Cooperative Decision Making in River Water Allocation Problems

Moes, N.

2013

COOPERATIVE DECISION MAKING IN RIVER WATER ALLOCATION PROBLEMS

ISBN 978 90 3610 321 3

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 544 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

This research was supported by the Netherlands Organization for Scientific Research (NWO) under the project number 400.07.159.

VRIJE UNIVERSITEIT

Cooperative Decision Making in River Water Allocation Problems

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Economische Wetenschappen en Bedrijfskunde
op vrijdag 11 januari 2013 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Nigel Moes
geboren te Amstelveen

promotor: prof.dr.ir. G. van der Laan
copromotor: dr. J.R. van den Brink

Contents

Introduction

Water is essential to life. The amount of water on Earth is finite, and less than 3% of it is fresh water. A rapidly growing world population combined with a global rise in living standards presents a major threat to sustainability of the world's water resources in the future. At present, however, it is not a lack of fresh water that is responsible for global water crises. Rather, it is the uneven distribution of fresh water over space and time.

Precipitation is the primary natural source of fresh water. It recharges groundwater aquifers and provides surface and subsurface runoff. The runoff flows downstream via streams and rivers into natural water bodies such as lakes, wetlands, seas and oceans, or into man-made reservoirs. Rivers, and the bodies of water they flow into, constitute the most important regional source of fresh water in the world. Withdrawals from surface water account globally for almost 75% of all fresh water withdrawals and slightly more than 80% of the world's population (4.9 billion people) is served by naturally occurring renewable water sources (United Nations, 2009).

The water that is withdrawn from rivers, and other surface water resources, is required for agriculture, industry, domestic use and ecosystem survival. In addition, rivers are used for fishing, transportation, electricity generation and waste discharge. The use of a river for these activities can have adverse impacts on downstream users and ecosystems. The simplest example of this is when an upstream user of a river withdraws such a large amount of water that downstream users are no longer able to withdraw their preferred amounts. In regions of water scarcity, it is clear that this can lead to tensions between upstream and downstream water users. Another example is that of pollution. The discharge of waste products in a river can lead to pollution, which in turn can cause environmental damage. When an upstream user of a river pollutes it, this can cause environmental damage on the territory of downstream users.

Asymmetric dependence on a water resource, as in the above examples, is often at the heart of disputes about the use of the resource. This is especially true if the *property rights* over it (the exclusive authority to determine how the resource is used) are not clearly defined. In a national setting, disputes about property rights over water resources are usually settled through a country's legal system. In an international setting, though, there typically is no third party that is able to enforce agreements.

History has shown that the absence of a supranational authority, able to establish and protect property rights over water resources, can lead to heated discussion and even military conflict over water resources. In 1948, shortly after the partitioning of the Indian subcontinent, India and Pakistan nearly went to war when India diverted water of the Indus river away from Pakistan. A conflict that actually escalated into war was that between Israel and its neighbors over access to the Jordan river and its accompanying ground water basins. From 1951 onwards this conflict led to a number of clashes that have been classified as water wars by authors on the subject (Barrett, 2003).

Global institutions, such as the United Nations, have tried to mitigate the impact of asymmetric dependence on water resources by promoting multilateral agreements between states sharing a resource. While this, so far, has not resulted in binding international laws (see Chapter 1), it did help in establishing over 400 agreements between countries sharing transboundary water resources (United Nations, 2009). These agreements normally stipulate rules and regulations, known as *water rights* and *water responsibilities*, that govern water use and protect water resources. Water rights and responsibilities have been documented in an extensive legal literature. In this dissertation, we will use this literature as one of the two main tools in modeling the outcomes of negotiations between water users (countries) sharing a river.

The other main tool that we use is game theory. Myerson (1991) describes *game theory* as "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers". Game theorists examine strategic interaction between (economic) *agents*, also called *players*, in a model, called a *game*, in which the success of a player's choices and actions depends on the choices and actions of the other players.

Game theory can roughly be divided into two subfields: *cooperative game theory* and *non-cooperative game theory*. The difference between the two lies in an institutional aspect. In a cooperative game players are able to sign enforceable contracts outside of those specifically modeled in the game (players can sign binding agreements). In a non-cooperative game they cannot.

In modeling international river water distribution agreements we will mostly work with concepts from cooperative game theory. Cooperative game theory revolves around two major issues: (1) the formation of coalitions of (a limited number of) cooperating players and (2) the division of the surplus of cooperation among these players. The question of which coalitions will form in a game, when players are able to sign binding agreements, is a positive economic question, i.e., in answering the question one wants to give a value-free description of, and prediction about, which coalitions of players will form. The question of how the benefits of cooperation have to be distributed among the players in a game is obviously more normative in character, i.e., in answering the question one is forced to make value judgments about the fairness of particular divisions.

What all the prominent solution concepts in cooperative game theory, e.g., the Shapley value (Shapley, 1953), core (Gillies, 1959) and nucleolus (Schmeidler, 1969), have in common is that they base (the fairness of) a particular division of cooperative surplus on the idea that a player in a coalition that has better possibilities outside the coalition (relative to the other players in the coalition) should not receive a smaller share in the surplus of the coalition. The two approaches that have been followed to implement this notion are the *strategic* and *axiomatic* approaches. Strategic solution concepts for cooperative games are based on non-cooperative ideas of what actions players will take when they face particular cooperative situations. In contrast, axiomatic solution concepts are obtained by stating a number of desirable properties (axioms) and showing that there exists a solution concept that satisfies these properties. Although the former approach seems more predictive and the latter approach more prescriptive, Shapley (1953) describes a *value*, which is usually defined axiomatically, as providing for each player an a priori assessment of the utility of becoming involved in a cooperative game. In this interpretation, the value concept can be seen as a predictive one that avoids the many complexities of modeling cooperation in a non-cooperative way. The value concept will play an important role in this dissertation.

As stated above, cooperative game theory models settings in which there is a small number of agents signing binding bi- or multilateral contracts to enforce cooperation. Since international river management normally takes place through the institution of a joint river basin committee and/or the signing of a water allocation agreement between a small number of countries (Ansink, 2009), cooperative game theory has been the principal method of choice of researchers modeling the management of international rivers. Over the last two decades a small, but growing, theoretical literature has emerged on the topic, see Parrachino, Dinar and Patrone (2006) for an overview. The primary aim of this dissertation is to extend this literature.

More specifically, this dissertation reports on three specific goals in extending the literature on cooperative decision making in (international) river water allocation problems. The first goal is to extend a single-stream international river water allocation model, originally introduced by Ambec and Sprumont (2002), to situations in which rivers are allowed to have several tributaries and distributaries (multiple streams merging into one main river and/or the main river splitting into a delta). The second goal is to allow the countries in an international river water allocation model to be composed of different water users (e.g., states, cities, or individual users). The third goal is to analyze the differences between the rival and non-rival use of water in international river water allocation models (the *rival* use of river water prevents use of the same water by other users, e.g in irrigation or as drinking water, the *non-rival* use does not; examples of non-rival use of water from an international river include pollution of it and the use of it in electricity generation).

Overview of the dissertation

We now give a short overview of this dissertation to explain which of the three goals mentioned above will receive attention in which chapter. The first two chapters of this dissertation have an introductory character.

In Chapter 1 we discuss international watercourse law, its underlying principles, and its historical development. Chapter 1 serves both as a further introduction into international river water allocation problems, as well as a specific discussion of the principles of international watercourse law that we use in the remainder of the dissertation to evaluate (the 'fairness' of) particular solutions to these problems.

In Chapter 2 we provide a technical introduction into cooperative game theory by discussing various transferable utility games. Most of the game-theoretic concepts that we use throughout the dissertation can be found in these preliminaries. In addition, we discuss the major findings of the literature on cooperative decision making in international river water allocation problems. In this literature authors have combined the international water law principles of Chapter 1 with concepts from cooperative game theory. This combination has resulted in a class of models that, on the one hand, is quite technical, but on the other hand, provides clear insight into how certain principles from international watercourse law can be made operational in actual river water allocation problems.

Chapter 3 is based on van den Brink, Estévez-Fernández, van der Laan and Moes (2011) and contains the first main contributions of this dissertation. In this chapter we consider the problem of sharing water among agents (countries) located from upstream to downstream along a single-stream river. Each agent has quasi-linear preferences over river water and money, where the benefit of consuming an amount of water is given by a continuous and concave benefit function. A solution to the river sharing problem efficiently distributes the river water over the agents and wastes no money. We introduce a number of (independence) axioms to characterize two new solutions and two solutions that were proposed in the literature discussed in Chapter 2. We apply all four solutions to the particular case that every agent along the river has constant marginal benefit of one up to a satiation point and marginal benefit of zero thereafter. In that case it follows that two of the solutions (one from the literature and a new one) can be implemented without monetary transfers between the agents.

Chapter 4 is based on van den Brink, van der Laan and Moes (2012a) and focuses on the first goal of this dissertation. Hence, in this chapter we examine the same model as in Chapter 3, except that the river now possibly has multiple springs. This means that there is a river that has several tributaries along which water users can be located. We consider two different assumptions on the benefit functions of the agents in the model. When the benefit functions of the agents are strictly increasing, each agent always wants to consume more water. The problem of finding a fair distribution of the welfare resulting from an optimal allocation of water among the agents can then be modeled by a transferable utility game in characteristic function form. When the benefit functions of the agents have satiation points, it could be that some agents experience externalities of coalition formation. The same problem then has to be modeled using a transferable utility game in partition function form. For both games we propose the class of weighted hierarchical solutions as a class of solutions satisfying the "territorial integration of all basin states" principle from international watercourse law (discussed in Chapter 1).

Chapter 5 is based on van den Brink, van der Laan and Moes (2012b) and provides a strategic implementation of the class of weighted hierarchical solutions introduced in Chapter 4. That is, in Chapter 5 we propose a non-cooperative mechanism of which the unique subgame perfect Nash equilibrium payoffs correspond to the weighted hierarchical solution payoffs of the cooperative game.

Chapter 6 is based on van den Brink, van der Laan and Moes (2011) and focuses on the second goal of this dissertation. Thus, in this chapter we further extend the model of Chapter 3 and Chapter 4 (with strictly increasing benefit functions) to allow countries to be composed of different water users. Moreover, in Chapter 6 rivers can have both multiple springs and multiple sinks. This means that there is a main river that can have several tributaries, but also several distributaries that form a delta. To take account of both the different water users within one country and the river that can have multiple springs and multiple sinks, we make use of transferable utility games in characteristic function form with both graph restricted communication and a priori unions. We introduce and characterize two new values for this type of games by applying the Shapley value to two associated transferable utility games in characteristic function form.

Chapter 7 is based on van der Laan and Moes (2012) and focuses on the third goal of this dissertation. So, in this chapter we study the non-rival use of river water by agents located along the river. We introduce a model in which the agents derive benefit while causing pollution, but also incur environmental costs of experiencing pollution from all upstream agents. It turns out that total pollution in the model decreases when the agents decide to cooperate. The resulting gain in social welfare can be distributed among the agents based on the property rights over the river. Using various principles from international watercourse law (of Chapter 1) we suggest 'fair' ways of distributing the property rights and therefore the cooperative gain in this model.

We end this dissertation with a short conclusion.

Chapter 1

On the law of international watercourses

1.1 Definitions and aim

There exists a rich legal literature on the law of international watercourses. It consists of bi- and multilateral treaties, judicial and arbitral decisions, declarations and resolutions by scholarly non-governmental organizations such as the Institute of International Law (IIL) and the International Law Association (ILA), and a United Nations (UN) convention. Prior to discussing this literature, it is necessary to establish some definitions.

In general, a watercourse can be defined as any flowing body of water. The definition of an international watercourse adopted by the UN in its "Convention on the Law of the Non-Navigational Uses of International Watercourses" is more formal and reads as follows[1]:

(a) "Watercourse" means a system of surface waters and groundwaters constituting by virtue of their physical relationship a unitary whole and normally flowing into a common terminus;

(b) "International watercourse" means a watercourse, parts of which are situated in different States.

