\infty)$, where

$$\pi(S_\infty) = c + \delta b + \sum_{t=3}^{T} \delta^{t-1} a \tag{13}$$

$$\pi(D_\infty) = c + \delta b + \delta\delta^2 m + \sum_{t=4}^{T} \delta^{t-1} f \tag{14}$$

	C_∞	C_2	P
C_∞	a, a	l, m	x, y
C_2	m, l	b, b	x, y
P	x, y	x, y	**e, f**

Figure 3.10.: The LT_∞ game in period $t=3$

	C_∞	C_2	P
C_∞	a, a	l, m	x, y
P	x, y	x, y	**e, f**

Figure 3.11.: The LT_∞ game in period $t \geq 4$
after defection in period $t=3$

	C_∞	P
C_∞	a, a	x, y
P	x, y	**e, f**

Figure 3.12.: The LT_∞ game in period $t \geq 4$
without defection in period $t=3$

To support S^2 as an equilibrium strategy for the LT game i.e., a strategy leading *immediately* to a high level of cooperation for all periods, defection in the *first* period must not be profitable to any of the players.

> $S^{2(LT)} \Rightarrow$ play C_2 in period $t=1$; then play C_2 as long as (C_2, C_2) has been played in period $t-1$ and $t<T$, otherwise play P from period t onwards; then play C_2 in the last period T if (C_2, C_2) has been played in period $T-1$, otherwise play P.

At the beginning of the *first* period, each player decides whether to play C_2 or whether to defect by playing $C < C_2$, knowing that (P, P) will be played from the following period. Each player plays C_2, if his or her payoffs from the equilibrium $\pi(S^2)$ are bigger than his or her payoffs from defecting $\pi(D^2)$. Therefore, *f* must satisfy $\pi(S^2) > \pi(D^2)$, where

$$\pi(S^2) = \sum_{t=1}^{T} \delta^{t-1} b \qquad (15)$$

$$\pi(D^2) = o + \sum_{t=2}^{T} \delta^{t-1} f \qquad (16)$$

	C_2	C_0	P
C_2	b, b	n, o	x, y
C_0	o, n	**d,d**	x, y
P	x, y	x, y	**e, f**

Figure 3.13.: The LT^2 game in period $t = 1$

	C_2	C_0	P
C_2	b, b	n, o	x, y
P	x, y	x, y	**e, f**

Figure 3.14.: LT^2 game in period $t \geq 2$
after defection in period $t = 1$

	C_2	P
C_2	b, b	x, y
P	x, y	**e, f**

Figure 3.15.: LT^2 game in period $t \geq 2$ without defection in period $t = 1$

To support S_2 as an equilibrium strategy for the LT game i.e., a strategy leading *gradually* to a high level of cooperation, defection in the *second* period must not be profitable to any of the players.

> $S_{2(LT)} \Rightarrow$ play C_1 in period $t = 1$; then play C_2 in period $t = 2$ if (C_1, C_1) has been played at $t = 1$, otherwise play P; then play C_2 in period $t > 2$ as long as (C_2, C_2) has been played in the previous period and $t < T$, otherwise play P from period t onwards; then play C_2 in the last period T if (C_2, C_2) has been played in period $T-1$,otherwise play P.

At the beginning of the first two periods, each player decides whether to increase the level of cooperation by playing $C_t > C_{t-1}$ for the first two periods and C_2 thereafter, or whether to defect by playing $C_t = C_{t-1}$ in the second period (we are only considering the most profitable defection), knowing that (P, P) will be played from the following period onwards. Each player plays $C_t > C_{t-1}$ for the first two periods and C_2 thereafter as long as his or her payoffs from the equilibrium $\pi(S_2)$ are bigger than his or her payoffs from defecting $\pi(D_2)$. Therefore, *f* must satisfy $\pi(S_2) > \pi(D_2)$, where

$$\pi(S_2) = c + \sum_{t=2}^{T} \delta^{t-1} b \qquad (17)$$

$$\pi(D_2) = c + \delta k + \sum_{t=3}^{T} \delta^{t-1} f \qquad (18)$$

	C_2	C_1	P
C_2	b, b	j, k	x, y
C_1	k, j	c, c	x, y
P	x, y	x, y	**e, f**

Figure 3.16.: The LT_2 game in period $t = 2$

	C_2	C_1	P
C_2	b, b	j, k	x, y
P	x, y	x, y	**e, f**

Figure 3.17.: The LT_2 game in period $t \geq 3$
after defection in period $t = 2$

	C_2	P
C_2	b, b	x, y
P	x, y	**e, f**

Figure 3.18.: The LT_2 game in period $t \geq 3$
without defection in period $t = 2$

To support S^1as an equilibrium strategy for the LT game i.e., a strategy leading *immediately* to a low level of cooperation for all periods, defection in the *first* period must not be profitable to any of the players. As the low level of cooperation is by assumption the first level of cooperation, the equilibrium strategy leading to the first level of cooperation *gradually* is equivalent to the equilibrium strategy leading to the first level of cooperation *immediately* i.e., $S_{1(LT)} = S^{1(LT)}$. Therefore, there is no need to describe the strategy leading *gradually* to a low level of cooperation separately.

> $S^{1(LT)} \Rightarrow$ play C_1 in period $t=1$, then play C_1 as long as (C_1, C_1) has been played in period $t-1$ and $t<T$, otherwise play P from period t onwards; then play C_1 in the last period T if (C_1, C_1) has been played in period $T-1$, otherwise play P.

The outcome of each period is observed before the next period begins. At the beginning of the *first* period, each player decides, whether to play C_1 or whether to defect by playing $C < C_1$, knowing that (P, P) will be played from the following period. Each player plays C_1, if his or her payoffs from the equilibrium $\pi(S^1)$ are bigger than his or her payoffs from defecting $\pi(D^1)$. Therefore, *f* must satisfy $\pi(S^1) > \pi(D^1)$, where

$$\pi(S^1) = \sum_{t=1}^{T} \delta^{t-1} c \tag{19}$$

$$\pi(D^1) = i + \sum_{t=2}^{T} \delta^{t-1} f \tag{20}$$

	C_1	C_0	P
C_1	c, c	h, i	x, y
C_0	i, h	**d,d**	x, y
P	x, y	x, y	**e, f**

Figure 3.19.: The LT^1 game in period $t=1$

	C_1	C_0	P
C_0	c, c	h, i	x, y
P	x, y	x, y	**e, f**

Figure 3.20.: The LT^1 game in period $t \geq 2$ after defection in period $t = 1$

	C_1	P
C_1	c, c	x, y
P	x, y	**e, f**

Figure 3.21.: The LT^1 game in period $t \geq 2$ without defection in period $t = 1$

3.5.1.3. The Equilibrium Conditions for the AI Game

To support S^{∞} as an equilibrium strategy for the AI game i.e., a strategy leading *immediately* to full cooperation in all periods but the last one, defection must not be profitable to any of the players.

> $S^{\infty(AI)} \Rightarrow$ play C_{∞} in period $t = 1$; then play C_{∞} in period t as long as (C_{∞}, C_{∞}) has been played in period $t-1$ and $t < T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_{∞}, C_{∞}) has been played in period $T-1$, otherwise play P.

As we are only considering the case with the most profitable defection, the game will be solved for the player with the reversible action i.e., the column player. At the beginning of each period the player with the reversible action decides whether to play C_{∞} or whether to defect by playing $C < C_{\infty}$, knowing that (P, P) will be played from the following period onwards. The player with the reversible action plays C_∞ as long as his or her payoffs from the equilibrium strategy $\pi(S^\infty)$ are bigger than his or her payoffs from defecting $\pi(D^\infty)$. Therefore, *f* must satisfy $\pi(S^\infty) > \pi(D^\infty)$, where

$$\pi(S^\infty) = \sum_{t=1}^{T-1} \delta^{t-1} a + \delta^{T-1} s \tag{21}$$

$$\pi(D^\infty) = s + \sum_{t=2}^{T} \delta^{t-1} f \tag{22}$$

	C_∞	C_0	P
C_∞	a, a	r, s	x, y
C_0	r, s	**d,.d**	x, y
P	x, y	x, y	**e,.f**

Figure 3.22.: The AI^∞ game in period $t = 1$

	C_∞	C_0	P
C_∞	a, a	r, s	x, y
P	x, y	x, y	**e, f**

Figure 3.23.: The AI^∞ game in period $t \geq 2$ after or without defection in period $t = 1$

To support S_∞ as an equilibrium strategy for the AI game i.e., a strategy leading *gradually* to full cooperation except for the last period, defection must not be profitable to any of the players.

> $S_{\infty(AI)} \Rightarrow$ play C_1 in period $t=1$; then play C_2 in period $t=2$ if (C_1, C_1) has been played at $t=1$, otherwise play P; then play C_∞ in period $t=3$ if (C_2, C_2) has been played in $t=2$, otherwise play P; then play (C_∞, C_∞) in period $t>3$ as long as (C_∞, C_∞) has been played in the previous period and $t<T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_∞, C_∞)has been played in period $T-1$,otherwise play P.

As before, the game will be solved for the player with the most profitable defection, i.e. the player with the reversible action. At the beginning of the first three periods, the player with the reversible action decides whether to increase the level of cooperation by playing $C_t > C_{t-1}$ for the first three periods and C_∞ thereafter, or whether to defect by playing $C_t = C_{t-1}$ in the *third* period (we are only considering the most profitable defection), knowing that (P, P) will be played from the following period onwards. He or she plays $C_t > C_{t-1}$ for the first three periods and C_∞ thereafter as long as his or her payoffs from the equilibrium $\pi(S_\infty)$ are bigger than his or her payoffs from defecting $\pi(D_\infty)$. Therefore, f must satisfy $\pi(S_\infty) > \pi(D_\infty)$, where

$$\pi(S_\infty) = c + \delta b + \sum_{t=3}^{T-1} \delta^{t-1} a + \delta^{T-1} s \tag{23}$$

$$\pi(D_\infty) = c + \delta b + \delta^2 s + \sum_{t=4}^{T} \delta^{t-1} f \tag{24}$$

	C_∞	C_2	C_1	C_0	P
C_∞	a, a	l, m	p, q	r, s	x, y
C_2	m, l	b, b	j, k	n, o	x, y
P	x, y	x, y	x, y	x, y	**e, f**

Figure 3.24.: The AI_∞ game in period $t = 3$

	C_∞	C_2	C_1	C_0	P
C_∞	a, a	l, m	p, q	r, s	x, y
P	x, y	x, y	x, y	x, y	**e, f**

Figure 3.25.: The AI_∞ in period $t \geq 4$
after or without defection in period $t = 3$

To support S^2 as an equilibrium strategy for the AI game i.e., a strategy leading *immediately* to a high level of cooperation for all periods, defection in all periods must not be profitable the player with the reversible action.

> $S^{2(AI)} \Rightarrow$ play C_2 in period $t = 1$; then play C_2 as long as (C_2, C_2) has been played in period $t-1$ and $t < T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_2, C_2) has been played in period $T-1$, otherwise play P.

At the beginning of the *each* period, the player with the reversible action decides, whether to play C_2 or whether to defect by playing $C < C_2$, knowing that (P, P) will be played from the following period. The player with the reversible action plays C_2, if his or her payoffs from the equilibrium strategy $\pi(S^2)$ are bigger than his or her payoffs from defecting $\pi(D^2)$. Therefore, f must satisfy $\pi(S^2) > \pi(D^2)$, where

$$\pi(S^2)=\sum_{t=1}^{T-1}\delta^{t-1}b+\delta^{T-1}o \qquad (25)$$

$$\pi(D^2)=o+\sum_{t=2}^{T}\delta^{t-1}f \qquad (26)$$

	C_2	C_0	P
C_2	b, b	n, o	x, y
C_0	o, n	**d, d**	x, y
P	x, y	x, y	**e, f**

Figure 3.26.: The AI^2 game in period $t=1$

	C_2	C_0	P
C_2	b, b	n, o	x, y
P	x, y	x, y	**e, f**

Figure 3.27.: The AI^2 game in period $t \geq 2$
after or without defection in period $t=1$

To support S_2 as an equilibrium strategy for the AI game i.e., a strategy leading *gradually* to a high level of cooperation, defection in the *second* period must not be profitable to the player with the reversible action.

$S_{2(AI)}\Rightarrow$ play C_1 in period $t=1$; then play C_2 in period $t=2$ if (C_1,C_1) has been played at $t=1$, otherwise play P; then play C_2 in period $t>2$ as long as (C_2,C_2) has been played in the previous period and $t<T$, otherwise play P from period t

onwards; then play C_0 in the last period T if (C_2, C_2) has been played in period $T-1$,otherwise play P.

At the beginning of the first two periods, the player with the reversible action decides whether to increase the level of cooperation by playing $C_t > C_{t-1}$ and C_2 thereafter except for the last period, or whether to defect by playing $C_t < C_{t-1}$ in the second period (we are only considering the most profitable defection) knowing that (P, P) will be played from the following period onwards. The player with the reversible action plays $C_t > C_{t-1}$ for the first two periods and C_2 thereafter, except for the last period, as long as his or her payoffs from the equilibrium strategy $\pi(S_2)$ are bigger than his or her payoffs from defecting $\pi(D_2)$. Therefore, *f* must satisfy $\pi(S_2) > \pi(D_2)$, where

$$\pi(S_2) = c + \sum_{t=2}^{T-1} \delta^{t-1} b + \delta^{T-1} o \tag{27}$$

$$\pi(S_2) = c + \delta o + \sum_{t=3}^{T} \delta^{t-1} f \tag{28}$$

	C_2	C_1	C_0	P
C_2	b, b	j, k	n, o	x, y
C_1	k, j	c, c	h, i	x, y
P	x, y	x, y	x, y	**e, f**

Figure 3.28.: The AI_2 game in period $t = 2$

	C_2	C_1	C_0	P
C_2	b, b	j, k	n, o	x, y
P	x, y	x, y	x, y	**e, f**

Figure 3.29.: The AI_2 game in period $t \geq 3$
after or without defection in period $t = 2$

To support S^1as an equilibrium strategy for the AI game i.e., a strategy leading *immediately* to a low level of cooperation for all periods, defection must not be profitable to any of the players. As the low level of cooperation is by assumption the first level of cooperation, the equilibrium strategy leading *gradually* to the first level of cooperation is equivalent to the equilibrium strategy leading *immediately* to the first level of cooperation i.e., $S_{1(AI)} = S^{1(AI)}$. Therefore, there is no need to describe the strategy leading *gradually* to a low level of cooperation separately.

> $S^{1(AI)} \Rightarrow$ play C_1 in period $t=1$, then play C_1 as long as (C_1, C_1) has been played in period $t-1$ and $t<T$, otherwise play P from period t onwards; then play C_0 in the last period T if (C_1, C_1) has been played in period $T-1$, otherwise play P.

At the beginning of each period, the player with the reversible action decides, whether to play C_1 or whether to defect by playing $C < C_1$ knowing that (P, P) will be played from the following period. The player with the reversible action plays C_1, if his or her payoffs from the equilibrium strategy $\pi(S^1)$ are bigger than his or her payoffs from defecting $\pi(D^1)$. Therefore, *f* must satisfy $\pi(S^1) > \pi(D^1)$, where

$$\pi(S^1) = \sum_{t=1}^{T-1} \delta^{t-1} c + \delta^{T-1} i \qquad (29)$$

$$\pi(D^1) = i + \sum_{t=1}^{T} \delta^{t-1} f \qquad (30)$$

	C_1	C_0	P
C_1	c, c	h, i	x, y
C_0	i, h	**d, d**	x, y
P	x, y	x, y	**e, f**

Figure 3.30.: The AI^1 game in period $t=1$

	C_1	C_0	P
C_1	c, c	h, i	x, y
P	x, y	x, y	**e, f**

Figure 3.31.: The AI^1 game in period $t \geq 2$ after or without defection in period $t = 1$

3.5.2. The Necessary Punishment (an Illustrative Example)

To determine the *necessary punishment* (f^*) that supports the different kinds and levels of cooperation in the three games, we use the lowest number of periods necessary to reach the target level of cooperation that we want to support. Whether the target level of cooperation is reached *immediately* or *gradually* makes a difference only in the LT game (except for the first level of cooperation) as will be shown.

1. To support $S^{\infty(PD)}$ as equilibrium strategy i.e., a strategy leading *immediately* to sustainable full cooperation in the PD game (T=2)

$$\pi\left(S^{\infty(PD)}\right) = a + \delta d > \pi\left(D^{\infty(PD)}\right) = s + \delta f$$

$$\therefore f \ < f^{*\infty(PD)} \equiv d - \frac{s-a}{\delta}$$

2. To support $S_{\infty(PD)}$ as equilibrium strategy i.e., a strategy leading *gradually* to sustainable full cooperation in the PD game (T=4)

$$\pi\left(S_{\infty(PD)}\right)=c+\delta b+\delta^{2}a+\delta^{3}d>\pi\left(D_{\infty(PD)}\right)=c+\delta b+\delta^{2}s+\delta^{3}f$$

$$\therefore f<f^{*}_{\infty(PD)}\equiv d-\frac{s-a}{\delta}$$

$$\therefore f^{*\infty(PD)}=f^{*}_{\infty(PD)}$$

3. To support $S^{\infty(LT)}$ as equilibrium strategy i.e., a strategy leading *immediately* to sustainable full cooperation in the LT game (T=2).

$$\pi\left(S^{\infty(LT)}\right)=a+\delta a>\pi\left(D^{\infty(LT)}\right)=s+\delta f$$

$$\therefore f<f^{*\infty(LT)}\equiv a-\frac{s-a}{\delta}$$

4. To support $S_{\infty(LT)}$ as equilibrium strategy i.e., a strategy leading *gradually* to sustainable full cooperation in the LT game (T=4)

$$\pi\left(S_{\infty(LT)}\right)=c+\delta b+\delta^{2}a+\delta^{3}a>\pi\left(D_{\infty(LT)}\right)=c+\delta b+\delta^{2}m+\delta^{3}f$$

$$\therefore f<f^{*}_{\infty(LT)}\equiv a-\frac{m-a}{\delta}$$

5. To support $S^{\infty(AI)}$ as equilibrium strategy i.e., a strategy leading *immediately* to sustainable full cooperation in the AI game (T=2)

$$\pi\left(S^{\infty(AI)}\right) = a + \delta s > \pi\left(D^{\infty(AI)}\right) = s + \delta f$$

$$\therefore f < f^{*\infty(AI)} \equiv s - \frac{s-a}{\delta}$$

6. To support $S_{\infty(AI)}$ as equilibrium strategy i.e., a strategy leading *gradually* to sustainable full cooperation in the AI game (T=4)

$$\pi\left(S_{\infty(AI)}\right) = c + \delta b + \delta^2 a + \delta^3 s > \pi\left(S_{\infty(AI)}\right) = c + \delta b + \delta^2 s + \delta^3 f$$

$$\therefore f < f^*_{\infty(AI)} \equiv s - \frac{s-a}{\delta}$$

$$\therefore f^{*\infty(AI)} = f^*_{\infty(AI)}$$

7. To support $S^{2(PD)}$ as equilibrium strategy i.e., a strategy leading *immediately* to a sustainable high level of cooperation in the PD game (T=2)

$$\pi\left(S^{2(PD)}\right) = b + \delta d > \pi\left(D^{2(PD)}\right) = o + \delta f$$

$$\therefore f < f^{*2(PD)} \equiv d - \frac{o-b}{\delta}$$

8. To support $S_{2(PD)}$ as equilibrium strategy i.e., a strategy leading *gradually* to a sustainable high level of cooperation in the PD game (T=3)

$$\pi\left(S_{2(PD)}\right) = c + \delta b + \delta^2 d > \pi\left(D_{2(PD)}\right) = c + \delta o + \delta^2 f$$

$$\therefore f < f^*_{2(PD)} \equiv d - \frac{o-b}{\delta}$$

$$\therefore f^{*2(PD)} = f^*_{2(PD)}$$

9. To support $S^{2(LT)}$ as equilibrium strategy i.e., a strategy leading *immediately* to a sustainable high level of cooperation in the LT game (T=2)

$$\pi\left(S^{2(LT)}\right) = b + \delta b > \pi\left(D^{2(LT)}\right) = o + \delta f$$

$$\therefore f < f^{*2(LT)} \equiv b - \frac{o-b}{\delta}$$

10. To support $S_{2(LT)}$ as equilibrium strategy i.e., a strategy leading *gradually* to a sustainable high level of cooperation in the LT game (T=3)

$$\pi\left(S_{2(LT)}\right) = c + \delta b + \delta^2 b > \pi\left(D_{2(LT)}\right) = c + \delta k + \delta^2 f$$

$$\therefore f < f^*_{2(LT)} \equiv b - \frac{k-b}{\delta}$$

11. To support $S^{2(AI)}$ as equilibrium strategy i.e., a strategy leading *immediately* to a sustainable high level of cooperation in the AI game (T=2)

$$\pi\left(S^{2(AI)}\right) = b + \delta o > \pi\left(S^{2(AI)}\right) = o + \delta f$$

$$\therefore f < f^{*2(AI)} \equiv o - \frac{o-b}{\delta}$$

12. To support $S_{2(AI)}$ as equilibrium strategy i.e., a strategy leading *gradually* to a sustainable high level of cooperation in the AI game (T=3)

$$\pi\left(S_{2(AI)}\right) = c + \delta b + \delta^2 o > \pi\left(D_{2(AI)}\right) = c + \delta o + \delta^2 f$$

$$\therefore f < f^*_{2(AI)} \equiv o - \frac{o-b}{\delta}$$

$$\therefore f^{*2(AI)} = f^*_{2(AI)}$$

13. To support $S^{1(PD)}$ (or $S_{1(PD)}$ as explained earlier) as equilibrium strategy i.e., a strategy leading *immediately* or *gradually* to a sustainable low level of cooperation in the PD game (T=2)

$$\pi\left(S^{1(PD)}\right) = c + \delta d > \pi\left(D^{1(PD)}\right) = i + \delta f$$

$$\therefore f < f^{*1(PD)} \equiv d - \frac{i-c}{\delta}$$

$$\therefore f^{*1(PD)} = f^*_{1(PD)}$$

14. To support $S^{1(LT)}$ (or $S_{1(LT)}$ as explained earlier) as equilibrium strategy i.e., a strategy leading *immediately* or *gradually* to a sustainable low level of cooperation in the LT game (T=2)

$$\pi\left(S^{1(LT)}\right) = c + \delta c > \pi\left(D^{1(LT)}\right) = i + \delta f$$

$$\therefore f < f^{*1(LT)} \equiv c - \frac{i-c}{\delta}$$

$$\therefore f^{*1(LT)} = f^{*}_{1(LT)}$$

15. To support $S^{1(AI)}$ (or $S_{1(AI)}$ as explained earlier) as equilibrium strategy i.e., a strategy leading *immediately* or *gradually* to a sustainable low level of cooperation in the AI game (T=2)

$$\pi\left(S^{1(AI)}\right) = c + \delta i > \pi\left(D^{1(AI)}\right) = i + \delta f$$

$$\therefore f < f^{*1(AI)} \equiv i - \frac{i-c}{\delta}$$

$$\therefore f^{*1(AI)} = f^{*}_{1(AI)}$$

Now, we have 11 critical values for the credible punishment f^{*}, which can be compared both across the games for the same level and/or type of cooperation, or across the levels and/or types of cooperation for the same game.

3.5.2.1. The Necessary Punishment Across the Games (for the Same Level of Cooperation)

As mentioned earlier, the superscript means that cooperation was reached *immediately* and the subscript means that the cooperation was reached *gradually*. In addition, we will use a two-star superscript to denote the cases where there is no difference between the values that are reached via *immediate* or *gradual* cooperation.
(For example: $f^{**}_{PD(C_\infty,C_\infty)} = f^{*\infty(PD)} = f^{*}_{\infty(PD)}$)

$f^{**}_{PD(C_\infty,C_\infty)} < f^{*LT(C_\infty,C_\infty)} < f^{**}_{AI(C_\infty,C_\infty)}$ $\qquad$ $f^{*LT(C_\infty,C_\infty)} < f^{*}_{LT(C_\infty,C_\infty)}$

$f^{*}_{LT(C_\infty,C_\infty)} < f^{**}_{AI(C_\infty,C_\infty)}$ $\qquad$ iff $\delta > \dfrac{s-m}{s-a}$

$f^{**}_{PD(C_2,C_2)} < f^{*LT(C_2,C_2)} < f^{**}_{AI(C_2,C_2)}$ $\qquad$ $f^{*LT(C_2,C_2)} < f^{*}_{LT(C_2,C_2)}$

$f^{*}_{LT(C_2,C_2)} < f^{**}_{AI(C_2,C_2)}$ $\qquad$ iff $\delta > \dfrac{o-k}{o-b}$

$f^{**}_{PD(C_1,C_1)} < f^{**}_{LT(C_1,C_1)} < f^{**}_{AI(C_1,C_1)}$ $\qquad$ for all $0 < \delta < 1$

3.5.2.2. The Necessary Punishment Across the Levels of Cooperation (for the Same Game)

In the PD game, the higher the level of cooperation, the higher the *necessary punishment* to sustain cooperation, whether the cooperation is *immediate* or *gradual*. This is because the payoffs from defection increase with higher levels of cooperation.

$$f^{**}_{PD(C_\infty,C_\infty)} < f^{**}_{PD(C_2,C_2)} < f^{**}_{PD(C_1,C_1)}$$

In the LT game, we have to differentiate between *immediate* and *gradual* cooperation. Whereas the payoffs from defection are expected to increase with higher levels of *immediate* cooperation, they do not necessarily increase with higher levels of *gradual* cooperation.

a) immediate cooperation

if $(s-a)>(o-b)>(i-c)$[22] $\Rightarrow \quad f^{*LT(C_\infty,C_\infty)} < f^{*LT(C_2,C_2)} \quad$ iff $\quad \delta < \frac{s-o}{a-b}-1$

$\Rightarrow \quad f^{*LT(C_2,C_2)} < f^{*LT(C_1,C_1)} \quad$ iff $\quad \delta < \frac{o-i}{b-c}-1$

b) gradual cooperation

if $(m-a)=(k-b)=(i-c)$[23] $\Rightarrow \quad f^{*}_{LT(C_\infty,C_\infty)} > f^{*}_{LT(C_2,C_2)} > f^{*}_{LT(C_1,C_1)}$

if $(m-a)>(k-b)>(i-c)$[24] $\Rightarrow \quad f^{*}_{LT(C_\infty,C_\infty)} < f^{*}_{LT(C_2,C_2)} \quad$ iff $\quad \delta < \frac{m-k}{a-b}-1$

$\Rightarrow \quad f^{*}_{LT(C_2,C_2)} < f^{*}_{LT(C_1,C_1)} \quad$ iff $\quad \delta < \frac{k-i}{b-c}-1$

---

[22] $(s-a)>(o-b)>(i-c)$ means that the benefits from deviation to the non-cooperation level C_0 increase with the level of cooperation achieved in the previous period.

[23] $(m-a)=(k-b)=(i-c)$ means that the benefits from deviation, in terms of keeping the level of cooperation achieved in the previous period constant, is equal for all levels of cooperation.

[24] $(m-a)>(k-b)>(i-c)$ means that the benefits from deviation, in terms of keeping the level of cooperation achieved in the previous period constant, increases with the level of cooperation achieved in the previous period.

In the AI game, *necessary punishment* might also decrease or increase with higher levels of cooperation, whether the cooperation is *immediate* or *gradual*.

$$f^{**}_{AI(C_\infty,C_\infty)} < f^{**}_{AI(C_2,C_2)} \quad \text{iff} \quad \delta < 1-\frac{a-b}{s-o} \qquad f^{**}_{AI(C_2,C_2)} < f^{**}_{AI(C_1,C_1)} \quad \text{iff} \quad \delta < 1-\frac{b-c}{o-i}$$

The results are summarised in table 3.1.

Table 3.1.: f^{*} and f^{**} across the games and across the levels of cooperation

	across the games *(same level of cooperation)*	***across the levels of cooperation*** *(same game)*
all δ	$f^{**}_{PD(C_\infty,C_\infty)} < f^{*LT(C_\infty,C_\infty)} < f^{**}_{AI(C_\infty,C_\infty)}$ $f^{**}_{PD(C_2,C_2)} < f^{*LT(C_2,C_2)} < f^{**}_{AI(C_2,C_2)}$ $f^{**}_{PD(C_1,C_1)} < f^{**}_{LT(C_1,C_1)} < f^{**}_{AI(C_1,C_1)}$	$f^{**}_{PD(C_\infty,C_\infty)} < f^{**}_{PD(C_2,C_2)} < f^{**}_{PD(C_1,C_1)}$
low δ	$f^{**}_{PD(C_\infty,C_\infty)} < f^{**}_{AI(C_\infty,C_\infty)} < f^{*}_{LT(C_\infty,C_\infty)}$ $f^{**}_{PD(C_2,C_2)} < f^{**}_{AI(C_2,C_2)} < f^{*}_{LT(C_2,C_2)}$	$f^{*}_{LT(C_\infty,C_\infty)} < f^{*}_{LT(C_2,C_2)} < f^{**}_{LT(C_1,C_1)}$ $f^{*}_{AI(C_\infty,C_\infty)} < f^{*}_{AI(C_2,C_2)} < f^{**}_{AI(C_1,C_1)}$
high δ	$f^{**}_{PD(C_\infty,C_\infty)} < f^{*}_{LT(C_\infty,C_\infty)} < f^{**}_{AI(C_\infty C_\infty)}$ $f^{**}_{PD(C_2,C_2)} < f^{*}_{LT(C_2,C_2)} < f^{**}_{AI(C_2 C_2)}$	$f^{*}_{LT(C_\infty,C_\infty)} > f^{*}_{LT(C_2,C_2)} > f^{**}_{LT(C_1,C_1)}$ $f^{*}_{AI(C_\infty,C_\infty)} > f^{*}_{AI(C_2,C_2)} > f^{**}_{AI(C_1,C_1)}$

3.5.3. The Interpretation of Credible Punishment in the Games

Partial or full cooperation can only be achieved if defection by any player can be punished by the other players. Therefore, threats and promises about future behaviour can only influence current behaviour if they are credible. The threat to punish can be exogenous to the game, as for example the threat to respond with military action to any defection of the agreement. In this case the punishment does not relate to the different irreversibility constraints of the players' actions, but depends on some other factors such as economic and military strength. In order to reflect the different irreversibility constraints, the punishment has to be endogenous to the game. One immediate result of defecting on the agreement could be foregoing the benefits from continuing and increasing cooperation between the players in the future. Therefore, one way to interpret punishment, that is endogenous to the game, could be as foregone benefits from increasing the levels of cooperation in the future until full cooperation is reached and thereafter keeping the level of cooperation constant. The foregone benefits will differ according to the irreversibility constraints on the players' actions as will be shown.

3.5.3.1. Credible Punishment in the PD Game

The PD game provides the harshest punishment for deviating of all three games, as actions are reversible at any period to the stage game NE of the first period. The *credible* punishment in terms of foregone benefits will also be the most negative one of all games and will increase with higher levels of cooperation, and will become more and more negative. There is no difference in the credible punishment whether the PD game is *immediate* or *gradual*.

The credible punishments for deviations from (C_∞, C_∞), $(C_2, C_2)$and(C_1, C_1) in the PD game are respectively:

$$f_{PD(C_\infty, C_\infty)} = \frac{d-a}{1-\delta}$$

$$f_{PD(C_2,C_2)} = (d-b) + \frac{\delta(d-a)}{1-\delta}$$

$$f_{PD(C_1,C_1)} = (d-c) + \delta(d-b) + \frac{\delta^2(d-a)}{1-\delta}$$

3.5.3.2. Credible Punishment in the LT Game

Whether cooperation is *gradual* or *immediate*, defection cannot be punished as severely in the LT game as in the PD game. The *credible* punishment will also decrease with higher levels of cooperation, becoming less and less negative until it actually turns positive close to full cooperation. Because actions are irreversible in the LT game, there is a big difference in the *credible* punishment according to the type of cooperation. The gains from deviating from *gradual* cooperation are less than the gains from deviating from *immediate* cooperation, except for the first level of cooperation. At the same time, the ability to punish is higher if the cooperation is *gradual* compared to the *immediate* cooperation.

The credible punishment for deviation from (C_∞, C_∞), $(C_2, C_2)$and(C_1, C_1) in the LT game are respectively:

a) *immediate cooperation*

$$f^{LT(C_\infty,C_\infty)} = \frac{s-a}{1-\delta}$$

$$f^{LT(C_2,C_2)} = o - b + \frac{\delta(o-a)}{1-\delta}$$

$$f^{LT(C_1,C_1)} = (i-c) + \delta(i-b) + \frac{\delta^2(i-a)}{1-\delta}$$

b) *gradual cooperation*

$$f_{LT(C_\infty,C_\infty)} = \frac{m-a}{1-\delta}$$

$$f_{LT(C_2,C_2)} = (k-b) + \frac{\delta(k-a)}{1-\delta}$$

$$f_{LT(C_1,C_1)} = (i-c) + \delta(i-b) + \frac{\delta^2(i-a)}{1-\delta}$$

3.5.3.3. Credible Punishment in the AI Game

The *credible* punishment in the AI game, whether *gradual* or *immediate* is exactly the same as the *credible* punishment in the *immediate* LT game. The actions in the AI game remain reversible for one of the players to any previous level up to the first level, while the other player's actions are completely irreversible at any achieved level of cooperation. Therefore, the punishment in terms of foregone benefits will also be the least negative one compared to the *gradual* LT game and the PD game. As in the LT game, *credible* punishment also decreases with higher levels of cooperation and becomes less and less negative until it actually turns positive close to full cooperation.

The credible punishment for deviation from (C_∞, C_∞), (C_2, C_2) and (C_1, C_1) in the AI game are respectively:

$$f_{AI(C_\infty,C_\infty)} = \frac{s-a}{1-\delta}$$

$$f_{AI(C_2,C_2)} = (o-b) + \frac{\delta(o-a)}{1-\delta}$$

$$f_{Ai(C_1,C_1)} = (i-c) + \delta(i-b) + \frac{\delta^2(i-a)}{1-\delta}$$

This time we have nine values for *credible punishment*, because *credible punishment* for *immediate* cooperation in the LT game is equivalent to the *credible punishment* for the AI game. These values can be compared both across the games for the same level of cooperation, and across the levels of cooperation for the same game.

3.5.3.4. Credible Punishment Across the Games (for the Same Level of Cooperation)

The PD provides the harshest credible punishment among all three games as both players can revert to any lower level of cooperation up to the non cooperation level any time. Therefore,

$$f_{PD(C_\infty,C_\infty)} < f_{LT(C_\infty,C_\infty)} < f^{LT(C_\infty,C_\infty)} = f_{AI(C_\infty,C_\infty)}$$

$$f_{PD(C_2,C_2)} < f_{LT(C_2,C_2)} < f^{LT(C_2,C_2)} = f_{AI(C_2,C_2)}$$

$$f_{PD(C_1,C_1)} < f_{LT(C_1,C_1)} = f^{LT(C_1,C_1)} = f_{AI(C_1,C_1)} \quad \text{iff } (C_1,C_1) \text{ is the first level of cooperation}$$

3.5.3.5. Credible Punishment Across Levels of Cooperation (for the Same Game)

While the *credible* punishment increases with higher levels of cooperation in the PD game, it decreases in both the LT game and the AI game. The *credible* punishment decreases much faster in the *immediate* LT game and the AI game compared to the *gradual* LT game, which has important implications for the sustainable levels of cooperation.