Apart from the term "international watercourse", the terms "international river" and "international drainage basin" are also frequently used in the literature on the subject. An international river is a specific type of international watercourse that, for instance, does not include lakes or aquifers. A drainage basin is a geographical area determined by the limits of a system of interconnected waters, including surface and underground waters, all of which flow into a common terminus. An international drainage basin extends to or over the territory of two or more states. In this chapter we do not focus on the distinction between watercourses, rivers and drainage basins and will therefore use the terms interchangeably.

[1] United Nations (1997).

International rivers fall into two categories: boundary (or contiguous) rivers and successive rivers. A boundary river flows between the territories of two or more states and hence forms the border between the states. A successive river flows from the territory of one state into the territory of another state. It is also possible that an international river is (partly) a boundary river and (partly) a successive river. Usually, international law does not draw legal distinctions between boundary and successive rivers (Lipper, 1967) although there exist documents in which a distinction between the two types is made (e.g., in the 1911 Madrid resolution of the IIL).

The use of water from an international watercourse can be segmented into consumptive use and non-consumptive use. Consumptive use diminishes the quantity of water in the watercourse and includes the use as drinking water, for irrigation, and in certain manufacturing processes. Non-consumptive use can diminish the quality of the water in the watercourse, but does not necessarily do so. The use of a watercourse for navigation or hydropower, for example, does not have to diminish the quality of water in the watercourse. Using a watercourse to discharge waste products or using water from a watercourse as cooling water in an industrial plant (causing thermal pollution) are instances of non-consumptive use that do diminish the quality of water in a watercourse.

Interesting is that it is not the distinction between consumptive and non-consumptive use of an international watercourse that divides the legal literature on the subject. It is the distinction between navigational and non-navigational uses. The reason for this is historical. The literature on navigational uses dates back over a millennium to a declaration by Emperor Charlemagne granting freedom of navigation to a monastery (Barrett, 1994). The literature on non-navigational uses is much younger as interests in formally establishing rights over non-navigational uses of international watercourses only arose after the industrial revolution (McCaffrey, 2001). The legal literature on the non-navigational uses of international watercourses is the primary subject of this chapter.

As stated before, the literature on the law of international watercourses is highly fragmented as it ranges from bilateral treaties to a UN convention. In this chapter we concentrate on that part of the literature in which international organizations have tried to codify the rules on the use of international watercourses in general treaties.[2] The reason why one would want to have generally binding rules on the use of international watercourses is twofold. One, in case there is disagreement between riparian states, a general treaty would prescribe rules or procedures to manage the disagreement. And two, a general treaty could provide a normative background against which negotiations over bi- or multilateral treaties between riparian states could take place.

One of the earliest attempts to determine general rules on the use of international watercourses was undertaken by the IIL in its 1911 Madrid resolution. It is, however, the work of ILA and the UN that have had the most impact historically. Salman (2007a, p.1) even calls the ILA's 1966 Helsinki rules "the first comprehensive set of rules dealing with international watercourses" and claims that they "have been widely accepted

[2]For an overview of the numerous (bilateral) treaties, declarations, adjudications, case law examples and state practice examples from international watercourse law see Garretson, Hayton and Olmstead (1967), Lammers (1984), McCaffrey (2001) and Boisson de Chazournes and Salman (2005).

as representing international water law". Important to note is that the declarations and resolutions by the IIL and ILA were not binding rules of international law because they were neither signed nor ratified by riparian states. Rather, they provided authoritative guidance in disputes over international watercourses based on the expertise of the members of the scholarly organizations. The UN is, in fact, the first international organization that is trying to install globally binding (signed and ratified) rules on the use of international watercourses with its "Convention on the Law of the Non-Navigational Uses of International Watercourses". At the time of writing, though, this convention has not (yet) entered into force because the large majority of the states that have adopted the convention in 1997 have not (yet) ratified it.

The aim of this chapter is to provide an overview of the most essential rules of international watercourse law that have been codified in general treaties, their underlying principles, and their historical development. This overview is not supposed to be complete, but to give an idea of the complications involved in (establishing) rules on the use of international watercourses. The chapter starts with a discussion of the Harmon doctrine. This doctrine has played an important role in forming the early opinions of (upstream) riparian states about international watercourse law. It exemplified the need for international rules that enforce cooperation among riparian states. This eventually led to the work of the ILA and the UN that we discuss subsequently.

1.2 The 1895 Harmon doctrine

The Rio Grande is a river that originates in the United States (US) and forms part of the border between the US and Mexico before emptying in the Gulf of Mexico. It flows for some 1000 kilometers through Colorado and New Mexico and then becomes the border between the US state of Texas and the Mexican states of Chihuahua, Coahuilan, Nuevo León and Tempaulipas for about 2000 kilometers.

In 1895 the Mexican government filed a complaint with the US government about the excessive use of water from the Rio Grande. Mexico claimed that some of its border communities, most notably Ciudad Juarez, were experiencing water shortages because the US was extracting too much water. They argued that water use for irrigation in the San Luis Valley, Colorado, was diminishing the flow of water in the Rio Grande to such an extent that there was not enough left to support the downstream Mexican frontier communities.

In dealing with the complaint of the Mexican government, the US Department of State asked the opinion of the Attorney General of the US Department of Justice, Judson Harmon. The central message of Harmon in a response to the State Department's request for his opinion was that "the United States is under no obligation to Mexico to restrain its use of the Rio Grande because its absolute sovereignty within its own territory entitles it to dispose of the water within that territory in any way it wishes, regardless of the consequences in Mexico" (McCaffrey, 1996, p.563). According to Harmon, the United States therefore had no obligations towards Mexico when it came to the distribution of water from the Rio Grande.

The US Department of State did not follow the opinion of Harmon in its actions and the Rio Grande dispute was resolved by an agreement between the US and Mexico that stipulated the equitable distribution of the Rio Grande water between the two countries. In its formal legal position, however, the US did follow Harmon by denying any liability on part of the US for depriving Mexico of water.

Absolute territorial sovereignty

The theory on which the US based its formal legal position in the Rio Grande dispute was not new. It was known as the principle of Absolute Territorial Sovereignty (ATS). Because the principle gained widespread familiarity through the dispute it is now also known as the Harmon doctrine. In its most general form the ATS principle, or Harmon doctrine, states that the use of a natural resource within a sovereign state is unrestricted. When applied to international watercourses the principle can be written as follows.

Principle 1.2.1 Absolute Territorial Sovereignty
A riparian state has absolute sovereignty over the portion of an international watercourse within its territory.

In practice, the ATS principle implies that a state could do whatever it wants with the portion of an international watercourse within its territory, irrespective of the harmful consequences this might have in downstream states. Kilgour and Dinar (1995) reason that the ATS principle therefore favors the use of river water by upstream states over that by downstream states. This claim seems to be supported by the fact that in other river disputes the initial legal position taken by the upstream state was often similar to the Harmon doctrine. In the early 1950s France asserted absolute territorial sovereignty when it wanted to divert water from the Carol river that flows from France into Spain (Wolf, 1999). Also in the 1950s Canada relied on a 1909 Boundary Waters Treaty, that embodied the Harmon doctrine, in a dispute with the US over the Columbia River (Lemarquand (1993), McCaffrey (1996)). In a dispute over the Indus river that flows from India into Pakistan, India initially invoked absolute territorial sovereignty until the dispute was resolved by the Indus Water Treaty of 1960 (McCaffrey, 2001). Likewise, in the long-lasting discussion over the division of the water from the river Nile, Ethiopia, the upstream country with the largest inflow of water, has more than once taken a position of absolute territorial sovereignty (Jovanovic (1985), McCaffrey (2001)).

While it is clear that the ATS principle mainly has been advocated by upstream states, McCaffrey (1996) questions whether it is actually possible to consider it a principle of international law:

It is axiomatic, for example, that a state is sovereign within its territory. In this sense, "sovereignty" implies complete and exclusive authority over that territory. However, does it necessarily follow that international law imposes no constraints upon a state's use within its territory of a river that flows into another state? Does a downstream state have no right to object to uses of a watercourse in an upstream state that results in harm to the former? And what of the sovereignty of the downstream state over its territory? Is it proper to regard that as having been infringed if actions in the upstream state unfavorably alter the characteristics of the portion of the watercourse in the downstream state? (pp.550-551).

McCaffrey is not the only one that has problems with regarding the ATS principle as a principle of international law. A survey of the view of publicists in international watercourse law reveals that, although there was some support for the ATS principle in the nineteenth and early twentieth century, the support declined sharply as the twentieth century progressed and the significance of non-navigational uses of international watercourses increased (McCaffrey, 2001). The view of contemporary commentators towards the ATS principle can even be called hostile:

As for the notion of territorial sovereignty, it was passed over in silence, its rejection being thus signaled (Bourne, 1996, pp.160-161).

In the case of the so-called "shared natural resources" such as, for example, the waters of an international watercourse which nature has put at the disposal of more than one State, such an approach is not only highly egoistic and bound to lead to great social and economic injustice, but also, from a legal point of view, self-contradictory. It is clear that the unrestricted disposal by State A of the waters of an international watercourse flowing from that State into State B based on the idea of State A's absolute territorial sovereignty is incompatible with the unrestricted disposal of those waters to which State B would be likewise entitled on the basis of its absolute territorial sovereignty over the natural resources which nature would ordinarily bring into its territory (Lammers, 1984, p.557).

Despite some support for the Harmon Doctrine, it has been nearly universally rejected (Moermond and Shirley, 1987, p.141).

However, this opinion and the principle it entailed were criticized and discredited, for obvious reasons, by subsequent decisions of international tribunals and writings of experts in this field. The basic principles of international law, contrary to the Harmon Doctrine, prohibit riparian states from causing harm to other states, and call for cooperation and peaceful resolution of disputes (Salman, 2007b, p.627).

Considering that this doctrine was immediately rejected by Harmon's successor and later officially repudiated by the US (McCaffrey, 1996); considering further that it was never implemented in any water treaty (with the rare exception of some internal tributaries of international waters), was not invoked as a source for judgment in any legal ruling regarding international waters, and was explicitly rejected by the international tribunal in the Lac Lanoux case in 1957 [...], *the Harmon Doctrine is wildly over-emphasized as a principle of international law* (Wolf, 1999, p.6).

The hostility towards the ATS principle concentrates on three main points. One, the ATS principle is considered to be unfair. The principle would allow upstream states to inflict enormous damage on downstream states (starving a downstream state and its population of water). Two, the ATS principle is considered to be self-contradictory. If an upstream state would invoke the ATS principle and divert water upstream it would immediately infringe on the sovereignty of a downstream state that is therefore no longer able to follow the same ATS principle. And three, the ATS principle is never used in treaties or agreements between nations. There even exists instances in which a state supported the ATS principle when it was the upstream state in a river dispute, but opposed the same principle when it was the downstream state in another (e.g., the US in the Rio Grande and Columbia cases, see Lammers (1984) and McCaffrey (1996)).

Whether one considers the ATS principle as a valid principle of international law or not, it does hold that, when there are no rules of international law that govern, states are free to do as they please. In this respect alone the ATS principle already has to be taken seriously. The ATS principle can thus be seen as some sort of status quo or default position that (upstream) states can take prior to negotiations of international water agreements. It has also played this role historically in the formation of the ILA's Helsinki rules and the UN convention on the law of the non-navigational uses of international watercourses, that are discussed in the next sections. In fact, one of the principal reasons why the UN convention has not yet been ratified by all its signatory states is that some of the states fear the loss of (absolute) sovereignty over shared waters. In a UN General Assembly discussion of a draft of the convention, a number of states criticized the convention for not taking into account the sovereignty of the watercourse states over their part of the international watercourse (Salman, 2007a).