$f_{PD(C_\infty,C_\infty)} < f_{PD(C_2,C_2)} < f_{PD(C_1,C_1)}$ *immediate* and *gradual* cooperation

$f_{LT(C_\infty,C_\infty)} > f_{LT(C_2,C_2)} > f_{LT(C_1,C_1)}$ *immediate* and *gradual* cooperation

$f_{AI(C_\infty,C_\infty)} > f_{AI(C_2,C_2)} > f_{AI(C_1,C_1)}$ *immediate* and *gradual* cooperation

3.5.4. Sustainable Cooperation in the PD, LT and AI Games

For sustainable cooperation, we require that $f_{Game(coop.level)} < f^*_{Game(coop.level)}$

But, if (from the nature of the game) $f_{Game(coop.level)} > f^*_{Game(coop.level)}$

then, $f_{Game(coop.level)}$ cannot support that level of cooperation in this specific game.

3.6. Results

The comparison between *necessary* and *credible* punishment in Table 3.2. shows the important role of the discount factor. A high δ increases the *credible* punishment (makes it more negative) and decreases the *necessary* punishment, thereby increasing the prospects of cooperation. A low δ, on the other hand, decreases the *credible* punishment (makes it less negative) and increases the *necessary* punishment, thereby decreasing the prospects of cooperation.

Table 3.2.: *Credible* and *necessary* punishment

	Credible punishment f		Necessary punishment (The critical value f*)	
	Immediate cooperation	Gradual cooperation	Immediate cooperation	Gradual cooperation
The PD game				
(C_∞, C_∞)	$\frac{d-a}{1-\delta}$	*same*	$d-\frac{s-a}{\delta}$	*same*
(C_2, C_2)	$(d-a)+\frac{\delta(d-a)}{1-\delta}$	*same*	$d-\frac{o-b}{\delta}$	*same*
(C_1, C_1)	$(d-c)+\delta(d-b)+\frac{\delta^2(d-a)}{1-\delta}$	*same*	$d-\frac{i-c}{\delta}$	*same*
The LT game				
(C_∞, C_∞)	$\frac{s-a}{1-\delta}$	$\frac{m-a}{1-\delta}$	$a-\frac{s-a}{\delta}$	$a-\frac{m-a}{\delta}$
(C_2, C_2)	$(o-b)+\frac{\delta(o-a)}{1-\delta}$	$(k-b)+\frac{\delta(k-a)}{1-\delta}$	$b-\frac{o-b}{\delta}$	$b-\frac{k-b}{\delta}$
(C_1, C_1)	$(i-c)+\delta(i-b)+\frac{\delta^2(i-a)}{1-\delta}$	*same*	$c-\frac{i-c}{\delta}$	*same*
The AI game				
(C_∞, C_∞)	$\frac{s-a}{1-\delta}$	*same*	$s-\frac{s-a}{\delta}$	*same*
(C_2, C_2)	$(o-b)+\frac{\delta(o-a)}{1-\delta}$	*same*	$o-\frac{o-b}{\delta}$	*same*
(C_1, C_1)	$(i-c)+\delta(i-b)+\frac{\delta^2(i-a)}{1-\delta}$	*same*	$i-\frac{i-c}{\delta}$	*same*

Table 3.3.: Equilibrium conditions for the games, levels and types of cooperation

	Equilibrium condition $f < f^*$	
	Immediate cooperation	**Gradual cooperation**
(C_∞, C_∞)		
PD game	$a > (1-\delta) + \delta^2 d$	*same*
LT game	$a > \frac{4}{5}s$	$a > \frac{4}{5}m$
AI game	$a > \frac{3}{4}s$	*same*
(C_2, C_2)		
PD game	$a > \left(-\frac{1}{\delta} + \frac{1}{\delta^2}\right)o + \left(1 - \frac{1}{\delta^2}\right)b + d$	*same*
LT game	$a > \frac{1}{\delta^2}o + \left(2 - \frac{1}{\delta} - \frac{1}{\delta^2}\right)b$	$a > \frac{1}{\delta^2}k + \left(2 - \frac{1}{\delta} - \frac{1}{\delta^2}\right)b$
AI game	$a > \left(1 - \frac{1}{\delta} + \frac{1}{\delta^2}\right)o + \left(1 - \frac{1}{\delta^2}\right)b$	*same*
(C_1, C_1)		
PD game	$a > \left(-\frac{1}{\delta^2} + \frac{1}{\delta^3}\right)i + \left(1 - \frac{1}{\delta}\right)b + \left(\frac{1}{\delta} - \frac{1}{\delta^3}\right)c + \frac{1}{\delta}d$	*same*
LT game	$a > \frac{1}{\delta^3}i + \left(1 - \frac{1}{\delta}\right)b + \left(\frac{2}{\delta} - \frac{1}{\delta^2} - \frac{1}{\delta^3}\right)c$	*same*
AI game	$a > \left(\frac{1}{\delta} - \frac{1}{\delta^2} + \frac{1}{\delta^3}\right)i + \left(1 - \frac{1}{\delta}\right)b + \left(\frac{1}{\delta} - \frac{1}{\delta^3}\right)c$	*same*

Table 3.3. provides the conditions, which ensure that credible punishment is more negative than necessary punishment. These conditions guarantee that sustainable cooperation is achieved for the different levels and types of cooperation in the three different games.

As long as δ is high enough, any level of cooperation whether *immediate* or *gradual* cannot be ruled out for the PD game. But, for both the LT game and the AI game, full cooperation, whether *immediate* or *gradual,* can under certain conditions be ruled out, no matter how close δ is to 1[25]. *Immediate* and *gradual* full cooperation can be ruled out in the LT game, if $a < 4/5s$ and $a < 4/5m$ respectively. *Immediate* and *gradual* full cooperation can be ruled out in the AI game if $a < 3/4s$. This is most likely to happen, because $(s, m > a)$[26]

Partial cooperation whether *immediate* or *gradual* cannot be ruled out in both the LT game and the AI game and depends on the discount factor and parameter values. While it makes no difference for the AI game whether the partial cooperation we want to support is *gradual* or *immediate*, matters are different for the LT game. For the LT game, *gradual* partial cooperation is always easier to sustain than *immediate* partial cooperation. On the one hand the *credible* punishment is higher (more negative) for *gradual* partial cooperation than for *immediate* partial cooperation, and on the other hand the *necessary* punishment needed to sustain *gradual* partial cooperation is less than the *necessary* punishment to sustain *immediate* partial cooperation for the same level of cooperation.

The most interesting part is the intermediate (or partial) levels of cooperation in both the LT and AI games i.e., all levels of partial cooperation except the first level of cooperation (in our example the high level of cooperation).

[25] Lockwoood and Thomas (2001) describe the *efficient symmetric equilibrium path* as a sequence of actions of rising levels of cooperation which converge to a limit value that cannot exceed full co-operation. They show that if payoffs are smooth functions of actions, the *efficient symmetric equilibrium path* converges to a level strictly below the full cooperation level, no matter how patient the players are. If payoffs are linear up to some joint cooperation level and constant or decreasing thereafter, then above some critical discount factor cooperation can converge asymptotically to the full cooperative level.
[26] See assumption (2).

Immediate partial levels of cooperation are always easier to sustain in the LT game than *immediate/gradual* partial levels of cooperation in the AI game, because

$$\frac{1}{\delta^2}o+\left(2-\frac{1}{\delta}-\frac{1}{\delta^2}\right)b<\left(1-\frac{1}{\delta}+\frac{1}{\delta^2}\right)o+\left(1-\frac{1}{\delta^2}\right)b \qquad \text{(because } o>b\text{)}$$

Gradual partial levels of cooperation are always easier to sustain in the LT game than *immediate/gradual* partial levels of cooperation in the AI game, because

$$\frac{1}{\delta^2}k+\left(2-\frac{1}{\delta}-\frac{1}{\delta^2}\right)b<\left(1-\frac{1}{\delta}+\frac{1}{\delta^2}\right)o+\left(1-\frac{1}{\delta^2}\right)b \qquad \text{(because } o>b\text{)}$$

In the first level of cooperation matters are much more straightforward[27]. First, there is no difference between *gradual* and *immediate* cooperation in all three games, and secondly, the *credible* punishment is the same for the LT game and the AI game. Since the levels of credible punishment are the same, but the necessary punishment is lower (i.e., f^* is higher) in the AI game than in the LT game, it is easier to sustain the first level of cooperation in the AI game than in the LT game.

3.7. Summary and Conclusion

Full cooperation, whether *immediate* or *gradual* can only be sustained under symmetric reversibility of the players' actions (the PD game) if δ is high enough. It can be ruled out

[27] Given that some tough restrictions apply for the full cooperation level ($a>4/5s$, $a>3/4s$ and $a>4/5m$), then sustaining full cooperation would be easier in the AI game than in the *immediate* LT game, *gradual* full cooperation in the LT game would be easier to sustain than *immediate* full cooperation, but nothing definitive could be said about the comparison of *gradual* full cooperation in the LT game and the AI game.

under both symmetric irreversibility (the LT game) and asymmetric irreversibility (the AI game) no matter how high δ is.

Partial cooperation can be sustained under both symmetric irreversibility (the LT game) and asymmetric irreversibility (the AI game). While *gradual* cooperation improves the prospects for cooperation for the LT game compared to *immediate* cooperation, the type of cooperation has no effect on the PD and AI game.

While *immediate* partial cooperation is always easier to sustain in the LT game than in the AI game for all intermediate levels of cooperation, and *gradual* partial cooperation is always easier to sustain in the *gradual* partial LT game than in the *immediate* partial LT game, therefore sustaining *gradual* partial cooperation in the LT game is easier than in the AI game.

Our results confirm the general intuition that despite obvious gains from cooperation, asymmetric irreversibility of the players' actions makes reaching even a partial agreement between countries sharing a water body with unidirectional externalities, such as a river, more difficult than reaching an agreement between countries sharing a water body with reciprocal externalities, such as a lake. Interpreting the unidirectional externality in terms of irreversible actions has not been covered in the economic literature as far as I know. Our approach to adapt an infinite dynamic game into a finite dynamic game allows tractable analysis of a very complicated, but highly relevant extension of the Lockwood and Thomas (2002) model.

Higher δ are conducive to cooperation for all games, levels of cooperation and types of cooperation. Discount factors give some indication of the valuation of long-run gains in a country. Countries with stronger concerns about the future are usually economically well developed. The importance they attach to long-run gains and losses is reflected in a relatively high discount factor. Countries with stronger concerns about immediate domestic pressures are usually less developed economically. The importance they attach to short-run gains and losses is reflected in a relatively lower discount factor. Therefore, it is expected that sustaining any level and type of cooperation will be easier for developed countries than for developing countries. Nevertheless, low discount factors are not responsible for the inability to reach full cooperation in finite time whenever at least one player's actions are irreversible.

One of the limitations of this study is the assumption that the countries not only share a water body, but also share the same discount factor. This might not be the case, especially if the countries are at different stages of economic development. Future research needs to consider what effect the asymmetry of economic development of the water sharing countries might have on reaching a water sharing agreement between them.

Chapter 4
Water Quality in Domestic and International Rivers

4.1. Introduction

Does the nature of a river affect the relationship between income and environmental quality? Can we identify a unidirectional externality effect of pollution from upstream countries to downstream countries? Is this unidirectional externality effect, if identified, income related or non-income related? The answers to these questions could have important implications for policy interventions, either on national level or on a more regional level through International Environmental Agreements (IEAs).

The theoretical debate about the relationship between environmental quality and levels of national income has sparked a lot of empirical studies. Many of these studies, such those of Cole *et al.* (1997), Grossman and Krueger (1995), Holtz-Eakin and Selden (1995), Roberts and Grimes (1997), Selden and Song (1994), Shafik (1994), and Stern *et al.* (1996) have found empirical evidence for the existence of an inverted-U shaped relationship between growth and environmental degradation, usually known as Environmental Kuznets curve (EKC), for at least some pollutants. The EKC hypothesises states that environmental degradation increases first with income at low levels of income, then decreases later with income at high levels of income. The estimated turning points of the income levels at which pollution is expected to decline has important implications for suggesting appropriate development strategies, especially for less developed countries. Although these critical levels of income vary among individual pollution indicators a general pattern for air quality indictors has emerged. Air pollutants with direct, local impacts seem to have turning points at lower levels of income than air pollutants with more indirect, regional or global impacts. Studies by Cole *et al.* (1997), Holtz-Eakin and Selden (1995), and Mason and Swanson (1998) confirmed the observed trend by finding relative high turning points for global pollutants such as carbon dioxide emissions, CO_2, a greenhouse gas responsible for global warming, and chloro-fluorocarbons, CFCs, ozone depleting substances. The reason for this could be that governments are under more stress from local communities to address local pollutants with obvious short-term health effects in contrast to global pollutants with less obvious long-term effects. Selden and Song (1994) have found that even for local air pollutants, urban concentrations start to decline at lower levels of income than total emission. One of the reasons they provide is that, in contrast to total emissions, urban concentrations can be easily reduced by just building taller chimneys. Relocating some industries to less polluted areas is also a possible way to reduce urban concentrations without affecting total emissions.

While the different behaviour of local and global air pollutants has been widely studied, the same has not been done for water pollutants. In this chapter our first objective is to determine the effect of the locality of water pollutants on the relationship between income and water quality. We are only concerned with pollutants that are discharged into rivers. Depending on the nature of the river, a pollutant is local if the river is confined to one country only, and is regional if the river flows through more than one country. We want to find evidence for the perception that whenever water quality is internalized, higher incomes tend to be associated with less water quality degradation. As for the case of air pollution, we believe that intra-country water pollution externalities whose impacts are confined to a domestic river are generally easier to internalize than inter-country water pollution externalities. These can be externalised by the upstream country, which has little incentive to improve water quality downstream of its borders. Conte Grand (1999) has introduced the externality effect of international water bodies, i.e., lakes and rivers, via the existence or non-existence of water treaties. Sigman (2002) has chosen to present the externality effect of international rivers through dummy variables, which specify whether a water quality monitoring station is an upstream, a downstream or a border station. In both studies the externality effect is not related to income. As there is little doubt in the economic literature that per capita income plays an important role in the income-environmental quality relationship, our second objective is to capture an income-related externality effect of upstream countries on downstream countries. Therefore, we include a new explanatory variable to the basic emission-growth relationship, namely the per capita income of the upstream country. While the country's own per capita income might be a sufficient explanatory variable for the country's own impact on its water pollution, the per capita income of the upstream country might be an indicator for the upstream country's impact on the country's water pollution.

The analysis in our study proceeds in the following way: First, we estimate a basic reduced form regression model relating seven water quality indicators and per capita incomes in three settings: all rivers, domestic rivers and international rivers. The purpose is to determine whether the income-water quality relationship is affected by the nature of the river. Secondly, we estimate an expanded reduced form regression model for the international rivers only to determine the effect of including the externality variables on the estimation results. Thirdly, we calculate the critical levels of income at which each water quality indicator is expected to reach its maximum or minimum, if such an income level can be identified. This facilitates the comparison among the different natures of the rivers and between the two regression models. Lastly, we hope to shed some light on the effects of ignoring inter-country pollution on the intra-country turning points of downstream countries. While the annual water quality indicators on rivers, lakes and underground aquifers for the period from 1979 until 1990 are made available by Grossman and Krueger (1995)[28], we derive both income and externality indicators from the central per capita GDP measured by Heston *et al.* (2002) and published as the Penn World Tables (Mark 6)[29]. Our analysis confirms that national income is an important determinant for river pollution especially for domestic rivers. It shows that upstream effects in general play an important role in international river pollution, and that income related upstream effects are even more important than non-income related upstream effects. It also shows that all water pollutants peak at lower levels of income in domestic than in international rivers and that five out of seven water quality indicators in international rivers peak at a lower level of income when upstream effects are considered.

[28] The data are available via ftp, at ftp://irs.princeton.edu in the environ directory. A summary of the data, which provides triennial data from monitoring stations on rivers, lakes and groundwater aquifers in 69 countries from 1979 until 1999 for 13 basic water pollutants, is published by the United Nations Environment Programme (GEMS/Water).

This study comprises three main contributions. First, the distinction between local water pollution and regional water pollution through the nature of the river has not been done in this context before as far as I know. Secondly, it captures maybe for the first time a possible income related externality effect from upstream countries to downstream countries. Although upstream effects, such as geographical location and treaties, that affect the income-water quality relationship have been considered before, they were all non-income related as far as I know. Finally, we confirm the observed pattern that local air pollutants tend to peak at lower income levels than regional or global pollutants, for water pollutants.

This chapter is organized in the following way. Section 4.2. contains a brief review of the empirical literature on the income-environmental quality relationship. The empirical literature on the environmental Kuznets curve can be broadly divided into three subgroups: the Basic Environmental Kuznets Curve in section 4.2.1., the Expanded Environmental Kuznets Curve in section 4.2.2., and alternative explanations for the Environmental Kuznets Curve in section 4.2.3..In section 4.3., the seven water quality indicators and the source of the data are described. Section 4.4. outlines the methodology that is employed to analyze the cross country and time series pollution data. Section 4.5. reports the main results and the conclusion follows in section 4.6.

4.2. Literature Review

For the last twenty years, a considerable number of empirical studies have emerged to debate whether environmental degradation increases steadily, decreases steadily, or increases first and deteriorates later, with economic growth. A lot of emphasis has been given to finding empirical evidence to support the inverted-U shaped relationship between environmental degradation and economic growth, usually termed as the 'Environmental Kuznets Curve' (EKC), after Simon Kuznets' work on growth and income distribution.

[29] The data are available at http://pwt.econ.upenn.edu/.

The EKC assumes that environmental quality worsens first with rising average income and then improves once a certain per capita income is reached.

Most empirical models[30] of the environmental quality-income relationship have used reduced form single-equation specifications to model an environmental quality indicator as a function of per capita income, the most important independent variable. The resulting regressions do not specify the causality between income and environmental quality but only reflect possible correlation between the two. Some of these studies have tried to provide plausible theoretical explanations for these results. Many empirical models have included other income-related or non-income-related variables in the emission-income relationship to identify possible causal relationship between these added variables and the environmental quality indicators. Another direction of research has used alternative models to explain the correlation between pollution and income.

Therefore, the empirical studies can be classified into three main groups: first, studies on the Basic Environmental Kuznets Curve models with the per capita income as the chief explanatory variable, and secondly, the Expanded Environmental Kuznets Curve models with additional explanatory variables that are equally important besides average income in explaining the pollution behaviour of different countries, and thirdly, models that provide alternative explanations for the environmental Kuznets Curve.

4.2.1. Basic Environmental Kuznets Curve

Grossman and Krueger (1995) demonstrate that economic and population growth does not always threaten the quality of the environment. They use panel data[31] for the time period from 1979 until 1990 to examine the reduced-form relationship between per capita GDP and four types of environmental indicators, one related to air pollution and the other three to the pollution of the world's fresh water resources. Their findings support the inverted U-

[30] For an excellent survey on both empirical models and theoretical macro-models on growth and environment see Panayotou (2000).

shape relation between environmental degradation and economic development for 12 out of 14 environmental quality indicators.

> *"...we find that while increases in GDP may be associated with worsening environmental conditions in very poor countries, air and water quality appear to benefit from economic growth once some critical level of income has been reached..." (p.370)*

Grossman and Krueger (1996) have always emphasized that the so-called 'environmental Kuznets curve' stems from a reduced-form relationship across countries and time i.e., it tells nothing about how growth affects the environment.

> *"...There is nothing automatic about the relationship between economic growth and the environment. If environmental improvements are mediated by changes in government policy, then growth and development cannot be a substitute for environmental policy..." (p.120)*

Shafik (1994) examines the relationship between income and ten environmental indicators using panel data for 149 countries over the time period from 1960-90. She finds three clear patterns between economic growth and environmental degradation. While clean water and sanitation tends to improve with rising income, particulates and sulfur oxides worsen first and then improve. A third group of environmental quality indicators, oxygen in rivers, municipal solid waste and carbon emissions worsen steadily. By looking at the turning points of different indicators, i.e., the income levels at which the emission-income relationship changes signs, she finds substantial variations. She explains that policy intervention and investments to reduce pollution do not come automatically, but depend on the changes in the relative costs and benefits that people and countries attach to environmental problems at different stages of economic development.

---

[31] The data covers 42countries for SO_2, 29 countries for heavy particles, 19 for smoke, 58 countries for dissolved oxygen, BOD, COD and nitrates, 42 countries for fecal coliform, 22 countries for total coliform, and10 countries for heavy metals.

"...Actions tend to be taken where there are generalized local costs and substantial private and social costs. Where the costs of environmental degradation are borne by others (by the poor or by other countries), there are few incentives to alter damaging behaviour..." (p.770)

Selden and Song (1994) examine the emission-income relationship of four air pollutants: suspended particulate matter, sulfur dioxide, nitrogen oxides and carbon monoxide. Their aim is to determine whether an inverted U-shaped relationship between air pollution and economic development exists, and whether urban emissions start to improve at lower income levels compared to national emissions. They use cross- country panel data for 30 mainly high-income countries from 1973-1984. Their results generally support the inverted-U relationship for the four indicators of air quality, and with substantially higher turning points as hypothesised. One possible reason they provide for lower turning points of urban air pollutants relative to national air pollutants is that urban air pollution can be easily shifted to other areas within the same country.

"...it seems likely that important own-country pollution effects and relatively low abatement costs makes these pollutants among the most likely to exhibit inverted U-shaped relationships with income with tuning points that are relatively low..." (p.155)

Holtz-Eakin and Selden (1995) examine the relationship between per capita income and carbon dioxide emissions, CO_2, a greenhouse gas responsible for global warming. They want to determine whether an environmental Kuznets curve exists for greenhouse gases whose effects are more diffuse and are much more difficult to internalize at a country level.

"...On the one hand, their [greenhouse gases] effects are substantially more costly to abate and less restricted to local areas. Thus, a free-rider problem argues against a tendency for greenhouse gas emissions to decline at higher per capita incomes. On the other hand, greenhouse gases may fall as a by-product of other abatement efforts (e.g. fuel efficiency standards directed at urban air quality, which

reduce greenhouse gases). These possibilities indicate that emissions may ultimately stop rising, or even fall, as economies develop..." (pp.86-87)

They use panel data on 130 countries for the years 1951 to 1986 and find relatively high turning points, as was expected. Nevertheless, they do not expect a decrease in total carbon dioxide emissions in the foreseeable future, because economic growth and population growth is expected to rise the fastest in lower income countries, which have the highest marginal propensity to emit carbon dioxide.

Stern *et al.* (1996) consider the implications of the unequal distribution of global income on global environmental quality. They claim that inferring, from the existence of an inverted U-shape relationship between environmental degradation and per capita income across countries, that growth would reduce environmental degradation globally is not true.

"...The existence of an EKC relationship across nations today would not guarantee that global environmental degradation would decline automatically with time and economic growth..." (p.1158)

The possible reason they supply is that the world per capita income is not normally distributed but very skewed, with the median income far below the mean income i.e., many more people are below the world's per capita income than above it. According to the simulation study they have carried out they indicate that ignoring the world wide distribution of income could lead to biased estimates for the turning point of the EKC. Therefore, they support the view that globally, the world median income might be more relevant in determining how growth affects environmental quality.

Cole *et al.* (1997) analyze the relationship between per capita income and a wide range of environmental indicators with direct and indirect local impact. They use a cross-country panel data set for mainly OECD countries for the time period from 1972 until 1992. Although all indicators, except for municipal waste and methane, show an inverted-U shaped relationship with income, an EKC within the observed income range exists only for local air pollutants.

> *"...meaningful EKCs exist only for local air pollutant, whilst indicators with a more global, or indirect, environmental impact either increase monotonically with income or else have high turning points with large standard errors..."(p.411)*

The relatively high turning point for nitrates in rivers indicates that nitrates will increase in the foreseeable future even in developed countries. They also anticipate that most global environmental indicators, such as carbon dioxide and total energy use, will increase monotonically within the observed income range. They think this is because individual countries have little incentive to take unilateral actions to reduce emissions, and that the impacts of global warming are mostly externalized to other countries and future generations.

Roberts and Grimes (1997;1998) are rather sceptical about the alleged inverted U-shaped relationship between per capita income and the National Carbon Intensity[32] (NCI). Therefore, they examine the National Carbon Intensity for 30 countries over the period from 1961 until 1991 to establish whether there exists an inverted-shaped relationship between NCI and per capita GDP. They show, with scatter diagrams, that the NCI/GDP relationship has changed from linear in 1965 to strongly curvilinear in 1990. To determine how countries have reached improved carbon dioxide efficiency, they divide the countries into high-income, middle-income and low-income countries[33]. They find that the apparent stage-based improvement of the CO_2 intensity has resulted from both the tremendous improvements in CO_2 intensity since 1970 in high-income countries and the worsening efficiency in middle-income and low-income countries.

> *"...These findings strongly suggest that the emergence of an inverted U-curve (the Environmental Kuznets curve) for carbon dioxide emissions intensity is the result not of individual countries passing through stages of development, but of a*

[32] The number of kg of carbon dioxide emission per unit of gross domestic product.

relatively small number of wealthy ones becoming more efficient since 1970 while the average for the rest of the world worsens...(p.196)

Mason and Swanson (2003) also find support for the observed trend of higher turning points for global pollutants relative to urban or local pollutants. They examine the relationship between countries' propensity to emit chloro-fluorocarbons (CFCs), ozone depleting substances, and per capita income using a dynamic estimation model on a panel of CFCs production in 60 countries over the period form 1976 until 1988. They find a significant inverted-U shaped relationship between CFCs emission and per capita income with a relatively high turning point in accordance with other studies on global pollutants (e.g. Holz-Eakin and Selden (1995)) and in contrast to other studies on urban or local pollutants (e.g. Selden and Song (1994) and Grossman and Krueger (1995)) with much lower turning points.

4.2.2. Expanded Environmental Kuznets Curve

In a theoretical study Boyce (1994) has developed the notion that environmental quality is a function of the balance of power between those who derive net benefits from pollution-generating activities and those who bear the costs of pollution. He hypothesizes that greater power and wealth inequalities will lead, ceteris paribus, to more environmental degradation. Later, Torras and Boyce (1998) set out to test this hypothesis. They examine the relationship between seven environmental air and water quality indicators on one side and average income and the distribution of income and power on the other side for the period from 1977 until 1991. The distribution of income is measured by the Gini index and the distribution of power is measured by literacy, political rights and civil liberties. They hypothesize that the more equitable the distribution of income and power, the larger the improvements of air and water quality for any given level of average income. They argue that those who have to bear the costs of pollution are more able to represent their interests and to influence economic policies against those who benefit from pollution-generating activities. While the income coefficients are not statistically significant for some

[33] As classified by the World Bank in 1970.

environmental quality indicators such as smoke, airborne heavy particles and faecal coliform, others such as sulphur dioxide, dissolved oxygen, access to clean water and access to sanitation show an inverted U-shaped relationship[34]. Their regression results generally support their hypothesis that more equitable distributions of income and power result in improved environmental quality. The results for the non-income inequalities (literacy, political rights and civil liberties) have an even stronger effect on environmental quality, especially in low-income countries. Panayotou (2000) comments on this in his survey on the income-environment relationship:

> *"... and hence, Kuznets' original hypothesis of an inverted-U relationship between inequality and income, could be an additional channel through which economic growth might help improve environmental quality..." (p.25)*

Selden and Song (1994) add population density as an independent variable in their emission-income model to test for EKC. Their sample covers four air quality indicators for 30 countries classified as 22 high-income, six middle-income and two low-income countries from 1973-1984. Population density has the expected negative sign and a possible explanation is that more densely populated countries face greater pressure to reduce per capita emissions at all income levels compared to less densely populated countries. Although the turning point estimates seem not to be very sensitive to the inclusion of the population density, they are lower for all pollutants, except for suspended particular matter, when population density is included in the model.

Panayotou (1997) has pointed out that the reduced form income-environment relationship captures only the net effect of income on environmental quality with no explanation of the underlying forces that cause this correlation.

[34] Significant positive coefficients on the income cubed variable imply that the environmental quality indicators will increase eventually with increasing income.

"...the conventional approach is basically a 'black box': it hides more than it reveals since income level is used as a catch-all surrogate variable for all the changes that take place with economic development..." (p.466)

He attempts to shed some light on the possible determinants of environmental quality by incorporating policy considerations and the rate of economic growth, first into two conventional reduced-form environmental-growth models, and secondly into a structural model of the growth-environment relationship[35] for sulfur dioxide. The sample covers 30 developed and developing countries for the period 1982-1994. The first reduced-form model includes per capita income and population density as explanatory variables. The second reduced-form model includes, in addition to the previous explanatory variables, the rate of economic growth and a policy variable (represented by government policies, social institutions and completeness and functioning of markets). Although both models result in an inverted-U shaped relationship between SO_2 and per capita income, the overall fit of the model improves in the second model[36]. While all the income terms become less significant in the second model, both the growth rate and the policy variable are highly significant. Therefore, although both reduced-form models confirm the inverted-U shaped emission-growth relationship, they show that the steepness of curve depends on government policies, social institutions and markets.

"...Better policies, such as removal of distortionary subsidies and introduction of more secure property rights over resources and pollution taxes (or other efficient instruments) to internalize externalities, are expected to reduce the 'environmental price' of economic growth, thereby flattening out the income-environment relationship and possibly achieving an earlier turning point..." (p.469)

The study of Panayotou (1997) is therefore another indication on the effect of the ability to internalize certain externalities on the empirical inverted-U shaped relationship between some pollutants and income.

---

[35] The structural EKC model will be discussed in section 1.3.
[36] R^2=0.15 in the first model improved to R^2=0.24 in the second model.

Vincent (1997) has criticized the fact that previous cross-country and panel data studies on the income-pollution relationship were conducted on a sample of developed and developing countries, whose data contained little or no overlap between the observations. The resulting inverted-U shaped income-pollution relationship could be merely the result of placing the positive environmental quality-income relationship in developed countries alongside the negative environmental quality-income relationship in developing countries. He sets out to analyze the relationship between pollution and income using panel data for a single country, Malaysia, which comprises 13 states. As Malaysia has one of the fastest growing economies since the seventies, it provides enough income variation to test for the EKC relationship. Moreover, because all states share the same pollution policies, Vincent (1997) thinks that the 13 states of the Malaysian federation are an ideal sample for analyzing the relationship between pollution and income. He adds population density, as done by Selden and Song (1994), as an explanatory variable to his reduced-form income-pollution model which covers the period from the late 1970s to the early 1990s. The analysis fails to find an inverted-U shaped relationship for all pollutants in the study, suspended particulates (TSP) and five water quality indicators, biological oxygen demand (BOD), chemical oxygen demand (COD), ammoniacal nitrogen, pH and suspended solids. While higher income is associated with higher concentrations of TSP, ammoniacal nitrogen and higher pH values, there is no significant relationship for BOD, COD and suspended solids. Holding income and time constant, higher population density is associated with worse air and water quality as measured by TSP and ammoniacal nitrogen, and better water quality as measured by suspended solids.

> *"...The lack of evidence of EKCs in Malaysia does not prove that EKCs do not exist anywhere. It does indicate, however, that policymakers in developing countries should not assume that economic growth will automatically solve air and water pollution problems..." (p.430)*

Sigman (2002) provides empirical evidence that externalities increase pollution, in terms of biological oxygen demand (BOD), in international rivers. She compares pollution levels of 291 water quality monitoring stations, at domestic and international rivers, in 49 countries for the period from 1979-1996, and finds that international rivers appear to have higher BOD levels than domestic rivers. A reduced-form pollution-income relationship, which includes dummy variables for the station type (upstream, downstream or border station), dummy variables for the country type (EU or non-EU), and upstream population as explanatory variables, is estimated. The empirical results support the inverted-U shaped income-pollution relationship and a positive upstream population impact. She attributes the success of reducing pollution in international rivers inside the EU mainly to its international institutions. She argues that normally an upstream country would not alter its water polluting activities, which affect a downstream country, unless there is some kind of agreement between them.

> *"…Stations upstream of borders within the EU have statistically lees free-riding than those upstream of non-EU borders … the results provide some evidence that EU institutions successfully curb free-riding…" (p.1158)*

4.2.3. Alternative Explanations for the Environmental Kuznets Curve

The inability of the reduced form environment-income relationship to explain the mechanism through which changes actually happen, has led researchers to look for alternative models. They want to gain more insight into why the observed correlation between per capita income and environmental quality changes from positive to negative once some critical level of per capita income is reached.

Moomaw and Unruh (1997) are sceptical about the observed inverted U-shaped relationship between income and carbon dioxide emissions (CO_2), an important atmospheric gas contributing to global warming. They argue that CO_2 emissions are a global problem and that individual countries have less incentive to take unilateral actions to reduce its emissions. Therefore, economic theory would not predict, *a priori*, an inverted U-shaped GDP/CO_2 relationship.

"...It is important to recognize, however, that there are theoretical problems with the presence of EKC-like behaviour in CO_2 emissions data...." (p.452)

In an attempt to find alternative explanations for the conflicting conclusions about the observed empirical income/CO_2 relationship, Moomaw and Unruh (1997) compare the conventional reduced form relationship model with structural transition models for 16 industrial OECD countries for the period from 1950 to 1994. They identify two separate time periods, 1950-1973 and 1974-1992, with a structural break in 1973[37]. While there is a strong positive correlation between CO_2 and per capita GDP in the first period, carbon dioxide emissions became de-linked from economic growth in the second period. There is a weaker negative CO_2/GDP correlation for six countries, a positive CO_2/GDP correlation for two countries and insignificant results for the remaining countries. The main findings suggest that the transition to negative CO_2 emissions elasticities does not happen once a certain income level is reached, but as a result of a specific exogenous event in 1973 and subsequent economic policies since then. The discovery of the ozone hole has also triggered international efforts that have a positive impact on CFC emissions in all industrial economies.

"...Hence, in the face of shocks, structural-transition models replace the 'inverted-U' of the EKC with an 'inverted-V'..." (p.461)

The rapid changes in response to external historical shocks have resulted in negative CO_2/GDP and CFC/GDP correlations. But because individual countries differ in their response to external shocks, the structural transition model identifies a different inverted V-shaped relationship for each individual country. This is in contrast to the conventional EKC models, where there is a unique income turning point for all countries under consideration. In another study Unruh and Moomaw (1998) point out that even if high-income levels are not the driving force behind improving the emission/growth relationship,

[37] The Arab-Israeli war in 1973 triggered an oil embargo, which caused many OECD countries to stabilize or decrease the fossil-fuel energy intensity in their economic activities since then.

a certain level of income and technological development makes the adaptations to the external shocks possible.

> *"...Wealth may be a conditioning factor that allows certain countries to be 'first movers', but low income need not be a barrier to other countries achieving both lower pollution levels and stronger economies in the future..." (p.228)*

In a further attempt of Panayotou (1997) to understand why the observed environmental quality-income relationships exists, he analyzes, in addition to reduced-from relationship models[38], a structural model that identifies the forces through which income affects the environment. These distinct effects are: first, the scale effect (i.e., the scale of economic activity per unit of area), secondly, the composition effect (i.e., the share of industry in GDP) and thirdly, the abatement or pure income effect (i.e., effect of income on the demand and supply of pollution abatement). The scale effect on pollution is expected to be an increasing function of income, assuming both the composition effect and pure income effect to be constant. The composition effect on pollution is expected be an inverted-U shaped function of income, assuming both the scale effect and the abatement effect to be constant. This is because the economic structure is represented by the share of industry in GDP, which is likely to rise first as income rises and declines later as the economy moves through the different stages of economic development. The pure income or abatement effect on pollution is expected to be a decreasing function of income, assuming the scale effect and composition effect to stay constant. In addition to decomposing the income effect into its different components, Panayotou (1997) includes population density, economic growth and a policy variable as additional variables into his model to explore the relationship between SO_2 and income more analytically. His panel data covers 30 countries from 1982-1994. His structural model improves considerably[39] compared to the conventional reduced form model. The significance of decomposing income into its different effects is apparent from the unexpected result, which indicates that better

[38] The extended conventional EKC reduced-form model of Panayotou (1997) was discussed in Section 1.2.
[39] R^2 rose from R^2=0.24 in the conventional reduced-form model to R^2=0.50 in the structural model.

environmental policies and institutions lead to larger environmental improvements at higher income levels compared to lower income levels.