1.3 The 1966 Helsinki rules

In 1954 the ILA, a scholarly non-governmental organization that promotes the study, clarification and development of international law, established the "committee on the uses of waters of international rivers", or rivers committee for short. The aim of this committee, that was composed of international academics, lawyers and policy makers, was to study the legal aspects of the uses of international rivers by sovereign states. The need for an authoritative statement on the matter had arisen in the early 1950s when a series of international river disputes broke out between upstream and downstream riparian states along some of the world's largest rivers. Specifically, India and Pakistan had a dispute over the Indus, Sudan and Egypt over the Nile, Canada and the US over the Columbia, and Israel and its neighbors over the Jordan. By reaching a consensus view on what comprised valid rules of international river use, the rivers committee hoped it could contribute to a peaceful resolution of these disputes and prevent international watercourse disputes in the future.

When the ILA rivers committee started its study there were four major views on the law of the use of international watercourses (Bourne, 1996): the ATS principle (discussed above), the principle of Unlimited Territorial Integrity (UTI) (also known as the principle of absolute territorial integrity or principle of riparian rights), the principle of prior appropriation (also known as the principle of historic rights), and the principle of equitable utilization.

Unlimited territorial integrity

The UTI principle is the antithesis of the ATS principle as it states that the use of a natural resource within a sovereign state is permitted only in so far it does not cause damage or injury in the territory of other sovereign states. The idea behind the UTI principle is that the use of a shared natural resource in one state would lead to an infringement of the territorial sovereignty of other states. For international watercourses the UTI principle has generally been presented as follows.

Principle 1.3.1 Unlimited Territorial Integrity
A riparian state has the right to demand the natural flow of an international watercourse into its territory that is undiminished in quantity and unchanged in quality by the upper riparian states.

It is clear that this version of the UTI principle can be incompatible when invoked by more than one riparian state. Consider, for instance, an international watercourse with one upstream state and one downstream state. If both the upstream and the downstream state adhere to this version of the UTI principle, then, according to the principle, both the upstream and the downstream state would have the right to the natural flow of the watercourse from the territory of the upstream state. Since the water originating in the upstream state can only be allocated once, this leads to an inconsistency. We will discuss this (theoretical) problem further below, when we introduce the principle of no (substantial) harm.

In practice, the UTI principle has only been advocated by downstream riparians. Lipper (1967) writes that an instance in which a state has tried to maintain a position of unlimited territorial integrity was during the Nile commission hearings in 1925. Egypt, the most downstream state in the Nile river basin, claimed that it had unlimited rights to the natural flow of the river Nile. The position was quickly rejected by the commission. Egypt, however, kept on supporting the UTI principle in international fora until as recently as 1981 (McCaffrey, 2001). McCaffrey (2001) gives a few other examples in which countries have taken a position of unlimited territorial integrity: it was the position of Spain in the France-Spain Carol river case mentioned earlier, the position of Pakistan in the India-Pakistan Indus dispute and the position of Bolivia in disputes concerning the Rio Mauri and the Rio Lauca, where Chile was upstream of Bolivia.

In the Chile-Bolivia disputes Bolivia referred to the 1933 Montevideo declaration of American States that, according to Bolivia, embodied the UTI principle. Lammers (1984) explains that the 1933 Montevideo declaration was mainly based on doctrinal views and not on state practice. Lammers, more generally, questions the validity of the UTI principle as a principle of international law and its value in dispute resolution:

The principle of absolute territorial integrity is far from attractive, not only because of the great difference in opportunity to make use of the water which it may create for the communities on both sides of the border in certain situations, but also because of the great restraint which it imposes on the use of territory and natural resources (Lammers, 1984, p.562).

Apart form Lammers (1984) there are other authors who question the UTI principle as a valid principle of international law. McCaffrey (2001) mentions that there were some supporters of the principle in the early twentieth century but that most modern texts firmly dismiss the UTI principle as a principle of international watercourse law. This can also be seen in the following quotes from the literature:

No case has been found in which the theory of territorial integrity has been applied by any tribunal in a dispute involving the rights of coriparian states in the uses of the waters of an international river. Nor is there evidence of a state having accepted a diplomatic settlement based upon this theory (Lipper, 167, p.18).

History has been no kinder to the doctrine of absolute territorial integrity than to its theoretical opposite, absolute territorial sovereignty. Both doctrines are in essence, factually myopic and legally 'anarchic': they ignore other states' need for and reliance on the waters of an international watercourse, and they deny that sovereignty entails duties as well as rights (McCaffrey, 2001, p.135).

This principle has also been criticized and, like the Harmon Doctrine, is not recognized as a part of contemporary international water law (Salman, 2007b, p.627).

The points of critique that exist against the ATS principle also apply against the UTI principle. It is considered to be unfair because it ignores the water needs of other states. It has never been used in treaties and agreements between nations. And, like the ATS principle, the UTI principle is considered by some authors to be self-contradictory. This, however, also depends on the interpretation of the UTI principle. The interpretation in Principle 1.3.1 assumes that upstream states along an international watercourse are only able to harm downstream states, and not the other way around. Salman (2007a) says the following about this:

It is a common mistaken belief among a large segment of lawyers and non-lawyers that harm can only "travel" downstream, and it is not recognized that upstream states can also be harmed by activities by downstream states. In other words, this mistaken notion is based on the assumption that only upstream riparians can harm downstream riparians. It is obvious, and clearer, that the downstream riparians can be harmed by the physical impacts of water quality and quantity changes caused by use by upstream riparians. It is much less obvious, and generally not recognized, that the upstream riparians can be harmed by the potential foreclosure of their future use of water caused by the prior use and the claiming of rights by downstream riparians. For example, a poor upstream country could be precluded from developing the water resources of an international waterway tomorrow if a richer downstream riparian, without consultation or notification, develops today (Salman, 2007a, p.9).

The example given by Salman draws upon the distinction between physical harm and legal harm. In the example the poor upstream country is able to inflict physical harm on the richer downstream country by depriving it of water. The downstream country, on the other hand, is able to inflict legal harm on the upstream country by developing the water resource before the upstream country does. By developing the water resource prior to the upstream country the downstream country places a legal claim on the use of the water resource which can prevent the upstream country from developing the water resource. It is however, also possible to think of an example in which a downstream state inflicts physical harm on an upstream state. A downstream state, for instance, can construct a dam on its territory that causes flooding upstream. A flood can clearly cause damage in the upstream state. Following the interpretation that a downstream state is able to cause (legal or physical) harm to an upstream state one could argue that when a downstream state would adhere to the classic interpretation of the UTI principle (as in Principle 1.3.1) it could cause damage or injury in the territory of upstream states which goes against the spirit of the same UTI principle.

Prior appropriation

The principle of prior appropriation gives the rights over the use of a natural resource to the state who's use existed prior in time. This has also been written shortly as "first in time, first in right" (Wolf, 1999, p.6). For international watercourses this means the following.

Principle 1.3.2 Prior Appropriation
A riparian state that first makes use of (a quantity of) water from an international watercourse has the right to the continued use of that (quantity of) water.

The principle of prior appropriation thus fully protects the use of water from an international watercourse that exists prior in time. No other considerations are relevant and, as in the ATS and UTI principles, no type of water use is superior to others. The principle of prior appropriation has primarily been referred to by downstream states, often in combination with some form of the UTI principle. Mexico asserted prior appropriation in the dispute with the US over the Rio Grande (McCaffrey, 1996) and Egypt has used the principle as its leading argument in defending its large use of Nile river water in comparison to some of its upstream states (Wolf, 2007). Although most authors reject the principle of prior appropriation in its absolute form, Lipper (1967) writes that there are some authors that feel that pre-existing use plays an important role in international watercourse disputes and therefore should, in certain circumstances, be given preferred treatment. These mixed opinions can also be observed in the following quotes from the literature:

The principle may work out as highly inequitable for a riparian State in which the exploitation of the water resources has, for reasons beyond its power, lagged behind (Lammers, 1984, p.364)

Consequently, it is not surprising that there is little support in the international community for the principle of prior use. However, many publicists indicate that international law demands compensation for injury to an existing use (Moermond and Shirley, 1987, p.143).

In contrast to the extreme rarity with which absolute principles are codified, prior uses are regularly protected (with one major exception described below), notably in every single boundary waters accord in our collection (Wolf, 1999, p.11).

Apart from constituting a stand-alone water distribution principle in international watercourse law, historical rights can also be an important factor in some of the other watercourse principles discussed below, such as the principle of equitable utilization and the principle of territorial integration of all basin states.

Equitable utilization

Out of the four principles mentioned by Bourne (1996), it was the principle of equitable utilization that eventually would form the basis of the 1966 Helsinki rules. In international watercourse disputes prior to 1966 the ATS principle was mostly supported by upstream riparian states and the UTI and prior appropriation principles by downstream riparian states. These opposing principles also divided the members of the ILA's rivers committee:

The ILA is a non-governmental organization composed of individuals. Those who serve on its committees do so in their individual capacities and not as representatives of their governments or any other person or organization. Nevertheless, some members of the Rivers Committee were evidently concerned about protecting the interests of their governments and thus opposed propositions that were contrary to positions taken by their governments in disputes with co-basin states (Bourne, 1996, p.158).

In meetings of the ILA rivers committee members from upstream states frequently stressed the importance of territorial sovereignty, while members from downstream states defended territorial integrity and existing use. It took the members of the rivers committee more than ten years and five meetings in Dubrovnik 1956, New York 1958, Hamburg 1960, Brussels 1962 and Tokyo 1964 to come up with a compromise. The compromise was eventually established in the 1966 Helsinki rules on the uses of the waters of international rivers. The core articles of the Helsinki rules are Articles IV and V[3]:

Article IV

Each basin State is entitled, within its territory, to a reasonable and equitable share in the beneficial uses of the waters of an international drainage basin.

Article V

1. What is a reasonable and equitable share within the meaning of Article IV is to be determined in the light of all the relevant factors in each particular case.

2. Relevant factors which are to be considered include, but are not limited to:

[3] International Law Association (1966).

(a) the geography of the basin, including in particular the extent of the drainage area in the territory of each basin State;
(b) the hydrology of the basin, including in particular the contribution of water by each basin State;
(c) the climate affecting the basin;
(d) the past utilization of the waters of the basin, including in particular existing utilization;
(e) the economic and social needs of each basin State;
(f) the population dependent on the waters of the basin in each basin State;
(g) the comparative costs of alternative means of satisfying the economic and social needs of each basin State;
(h) the availability of other resources;
(i) the avoidance of unnecessary waste in the utilization of waters of the basin;
(j) the practicability of compensation to one or more of the co-basin States as a means of adjusting conflicts among uses; and
(k) the degree to which the needs of a basin State may be satisfied, without causing substantial injury to a co-basin State.

3. The weight to be given to each factor is to be determined by its importance in comparison with that of other relevant factors. In determining what is a reasonable and equitable share, all relevant factors are to be considered together and a conclusion reached on the basis of the whole.

Article IV, in fact, has become the standard formulation of the principle of equitable utilization.

Principle 1.3.3 Equitable Utilization
Each basin State is entitled, within its territory, to a reasonable and equitable share in the beneficial uses of the waters of an international drainage basin.

A more precise definition is given by Lipper (1967, p.63): "Equitable utilization is the division of the waters of an international river among the coriparian states in accordance with the legitimate economic and social needs of each, in such a manner as to achieve the maximum benefit for all with the minimum detriment to each". Lipper writes that the cornerstone of the principle of equitable utilization (and therefore of the Helsinki rules) is equality of rights. This means that each riparian state has an equal right to use the water from an international watercourse in accordance with their needs. It is important to note that equality of rights is not synonymous with the equal division of water. The right of a riparian state depends on the economic and social needs of the state, and possibly on other factors. In the Helsinki rules some factors that determine the right of each riparian state to a share of the water from an international watercourse are given in Article V.

The principle of equitable utilization has some remarkable features. Fundamental to the principle is that implementation of it requires a case-by-case approach. The principle does not provide a clear-cut rule that can be applied to each international watercourse. Instead, different factors can have different weights in different international watercourse disputes and there is no a priori hierarchy among the factors. McCaffrey (2001) writes that equitable utilization is therefore best understood as a dynamic process, which depends heavily upon active cooperation between states sharing water resources, rather than a fixed state of affairs. Also, the principle of equitable utilization does not require that the water of an international watercourse is put to its most productive use, i.e., it does not require efficiency. For the use of water from an international watercourse to be protected by the principle of equitable utilization the use has to be beneficial, but not necessarily the most beneficial it could be. Efficiency, nevertheless, could be a factor in determining what are reasonable and equitable shares within the meaning of the principle. Another feature of the principle of equitable utilization is that no type of water use is inherently superior to any other type. Historically, navigation had always been given priority over non-navigational uses of (international) watercourses. Later, some authors have argued that water for direct consumption and irrigation should be given priority over other uses. The principle of equitable utilization does not require this.

The most remarkable feature of the principle of equitable utilization, and its relatively general acceptance in the 1960s and 70s, is that it does not prohibit the causing of harm to other riparian states. Especially because the principle of no substantial harm has played a significant role in (debates about) the UN Convention on the Law of the Non-Navigational Uses of International Watercourses in the 1980s and 90s.

Principle 1.3.4 No (Substantial) Harm
A riparian state is free to use the water of an international watercourse provided that this use does not cause (substantial) harm to other riparian states.

Since the principle of equitable utilization does not explicitly prohibit the causing of harm to other riparian states, it is clear that the principles of equitable utilization and of no substantial harm can be incompatible in certain cases. The friction between the two principles and the role they play in the UN convention are discussed at length in the next section.

As a short intermezzo, consider the similarity between the no substantial harm principle and the UTI principle. The most general form of the UTI principle says that the use of a natural resource within a sovereign state is permitted only in so far it does not cause damage or injury in the territory of other sovereign states. This is, in essence, the same as the no substantial harm Principle 1.3.4. For international watercourses, however, the UTI principle has evolved to a statement about a riparian state and all its upstream states. That is, the UTI principle says that a riparian state has the right to demand the natural flow of an international watercourse into its territory that is undiminished in quantity and unchanged in quality by the upper riparian states. In contrast, the no substantial harm principle says that a riparian state is not allowed to cause (substantial) harm to other riparian states, both upstream and downstream. In the subsequent chapters we will consider the UTI principle to be about a riparian state and all its upstream states and the no substantial harm principle to be about a riparian state and all states both upstream and downstream to it.

Let's now return to the principle of equitable utilization. Apart from the incompatibility with the no substantial harm principle, the main point of critique that has been raised against this principle is that it is too vague. Unlike the ATS, UTI and prior appropriation principles, the concept of equitable utilization does not lend itself to a precise formulation in terms of water division. Lammers (1984) and Caponera (1985) argue that the principle of equitable utilization merely changes the question of how to share the waters of an international watercourse to the question what can be considered reasonable and equitable shares in the beneficial uses of an international watercourse. Wolf (1999), on the other hand, maintains that the Helsinki rules, and with it the principle of equitable utilization, created a key shift in legal thinking from the division of water per se to the division of the wealth resulting from the beneficial use of water. A fact remains that the implementation of the principle of equitable utilization depends crucially on how one weighs the different factors on which the principle relies.

In Article V of the Helsinki rules eleven factors are mentioned that can be relevant when determining what comprise reasonable and equitable shares within the meaning of the principle of equitable utilization. Of the eleven factors, factors (d) and (k) are of particular interests. Factor (d), on its own, is equal to the principle of prior appropriation and factor (k), on its own, is equal to the principle of no substantial harm. Hence, the Helsinki rules treat the prior appropriation and no substantial harm principles as factors subordinate to the principle of equitable utilization. Although the principles of prior appropriation and no substantial harm do not appear in the Helsinki rules on a stand-alone basis, it is thus recognized that they do have some merit.

In addition to a chapter on the principle of equitable utilization, the Helsinki rules also contain chapters on pollution, navigation, timber floating and procedures for the settlement and prevention of disputes. The articles in these chapters are all in accordance with the principle of equitable utilization (Bourne, 1996). According to Salman (2007b) the Helsinki rules were the first international legal instrument to include rules for both navigational as well as non-navigational uses of international watercourses. The decline in importance of navigation in law was, however, also confirmed by the Helsinki rules as Article VI states that "a use or category of uses is not entitled to any inherent preference over any other use or category of uses".[4]

Bourne (1996) writes that the Helsinki rules were quickly accepted by the international community as customary international law. What could have helped in this process is that many international watercourse treaties were already based on some form of the principle of equitable utilization. For instance, a 1927 treaty between the Soviet Union and Turkey, the 1929 Nile agreement between Egypt, Sudan and Great Britain (the colonial power at that time) and a 1930 agreement between the Dominican Republic and Haiti were all based on the principle of equitable utilization (Moermond and Shirley, 1987). Also after the publication of the Helsinki rules the principle of equitable utilization played a major role in many agreements. Moermond and Shirley (1987) mention the 1969 Plata river basin treaty and the 1978 Amazon cooperation treaty, Wolf (1999) mentions a Mekong committee agreement and Salman (2007b) the 1973 Asian-African legal consultative committee agreement, a 1992 agreement between Namibia and South Africa on the establishment of a permanent water commission and the 1995 protocol on

[4] International Law Association (1966).

shared watercourse systems in the Southern African development community. It should not come as a surprise that because of this the opinions of contemporary commentators on the Helsinki rules and the principle of equitable utilization are mostly positive:

The vast support for the principle of equitable utilization has lead publicists to state that of the principles of international river law, the principle of equitable utilization comes the closest to being a general principle of international law (Moermond and Shirley, 1987, p.152).

The Helsinki Rules obviously constitute a monumental work. They have had a major impact upon the development of the law of international watercourses, and reflect many principles and trends that later found expression in the UN Convention (McCaffrey, 2001, p.321).

Like other IIL and ILA rules and resolutions, the Helsinki Rules have no formal standing or legally binding effect per se. However, until the adoption of the UN Convention 30 years later, they remained the single most authoritative and widely quoted set of rules for regulating the use and protection of international watercourses. Indeed, those Rules are the first general codification of the law of international watercourses (Salman 2007b, p.630).

Despite the major role the Helsinki rules and the principle of equitable utilization have played in (the development of) international watercourse law, it is important to remember that their acceptance has never been universal. There are many instances in which the Helsinki rules were ignored and riparian states referred back to the ATS, UTI and prior appropriation principles. This is one of the primary reasons why the work on international watercourse law has continued after the publication of the Helsinki rules and why the UN set out to formalize the law on international watercourses in 1970.

1.4 The 1997 UN convention

On December 8, 1970 the UN General Assembly adopted a resolution that asked the International Law Commission (ILC), a UN committee of legal experts that promotes the progressive development and codification of international law, to take up the study of the law of international watercourses. The main purpose of this study was to investigate whether, and how, it was possible to codify the law of the non-navigational uses of international watercourses in an international convention.

It took the ILC twenty-three years to agree on a draft article that set out the fundamental rules of international watercourse law. Some of the complexities that caused long debates within the ILC were the definition of the term "international watercourse", the status of existing watercourse agreements, and the procedures for dispute settlement (Salman, 2007a). But, it was the relationship between the principle of equitable utilization and the no substantial harm principle that was responsible for the longest delays.

When the draft convention was submitted to the General Assembly in 1994 it took another three years, and several revisions, before the convention was finally put up for vote in the same General Assembly. On May 21, 1997 the UN convention on the law of the non-navigational uses of international watercourses was adopted with 103 countries voting in favor of the convention, 3 countries voting against, 27 abstentions and 52 countries not participating in the vote.

Countries can become parties to the convention by ratifying it (accepting it through their constitutional process). The convention requires thirty-five ratifications to enter into force. At the time of writing this number has not (yet) been reached.

The main goal of the UN convention is to ensure the utilization, development, management and protection of international watercourses and to promote their optimal and sustainable use (Salman, 2007a). It can be seen as a framework convention that has been largely based on the work of the ILA, in particular the Helsinki rules. The concept of a framework convention should be understood as one that sets out general principles and rules that may be tailored to suit the conditions of specific watercourses (McCaffrey, 2001). The focus of the convention therefore lies more on procedural aspects and less on substantive ones. It leaves out many of the details for specific watercourses and allows riparian states to fill in these details in agreements among themselves. Thus, the idea is that, once in force, the UN convention provides a framework that riparian states may apply and adjust to particular watercourses (by mutual consent). Interesting in this respect is also that the UN convention does not affect the rights or obligations under pre-existing agreements. It calls upon riparian states to harmonize their current agreements with the basic principles of the convention but does not force them to do so. Hence, it seems that it is not the aim of the UN convention to replace the fragmented international watercourse law but to complement it by promoting the creation of specific watercourse agreements.

McCaffrey (2001) writes that the four key elements of the UN convention are the principle of equitable utilization, the principle of no substantial harm, the notion of prior notification about planned measures and the protection of ecosystems. We discuss each of these elements starting with the principle of equitable utilization, the principle of no substantial harm and (the debate about) their relation in the convention. The principles of equitable utilization and no substantial harm appear in the convention as (part of) Articles 5, 6 and 7:

Article 5

Equitable and reasonable utilization and participation

1. Watercourse States shall in their respective territories utilize an international watercourse in an equitable and reasonable manner. In particular, an international watercourse shall be used and developed by watercourse States with a view to attaining optimal and sustainable utilization thereof and benefits therefrom, taking into account the interests of the watercourse States concerned, consistent with adequate protection of the watercourse.

2. Watercourse States shall participate in the use, development and protection of an international watercourse in an equitable and reasonable manner. Such participation includes *both the right to utilize the watercourse and the duty to cooperate in the protection and development thereof, as provided in the present Convention.*

Article 6

Factors relevant to equitable and reasonable utilization

1. Utilization of an international watercourse in an equitable and reasonable manner within the meaning of article 5 requires taking into account all relevant factors and circumstances, including:
(a) Geographic, hydrographic, hydrological, climatic, ecological and other factors of a natural character;
(b) The social and economic needs of the watercourse States concerned;
(c) The population dependent on the watercourse in each watercourse State;
(d) The effects of the use or uses of the watercourses in one watercourse State on other watercourse States;
(e) Existing and potential uses of the watercourse;
(f) Conservation, protection, development and economy of use of the water resources of the watercourse and the costs of measures taken to that effect;
(g) The availability of alternatives, of comparable value, to a particular planned or existing use.

2. In the application of article 5 or paragraph 1 of this article, watercourse States concerned shall, when the need arises, enter into consultations in a spirit of cooperation.

3. The weight to be given to each factor is to be determined by its importance in comparison with that of other relevant factors. In determining what is a reasonable and equitable use, all relevant factors are to be considered together and a conclusion reached on the basis of the whole.

Article 7

Obligation not to cause significant harm

1. Watercourse States shall, in utilizing an international watercourse in their territories, take all appropriate measures to prevent the causing of significant harm to other watercourse States.

2. Where significant harm nevertheless is caused to another watercourse State, the States whose use causes such harm shall, in the absence of agreement to such use, take all appropriate measures, having due regard for the provisions of articles 5 and 6, in consultation with the affected State, to eliminate or mitigate such harm and, where appropriate, to discuss the question of compensation.

The first thing to note in these articles is the striking similarity between articles IV and V of the Helsinki rules and articles 5 and 6 of the UN convention. Articles IV and 5 define the principle of equitable utilization and articles V and 6 provide a list of factors relevant in determining what is equitable. When comparing the factors of the Helsinki rules to the factors of the UN convention one can conclude that those of the UN convention are primarily based on those in the Helsinki rules. During the negotiations on the factors to be included in the UN convention's list, some suggestions were made by riparian states that did not make it into the final draft of the convention: India wanted that the contribution to the watercourse of each watercourse state was mentioned explicitly, as it was in the Helsinki rules, Egypt suggested a sentence on the availability of other water resources and Finland called for the sustainable development of the watercourse and the recognition of the needs and interests of future generations (McCaffrey and Sinjela, 1998).