> "*...when all effects* [scale effect, composition effect and abatement effect] *are considered, the relationship between growth and the environment turns out to be much more complex with wide scope for active policy intervention to bring about more desirable and more efficient economic and environmental outcomes...*" *(p.483)*

To gain more information on the underlying forces that determine the estimated U-shaped relationship between some pollutants and income, De Bruyn (1997) decomposes SO_2 emissions for the Netherlands and West Germany during the 1980's into its main sources. First, the scale effect of economic activity on emissions, which is positive, provided that there are no changes in the other effects. Secondly, the structural effect (also called inter-sectoral effect) which detects changes of the composition of economic activity on emissions and can be either positive or negative, assuming that there are no changes in the other effects. Thirdly, the technological effect (also called intra-sectoral effect) which detects changes of emission intensities on emissions within sectors. It appears from the empirical study that the decline in SO_2 emissions in both the Netherlands and West Germany since the seventies and especially during the eighties can be explained by the interplay of two opposing effects, namely the scale effect and the technological effect. There is no evidence of any structural effect on SO_2 emissions and the increases in emissions due to the scale effect are more than compensated by the decline in emissions due to technological improvements in production techniques. It is also known that SO_2 causes transboundary air pollution problems in neighbouring countries through wet and dry deposition, and that there is generally little motivation for individual countries to engage in unilateral actions to reduce air pollution beyond their borders.

> "*...technological changes, probably triggered by the rapid energy price increases after 1973, resulted in a decrease of energy consumption...*" *(p.492)*

"...the largest part of reduction in SO_2 emissions due to technological change could hence be attributed to the instalment of end-of-pipe technology..." (p.423)

De Bruyn (1997) argues that technological change might have come about through environmental policies, which had their driving force from historical events, such as the oil crisis in the 1970's, and from negotiations on transboundary air pollutants, such as the First and Second Sulphur Protocols[40].

All reduced form EKC models assume a unidirectional causality between income and environmental degradation without addressing the possibility that a lot of environmental degradation might be irreversible in the future. De Bruyn *et al.* (1998) introduce dynamics into the reduced-form EKC-model to estimate the impacts of income, technological and structural change, and environmental policies on the variation in the emission intensity for individual countries over time. Their sample covers carbon dioxide (CO_2), nitrogen oxide (NO_x) and sulfur dioxide (SO_2) for the Netherlands, the United Kingdom, the United States, and West Germany for various time intervals between 1961 and 1993. Their results, as summarized by Panayotou (2000), indicate that the time pattern of emissions correlates positively with economic growth and that reductions in emissions can be attributed to structural and technological change. Their model enables them to determine 'the sustainable growth rate' for each individual country i.e.,

"...the rate of economic growth that can be associated with a non-increasing level of emissions..."(p.171)

As in only half of the investigated cases, income accumulation can explain the reduction in the level of emissions, the authors conclude that the assumption that economic growth leads to improvements in environmental quality cannot be supported by evidence for the pollutants and countries in their analysis.

---

[40] The First Sulphur Protocol was signed in 1985 in Helsinki by 20 countries agreeing to reduce annual SO_2 emissions in 1993 uniformly by 30% compared to 1980. The Second Sulphur Protocol was signed in 1994 in Oslo by 27 countries agreeing to non-uniform reductions percentages in their annual SO_2 emissions.

Conte Grand (1999) analyzes the link between environmental agreements and pollution levels in international rivers and lakes to determine the effects of such agreements on the actual levels of pollution in international water bodies. She finds evidence for a relationship between water quality in some international rivers and their respective agreements:

> *"...the more the watershed is under the influence of specific treaties which deal with its pollution, the better is the state of the water..." (p.14)*

From the presented literature review on the relationship between income and environmental quality, two points are of interest to our study. First, evidence suggests that there exists an Environmental Kuznets curve for at least some air and water pollutants, with pollution levels rising with income at low levels of income and pollution levels declining with income at high levels of income. Secondly, evidence suggests that the income levels at which air pollution starts to decline is generally higher for global air pollutant than for local air pollutants. This second point gives rise to an important question: Is the relationship between income and water quality also affected by the nature of the water pollutant? As water is the medium that carries water pollutants, the nature of the water body determines whether a pollutant is local or regional. Therefore, we first analyse the relationship between income and water pollution in domestic and international rivers. If it becomes evident that the nature of the river has some influence on the income-water quality relationship, due to upstream-downstream externalities, the search continues to look for possible underlying factors. The current economic literature tries to capture this externality effect through variables that are all non-income related, such as geographical location and the existence or non-existence of an agreement. Our study tries to capture this externality through an income-related variable, namely the income of the upstream country. We argue that if national income is an important determinant for water quality in domestic rivers, then national income could also have an effect on the water quality in international rivers. The water quality of international rivers in downstream countries can therefore be unilaterally affected by the income of the upstream country through the externality effect.

4.3. The Data

In this section we provide information on the source and the characteristics of the data used in the analysis. We start with the water quality indicators in section 4.3.1, continue with the source of income data in section 4.3.2. and conclude with the externality indicators in section 4.3.3..

4.3.1. Water Quality Indicators

The annual water quality indicators are made available by Grossman and Krueger (1995) for the period from 1979 until 1990[41]. A summary of the data, which is published by the United Nations Environment Programme (GEMS/Water), provides triennial data from monitoring stations on rivers, lakes and groundwater aquifers in 69 countries from 1979 until 1999 for 13 basic water pollutants. These data are location-specific to the water quality monitoring station, as the GEMS database does not provide aggregate country-level measures of pollution. The water pollutants in this chapter can be divided into three broad categories: the oxygen balance in rivers and nitrates, the pathogenic contamination, and heavy metals.

4.3.1.1.The Oxygen Balance in Rivers and Nitrates

The oxygen balance in rivers is crucial for aquatic plants and animal life. Organic matter such as dead plants and leaves, manure, sewage and food waste is decomposed by micro-organisms such as bacteria using much of the available oxygen in rivers, thereby depriving fish and other aquatic organism of the oxygen they need to live. Dissolved Oxygen (DO) is the amount of oxygen that is dissolved in water. It comes from the oxygen in the air that has dissolved in the water or from photosynthesis of aquatic plants. Generally, a higher concentration of dissolved oxygen indicates better water quality. The concentration of dissolved oxygen, should ideally be between 9 and 11 mg/l depending on the water temperature[42]. Biological Oxygen Demand (BOD) and Chemical Oxygen (COD) demand are both indirect measures of the oxygen balance of rivers. While BOD measures the amount of oxygen that is needed by organisms to decompose organic waste, COD measures the amount of oxygen consumed when a chemical oxidant is added to a water sample. The GEMS study (1995) considers that BOD levels of 2 mg/l indicate low levels or organic pollution. Higher BOD levels are usually measured at locations, which are downstream of municipal wastewater discharges and/or industrial effluents. While there is an inverse relationship between BOD levels and dissolved oxygen, there is a positive relationship between BOD and COD levels. When BOD levels are high, dissolved oxygen levels decrease, because the available dissolved oxygen is consumed by bacteria. Nitrates may come from fertilisers, sewage, animal wastes and industrial waste. High levels of nitrates in a river cause plant life and algae to flourish. When plants grow quickly they usually die quickly as well, thereby contributing to the organic waste in rivers, which then needs dissolved oxygen to be decomposed by bacteria. Therefore, rivers with high concentrations of nitrates usually have high levels of BOD and consequently low levels of dissolved oxygen.

4.3.1.2.The Pathogenic Contamination

(GEMS 1995) reports also that rivers polluted by untreated sewage carry a range of harmful pathogens such as bacteria[43], viruses and protozoans, accountable for 5 million deaths globally every year. As these pathogens are difficult to detect, most water quality surveys rely on total coliform and faecal coliform as indicators of the existence of faecal contamination. These are measured in *counts/100ml* and are considered the most variable

---

[41] The data are available via ftp, at ftp://irs.princeton.edu in the environ directory

[42] Colder water can hold more oxygen in it than warm water, because the metabolic rates of aquatic organisms increase in warm waters.

[43] These include: Shigella, Salmonella, Cholera Vibrio, and Escherichia (see GEMS (1995)).

of all GEMS water quality indicators, even for the same water quality monitoring station. The counts range from less than 100/100ml in relatively uncontaminated rives, to between 1000 and 10,000/100ml in moderately contaminated rivers, to more than 100,000/100ml in rivers that receive untreated sewage. Grossman and Krueger (1995) consider faecal coliform a better water quality indicator than total coliform, because the latter includes some organism which are found naturally in the environment. Therefore, we consider it sufficient to estimate only the relationship between faecal coliform and economic growth in this chapter. We follow Grossman and Krueger (1995) in measuring faecal coliform in log (1+concentration level)[44].

4.3.1.3.Heavy Metal Contamination

Heavy metals, such as lead, cadmium, arsenic, mercury and nickel, which are discharged by industrial and agricultural activities into rivers, cause many damaging diseases[45] through drinking water and eatable fish. In this chapter, we restrict our attention to lead and cadmium, because the sample sizes of the other heavy metals are not large enough for a separate analysis on domestic and international rives.

Table 4.1. provides the summary statistics for the water quality indicators examined in this chapter according to the different types of the rivers.

[44] log was chosen because faecal coliform grows exponentially and its distribution is highly positively skewed; one had to be added to the concentration level, as log cannot be taken from a zero measurement.

[45] According to Grossman and Krueger (1995), lead might bring about convulsions, anaemia, kidney damage brain damage, cancer and birth defects; cadmium might result in tumours, renal dysfunction, hypertension, and arteriosclerosis; arsenic leads to vomiting, poisoning, liver damage and kidney damage; mercury contributes to irritability, depression, kidney and liver damage, and birth defects; nickel leads to gastrointestinal and central nervous system damage and cancer

Table 4.1.

Statistics Summary of the Dependent Variables (annual mean concentration of pollutants) and per Capita Income (in thousands of 1996 US$)

Variables	All rivers BS-model	Domestic rivers BS-model	International rivers BS-model	International rivers UD-model
Dissolved Oxygen *(mg/l)*				
mean	8.1116	8.2386	7.7934	7.6504
st.deviation	3.0737	3.4815	2.7052	2.8311
per capita GDP				
mean	10.0021	8.9908	12.5349	12.7342
st.deviation	7.3704	7.2454	7.0732	7.2897
sample size	1563	1117	446	392
BOD *(mg/l)*				
mean	6.6946	5.3790	9.9908	11.1583
st.deviation	22.9830	15.1395	35.5661	38.5026
per capita GDP				
mean	8.0149	6.8573	10.9155	10.8847
st.deviation	6.4744	6.0745	6.5442	6.7819
sample size	1248	892	356	302
COD *(mg/l)*				
mean	49.1705	37.0130	79.5126	87.5568
st.deviation	121.2475	68.4242	196.3098	207.8481
per capita GDP				
mean	7.5895	5.6792	12.35727	11.96886
st.deviation	6.6088	5.4900	6.7643	7.0708
sample size	825	589	236	208

To be continued

Table 4.1.continued: Statistics Summary of the Dependent Variables and per Capita Income (in thousands of 1996 US$)

Variable	All rivers BS-model	Domestic rivers BS-model	International rivers BS-model	International rivers UD-model
Nitrates ***(mg/l)***				
mean	1.1183	1.0947	1.1918	0.7880
st.deviation	2.8510	3.1629	1.5086	1.2233
per capita GDP				
mean	9.3277	9.2019	9.7209	9.3229
st.deviation	7.3597	7.5136	6.8563	7.2237
sample size	982	774	238	186
Faecal Coliform[(a)]				
mean	6.8133	7.0258	6.3286	6.1185
st.deviation	3.1866	3.2552	2.9721	3.0378
per capita GDP				
mean	10.4175	9.2559	13.0674	13.0856
st.deviation	7.5738	7.4969	7.0734	7.4089
sample size	1260	876	384	340
Cadmium ***(μg/l)***				
mean	0.0389	0.0571	0.0045	0.0051
st.deviation	0.1587	0.1931	0.0237	0.0255
per capita GDP				
mean	15.1675	14.0821	17.2232	17.6526
st.deviation	5.8996	6.4895	3.8220	3.8021
sample size	628	411	217	186
Lead ***(μg/l)***				
mean	0.0314	0.0242	0.0454	0.0515
st.deviation	0.2925	0.0607	0.4963	0.5380
per capita GDP				
mean	15.0931	13.8566	17.5125	18.0144
st.deviation	5.9907	6.6590	3.2545	3.0642
sample size	615	407	208	177

(a) measured as log (1+ mean concentration level in count/100ml)

There seems to be some differences both in the average pollution levels and in the average per capita income levels among the different river types, especially between the domestic rivers and the international rivers.

Domestic rivers appear to have higher average concentrations of dissolved oxygen and lower average concentrations of BOD, COD, nitrates and lead compared to international rivers. This gives some indication that domestic rivers are on average less affected by some pollutants than international rivers. The exceptions are faecal coliform and cadmium. According to the GEMS study (1995), if water quality monitoring stations are located far enough downstream of sewage discharges, the counts of faecal contamination will be less than 100/100ml. It could be that international water quality monitoring stations are on average closer to political borders, while domestic water quality monitoring stations are on average closer to highly populated areas. If this is actually the case, then it can be expected that domestic rivers appear to be more affected by faecal contamination than international rivers. As for cadmium, either the sample is not large enough to get accurate averages of cadmium contamination in domestic and international rivers, or there is some other explanation for this unexpected result.

4.3.2. Income Indicators

Income per capita, or more precisely per capita gross domestic product (GDP), is the most widely used independent variable in exploring the reduced form relationship between economic development and environmental degradation. In many cases, this per capita GDP is measured in real purchasing power parity (PPP), where the national incomes are transformed using a common set of international prices to make comparisons among the different countries possible. In this chapter, the central per capita GDP variable is measured by Heston *et al.* (2002) in PPP with the base year 1996, and published as the Penn World Tables (Mark6)[46]. All other income related independent variables in this study are derived from this basic form of per capita GDP.

[46] The data are available at http://pwt.econ.upenn.edu/

4.3.3. Externality Indicators

This study expands the basic income-environmental quality relationship by adding 'upstream per capita income' as an explanatory variable. This new variable takes into account explicitly the effects of economic growth in an upstream country on the water quality of the river in the downstream country. This effect will be called the income related externality effect throughout this chapter and as far as I know has not been analyzed before in this context. The upstream per capita income is derived from the same central per capita GDP as the income indicators. Specifying whether the international water quality monitoring station lies in an upstream or in a downstream country provides us with the non-income related upstream effect. Taking both the income related and the non-income related externality effects into consideration is a modest attempt to determine the forces that affect the unidirectional externality in river sharing countries.

4.4. Methodology

To gain more insight into the relationship between water quality in rivers and income, we basically estimate two reduced form regressions, which relate a chosen water quality indicator to a function of current and lagged per capita income of the country where the measurement is taken. The main advantage of the reduced form approach is that it captures the net effect of a country's per capita income on water quality in rivers directly; moreover the necessary data are usually available and reliable. The main disadvantage of the reduced form approach is that it reveals nothing about the underlying determinants that cause the relationship to exist[47], i.e., it cannot explain why income affects pollution in this manner.

[47] An alternative to the reduced form approach is a model of structural equations, where environmental regulations, technology and industrial composition are first linked to per capita income and then to the level of pollution. Although the net effect of income on pollution can then be determined by solving back, the implied estimates might not be very reliable, as it is usually very difficult to obtain reliable data on environmental regulation and technology.

The first regression, hereafter named the Basic mode (BS-model) relates the water quality indicator to the current and lagged per capita income of the country where the water quality is measured. The second regression, hereafter named the Upstream-downstream model (UD-model) relates the water quality indicator not only to the current and lagged per capita income of the country where the water quality is measured, but also to the current and lagged per capita income of the upstream country. While the Basic model will estimate the relationship between water pollution and per capita income for both domestic and international rivers, the Upstream-downstream model will only estimate the relationship between water pollution and per capita income in international rivers.

The study will estimate the following reduced form regressions:

The Basic Model (BS-model)

$$p_{it} = \alpha_i + \beta_t + \sum_{\tau=0}^{1}\sum_{n=1}^{3}\Gamma_{n\tau} y_{i,t-\tau}^{n} + \kappa\eta_t + \phi_{it} \qquad (1)$$

The Upstream-downstream Model (UD-model)

$$p_{it} = \alpha_i + \beta_t + \sum_{\tau=0}^{1}\sum_{n=1}^{3}\Gamma_{n\tau} y_{i,t-\tau}^{n} + (1-\Delta_i)\sum_{\tau=0}^{1}\sum_{\eta=1}^{3}\mathrm{E}_{n\tau} y_{j,t-\tau}^{n} + \zeta\Delta_i + \kappa\eta_t + \phi_{it} \qquad (2)$$

where p_{it} = water quality indicator at station i in year t

α_i = station specific effect (not upstream/downstream effect)

β_t = year-specific/time-trend effect

Γ = coefficients on income-related pollution effect (income and/or lagged income)

y_{it} = income of the country where the station i is located in year t

E = coefficients on income-related upstream/downstream effect

y_{jt} = income of country j upstream of the country where station i is located

ζ = coefficient on non-income-related upstream/downstream effect

$\Delta_i = 1$ if the country where the station i is located is an upstream country

$\Delta_i = 0$ if the country where the station i is located is not an upstream country

κ = coefficient of mean temperature

η_{it} = mean temperature of water at station i in year t

ϕ_{it} = error term

For all water quality indicators, except faecal coliform, the dependent variable (p_{it}) is measured as the mean concentration level at a certain water quality monitoring station over the course of one year. For faecal coliform the dependent variable is measured as log (1+ mean concentration level) for the same reasons mentioned by Grossman and Krueger (1995)[48].

We include a linear time trend (β_t) as a separate regressor in both models to adjust for the year in which the measurements of the water quality indicators were taken. This time trend serves as a proxy for environmental improvements that are not caused by increased per capita incomes, but by global environmental improvements. These are often the result of increasing global concern over current and prospective environmental problems, and/or of the many global advances in environment-related technology that make improvements affordable.