Article 7 of the UN convention expresses the principle of no substantial harm. Tanzi and Arcari (2001, pp.175-176) state that the no substantial harm principle is primarily seen as a principle applicable to environmental issues:

From an historical perspective, it is true that the concept of equitable utilisation, as a development of the principle of 'equitable apportionment', has grown out of the need for the regulation of conflicting claims over water quantity issues prior to the no harm rule which addresses primarily transboundary pollution issues.

This also explains why there is no separate article on the principle of no substantial harm present in the Helsinki rules. Environmental concerns about international watercourses only matured after its publication, in the 1970s and 80s.

Although historically the interpretation of the principle of equitable utilization has been one about water quantity and the interpretation of the no substantial harm principle one about water quality, they do not play this restrictive role in the UN convention. This also explains why Article 7 is the most controversial article of the UN convention (McCaffrey, 2001). The inclusion of both principles in the convention opens the doors to a debate about which of the principles, the principle of equitable utilization or the principle of no substantial harm, takes priority over the other. In the revision process prior to the adoption of the convention it quickly became clear that upper riparians tended to favor the principle of equitable utilization and lower riparians the principle of no substantial harm. After years of debate a compromise was eventually reached on the text that upper riparians believed to support the subordination of the principle of no substantial harm to the principle of equitable utilization and lower riparians considered neutral enough not to suggest such a subordination. Balancing the two principles allowed both sides to claim victory (McCaffrey, 2001). Although the compromise led to the adoption of the convention, it did not end the debate about the issue. According to Salman (2007a) the ongoing debate about the priority of the principles is the key reason why the convention has not (yet) entered into force. Salman writes that, on the one hand, upper riparians still feel that the convention is biased in favor of lower riparians because the obligation not to cause significant harm is specifically mentioned. On the other hand, downstream riparians such as Egypt, Pakistan and Peru abstained from voting on the convention because they were afraid that the convention favors upstream riparians by the subordination of the no harm principle to the principle of equitable utilization.

Not only riparian states are divided over the subordination question, there is also no consensus among publicists on the topic. Salman (2007a, p.6), for instance, states:

Accordingly, a careful reading of Articles 5, 6 and 7 of the Convention should lead to the conclusion that the obligation not to cause harm has indeed been subordinated to the principle of equitable and reasonable utilization.

Other authors are more careful in their conclusions. McCaffrey (2001, p.308) writes:

[I]t seems reasonable to conclude that the 'no-harm' rule would not automatically override that of equitable utilization in case the two came into conflict.

McCaffrey (2001) further stresses that the version of the no substantial harm principle adopted in the convention is not an absolute or strict one but rather one of 'due diligence'. Article 7 only requires states to "take all appropriate measures" to prevent the causing of significant harm and does not outright forbid it. Tanzi and Arcari (2001) argue that there is no conflict between the principles of equitable utilization and no substantial harm in the UN convention, or at least that the convention sufficiently balances the two principles to keep it internally consistent. One of their arguments to support this claim is that the principle of equitable utilization and the principle of no substantial harm are both part of the same normative framework. This framework can be stated as a covering principle known as the principle of limited territorial sovereignty.

Principle 1.4.1 Limited Territorial Sovereignty
There exist legal restrictions on every state's use of an international watercourse.

While definitions and interpretations of the principle of limited territorial sovereignty have varied, it is clear that the principle of equitable utilization and the principle of no substantial harm are both special cases of it. A number of authors have expressed the view that the principle of limited territorial sovereignty is the most general hybrid of the ATS and UTI principles that one can think of (e.g., Lipper (1967), McCaffrey (2001)). The principle of limited territorial sovereignty is almost universally accepted. This is not very strange because the principle is such a broad one. It only says that states are not allowed to do whatever they like with an international watercourse but must respect the fact that they are sharing it with other states. The wide support for the principle of limited territorial sovereignty has been illustrated, among others, by Lipper (1967, p.28):

An examination of governmental statements shows that those nations once in the forefront as proponents of the Harmon Doctrine no longer subscribe to it; also, among those governments which have publicly taken a position on the issue, the limited territorial sovereignty principle overwhelmingly predominates, and some have even moved beyond it.

Apart from the section on the principles of international watercourse law the UN convention also contains parts on "planned measures", "protection, preservation and management", "harmful conditions and emergency situations" and an annex on "arbitration".

The section on "planned measures" is about the notification of such planned measures and is the longest of the convention. It lays down rules about the obligation of notifying of planned measures, the period for reply, absence of reply, consultations, procedures in absence of notification and urgent implementation of planned measures (Salman, 2007a). It requires the notification of a project (use of water from the international watercourse) in case this project has the potential to cause significant adverse effects to other riparian states. This means that when a riparian state uses water from an international watercourse and this use causes no harm to any other riparian states it also does not have to notify them of the use. If the use does cause harm to another riparian state and this state objects against the use, then the convention lays down procedures that have to be followed: riparian states are obliged to enter into consultations and, if necessary, negotiations to arrive at a reasonable and equitable resolution of the problem. McCaffrey (2001) explains that the fact that the convention requires prior notification of (detrimental) changes is significant because this signifies that the international community as a whole seems to reject the notion of absolute territorial sovereignty. By accepting that it is required to inform another riparian state of any (detrimental) change a state gives up the idea that it can do whatever it pleases with that part of an international watercourse within its territory without taking into account the effects this has on other riparian states.

The first article of the section on "protection, preservation and management", Article 20, says that watercourse states have to protect and preserve the ecosystems of international watercourses. The use of the term 'ecosystem' is remarkable because it is only used twice in the entire document and implies a wider responsibility than the protection of only the watercourse itself. McCaffrey writes that "the 'ecosystem' of an international watercourse should be understood to include not only the flora and fauna in and immediately adjacent to a watercourse, but also the natural features within its catchment that have an influence on, or whose degradation could influence, the watercourse" (McCaffrey, 2001, p.393). As an example McCaffrey mentions the logging of trees upstream along a river that could cause soil erosion, which in turn could result in downstream floods or mudslides. The use of the 'modern' concept of ecosystem is slightly out of line with the rest of the convention, also because the full potential of the concept does not seem to be employed in the subsequent articles. Since the protection of ecosystems plays a larger role in the ILA's 2004 Berlin rules than it does in the UN convention this topic is discussed further in the following section.

Although the UN Convention on the Law of the Non-Navigational uses of International Watercourses has not (yet) entered into force, most authors agree that its development has been a success. Salman (2007a, p.13) concludes that "even if the Convention does not enter into force, it has received broad endorsements, and it is widely agreed that it reflects and embodies the basic principles of international water law". Also Salman states that "the Convention is and will continue to be the most authoritative instrument in the field of International Water Law" (Salman, 2007a, p.13). The development of the convention is seen as significant because most articles of the convention reflect the current views of the international community on international watercourse law. The reason why one can make this statement is because it was globally negotiated by almost all interested riparian nations. Even if the convention would never receive enough ratifications to enter into force, it still has a use as a normative benchmark in negotiations about bi- or multilateral treaties between riparian states. In fact, the convention has already fulfilled this role as it has influenced many agreements before and after its adoption in 1997 (see Salman (2007a)).

1.5 The 2004 Berlin rules

After the publication of the Helsinki rules in 1966 the ILA did not stop its work on the law of international watercourses. On the contrary, in 1972 it published articles on flood control, in 1976 it agreed on rules about administration of international watercourses, in 1980 it adopted two sets of rules on international watercourses at its Belgrade conference, in 1982 at the Montreal conference it accepted articles dealing with the pollution of international watercourses and in 1986 the "Complementary Rules Applicable to International Water Resources", which included rules on transboundary groundwater, were adopted at the Seoul conference.

In the 1990s the members of the ILA realized that the rules they had created over the previous three decades were highly fragmented and dealt with a large number of related issues. They decided to deal with this problem in two ways. The first was to compile an overview of the ILA's work in a single document. This led to the 1999 Campione consolidation that included all of the ILA's 1966-1999 rules but no new ones. The second way was to commence on a new project that would eventually lead to the 2004 Berlin rules on water resources.

The 2004 Berlin rules were created with the idea to revise and update the 1966 Helsinki rules. The main difference between the two sets of rules is that the Berlin rules apply to a much wider set of topics than the Helsinki rules. This is also reflected in the titles of the two documents. Whereas the Helsinki rules are on "the uses of the waters of international rivers" the Berlin rules are simply on "water resources". The Berlin rules contain provisions that are applicable to national as well as international waters and deal with surface as well as groundwaters. Salman (2007b) calls the Berlin rules a comprehensive and detailed set of rules that go beyond the Helsinki rules and the UN convention. Salman also writes that, in contrast to the Helsinki rules and the UN convention, the Berlin rules are controversial in the sense that they do not really reflect the global consensus view on international water law. Rather, the rules reflect the judgment of experts in the field of what international water resource law should be. Among others, the Berlin rules contain chapters on the principles of international water resource law, the right of persons, the protection of aquatic environments, groundwater, navigation, administration and the settlement of disputes.

On the principles of international water resource law Salman (2007b) writes the following:

The major distinction between the Helsinki Rules and the UN Convention on the one hand, and the Berlin Rules on the other, is that the former establish and emphasize the right of each basin state to a reasonable and equitable share. This is based on the concept of equality of all riparian states in the use of the shared watercourse. On the other hand, the Berlin

Rules obliges each basin state to manage the waters of an international drainage basin in an equitable and reasonable manner. The term 'manage' is defined in Article 3(14) of the Berlin Rules to include the development, use, protection, allocation, regulation, and control of the waters. Thus, whereas the Helsinki Rules and the UN Convention establish and emphasize the right of each of the riparian states to a reasonable and equitable share, the Berlin Rules emphasize the obligation to manage the shared watercourse in an equitable and reasonable manner (Salman, 2007b, p.636).

This distinction should be seen in the light of Articles 12(1) and 16 of the Berlin rules. Article 12(1) is a representation of the principle of equitable utilization and Article 16 of the principle of no substantial harm[5]:

Article 12

Equitable Utilization

1. Basin States shall in their respective territories manage the waters of an international drainage basin in an equitable and reasonable manner having due regard for the obligation not to cause significant harm to other basin States.

Article 16

Avoidance of Transboundary Harm

Basin States, in managing the waters of an international drainage basin, shall refrain from and prevent acts or omissions within their territory that cause significant harm to another basin State having due regard for the right of each basin State to make equitable and reasonable use of the waters.

What becomes clear from these articles is that in the Berlin rules the principles of equitable utilization and no substantial harm are subordinated to each other. This means that neither of the principles dominates the other in the absolute sense. It is the current view of the ILA that with the right to an equitable share in the beneficial uses of an international watercourse comes the obligation not to cause significant harm to other users of the watercourse. In the commentaries of the ILA on articles 12 and 16 it is explained that the interrelation between the principle of equitable utilization and the principle of no substantial harm must be worked out in each individual case.

The view of the ILA on the principle of equitable utilization as the leading principle of international watercourse law has changed dramatically from the publication of the Helsinki rules to the publication of the Berlin rules. In the Helsinki rules the principle of equitable utilization is considered as the only valid principle of international watercourse law. In the Berlin rules the principle of equitable utilization stands on equal footing with the principle of no substantial harm. This change reflects the growing role of watercourse

[5]International Law Association (2004).

pollution and environmental protection in international watercourse law. Over the past five decades quantitative harm caused by the use of an international watercourse has mainly been addressed through the principle of equitable utilization and qualitative harm mainly through the principle of no substantial harm. The increasing focus of the international community on environmental issues led to increasing support for the principle of no substantial harm.