We also include the mean annual water temperature (η_{it}) for the river, where the water quality monitoring station is located, as a separate regressor to account for site-specific effects on the water quality that are not related to per capita income. The mean annual water temperature is considered an appropriate site-specific explanatory variable, as the rate of chemical and biological processes are strongly influenced by water temperature. For

[48] log was chosen because faecal coliform grows exponentially and its distribution is highly positively skewed; one had to be added to the concentration level, as log cannot be taken from a zero measurement.

example, warmer water speeds up the rate of growth and death of aquatic plants and organisms. This results in higher BOD levels, because bacteria require oxygen to decompose the organic matters. The missing mean annual water temperatures are estimated by regressing temperature on the latitude of the location of the water monitoring stations, as done by Grossman and Krueger (1995).

One of the important objectives of this study is to determine the shape of the relationship between income and environmental degradation. Therefore, we follow Grossman and Krueger (1995) in choosing a cubic functional form for the income variables (y_{it}) and (y_{jt}) to allow for the needed flexibility. As most other empirical studies we use country level per capita GDP, even if all the water quality indicators are on a very local level, namely the water quality monitoring station. This seems reasonable, because on one hand local per capita GDP is not easily available and is not comparable among different countries as the data from Penn World Tables; and on the other hand, environmental regulations are usually applied at country level.

Grossman and Krueger (1995) include a cubic term of the average per capita GDP of the three previous years as a proxy for permanent income, which they consider an important factor in current environmental quality. They argue that because current and lagged per capita GDP are highly correlated, including only current income or only permanent income would not qualitatively change the estimation results. As we are interested in the overall effect of income we follow Grossman and Krueger (1995) in using a proxy for permanent income; but we use the per capita GDP for the previous year only. The reason for this is the limited number of available observations, especially for international rivers which account for roughly one third of all the rivers in our sample, and our wish to keep the sample sizes as large as possible.

While we apply the Basic model to three of our four settings (all rivers, domestic rivers, and international rivers), we apply the Upstream-downstream model only to international rivers. The UD-model takes into account explicitly the effect of the per capita income of the upstream country on the pollution level of the downstream country. This requires the construction of a dummy variable (Δ_i) indicating whether a station is located in an upstream country ($\Delta_i = 1$), or in a downstream country ($\Delta_i = 0$). This dummy variable enables us to differentiate between the income related upstream effects (E) on pollution and the non-income related upstream effects (ζ) on pollution.

We follow Grossman and Krueger (1995) in choosing the random effects approach to estimate equation (1) and (2). This is because there are probably some individual characteristics for each water quality station, which affect pollution levels but are unaccounted for in the explanatory variables. These omitted variables could result in temporal correlation of the error term. To avoid this problem, we employ a random effects model, where $\phi_{it}' = \alpha_i + \phi_{it}$ is considered an error term with two components: a site-specific random component that does not change over time (α_i), and a remainder error that is uncorrelated over time (ϕ_{it}). We then derive the maximum likelihood random-effects estimator under the assumption that the explanatory variables are strictly exogenous, i.e., the error terms are uncorrelated with the past, current or future values of the explanatory variables. Verbeek (2000) suggests the use of the random effects approach, if the study is not primarily interested in the particular value of the individual error component. Hsiao, (2003) suggests to treat the effects of omitted variables as random, if we want to make inferences beyond the effects in the model, i.e., inferences about a population of effects, from which the effects in our model are a random sample. This usually happens if the number of individuals is large, and we are not interested in each individual effect, but rather in the characteristics of the population where the individuals were drawn from.

The analysis in our study proceeds in the following steps: First, we estimate the Basic reduced form regression model relating seven water quality indicators and per capita incomes in three settings: all rivers, domestic rivers and international rivers. The purpose is to determine whether the income-water quality relationship is affected by the nature of the river. Secondly, we estimate the Upstream-downstream reduced form regression model for the international rivers only, to determine the effect of including the externality variables on the estimation results for international rivers. Thirdly, we calculate the critical levels of income at which each water quality indicator is expected to reach its maximum (or minimum for dissolved oxygen), if such an income level can be identified. This facilitates the comparison among the different natures of the rivers and between the two regression models. Lastly, we hope to shed some light on the effects of ignoring inter-country pollution on the intra-country turning points of water pollution in downstream countries.

4.5. Results

We have estimated regression (1) and regression (2) for all seven pollutants: dissolved oxygen, biological oxygen demand, chemical oxygen demand, nitrates, faecal coliform, cadmium and lead. The following two tables, table 4.2. and table 4.3, show the significance[49] of three different combinations of income coefficients: a) the six income and lagged income variables taken together b) the six upstream income and lagged upstream income variables taken together and c) the 12 income, lagged income, upstream income and lagged upstream income variables taken together. The lower (higher) the probability value, the higher (lower) the level of significance, i.e., the lower (higher) the probability of rejecting the true null hypothesis. If, for example, the probability value is 0.05, there is a chance of being wrong 5% of the time, if the true null hypothesis is rejected. The variables had to be combined, because income coefficients suffer from strong multicollinearity[50] among them, which makes it difficult to infer anything from individual income coefficients. The detailed MLE estimators for each individual pollutant are reported in Appendixes 1.

In tables 4.2. and 4.3. we examine the following question: is there a significant relationship between concentrations of water pollutants and per capita income irrespective of the river type? First, for *all rivers*, the tables show that national income is jointly significant at less than 5 percent in four cases (dissolved oxygen, COD, faecal coliforms and cadmium), at

[50] There is expected to be a strong multicollinearity between current and lagged income on one hand, and between the different powers of income on the other hand.

Table 4.2.

Significance ($prob > \chi^2$) for Income-related Effects of the Oxygen Balance and Nitrates Contamination in Rivers

	Dissolved Oxygen	Biological Oxygen Demand	Chemical Oxygen Demand	Nitrates
All Rivers (BS-Model)				
Income & lagged income combined	0.0020	0.3700	0.0479	0.0865
Domestic Rivers (BS-Model)				
Income & lagged income combined	0.0121	0.6549	0.0003	0.1568
International Rivers (BS-Model)				
Income & lagged income combined	0.0004	0.4040	0.1218	0.0142
International Rivers (UD-Model)				
-Income and lagged income combined	0.0091	0.3398	0.0028	0.0002
-Upstream income & lagged upstream income combined	0.0636	0.9800	0.2295	0.0001
-All income variables	0.0009	0.8437	0.0507	0.0000
-Non-income upstream effect	0.6247	0.5042	0.3608	0.0004
-All upstream variables	0.0154	0.2220	0.0173	0.0000

less than 10 percent in one case only (nitrates) and insignificant in two cases (BOD and Lead). Secondly, for *domestic rivers*, the results show that national income is jointly significant at less than 5 percent in the same previous four cases (dissolved oxygen, COD, faecal coliforms and cadmium), at less than 10 percent in one case only (lead), at 15.7 percent in one case (nitrates) and insignificant in the remaining case (BOD). Thirdly, for *international rivers (BS-model)*, the results indicate that national income is jointly significant at less than 5 percent in two cases only (dissolved oxygen and nitrates), at 12.2 percent in one case (COD) and insignificant in the remaining cases (BOD, COD, faecal coliforms, cadmium and lead). Lastly, for *international rivers (UD-model)*, the results improve somewhat again and show that national income is jointly significant at less than 5 percent in four cases (dissolved oxygen, COD, nitrates and cadmium), at less than 10 percent in one case only (faecal coliforms), and insignificant in the two remaining cases (BOD and lead).

Table 4.3.

Significance ($prob > \chi^2$) for Income-related Effects of Faecal Contamination and Heavy Metals Contamination in Rivers

	Faecal Coliforms	Cadmium	Lead
All Rivers (BS-Model)			
Income & lagged income combined	0.0245	0.0002	0.9355
Domestic Rivers (BS-Model)			
Income & lagged income combined	0.0006	0.0001	0.0812
International Rivers (BS-Model)			
Income & lagged income combined	0.4980	0.8918	0.9649
International Rivers (UD-Model)			
-Income & lagged income combined	0.0677	0.0169	0.9893
-Upstream income & lagged upstream income combined	0.0001	0.0738	0.9999
-All income variables	0.0001	0.1833	0.9991
-Non-income upstream effect	0.0771	0.9417	0.9684
-All upstream variables	0.0000	0.0075	0.9980

It appears therefore that national income could be an important determinant for river pollution, especially for domestic rivers, and that the UD-model is better suited for international rivers than the BS-model.

We also want to explore the central question of our study: is there a significant relationship between concentrations of water pollutants and upstream effects? Looking at tables 4.2 and 4.3., we find that upstream effects (income related and non-income related) are jointly significant at less than 5 percent in five cases (dissolved oxygen, COD, nitrates, faecal coliforms and cadmium) and insignificant in the remaining cases (BOD and lead). It has to be noticed that the non-income related upstream effect is significant at less than 5 percent in one case only (nitrates), at less than 10 percent in one case (faecal coliform) and insignificant in the remaining five cases (dissolved oxygen, BOD, COD, cadmium and lead). On the other hand, the results for the income related upstream effects indicate that upstream income is jointly significant at less than 5 percent in two cases (nitrates and faecal coliforms), at less than 10 percent in two cases (dissolved oxygen and cadmium), and insignificant in the remaining cases (BOD, COD and lead). It therefore appears that, although upstream effects in general could play an important role in river pollution of international rivers, income related upstream effects seem to play a more important role than non-income related upstream effects.

Table 4.4. answers the following questions: what is the shape of the income-water pollution relationship? Is this relationship for *international rivers* affected by upstream unidirectional externalities? The table shows the level of national incomes at which dissolved oxygen reaches its minimum, and all other water pollutants reach their maximum. While dissolved oxygen is a direct measure of water quality, all other water quality indicators are inverse measures, because they measure the level of water pollutants. The turning points, the peaks and troughs, for the Basic model were calculated under the assumption that current income equals lagged income[51]. The turning points for the

[51] By multiplying the income, squared income and cubed income by the sum of the estimated coefficients for current and lagged income.

Upstream-downstream model were calculated under the additional assumption that income equals upstream income[52]. The standard errors were calculated using the Delta method.

The results indicate that, in accordance with many other empirical studies[53], dissolved oxygen follows a U-shaped relationship with per capita income, while all other water quality indicators display an inverted U-shaped relationship with per capita income for all river types, except for faecal coliforms in *all rivers*, which decreases steadily.

18 of the estimated income turning points are statistically significant at the 2 percent level, two estimates are significant at the 5 percent level, one estimate is significant at the 10 percent level, and six estimates are statistically insignificant[54]. This allows us to comment on an observed pattern among the income turning points. Table 4.4. shows that each water quality indicators peaks[55] at a lower level of income in domestic rivers than in international rivers. The range of significant income turning points for the pollutants in domestic rivers is from US$ 5,249 to $ 8.988 in 1996 prices, in international rivers from US$ 12,996 to US$ 20,767 in 1996 prices (for the BS-model), and from US$ 9,016 to US$ 18,467 in 1996 prices (for the UD-model).

To facilitate comparisons we adopt the country classification of the World Bank[56], where economies are divided according to their GNI per capita in the 2001 into four different income groups:

(a) low income countries (LI) for incomes of $745 or less
(b) lower middle income countries (LMI) for incomes between $746 and $2,975
(c) upper middle income countries (UMI) for incomes between $2,976 and $9,205
(d) high income countries (HI) for incomes of $9,206 and more

[52] By multiplying the income, squared income and cubed income by the sum of the estimated coefficients for current income, lagged income, upstream income and upstream lagged income
[53] See literature review in section 4.2.
[54] There is no turning point estimate for one out of the 28 cases.
[55] Dissolved oxygen reaches its minimum.

Table 4.4.

Estimated Critical per Capita Incomes at Peak Pollution Levels (in thousands of 1996 US$)
Standard Errors[(a)] in Parentheses
Average Income in Italics

Pollutant	Critical GDP all rivers BS-Model	Peak GDP domestic rivers BS-Model	Peak GDP intern'l rivers BS-Model	Peak GDP intern'l rivers UD-Model[(c)]
Dissolved Oxygen[(b)]	1.772 LMI (153.498) *10.002*	2.323 LMI (12.912) *8.991*	12.996 HI (6.424) *12.535*	10.592 HI (8.955) *12.734*
BOD	11.520 HI (3.998) *8.015*	7.488 UMI (2.236) *6.857*	14.440 HI (3.579) *10.916*	15.532 HI (2.357) *10.885*
COD	11.917 HI (3.612) *7.590*	8.988 UMI (1.746) *5.679*	16.053 HI (6.879) *12.357*	9.016 UMI (2.871) *11.969*
Nitrates	9.444 HI (3.322) *9.328*	8.093 UMI (1.918) *9.202*	15.657 HI (0.513) *9.721*	14.473 HI (1.334) *9.329*
Faecal Coliform	no turning point *10.416*	5.249 UMI (2.508) *9.256*	20.767 HI (1.849) *13.067*	18.467 HI (0.998) *13.086*
Cadmium	16.182 HI (2.931) *15.168*	15.106 HI (11.956) *14.082*	15.056 HI (3.674) *17.223*	27.714 HI (60.041) *17.653*
Lead	15.650 HI (2.264) *15.093*	7.346 UMI (4.049) *13.857*	14.652 HI (6.260) *17.513*	14.421 HI (13.639) *18.014*

(a) the standard errors were calculated using the Delta method
(b) for dissolved oxygen the trough is reported rather than the peak
(c) assuming upstream current and lagged incomes are equal to current and lagged incomes

[56] For more detailed information on the World Bank country classifications and methods used to calculate GNI per capita, see http://www.worldbank.org/data/countryclass/countryclass.html

We also include the average per capita income of each sample group in table 4.4., because relatively more high-income countries provide data on international river pollution compared to low-income and middle-income countries, which could affect the results. While all significant income turning points for domestic rivers are in the income range of upper middle-income countries, 10 out of 11 significant income turning points for international river in both models are in the income range of high-income countries. These results suggest that upstream effects influence the critical level of income at which water quality starts to improve.

Our second observation is that for three out of four significant cases[57], the concentrations of pollutants in international rivers peak at a lower level of income when upstream effects are considered. Because the difference in income-turning points is not significant, it is too early to infer from this that ignoring inter-country externality effects leads to an upward bias in the intra-country income turning points for water pollutants in international rives. This could have happened because most data on international rivers comes from high-income countries.

4.6. Summary and Conclusion

We have examined two reduced form relationships between per capita income and water contamination in rivers. Our purpose is to detect the unidirectional externality effect of water pollutants that are discharged by upstream countries and borne mainly by downstream countries. The first reduced form relationship, the Basic model, relates water quality indicators to national per capita GDP of the countries where the monitoring station is located. This was done for rivers in general, for domestic rivers only, and for international rivers only. The second reduced form relationship, the Upstream-downstream model, relates the same water quality indicators, but for international rivers only, to both the national per capita of the country where the monitoring station is located and to the per capita GDP of the upstream country, where available.

[57] COD, nitrates and faecal coliforms.

In accordance with other empirical studies on the income-environmental quality relationship, our results confirm that national income could be an important factor in determining water quality in rivers, especially for *domestic rivers* and for *international rivers* if the upstream effects are considered. With these results in mind, we ask the following question: are international environmental agreements (with appropriate compensation for the upstream country) the only way to internalize unidirectional externality effects in international rivers? To gain more insight into how this unidirectional externality might operate and affect water quality in downstream countries, we break up the upstream effect into an income related upstream effect and a non-income related upstream effect. While there is a significant relationship between water quality and the overall upstream effect for five out of seven pollutants (dissolved oxygen, COD, nitrates, faecal coliforms and cadmium), and a significant relationship between water quality and the income related upstream effect for four out of seven pollutants (dissolved oxygen, nitrates, faecal coliforms and cadmium), there is a significant relationship between water quality and the non-income related upstream effect for two pollutants only (nitrates and faecal coliforms). These results suggest that income related unidirectional externality effects might be more important than non-income related unidirectional externality effects. As with all reduced form approaches, we can only identify variables that affect the dependent variable, without making any inferences on how this affect actually works. All that can be said is that while national per capita income has a significant effect on five out of seven water quality indicators in *all rivers* and *domestic rivers*, upstream per capita has a significant effect on four out of seven water quality indicators in international rivers.

Our next results on income turning points suggest that the type of the water quality indicator, whether it is local or regional, affects the income-water quality relationship. If the river is domestic and therefore confined to one country only, then the water pollutant is considered local, because it only affects this one country. If the river is international and therefore flows through more than one country, then the water pollutant is regional, because it affects more than one country. Many empirical studies were able to show that air pollutants with local or direct impacts reach their maximum levels at lower levels of income than air pollutants with more indirect, regional or global impacts. Our results point in the same direction for local and regional water pollutants in rivers. Without exception, all river pollutants peak[58] at a lower level of income in domestic rivers than in international rivers. Even if some[59] of the income turning points are not statistically significant, our results lay the foundation for additional empirical research in the future with larger data sets. It is too early to tell whether ignoring the unidirectional inter-country externality effect leads to an upward bias in the estimated intra-county income turning points, where the pollutants reach their maximum (minimum for dissolved oxygen) level. This result could also happen because for all pollutants, we have more data on international rivers in higher income countries than on international rivers in lower income countries. We need more comparable data from more countries on all types of rivers, but especially on international rivers to get more accurate estimates and more precise results. Our analysis is restricted to comparing domestic to international rivers, excluding rivers that change from domestic to international or vice versa. From an empirical point of view, it would be interesting to analyze the effect of changes of political borders on pollution levels of domestic rivers that have turned international (as result of the break-up of the former Soviet Union) and on international rivers that have turned domestic (as result of the unification of Germany). As these changes in the political borders have happened only in the last ten to fifteen years, we might have to wait until sufficient data becomes available.

Although more data, especially on international rivers, would benefit this analysis, we contribute to the literature in three ways. First, we distinguish between local and regional water pollutants through the nature of the river, which has not been done in this context before. Secondly, we are able to identify an income related externality effect and separate it from the overall upstream effect, while all previous studies that have identified upstream effects have looked only at non-income related explanations. Thirdly, we confirm the observed pattern that local air pollutants tend to peak at lower income levels than regional or global pollutants, for water pollutants. This has not been done for water quality indicators as far as I know.