In the Helsinki rules there is a single chapter, consisting of two articles, on the pollution of international watercourses. These articles, in short, present the view that pollution of international watercourses should be dealt with through the principle of equitable utilization. As mentioned in the previous section, the UN convention contains a chapter on "protection, preservation and management". One of the articles in this chapter is on "prevention, reduction and control of pollution". Another article is on the "protection and preservation of ecosystems". McCaffrey (2001, p.396) writes the following about the UN obligation to protect and preserve ecosystems:

While this obligation may be described as 'new' or 'emerging', its basic elements are already part of general international law. The obligation, as formulated in Article 20 of the UN Convention, simply reflects advances in scientific knowledge about the interrelationships of natural systems.

As stated by McCaffrey (2001) the protection of ecosystems is an emerging obligation in the field of international watercourse law. In the Berlin rules this obligation plays a central role and encompasses the prevention, reduction and control of pollution of international watercourses. Article 22 of the Berlin rules states a new fundamental principle of international watercourse law, the principle of ecological integrity.

Principle 1.5.1 Ecological Integrity
States shall take all appropriate measures to protect the ecological integrity necessary to sustain ecosystems dependent on particular waters.

The term ecological integrity is defined in Article 3 of the Berlin rules as follows: "'Ecological integrity' means the natural condition of waters and other resources sufficient to assure the biological, chemical, and physical integrity of the aquatic environment".[6] According to the ILA commentary accompanying the article the definition of the term reflects a balance whereby ecological integrity does not require the absolute protection of waters but that level of integrity necessary for the survival of ecosystems.

In the past pollution was seen as the only environmental watercourse problem, see, for instance, Lammers (1984). These days the degradation of complete watercourse ecosystems is considered as the major environmental threat to international watercourses. One of the key reasons behind this change is that scientific development has revealed that physical, chemical and biological elements are often part of an integrated system. Complex interactions within a watercourse ecosystem can lead to unforeseen consequences. This can threaten essential services provided by watercourse ecosystems, such as the controlling

[6] International Law Association (2004).

of floods, purifying of water, recharging of aquifers, restoring of soil fertility, nurturing of fisheries and supporting of recreation (McCaffrey, 2001).

The Berlin rules try to prevent degradation of ecosystems by demanding that riparian states take into account the consequences of their actions on the entire ecosystem. Tanzi and Arcari (2001) argue that, while the protection of ecosystems is a laudable goal, there are difficulties in defining the boundaries of a watercourse ecosystem. Because of the complex interactions within an ecosystem it is very hard to predict what all the consequences of a particular action might be. It could, for example, be the case that an activity that takes place miles away from a watercourse, and that at first sight appears to have no relation to the watercourse ecosystem, turns out to do great harm to it through a complex chain of events. The question then is whether the activity causing the harm can be detected. And, if it is detected, whether the Berlin rules are the right tool to deal with the problem.

The Berlin rules still draw heavily on the Helsinki rules and the UN convention. Salman (2007b), nevertheless, concludes that there are three main differences. One, some of the Berlin rules apply to both national as well as international waters. Two, the Berlin rules do not only include established principles but also emerging ones. And three, the Berlin rules downgrade the principle of equitable utilization so that it is on equal footing with the principle of no substantial harm. Especially this last point is significant from a historical perspective. The work on international watercourse law over the past five decades has been dominated by the debate about the relation between the principle of equitable utilization and the principle of no substantial harm. Salman (2007b, p.639) also confirms this when he writes that “the recent history of, and the work on, international water law has been occupied largely by the relationship between those principles”. Although the Berlin rules try to put an end to this debate by subordinating the principles to each other, it is unlikely that the last word has been said about it.

1.6 Territorial integration of all basin states

Consider the following example by Lipper (1967, p.38):

The ideal location for a necessary installation, such as a dam for harnessing basin waters for hydroelectric use, may be within the territory of a riparian state uninterested in such a use, while only a less desirable location would be available in the interested coriparian state. In such a case, the principle of equitable utilization may not permit the most beneficial development of the basin. Moreover, parallel independent development of a river by each riparian is likely to prove economically wasteful.

An international watercourse principle that attempts to solve the problem sketched in this example, and that can serve as a basis for international watercourse law, is the principle of Territorial Integration of all Basin States (TIBS). This principle is also known as the principle of community (of interests) in the waters, the principle of common management or the drainage basin approach.

Principle 1.6.1 Territorial Integration of all Basin States
The water of an international watercourse belongs to all basin states combined, no matter where it enters the watercourse. Each riparian state is entitled to a reasonable and equitable share in the optimal use of the waters of the international watercourse.

The TIBS principle considers the entire river basin as an economic unit and ascribes the rights over the waters to the collective body of all riparian states. Because the geography of an international drainage basin often has no relationship to its political borders, efficiency of the water use in the basin requires an integrated approach by all riparian states. The TIBS principle requires the development of the drainage basin without reference to state borders. Lipper (1967) writes that this includes the joint planning, construction and management of projects and the sharing of the burdens of maintenance. More generally, TIBS requires the full sharing of both the benefits and costs of the management of an international watercourse. The implementation of the TIBS principle can be summarized in three steps: (1) the water rights over an international watercourse belong to all basin states combined, (2) the basin states are obliged to put the water from the international watercourse to its most productive use, i.e., the water of the international watercourse has to be used efficiently, and (3) each basin state has a right to a reasonable and equitable share of the benefit (wealth) that results from the optimal use of the water from the international watercourse, these reasonable and equitable shares possibly require a redistribution of wealth through (monetary) transfers.

The TIBS principle is very similar to the principle of equitable utilization but differs in one important aspect: it demands that the water of an international watercourse is used efficiently. While the principle of equitable utilization assigns a reasonable and equitable share in the beneficial uses of the waters *within its territory*, the TIBS principle assigns a reasonable and equitable share in *the optimal use of the waters of the international watercourse*. The TIBS principle therefore encompasses an international watercourse principle that is known as the principle of optimal use.

Principle 1.6.2 Optimal Use
Riparian states must together make optimal use of the water of an international watercourse as if they were not intersected by state boundaries.

In contrast to the other principles mentioned in this chapter, the principle of optimal use says something about the efficiency of the water use of an international watercourse and not about the division of the (benefit of) water. While the principles on the division of the water of an international watercourse all have a normative character, the principle of optimal use is a positive principle. The principle of optimal use is highly desirable from an economic point of view. Nonetheless, the political reality is different and the optimal use principle is hardly found in international watercourse treaties. In the literature it is recognized that optimal use of the international watercourse is something that basin states should strive towards, but also that it is something that by no means is part of contemporary international water law:

Highly desirable from an overall hydroeconomic point of view, it cannot be said that general international law has already so far developed that basin states are legally obliged to strive *at the optimum rational development of common water resources on a basin-wide scale* (Lammers, 1984, p.560).

Since the TIBS principle encompasses the principle of optimal use, it is obvious that also the TIBS principle has not been used often in state practice. This is unfortunate because compared to a situation in which there is no cooperation among riparian states, there could exist implementations of the TIBS principle that lead to a (Pareto) superior outcome for all riparian states. This is one of the reasons why the TIBS principle has been embraced by publicists in the field of international watercourse law. They, however, also realize that there exists a large gap between what is optimal theoretically and what is attainable politically:

But the assertion of the unity of an international drainage basin, which is really the basis of this principle, was and still is a doubtful proposition of law (Bourne, 1996, p.175).

Although this is no doubt a laudable goal for States to pursue, it is not yet required by general international law (Lammers, 1984, p.371).

This principle did not gain wide acceptance because riparian states believe that it forces them into reaching an agreement (Salman, 2007b, p.627).

In addition to its desirable efficiency property, the TIBS principle seems the most suitable international watercourse principle to deal with environmental issues. That extensive cooperation among riparian states is required for the protection of complete watercourse ecosystems is self-evident. The TIBS principle demands such close cooperation and therefore seems the best complement to the emerging principle of ecological integrity when it comes to protection of international watercourses.

Encouraging for proponents of the TIBS principle is that cooperation in international river basin committees that address both water distribution issues, as well as environmental protection issues, is on the rise. Perhaps the best example of cooperation in a basin-wide committee is that of the Nile Basin Initiative. The Nile Basin Initiative is an inter-governmental organization that includes representatives from Burundi, Democratic Republic of Congo, Egypt, Ethiopia, Kenya, Rwanda, Sudan, Tanzania and Uganda. Its main goal is to ensure the equitable and sustainable management and development of the shared water resources of the Nile river basin. Although cooperation in the Nile Basin Initiative has had mixed success, this type of close cooperation in specific river basin committees seems the only way forward from the still present practice of absolute claims and principles.

Finally, observe that it is crucial for the implementation of the TIBS principle that the countries sharing a river have the possibility to make monetary transfers to each other. While, for instance, the ATS and UTI principles are stated purely in terms of water, the TIBS principle is stated in terms of wealth. Each basin state has the right to a reasonable and equitable share of the wealth that results from the optimal use of the water from an international watercourse. A redistribution of this wealth possibly requires monetary transfers between countries.

To see this, consider the two-state example in which an upstream country consumes a large amount of water that can be put to a much more productive use in a downstream country. The TIBS principle then requires that the upstream country gives up its water for consumption in the downstream country. From a non-legal, non-cooperative, point of view, the upstream country would only be willing to do this if it is compensated for its water loss. The compensation can take place through a monetary transfer from the downstream country to the upstream country. The TIBS principle from international watercourse law recognizes that the upstream country would not be willing to give up water if it is not compensated for its loss by the downstream country. How large this compensation must be all depends on the reasonable and equitable shares in the definition of the principle. We return to this issue in Chapter 4.

In the rest of this dissertation we discuss river water allocation models in which the agents are able to make monetary transfers to each other. These models, in which we apply some of the international watercourse principles from this chapter, will make heavy use of cooperative game theory. For this reason we first discuss some preliminaries on cooperative game theory in the next chapter.

Chapter 2

Preliminaries

2.1 Cooperative games

TU-games

A cooperative game with transferable utility in characteristic function form is a rudimentary model of cooperation among (economic) agents. TU-games were introduced by von Neumann and Morgenstern (1944) and have since become a central object of study in the field of cooperative game theory. A *cooperative game with transferable utility in characteristic function form*, or *TU-game*, is a pair (N, v), where N is a finite set of $n = |N|$ *players* (we write the cardinality of a set A as $|A|$) and $v : 2^N \to \mathbb{R}$ is a *characteristic function* on N such that $v(\emptyset) = 0$ (we write the power set of a set A as 2^A).[1] A subset $S \subseteq N$, $S \neq \emptyset$, is called a *coalition*. For any coalition S, $v(S)$ displays the *worth* of that coalition. The worth of a coalition can be interpreted as the wealth, measured in units of transferable utility, which the members of that coalition are able to divide among themselves when they decide to cooperate. For $S \subset N$, the TU-game (S, v_S) denotes the *subgame* restricted to S with v_S given by $v_S(T) = v(T)$ for every $T \subseteq S$.[2] The collection of all TU-games is denoted by $\mathcal{G}$. Given a fixed player set N, the collection of all TU-games on N is denoted by $\mathcal{G}^N$. For a general introduction to TU-games see, for instance, Peleg and Sudhölter (2003).

A TU-game $(N, v) \in \mathcal{G}$ is *non-negative* if $v(S) \geq 0$ for all $S \subseteq N$, *monotone* (or *monotonic*) if $v(S) \leq v(T)$ for all $S \subseteq T \subseteq N$, and *zero-monotone* (or *zero-monotonic*), if $v(S) + \sum_{i \in T \setminus S} v(\{i\}) \leq v(T)$ for all $S \subseteq T \subseteq N$. It is *superadditive* if $v(S) + v(T) \leq v(S \cup T)$ for all $S, T \subseteq N$ such that $S \cap T = \emptyset$, and *convex* if $v(S) + v(T) \leq v(S \cup T) + v(S \cap T)$ for all $S, T \subseteq N$.

A special class of TU-games is the class of unanimity games. For each nonempty $T \subseteq N$, the *unanimity game* (N, u^T) is given by the player set N and the characteristic function $u^T(S) = 1$ if $T \subseteq S$ and $u^T(S) = 0$ otherwise. Shapley (1953) has shown that the unanimity games (N, u^T), $T \subseteq N$, form a basis for $\mathcal{G}^N$. Thus, for every $(N, v) \in \mathcal{G}^N$ there exist numbers $\Delta^T(v)$, $T \in 2^N \setminus \{\emptyset\}$, called (Harsanyi) *dividends*,

[1] Given a fixed player set N, a TU-game is sometimes denoted by its characteristic function v.

[2] The subgame (S, v_S) is sometimes written as (S, v).

so that $v = \sum_{T\in 2^N\setminus\{\emptyset\}} \Delta^T(v)u^T$. By definition of the unanimity games, it holds that $v(S) = \sum_{T\subseteq S} \Delta^T(v)$, i.e., the worth $v(S)$ is equal to the dividend of S plus the sum of the dividends of all its proper subcoalitions. The dividend of S can therefore be interpreted as the additional contribution of cooperation among the players in S, that they did not already realize by cooperating in smaller coalitions, see Harsanyi (1963). Using the Möbius transform it follows that

$$\Delta^T(v) = \sum_{S\subseteq T}(-1)^{|T|-|S|}v(S),\ T \in 2^N \setminus \{\emptyset\}.$$

In the sequel, for $S \subseteq N$, $\mathbb{R}^S$ denotes the $|S|$-dimensional Euclidean space with elements $x \in \mathbb{R}^S$ having components x_i, $i \in S$. The vector $\mathbf{0} \in \mathbb{R}^S$ denotes the null-vector with all components equal to zero. A payoff vector for a TU-game (N, v) is a vector $x \in \mathbb{R}^N$, assigning a payoff x_i to every $i \in N$. A payoff vector is *efficient*[3] if $\sum_{i\in N} x_i = v(N)$ and *individually rational* for player $i \in N$ if $x_i \geq v(\{i\})$. Let $\mathcal{F} \subseteq \mathcal{G}$ be a class of TU-games. A (set-valued) *solution* F on $\mathcal{F}$ assigns a set $F(N, v) \subset \mathbb{R}^N$ of payoff vectors to every TU-game $(N, v) \in \mathcal{F}$. A solution F on $\mathcal{F}$ is single-valued if it assigns to every $(N, v) \in \mathcal{F}$ a unique payoff vector. A single-valued solution is also called a *value* and is sometimes denoted by f. Hence, f on $\mathcal{F}$ assigns precisely one payoff vector $f(N, v) \in \mathbb{R}^N$ to every TU-game $(N, v) \in \mathcal{F}$. As mentioned before, Shapley (1953) describes a value as providing for each player in a TU-game an a priori assessment of the utility of becoming involved in a game.

The most applied set-valued solution on $\mathcal{G}$ is the *core* (Gillies, 1959). It assigns to every TU-game $(N, v) \in \mathcal{G}$ the set

$$\text{core}(N, v) = \{x \in \mathbb{R}^N \mid \sum_{i\in N} x_i = v(N),\ \sum_{i\in S} x_i \geq v(S) \text{ for all } S \subset N\},$$

i.e., the set of all efficient payoff vectors that are stable in the sense that no coalition can do better (obtain a larger total payoff) by separating from the grand coalition N. The core of a TU-game is nonempty if and only if the game is balanced, see e.g., Bondareva (1963) or Shapley (1967). Since every convex game is balanced, it follows that the core of a convex game is nonempty, see Shapley (1971).

The best-known single-valued solution on $\mathcal{G}$ is the *Shapley value* (Shapley, 1953). It assigns to every TU-game $(N, v) \in \mathcal{G}$ the payoff vector $Sh(N, v)$ given by

$$Sh_i(N, v) = \sum_{\{S\subseteq N \mid i\in S\}} \frac{(|N| - |S|)!(|S| - 1)!}{|N|}(v(S) - v(S \setminus \{i\})),\ i \in N.$$

In this formula, and in general, $v(S) - v(S\setminus\{i\})$, $i \in S$, is called the *marginal contribution* of player i to coalition S.

There exist various alternative definitions of the Shapley value, we give two. Let $\pi = (i_1, i_2, \ldots, i_n)$ be an ordering of the player set N and let Π^N denote the set of all orderings of N. That is, $\pi \in \Pi^N$ is a one to one mapping of N onto $\{1, \ldots, n\}$. Then,

[3] Formally, the use of this term is only correct if $v(N) \geq \sum_{P_j\in P} v(P_j)$ for every partition P of N.

given the ordering π, the *marginal vector* of the TU-game (N,v) is the payoff vector $m^\pi(N,v)$ given by $m^\pi_{i_k}(N,v) = v(i_1,\ldots,i_k) - v(i_1,\ldots,i_{k-1})$, $k \in N$.[4] Alternatively, given an ordering π from the set Π^N assigning rank number $\pi(i) \in N$ to any player $i \in N$, define $\pi^i = \{j \in N \mid \pi(j) \leq \pi(i)\}$. So, π^i is the set of all players with rank number at most equal to the rank number of player i, including i itself. The marginal vector $m^\pi(N,v)$ of TU-game (N,v) and ordering π is given by $m^\pi_i(N,v) = v(\pi^i) - v(\pi^i \setminus \{i\})$, $i \in N$. The Shapley value now is equal to the average of all marginal vectors of (N,v):

$$Sh_i(N,v) = \frac{1}{|N|!} \sum_{\pi \in \Pi^N} m^\pi_i(N,v),\ i \in N,\ (N,v) \in \mathcal{G}.$$

The Shapley value can also be written as the single-valued solution that distributes the dividends of all coalitions equally among the players in those coalitions:

$$Sh_i(N,v) = \sum_{\{T \subseteq N | i \in T\}} \frac{\Delta^T(v)}{|T|},\ i \in N,\ (N,v) \in \mathcal{G}.$$

TU-games with coalition structure

Aumann and Drèze (1974) were one of the first to consider restrictions on cooperation possibilities of the players in a TU-game by partitioning the set of players in a number of a priori unions (elements of the partition). Nowadays, TU-games with a partition of the set of players are known as TU-games with coalition structure, or TU-games with a priori unions. Let $\mathcal{P}^N$ be the set of partitions of N. Hence, for some $m \leq |N|$, $P = \{P_1,\ldots,P_m\} \in \mathcal{P}^N$ if and only if $\bigcup_{i=1}^m P_i = N$, $\forall k\ P_k \neq \emptyset$ and $\forall l\ P_k \cap P_l = \emptyset$ if $k \neq l$. For a given $P = \{P_1,\ldots,P_m\}$, denote $M = \{1,\ldots,m\}$. Then $P = \{P_j \mid j \in M\}$ is called a *coalition structure*, or a *system of a priori unions*, and any element P_j, $j \in M$, is called a *union* of P. A *TU-game with coalition structure*, or *TU-game with a priori unions*, is a triple (N,v,P) with $(N,v) \in \mathcal{G}^N$ and $P \in \mathcal{P}^N$ a partition of N. The collection of all TU-games with coalition structure is denoted by $\mathcal{CG}$. Given a fixed player set N, the collection of all TU-games with coalition structure on N is denoted by $\mathcal{CG}^N$.

A value f on $\mathcal{CG}$ assigns a unique payoff vector $f(N,v,P) \in \mathbb{R}^N$ to every TU-game with coalition structure $(N,v,P) \in \mathcal{CG}$. To obtain a value for TU-games with coalition structure, Aumann and Drèze (1974) assumed that the players in the game are only allowed to cooperate within their own union. They then applied (among other solutions), for each union P_j, $j \in M$, the Shapley value to the subgame within the union (P_j, v_{P_j}). So, the *Aumann-Drèze value* for TU-games with coalition structure AD is given by

$$AD_i(N,v,P) = Sh_i(P_j, v_{P_j}),\ i \in P_j,\ j \in M.$$

Since the Shapley value provides an efficient payoff vector, the total payoff assigned to the players in P_j according to the Aumann-Drèze value is equal to $v(P_j)$. However, because $\sum_{j \in M} v(P_j)$ does not have to be equal to $v(N)$, the Aumann-Drèze value in general does not provide an efficient payoff vector.

[4] Observe that given an ordering π, a marginal vector itself can be seen as a single-valued solution for a TU-game.

Owen (1977) proposed a different value for TU-games with coalition structure which does provide an efficient payoff vector. He considered the situation in which all the players in the game can cooperate, but a subset of players within a union is only allowed to cooperate with complete other unions. For every $j \in M$ and every $S \subseteq P_j$, let $(M, v_S^j) \in \mathcal{G}^M$ be the TU-game with player set M and characteristic function v_S^j defined by

$$v_S^j(Q) = \begin{cases} v(\bigcup_{k\in Q} P_k), & j \notin Q, \\ v(\bigcup_{k\in Q\setminus\{j\}} P_k \cup S), & j \in Q, \end{cases} \qquad \text{for all } Q \subseteq M.$$

So, for some $j \in M$ and $S \subseteq P_j$, the worth of subset Q of M in TU-game (M, v_S^j) is equal to the worth in TU-game (N, v) of the players in subset Q, except that the players in P_j are replaced by the players in S. Observe that $v^P := v_{P_j}^j = v_{P_k}^k$ for all $j, k \in M$, which gives the *quotient game* (M, v^P) of Owen (1977). The *Owen value* for TU-games with coalition structure can now be obtained from a two-step procedure in which the Shapley value is applied twice. First, let (P_j, v^j) be the TU-game with player set P_j and characteristic function $v^j(S) = Sh_j(M, v_S^j)$, $S \subseteq P_j$. In this game the worth of coalition $S \subseteq P_j$ is the Shapley value payoff of player $j \in M$ in the TU-game (M, v_S^j). Next, the Owen value Ow can be defined as

$$Ow_i(N, v, P) = Sh_i(P_j, v^j), \; i \in P_j, \; j \in M.$$

Since the Shapley value provides an efficient payoff vector, it holds that

$$\sum_{j\in M}\sum_{i\in P_j} Ow_i(N, v, P) = \sum_{j\in M}\sum_{i\in P_j} Sh_i(P_j, v^j) = \sum_{j\in M} v^j(P_j) = \sum_{j\in M} Sh_j(M, v_{P_j}^j) = v(N).$$

This implies that the Owen value indeed provides an efficient payoff vector.

Recently, Kamijo (2011) has introduced another value for TU-games with coalition structure. The main difference between this value and the Owen value is that unions are only allowed to cooperate when all players in the unions agree. This means that in a TU-game with coalition structure individual players can cooperate within their union, and complete unions can cooperate (when complete unions cooperate, they force all their constituent players to cooperate), but proper subsets of different unions cannot cooperate. Thus, given a partition $P \in \mathcal{P}^N$, players in any coalition $S \subseteq P_j \in P$ can cooperate with each other and obtain the worth of the coalition $v(S)$. In addition, there is the possibility of cooperation among players in different unions, but only if all players in these unions agree. Let $S \subset P_j \in P$ and $P_k \in P$, $P_k \neq P_j$. Then P_j and P_k can obtain their worth $v(P_j \cup P_k)$ when they decide to cooperate. However, S and P_k can obtain only $v(S) + v(P_k)$ because all players in P_j and P_k are necessary in establishing cooperation between these unions. Given $P = \{P_j \mid j \in M\} \in \mathcal{P}^N$, for all $S \subseteq N$, $S \neq \emptyset$, denote $S/P = \{S \cap P_k \mid S \cap P_k \neq P_k\} \cup \{\bigcup_{\{k \mid S\cap P_k = P_k\}} P_k\}$. That is, S/P consists of the sets $S \cap P_k$ for every k with $S \cap P_k \neq P_k$ and the single set being the union of all sets P_k that are contained in S. Notice that $N/P = \{N\}$. Given a TU-game $(N, v) \in \mathcal{G}^N$ and a partition $P \in \mathcal{P}^N$ the corresponding game induced by coalition structure P now is the

TU-game $(N, v|_P)$ with player set N and characteristic function

$$v|_P(S) = \sum_{T \in S/P} v(T), \quad \text{for all} \quad S \subseteq N.$$

We call this game the *partition restricted game*. In a partition restricted game the worth of an arbitrary coalition of players is equal to the worth of the union of all complete unions within the coalition, plus the sum of the worths of all remaining parts of the coalition that are not complete unions. The value for TU-games with coalition structure proposed by Kamijo (2011), called the *collective value* and denoted by Ka, is equal to the Shapley value of the corresponding partition restricted game. Thus, for every $(N, v, P) \in \mathcal{CG}$ the collective value is given by

$$Ka(N, v, P) = Sh(N, v|_P).$$

Since the Shapley value provides an efficient payoff vector, it holds that

$$\sum_{i \in N} Ka_i(N, v, P) = \sum_{i \in N} Sh_i(N, v|_P) = v|_P(N) = \sum_{T \in N/P} v(T) = \sum_{T \in \{N\}} v(T) = v(N).$$

This implies that the collective value also provides an efficient payoff vector.