[58] For dissolved oxygen, the only direct measure of water quality we report the trough rather than the peak.
[59] Six out of 27 turning point estimates are not statistically significant.

APPENDICES

All econometric results are obtained using STATA Version 7.0.

Appendix 4.1.: Random Effects ML Estimation for Water Quality Indicators

The following tables show the MLE estimates for each water quality indicator. They also show the significance levels ($prob > \chi^2$) for: a) the six income and lagged income variables taken together, b) the six upstream income and lagged upstream income variables taken together, and c) the 12 income, lagged income, upstream income and lagged upstream income variables taken together. Because of the expected strong multicollinearity between current and lagged income on one hand, and between the different powers of income on the other hand, individual coefficients are not expected to be significant. But as can be seen, the combinations of the variables as detailed above give highly significant results in many cases.

Table 4.5.

Random-effects ML Estimation for Dissolved Oxygen (DO)

Dependent Variable is Annual Mean Concentration
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	.9189 (.9639)	1.0030 (1.2552)	.3643 (1.0544)	.6382 (1.1689)
Income squared	-.0621 (.0794)	-.0683 (.1090)	-.0312 (.0802)	-.0512 (.0879)
Income cubed	.0014 (.0019)	.0015 (.0026)	.0008 (.0018)	.0013 (.0019)
Lagged income	-.9284 (.9952)	-1.0830 (1.2918)	-.3819 (1.0973)	-.7880 (1.1723)
Lagged income sq.	.0645 (.0852)	.08715 (.1164)	.0139 (.0870)	.0471 (.0927)
Lagged income cu.	-.0013 (.0021)	-.0020 (.0029)	.0001 (.0020)	-.0008 (.0022)
Upstream income				1.4693 (1.2367)
Upstream income sq.				-.1941 (.0937)
Upstream income cu.				.0051 (.0022)

To be continued

Table 4.5. continued: Random-effects ML Estimation for Dissolved Oxygen (DO)

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Lagged upstr. income				-1.3605 (1.3066)
Lagged upstr. income squ.				.1919 (.1036)
Lagged upstr. income cu.				-.0052 (.0025)
Upstream effect				-1.1253 (2.3005)
Mean temperature	-.0923 (.0226)	-.0875 (.0308)	-.1144 (.0201)	-.1144 (.0229)
Year	-.0730 (.0207)	-.0655 (.0281)	-.0861 (.0190)	.0943 (.0223)
χ^2 (income & lagged income combined)	20.76 **.0020**	16.32 **.0121**	24.66 **.0004**	17.06 **.0091**
χ^2 (upstr. income & lagged upstr. income combined)				11.93 **.0636**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				33.13 **.0009**
Sample size	1563	1117	446	392
Log likelihood	-3549.6530	-2670.8093	-715.1588	-629.1976

*Equations also include an intercept.

Table 4.6.

Random-effects ML Estimation* for Biological Oxygen Demand (BOD)

Dependent Variable is Annual Mean Concentration
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	-.8142 (6.2107)	2.1053 (6.2860)	-18.0030 (18.7174)	-22.9526 (22.1489)
Income squared	.2184 (.5639)	-.0168 (.6126)	1.5665 (1.5217)	2.0914 (1.7979)
Income cubed	-.0044 (.01449)	-.0014 (.0161)	-.0333 (.0368)	-.0455 (.0429)
Lagged income	1.5294 (6.4467)	-.3679 (6.4861)	17.5674 (19.4318)	21.2535 (22.4791)
Lagged income sq.	-.2283 (.6087)	-.1430 (.6557)	-1.3968 (1.6442)	-1.7323 (1.9119)
Lagged income cu.	.0032 (.01637)	.0053 (.0180)	.0261 (.0413)	.0320 (.0477)
Upstream income				7.7439 (21.6670)
Upstream income sq.				-1.1735 (1.9018)

To be continued

Table 4.6. continued: Random-effects ML Estimation* for (BOD)

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Upstream income cu.				.0307 (.0463)
Lagged upstr. income				-10.2549 (23.3280
Lagged upstr. income squ.				1.3541 (2.1125)
Lagged upstr. income cu.				-.0346 (.0526)
Upstream effect				21.5061 (32.2033)
Mean temperature	.2195 (.1376)	.2432 (.1349)	.1536 (.3308)	.1593 (.4151)
Year	.0100 (.1198)	.1237 (.1203)	-.3282 (.2922)	-.2415 (.3949)
χ^2 (income & lagged income combined)	6.50 **.3700**	4.16 **.6549**	6.17 **.4040**	6.80 **.3398**
χ^2 (upstr. income & lagged upstr. income combined)				1.12 **.98**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				7.21 **.8437**
Sample size	1248	892	356	302
Log likelihood	-4932.8016	-3359.8026	-1501.8175	-1292.2348

*Equations also include an intercept.

Table 4.7.

Random-effects ML Estimation* for Chemical Oxygen Demand (COD)

Dependent Variable is Annual Mean Concentration
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	-37.9159 (42.6360)	-15.5937 (41.0315)	37.1883 (196.7878)	46.2916 (223.6648)
Income squared	.6660 (4.0552)	-2.9618 (3.9622)	-2.2247 (17.7543)	-2.6499 (20.3266)
Income cubed	.0355 (.1091)	.1189 (.1052)	.0907 (.4677)	.1191 (.5349)
Lagged income	42.4963 (44.3595)	28.9531 (42.1964)	-28.5221 (204.0564)	13.7033 (232.3509)
Lagged income sq.	-.6776 (4.3607)	2.0604 (4.2060)	2.2676 (18.9638)	-1.1067 (21.7262)
Lagged income cu.	-.0456 (.1215)	-.1072 (.1157)	-.1037 (.5155)	-.0366 (.5891)
Upstream income				154.695 (264.1412)
Upstream income sq.				-15.0257 (15.0701)
Upstream income cu.				.3521 (.2965)

To be continued

Table 4.7. continued : Random-effects ML Estimation* for (COD)

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Lagged upstr. income				-160.2759 (274.2759)
Lagged upstr. income squ.				14.4248 (16.1767)
Lagged upstr. income cu.				-.3355 (.3281)
Upstream effect				118.4136 (129.5851)
Mean temperature	.8727 (.8529)	.9410 (.8276)	1.6286 (1.6731)	1.5353 (2.1087)
Year	.7128 (.7339)	1.2765 (.7515)	-2.3054 (1.8023)	-3.1942 (2.3452)
χ^2 (income & lagged income combined)	12.71 **.0479**	25.39 **.0003**	10.07 **.1218**	19.97 **.0028**
χ^2 (upstr. income & lagged upstr. income combined)				8.12 **.2295**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				20.98 **.0507**
Sample size	825	589	236	208
Log likelihood	-4573.4457	-3162.9687	-1344.4748	-1187.8497

*Equations also include an intercept.

Table 4.8.

Random-effects ML Estimation* for Nitrates

Dependent Variable is Annual Mean Concentration
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	-2.0064 (1.5024)	-2.1393 (1.8350)	-1.5333 (1.6341)	-.0332 (1.9677)
Income squared	.1732 (.1229)	.1848 (.1542)	.1569 (.1285)	.0417 (.1567)
Income cubed	-.0040 (.0030)	-.0042 (.0038)	-.0042 (.0030)	-.0016 (.0037)
Lagged income	2.1589 (1.5132)	2.5470 (1.8463)	.8632 (1.6311)	-1.0470 (1.9412)
Lagged income sq.	-.1808 (.1275)	-.2184 (.1608)	-.0657 (.1299)	.0365 (.1542)
Lagged income cu.	.0040 (.0031)	.0049 (.0040)	.0012 (.0031)	-.0004 (.0036)
Upstream income				-3.2808 (1.6420)
Upstream income sq.				.2680 (.1476)
Upstream income cu.				-.0060 (.0034)

To be continued

Table 4.8. continued: Random-effects ML Estimation* for Nitrates

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Lagged upstr. income				4.1605 (1.7092)
Lagged upstr. income squ.				-.3097 (.1584)
Lagged upstr. income cu.				.0065 (.0038)
Upstream effect				5.3547 (1.5110)
Mean temperature	-.0015 (.0246)	.0101 (.0317)	.0360 (.0236)	.0866 (.02521)
Year	-.0601 (.0330)	-.0680 ().0431	-.0306 (.0250)	-.01435 (.0328)
χ^2 (income & lagged income combined)	11.06 **.0865**	9.31 **.1568**	15.91 **.0142**	26.22 **.0002**
χ^2 (upstr. income & lagged upstr. income combined)				28.36 **.0001**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				33.06 **.0000**
Sample size	982	744	238	186
Log likelihood	-2357.6517	-1869.1919	-348.1449	-260.4146

*Equations also include an intercept.

Table 4.9.

Random-effects ML Estimation* for Faecal Coliforms

Dependent Variable is log (1+Annual Mean Concentration)
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	-1.3736 (.9139)	-1.7375 (1.0869)	-.0496 (1.8239)	2.3423 (1.9900)
Income squared	.1313 (.0761)	.1627 (.0966)	.0163 (.1381)	-.1473 (.1495)
Income cubed	-.0031 (.0018)	-.0039 (.0023)	-.0004 (.0031)	.0025 (.0033)
Lagged income	1.2447 (.9424)	1.9388 (1.1201)	-1.6699 (1.9281)	-1.7117 (2.0063)
Lagged income sq.	-.1226 (.0815)	-.1850 (.1030)	.1304 (.1522)	.1405 (.1578)
Lagged income cu.	.0028 (.0020)	.0043 (.0025)	-.0030 (.0035)	-.0032 (.0036)
Upstream income				-6.0863 (2.1755)
Upstream income sq.				.5126 (.1710)
Upstream income cu.				-.0115 (.0040)

To be continued

Table 4.9.continued: Random-effects ML Estimation* for Faecal Coliforms

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Lagged upstr. income				3.4832 (2.3414)
Lagged upstr. income squ.				-.3128 (.1926)
Lagged upstr. income cu.				.0072 (.0047)
Upstream effect				-5.7863 (3.2734)
Mean temperature	.0360 (.0210)	.0022 (.0259)	.0789 (.0374)	.1752 (.0422)
Year	.1090 (.0192)	.0971 (.0233)	.1054 (.0388)	.2379 (.0387)
χ^2 (income & lagged income combined)	14.50 **.0245**	23.53 **.0006**	5.36 **.4980**	11.75 **.0677**
χ^2 (upstr. income & lagged upstr. income combined)				28.22 **.0001**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				40.49 **.0001**
Sample size	1260	876	384	340
Log likelihood	-2638.4069	-1829.7058	-793.7762	-691.1880

*Equations also include an intercept

Table 4.10.

Random-effects ML Estimation* for Cadmium

Dependent Variable is Annual Mean Concentration
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	-.2153 (.1180)	-.3525 (.1618)	.0146 (.0552)	.0646 (.0658)
Income squared	.0256 (.0085)	.0411 (.0125)	-.0015 (.0035)	-.0054 (.0043)
Income cubed	-0.007 (.0002)	-.0011 (.0003)	.0000 (.0001)	.0001 (.0001)
Lagged income	.2014 (.1228)	.3412 (.1684)	-.0163 (.0581)	-.0310 (.0973)
Lagged income sq.	-.0242 (.0091)	-.0401 (.0136)	.0017 (.0037)	.0025 (.0058)
Lagged income cu.	.0006 (.0002)	.0010 (.0003)	-.0000 (.0001)	-.0001 (.0001)
Upstream income				-.0100 (.1403)
Upstream income sq.				.0015 (.0085)
Upstream income cu.				-.0000 (.0075)

To be continued

Table 4.10. continued: Random-effects ML Estimation* for Cadmium

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Lagged upstr. income				-.0266 (.1432)
Lagged upstr. income squ.				.0016 (.0079)
Lagged upstr. income cu.				-.0000 (.0001)
Upstream effect				-.0212 (.2893)
Mean temperature	-.0020 (.0019)	-.0032 (.0028)	.0005 (.0006)	.0008 (.0007)
Year	.0046 (.0020)	.0076 (.0031)	-.0009 (.0005)	-.0013 (.0006)
χ^2 (income & lagged income combined)	26.82 **.0002**	27.53 **.0001**	2.28 **.8918**	15.47 **.0169**
χ^2 (upstr. income & lagged upstr. income combined)				11.51 **.0738**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				16.18 **.1833**
Sample size	628	411	217	186
Log likelihood	380.2141	171.2093	523.2210	443.9787

*Equations also include an intercept.

Table 4.11.

Random-effects ML Estimation* for Lead

Dependent Variable is Annual Mean Concentration
(standard errors are in parentheses)
$prob > \chi^2$ are in bold

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Income (thousands)	.0830 (.2655)	.0824 (.0544)	-.0782 (1.418)	-.1263 (1.7631)
Income squared	-.0098 (.0195)	-.0078 (.0043)	.0068 (.0893)	.0105 (.1190)
Income cubed	.0002 (.0004)	.0002 (.0001)	-.0002 (.0019)	-.0003 (.0026)
Lagged income	-.1021 (.2713)	-.0772 (.0567)	.1220 (1.3258)	-.0065 (3.2620)
Lagged income sq.	.0122 (.0204)	.0073 (.0047)	-.0073 (.0857)	.0048 (.1886)
Lagged income cu.	-.0003 (.0005)	-.0002 (.0001)	.0001 (.0018)	-.0001 (.0036)
Upstream income				.1767 (4.2034)
Upstream income sq.				-.0045 (.2215)
Upstream income cu.				.0000 (.0039)

To be continued

Table 4.11. continued: Random-effects ML Estimation* for Lead

Variable	All Rivers BS-Model	Domestic Rivers BS-Model	Inter'l Rivers BS-Model	Inter'l Rivers UD-Model
Lagged upstr. income				.0301 (5.2483)
Lagged upstr. income squ.				-.0120 (.2819)
Lagged upstr. income cu.				.0003 (.0051)
Upstream effect				.4642 (11.7116)
Mean temperature	.0033 (.0032)	.0013 (.0009)	.0117 (.0114)	.0121 (.0149)
Year	-.0005 (.0043)	.0023 (.0011)	-.0064 (.0119)	-.0067 (.0178)
χ^2 (income & lagged income combined)	1.82 **.9355**	11.24 **.0812**	1.42 **.9649**	0.90 **.9893**
χ^2 (upstr. income & lagged upstr. income combined)				0.19 **.9999**
χ^2 (income, lagged income, upstr. income & lagged upstr. income comb.)				2.17 **.9991**
Sample size	615	407	208	177
Log likelihood	-114.4858	608.2727	-141.8750	-139.4187

*Equations also include an intercept.

Appendix 4.2.: List of International Rivers

The original sample from GEMS contains data from national and international rivers, lakes and groundwater. This study is confined to rivers, both domestic and international rivers. We have derived our sample of international rivers from (Conte Grand, Mariana 1999) classification of international rivers, but with a stricter definition for international rivers. Rivers are only classified as international rivers if they contribute more than 3% of the basin area in at least two countries. Therefore, some international rivers which were included in Conte Grand's study, such as the Ebro river, the Garonne river, the Seine river and the Rhone river are not included in this study.

Some international rivers do not have international monitoring stations in the GEMS dataset and had to be excluded form the study. This was the case for all international rivers between the United Kingdom and Ireland, and many rivers crossing countries such as Estonia, Latvia, Lithuania, Belarus, that were part of the former Soviet Union. There was also no income data for the former Soviet Union and pre-unification Germany in 1996 prices. Therefore, all Russian and German stations were excluded from the study. Some other some international rivers do have international monitoring stations and national income data, but no income data for the upstream country, as for example Mongolia and the Democratic People's Republic of Korea (North Korea).

The international rivers included in this study are classified according to the countries where the river stations are located. The number in between the brackets is the station number according to the GEMS dataset[60].

EUROPE

Belgium: Scheldt River (51009, 51010, 51015), Meuse River (51012, 51013), Escaut River (51001), Espierre River (51008), Lys River (51007), Sambre River (51011), Sure River

[60] The stations numbers are available at http://www.gemswater.org/datareporting-e.html

(51014). Finland: Tornionjoko River (65001) Greece: Nestos River (30001). Hungary: Tisza River (66001), Danube River (66002). Italy: Po River (68006, 68008). Luxembourg: Sure River (16001). Netherlands: Rhine River (46001, 46002, 46003, 46004), Maas River (46005). Poland: Odra River (21004, 21005, 21006). Portugal: Tajo River (73001). Switzerland: Rhine River (200001, 200002, 200003). Spain: Miño River (75002), Tajo River (75004), Guadiana River (75005)

NORTH and SOUTH AMERICA

Canada: St. Lawrence River (39003), Saskatchewan River (39004), Roseau River (39006), St. John River (39012). Mexico: Colorado River (37001), Río Bravo/Río Grande (37002, 37003), Río Usumacinta (37015), Río Grijalba (37016). USA: Columbia River (28002), Yukon River (28003), Colorado River (2806), Río Bravo/Río Grande (28010), St. Marys River (28018), St. Clair River (28019), Niagara River (28020), St. Lawrence River (28021). Argentina: Río Paraná (1001, 1002, 1004), Paraguay River (1003), Río Uruguay (1006). Uruguay: Río Uruguay (48015, 48038), Río de La Plata (48039)

MIDDLE EAST and SOUTH-EAST ASIA

Egypt: Nile River (10002, 10003, 10004, 10005, 10006, 10007 10008). Iran: Karun River (14010). Pakistan: Ravi river (56003, 56004), Lower Chenab River (56005), Indus River (56006). Sudan: Nile River (78002). Bangladesh: Lower Ganges River (136002), Brahmaputra River (136003), Meghna River (136004), Surma River (136005).

AFRICA

Ghana: Volta River (81003). Mali: Niger River (99001, 99002, 99004, 99006, 99007, 99008, 99009). Senegal: Senegal River (100002, 100005, 100006, 100007, 100008). Tanzania: Kagera River (104003). Uganda: Nile River (110004, 110005, 110007, 110009, 110010,110011, 110017). Zaire: River Zaire (98001).

Chapter 5
Concluding Remarks

We have attempted in this thesis to explore the effects of the unidirectional externality, which upstream countries exert on downstream countries through a shared international river, on two different issues, which are equally important: self-enforcing water sharing agreements and water pollution. The inability of downstream countries to adversely affect upstream countries affects the credibility of any water-related punishments on the side of the downstream countries. Our models in chapter 2 and 3 show that self-enforcing sustainable water sharing agreements are possible despite the obvious obstacles of the asymmetry of enforcement power between the upstream and downstream country once an agreement is reached. While our model in chapter 2 concentrates solely on the quantity aspect of the unidirectional externality, our model in chapter 3 can be applied to both the quantity and the quality aspect of the unidirectional externality. Our estimations in chapter 4 concentrate exclusively on the quality aspect of the unidirectional externality of international rivers. We not only identify an income related upstream effect from upstream countries on the water quality of downstream countries, but also show that local water pollutants peak at a lower income level than regional water pollutants, which is in accordance with other studies done on local and global air pollutants.

Our purpose in chapter 2 was to model strategic behaviour in reaching an agreement on sharing the waters of the River Nile between Egypt and Ethiopia. Our dynamic game of one-sided incomplete information shows, that incomplete information about the military strength of the naturally disadvantaged downstream country (Egypt) on the side of the upstream country (Ethiopia) can, under certain conditions, work to the benefit of both countries, because it can make sustainable full cooperation between them possible. The perceived threat of using military power to enforce an agreement can lead to sustainable full cooperation between the perceived-as-strong, but vulnerable downstream country and the upstream country that has incomplete information about the actual military strength of the downstream country. The main results of our analysis are: first, beliefs do matter and Ethiopia's beliefs about Egypt' military strength affect the behaviour of both Ethiopia and Egypt. Secondly, conflict between Egypt and Ethiopia can be avoided if Ethiopia perceives Egypt as strong enough. Thirdly, more accurate information about Egypt cannot prevent conflict between Ethiopia and Egypt, but can prevent war between them; and finally, being perceived as strong is even more important for Egypt when its vulnerability is increased as is expected to happen after an agreement is reached.

Most studies on the River Nile conflict concentrate on the significant gains of cooperation, which should motivate downstream Egypt to accept upstream water development projects. These studies do not provide an answer to the greatest obstacle to cooperation, namely the loss of Egypt's control over its water resources, if Ethiopia is to build its own over-year storage on the Blue Nile. Our study contributes to the economic literature by providing an analysis, which takes strategic considerations, beliefs and perceptions explicitly into account. Our model demonstrates with a Bayesian game that perceived credible threats are crucial in reaching self-enforcing water agreements.

The purpose of chapter 3 was to model the effect of the different irreversibility constraints on the players' actions on the feasibility of self-enforcing (IEA's). Our dynamic game of one-sided irreversibility, shows that the desire of the upstream country to reap continuing future benefits from increasing cooperationn over an international river with a downstream country helps to achieve partial cooperation. The credible threat of not cooperating in the future gives the downstream country a limited enforcement power over the upstream country. Our results confirm the general intuition that despite obvious gains from cooperation, asymmetric irreversibility of the players' actions makes reaching even a partial agreement between countries sharing a water body with unidirectional externalities, such as a river, more difficult than reaching an agreement between countries sharing a water body with reciprocal externalities, such as a lake. The main results of the analysis are: first, full co-operation can only be sustained if both players' actions are reversible. Secondly, higher levels of partial co-operation can be sustained under symmetric irreversibility of the players' actions than under asymmetric irreversibility of the players' actions. Thirdly, while *gradual* co-operation improves the situation for the symmetric irreversibility case compared to *immediate* co-operation, the type of cooperation has no effect on the other cases; and finally, a higher discount factor is conducive to cooperation for all irreversibility constraints, for all levels and for all types of co-operation.

Interpreting the unidirectional externality in terms of irreversible actions has not been covered in the economic literature as far as I know; and our approach to adapt an infinite dynamic game into a finite dynamic game allows tractable analysis of a very complicated, but highly relevant extension of the Lockwood and Thomas (2002) model.

Our purpose in chapter 4 was to investigate whether the nature of a river affects the relationship between income and environmental quality. Our analysis provides some evidence via *reduced-form relationships between water quality and per capita income* that inter-country pollution affects the intra-country income-water quality relationship i.e., the Environmental Kuznets' Curve, and that upstream income might be one of the determinants of the water quality in downstream countries. The main results on the analysis are: first, national income is an important determinant for river pollution, especially for domestic rivers. Secondly, the Upstream-downstream model is better suited for international rivers than the Basic model, because it takes upstream effects explicitly into account. Thirdly, upstream effects and especially income related upstream effects, seem to play an important role in river pollution. Fourthly, all water pollutants peak at lower level of income in domestic than in international rivers; and finally most of the water quality indicators in international rivers peak at a lower level of income when upstream effects are considered. Although more data, especially on international rivers, would benefit this analysis, we contribute to the literature in three ways: first, we distinguish between local and regional water pollutants through the nature of the river, which has not been done in this context before. Secondly, we are able to identify an income related externality effect and separate it from the overall upstream effect, while all previous studies that have identified upstream effects have looked only at non-income related explanations. Thirdly, we confirm the observed pattern that local air pollutants tend to peak at lower income levels than regional or global pollutants, for water pollutants. This has not been done for water quality indicators as far as I know.

Our thesis not only highlights the specific difficulties caused by the unidirectional externalities in international rivers, but also provides two different models to tackle the main obstacle of cooperation, namely the limited ability of the downstream country to enforce water sharing agreements. Identifying an income related upstream effect, which affects water quality in downstream countries, might improve future estimations of the critical levels of income at which water quality starts to improve. Overall, our results indicate that both the quantity aspect and the quality aspect of the unidirectional externality in international rivers are important and need to be taken into consideration when recommending any policy interventions, whether on national level or on a more regional level through international environmental agreements

References

Admati, Anat and Motty Perry. 1991. "Joint Projects without Commitment." *Review of Economic Studies*, 58, pp. 259-76.

Allan, J. Anthony. 1997. "'Virtual water': a long term solution for water short Middle Eastern economies?" *Paper presented at the 1997 British Association Festival of Science*: University of Leeds.

Allan, J. Anthony. 1999. "The Nile Basin: Evolving Approaches to Nile Waters Management." Occasional Paper 20, SOAS Water Issue Group.

Barrett, Scott. 1990. "The Problem of Global Environmental Protection." *Oxford Review of Economic Policy*, 6:1, pp. 60-79.

Barrett, Scott. 1994. "Conflict and Cooperation in Managing International Water Resources." CSERGE Working Paper WM 94-04.

Bennett, Lynne L., Shannon E. Ragland, and Peter Yolles. 1998. "Facilitating International Agreements through an Interconnected Game Approach: The Case of River Basins," in *Conflict and Cooperation on Trans-Boundary Water Resources*. Richard E. Just and Sinaia Netanyahu eds. Boston: Kluwer Academic Publishers.

Biswas, Asit B. 1999. "Management of International Waters: Opportunities and Constraints." *International Journal of Water Resources Development*, 15:4, pp. 429-41.

Boyce, James K. 1994. "Inequality as a Cause of Environmental Degradation." *Ecological Economics*, 11, pp. 169-78.

Cesar, Herman and Aart De Zeeuw. 1996. "Issue Linkage in Global Environmental Problems," in *Economic Policy for the Environment and Natural Resources*. Anastasios Xepapadeas ed. Cheltenham UK: Edward Elgar.

Cole, M. A., A. J. Rayner, and J. M. Bates. 1997. "The environmental Kuznets curve: an empirical analysis." *Environment and Development Economics*, 2:4, pp. 401-16.

Conte Grand, Mariana. 1999. "Do Regional Agreements Have Any Effect on Water Quality?" Working Paper No. 141: Universidad del CEMA.

Dasgupta, Partha. 1993. *An Inquiry into Well-Being and Destitution*. Oxford: Clarendon Press.

De Bruyn, S. M. 1997. "Explaining the environmental Kuznets curve: structural change and international agreements in reducing sulphur emissions." *Environment and Development Economics*, 2:4, pp. 485-504.

De Bruyn, S. M., J. C. Van Den Bergh, and J. B. Opschoor. 1998. "Economic growth and emissions: reconsidering the empirical basis of environmental Kuznets curves." *Ecological Economics*, 25:2, pp. 161-76.

Dinar, Ariel and Aaron Wolf. 1994. "International Markets for Water and the Potential for Regional Cooperation: Economic and Political Perspectives in the Western Middle East." *Economic Development and Cultural Change*, 43:1, pp. 43-66.

Finus, M. 2000. "Game Theory and Interntatioal Enviromental Co-operation: A Survey with an Application to the Kyoto-Protocol." NOTA DI LAVARO 86.2000, Fondazione Eni Enrico Mattei: Milano.

Finus, M. 2001. *Game Theory and International Environment Cooperation*. Cheltenham, UK: Edward Elgar Publishing Limited.

Frey, Frederick W. 1993. "The Political Context of Conflict and Cooperation over International River Basins." *Water International*, 18, pp. 54-68.

Frisvold, George B. and Margriet F. Caswell. 1995. "A Bargaining Model for Water Transfers Including Water Quality," in *Water Quantity/Quality Management and Conflict Resolution.* Ariel Dinar and Edna Tusak Loehman eds. Westport: Praeger, pp. 399-407.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory.* Cambridge: The MIT Press.

GEMS. 1995. "Water Quality of World River Basins." 1-39. United Nations Environment Programme: Nairobi.

Gibbons, Robert. 1992. *A Primer in Game Theory.* London: Prentice Hall.

Gleick, Peter H. 1992. "Effects of Climate Change on Shared Fresh Water Resources," in *Confronting Climate Change. Risks, Implications and Responses.* Irving M. Mintzer ed: Cambridge University Press, pp. 127-40.

Gleick, Peter H. 1998. *The World's Water 1998-1999. The Biennial Report on Freshwater Resources.* Washington D.C.: Island Press.

Grossman, G. M. and A. B. Krueger. "Annual Water Quality Indicators." ftp://irs.princeton.edu *in the environ directory.*

Grossman, G. M. and A. B. Krueger. 1995. "Economic Growth and the Environment." *Quarterly Journal of Economics*, 110:2, pp. 353-77.

Grossman, G. M. and A. B. Krueger. 1996. "The inverted-U: What does it mean?" *Environment and Development Economics*: February, pp. 119-22.

Guariso, Giorgio and Dale Whittington. 1987. "Implications of Ethiopian Water Development for Egypt and Sudan." *International Journal of Water Resources Development*, 3:2, pp. 105-14.

Güner, Serdar. 1998. "Signalling in the Turkish-Syrian Water Conflict." *Conflict Management and Peace Science*, 16:2, pp. 185-206.

Heston, Allan, Robert Summers, and Bettina Aten. 2002. "Penn World Table (Mark 6)." http://pwt.econ.upenn.edu/.

Holtz-Eakin, D. and T. M. Selden. 1995. "Stoking the fires? CO~2 emissions and economic growth." *Journal of Public Economics*, 57:1, pp. 85.

Hsiao, Cheng. 2003. *Analysis of Panel Data*. Cambridge: Cambridge University Press.

International, Court of Justice. 1988. "Egypt-Israel Arbitration Tribunal: Award in Boundary Dispute concerning the Taba Area," in *International Legal Materials vol. XXVII*, pp. 1469.

Jovanovic, D. 1985. "Ethiopian Interests in the Division of the Nile River Waters." *Water International*, 10, pp. 82-85.

Kemp, A. and U. Ben-Eliezer. 2000. "Dramatizing Sovereignty: The Construction of Territorial Dispute in the Israeli-Egyptian Border at Taba." *Political Geography*, 19, pp. 315-44.

Kilgour, Marc D. and Ariel Dinar. 1995. "Are Stable Agreements for Sharing International River Waters Now Possible?" Policy Research Working Paper No. 1474, World Bank: Washington.

Lekakis, Joseph N. 1998. "Bilateral Monopoly: A Market for Intercountry River Water Allocation." *Environmental Management*, 22:1, pp. 1-8.

LeMarquand, David. 1978. "Politics of International River Basin Cooperation and Management," in *Water in a Developing World-The Management of a Critical Resource*. Albert E. Utton and Ludwik Teclaff eds. Colorado: Westview Press, pp. 147-65.

Lockwood, B. and J. P. Thomas. 2002. "Gradualism and Irreversibility." *Review of Economic Studies*, 69:2, pp. 339-56.

Mäler, Karl-Göran. 1990. "International Environmental Problems." *Oxford Review of Economic Studies*, 6:1, pp. 80-108.

Marx, Leslie M. and Steven A. Matthews. 1997. "Dynamic Contribution to A Public Project." Discussion Paper No. 1188, Center for Mathematical Studies in Economics and Management Science: Northwestern University.

Mason, R. and T. Swanson. 2003. "A Kuznets curve analysis of ozone-depleting substances and the impact of the Montreal Protocol." *Oxford Economic Papers*, 55:1, pp. 1-24.

Moomaw, W. R. and G. C. Unruh. 1997. "Are environmental Kuznets curves misleading us? The case of CO~2 emissions." *Environment and Development Economics*, 2:4, pp. 451-64.

Panayotou, T. 1997. "Demystifying the environmental Kuznets curve: turning a black box into a policy tool." *Environment and Development Economics*, 2:4, pp. 465-84.

Panayotou, T. 2000. "Economic Growth and the Environment." CID Working Paper No. 56. Center for International Development at Harvard University.

Pearce, David W. and R. Kerry Turner. 1990. *Economics of National Resources and the Environment*. New York: Harvester Wheatsheaf.

Roberts, J. T. and P. E. Grimes. 1997. "Carbon Intensity and Economic Development 1962-91: A Brief Exploration of the Environmental Kuznets Curve." *World Development*, 25:2, pp. 191-98.

Roberts, J. T. and P. E. Grimes. 1998. "Reply to McNaughton and Lee." *World Development*, 26:12, pp. 2221.

Rogers, Peter. 1969. "A Game Theory Approach to the Problems of International River Basins." *Water Resources Research*, 5:4, pp. 749-60.

Rogers, Peter. 1993. "The Value of Cooperation in Resolving International River Basin Disputes." *Natural Resource Forum*, 17:2, pp. 117-31.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, Massachusetts: Harvard University Press.

Selden, T. M. and D. Song. 1994. "Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions?" *Journal of Environmental Economics and Management*, 27:2, pp. 147.

Shafik, N. 1994. "Economic Development and Environmental Quality: an Econometric Analysis." *Oxford Economic Papers*, 46:4//Sup, pp. 757.

Shapland, Greg. 1995. "Policy Options for Downstream States in the Middle East," in *Water in the Middle East: Legal, Political and Commercial Implications*. J. Anthony Allan and Chilibi Mallat eds. London: Taurus Academic Studies, pp. 301-23.

Shapland, Greg. 1997. *Rivers in Discord. International Water Disputes in the Middle East*. London: Hurst & Company.

Sigman, H. 2002. "International Spillovers and Water Quality in Rivers: Do Countries Free Ride?" *American Economic Review*, 92:4, pp. 1152-59.

Smith, S. E. and H. M. Al-Rawahy. 1990. "The Blue Nile: Potential for Conflict and Alternatives for Meeting Future Demands." *Water International*, 15, pp. 217-22.

Spagnolo, Giancarlo. 1996. "Issue Linkage, Delegation and International Policy Cooperation." Working Paper Economics, Energy, Environment No. 49.96, Fondazione Eni Enrico Mattei: Milano.

Stern, D. I., M. S. Common, and E. B. Barbier. 1996. "Economic Growth and Environmental Degradation: The Environmental Kuznets Curve and Sustainable Development." *World Development*, 24:7, pp. 1151-60.

Swain, Ashok. 1997. "Ethiopia, the Sudan and Egypt: The Nile River Dispute." *The Journal of Modern African Studies*, 35:4, pp. 675-94.

Torras, M. and J. K. Boyce. 1998. "Income, inequality, and pollution: a reassessment of the environmental Kuznets Curve." *Ecological Economics*, 25:2, pp. 147-60.

Unruh, G. C. and W. R. Moomaw. 1998. "An Alternative Analysis of Apparent EKC-type Transitions." *Ecological Economics*, 25, pp. 221-29.

Verbeek, Marno. 2000. *Modern Econometrics*. Chichester, England: John Wiley&Sons, Ltd.

Vincent, Jeffrey R. 1997. "Testing for Environmental Kuznets Curves within a Developing Country." *Environment and Development Economics*, 2, pp. 417-31.

Waterbury, John. 1987. "Legal and Institutional Arrangements for Managing Water Resources on the Nile Basin." *International Journal of Water Resources Development*, 3:2, pp. 92-104.

Waterbury, John. 1994. "Transboundary Water and the Challenge of International Cooperation in the Middle East," in *Water in the Arab World. Perspectives and Prognoses.* Peter Rogers and Peter Lydon eds: Harvard University Press, pp. 39-64.

Whittington, Dale and Kingsley E. Haynes. 1985. "Nile Waters for Whom? Emerging Conflicts in Water Allocation for Agricultural Expansion in Egypt and Sudan," in *Agricultural Development in the Middle East.* Peter Beaumont and Keith McLachlan eds. Chichester: John Wiley & Sons Ltd., pp. 125-49.

Whittington, Dale and Elizabeth McClelland. 1992. "Opportunities for Regional and International Cooperation in the Nile Basin." *Water International*, 17:3, pp. 144-54.

Whittington, Dale, John Waterbury, and Elizabeth McClelland. 1995. "Towards a New Nile Water Agreement," in *Water Quantity/Quality Management and Conflict Resolution.* Ariel Dinar and Edna Tusak Loehman eds. Westport: Praeger, pp. 167-78.

Wolf, Aaron T., Jeffrey A. Natharius, Jeffrey J. Danielson, Brian S. Ward, and Jan K. Pender. 1999. "International River Basins of the World." *International Journal of Water Resources Development*, 15:4, pp. 387-427.

Wouters, Patricia. 1997. "Present Status of International Water Law," in *International Water Law. Selected Writings of Professor Charles B. Bourne.* Patricia Wouters ed. London: Kluwer Law International.