TU-games with graph structure

A different form of restrictions on TU-games was considered by Myerson (1977). In his model the restrictions in the game are not given by a partition of the set of players, but by the links in an undirected (communication) graph. An undirected *graph* is a pair (N, L) where N is a set of *nodes* and $L \subseteq \{\{i, j\} \mid i, j \in N, i \neq j\}$ is a set of unordered pairs of distinct elements of N.[5] In this dissertation the nodes in a graph will represent the players in a game, we therefore refer to the nodes as players. Two players $i, j \in N$ are called *neighbors* in (N, L) if $\{i, j\} \in L$. The elements of L are called *links* or *edges*. We denote the collection of all undirected graphs on N by $\mathcal{L}^N$.

For $S \subseteq N$, the graph $(S, L(S))$ with $L(S) = \{\{i, j\} \in L \mid i, j \in S\}$ is called the *subgraph* of L on S. Given $(N, L) \in \mathcal{L}^N$, a sequence of k different players $(i_1, \ldots, i_k)$ is a *path* in $(S, L(S))$ if $\{i_l, i_{l+1}\} \in L(S)$ for $l \in \{1, \ldots, k-1\}$. A path $(i_1, \ldots, i_k)$ with $k \geq 3$ is called a *cycle* in $(S, L(S))$ if $\{i_k, i_1\} \in L(S)$. A graph (N, L) is *cycle-free* if it does not contain any cycle. Two different players $i, j \in S$ are called *connected* in $(S, L(S))$ if there exists a path $(i_1, \ldots, i_k)$ in $(S, L(S))$ with $i_1 = i$ and $i_k = j$. A graph (N, L) is *connected* if any two different players $i, j \in N$ are connected in (N, L). When a graph is both cycle-free and connected it is called a *tree*. We denote the collection of all trees on N by $\mathcal{L}_T^N$.

Given a graph $(N, L) \in \mathcal{L}^N$, a set of players $S \subseteq N$ is said to be *connected* (in (N, L)) when the subgraph $(S, L(S))$ is connected; such a coalition is called a *connected coalition*. A set of players $K \subseteq N$ is a *component* of (N, L) if and only if (1) K is connected in

[5] If there is no possible confusion about the player set, we sometimes denote a graph (N, L) just by its set of links L.

(N, L), and (2) $K \cup \{i\}$ is not connected in (N, L) for every $i \in N \setminus K$. The set of components of $(S, L(S))$ is denoted by $C^L(S)$. Note that every player in S that is not linked with any other player in S is a (singleton) component in $(S, L(S))$.

A *TU-game with graph structure*, also known as a *TU-game with communication structure* or *graph game*, is a triple (N, v, L) with $(N, v) \in \mathcal{G}^N$ and $(N, L) \in \mathcal{L}^N$. The collection of all TU-games with graph structure is denoted by $\mathcal{GG}$. Given a fixed player set N, the collection of all TU-games with graph structure on N is denoted by $\mathcal{GG}^N$.

A value f on $\mathcal{GG}$ assigns a unique payoff vector $f(N, v, L) \in \mathbb{R}^N$ to every TU-game with graph structure $(N, v, L) \in \mathcal{GG}$. Myerson (1977) assumed that the players in a TU-game with graph structure are only allowed to cooperate when they are connected in a graph, thus, when there exists a set of links in the graph that connects the cooperating players. A graph therefore provides a cooperation structure on the set of players, that restricts the cooperation possibilities of the players in a cooperative game. Under the assumption of Myerson (1977), a coalition S is only able to realize its worth $v(S)$ if S is connected in (N, L). When S is not connected in (N, L), the players in S can realize the sum of the worths of the components of the subgraph $(S, L(S))$. Given a TU-game $(N, v) \in \mathcal{G}^N$ and a graph $L \in \mathcal{L}^N$, the corresponding *graph restricted game* (Myerson, 1977) induced by graph L, is the TU-game (N, v^L) with player set N and characteristic function

$$v^L(S) = \sum_{T \in C^L(S)} v(T) \text{ for all } S \subseteq N.$$

The *Myerson value* (Myerson, 1977) of a TU-game with graph structure, denoted by My, is defined as the Shapley value of the corresponding graph restricted game (N, v^L). That is, for every TU-game with graph structure $(N, v, L) \in \mathcal{GG}$ the Myerson value is defined as

$$My(N, v, L) = Sh(N, v^L).$$

An alternative value for TU-games with graph structure is the average tree solution. This solution was originally introduced in Herings, van der Laan and Talman (2008) on the class of cycle-free graph games. A cycle-free graph game is just a TU-game with graph structure in which the graph is cycle-free. The rationale for the average tree solution comes from adapting the 'fairness' axiom that Myerson (1977) used to characterize the Myerson value. Fairness says that deleting a link between two players in a graph (game) should yield both players the same change in payoff. Herings, van der Laan, Talman (2008) replace Myerson's fairness by an alternative fairness property, called 'component fairness'. Component fairness says that deleting a link between two players in a cycle-free graph (game) should yield for both resulting components (subgraphs remaining after deleting the link) the same average change in payoff, with the average taken over the players in the component. Hence, the loss associated with deleting a link in a graph (game) is attributed to the two resulting components proportional to the size of the component, rather than to the two individual players whose link is deleted. Depending on whether one considers the 'fairness' or 'component fairness' axiom (in a particular application) one can combine it with an axiom known as 'component efficiency' to obtain a characterization of the

Myerson value, and the average tree solution respectively, on the class of cycle-free graph games.

To introduce the average tree solution on the class of cycle-free graph games, we first define the concept of a directed graph. A *directed graph* is a pair (N, D) where N is a set of nodes (players) and $D \subseteq \{(i,j) \in N \times N \mid i \neq j\}$ is a set of ordered pairs of distinct elements of N. The elements of D are called *directed links* or *arcs*. If $(i,j) \in D$, then player j is called a *successor* of player i and player i is called a *predecessor* of player j. Given a directed graph (N, D), a sequence of $k \geq 2$ different players $(i_1, \dots, i_k)$ is a *directed path* in (N, D) if $(i_\ell, i_{\ell+1}) \in D$ for $\ell \in \{1, \dots, k-1\}$. A sequence of players $(i_1, \dots, i_{k+1})$ is called a *directed cycle* in (N, D) if (1) $(i_1, \dots, i_k)$ is a directed path, (2) $(i_k, i_{k+1}) \in D$ and (3) $i_{k+1} = i_1$. A directed graph (N, D) is *acyclic* when it does not contain any directed cycle. We say that $j \neq i$ is a *subordinate* of i (and i is a *superior* of j) in (N, D) if there exists a directed path in (N, D) from i to j. A directed graph (N, D) is called a *rooted tree* if there is one player $i \in N$, called the *root*, that has no predecessors in (N, D) and there exists a unique directed path from this root to any other player in (N, D). Let $(N, \widehat{D})$ be the undirected graph induced by (N, D), i.e., $\widehat{D} = \{\{i,j\} \mid (i,j) \in D\}$. We call (N, D) connected when it induces a connected undirected graph $(N, \widehat{D})$. Observe that a rooted tree is connected. Given a tree $(N, L) \in \mathcal{L}_T^N$ and a player $i \in N$, let (N, L^i) be the unique rooted tree with root i induced by the tree (N, L) and the player i, i.e., $L^i = \{(j,k) \mid \{j,k\} \in L, j \text{ is on the path from } i \text{ to } k\}$. Further, let S_j^i denote the set of successors of j in the rooted tree (N, L^i) and let $\widehat{S}_j^i$ denote the set containing agent j itself and all its subordinates in the rooted tree (N, L^i).

A tree game is a TU-game with graph structure in which the graph (N, L) is tree. For a tree game (N, v, L) with $(N, v) \in \mathcal{G}^N$ and $(N, L) \in \mathcal{L}_T^N$, Demange (2004) defines for every player $i \in N$ the hierarchical outcome $h^i(N, v, L)$ as the marginal vector $m^\pi(N, v^L)$ of the graph restricted game that is obtained for an ordering of the players π that is consistent with the rooted tree (N, L^i), i.e., if the path from i to g in (N, L) contains j then $\pi(g) < \pi(j)$. In other words, it is the marginal vector of the TU-game (N, v^L) corresponding to any ordering where the players enter 'from the bottom up' relative to player i. More formally, the hierarchical outcome $h^i(N, v, L)$, $i \in N$, of a tree game (N, v, L) is given by

$$h_j^i(N, v, L) = v(\widehat{S}_j^i) - \sum_{k \in S_j^i} v(\widehat{S}_k^i), \ j \in N.$$

The *average tree solution* on the class of tree games AT assigns to every tree game (N, v, L) the average of the $|N|$ hierarchical outcomes $h^i(N, v, L)$, $i \in N$. So, on the class of tree games it holds that

$$AT(N, v, L) = \frac{1}{|N|} \sum_{i \in N} h^i(N, v, L).$$

This definition of the average tree solution for tree games can be extended to the average tree solution for cycle-free graph games of Herings, van der Laan and Talman (2008) by determining the average payoff of a player only over the hierarchical outcomes in its own component.

A generalization of the average tree solution to arbitrary graph games was suggested by Herings, van der Laan, Talman and Yang (2010). Given a graph $(N, L) \in \mathcal{L}^N$, an n-tuple $H = (H_1, \dots, H_n)$ of n subsets of N is *admissible* if it satisfies (1) for all $i \in N$, $i \in H_i$, and for some $j \in N$, $H_j = N$ and (2) for all $i \in N$ and $K \in C^L(H_i \setminus \{i\})$ it holds that $K = H_j$ and $\{i, j\} \in L$ for some $j \in N$. Given a graph $L \in \mathcal{L}^N$, we denote the collection of all admissible n-tuples H by $\mathcal{H}^L$. Given $H \in \mathcal{H}^L$, the vector $m^H(N, v, L)$ of a game with graph structure (N, v, L) is the vector of payoffs given by $m_i^H(N, v, L) = v(H_i) - \sum_{K \in C^L(H_i \setminus \{i\})} v(K)$, $i \in N$. For every TU-game with graph structure $(N, v, L) \in \mathcal{GG}$ the average tree solution assigns the payoff vector $AT(N, v, L)$ given by

$$AT(N, v, L) = \frac{1}{|\mathcal{H}^L|} \sum_{H \in \mathcal{H}^L} m^H(N, v, L).$$

TU-games with coalition and graph structure

Vázquez-Brage, García-Jurado and Carreras (1996) combined the ideas of Aumann and Drèze (1974) and Myerson (1977) in TU-games with coalition and graph structure. A *TU-game with coalition and graph structure*, or *TU-game with graph restricted communication and a priori unions*, is a quadruple (N, v, L, P) with player set N, characteristic function v, $L \in \mathcal{L}^N$ a graph on N and $P \in \mathcal{P}^N$ a partition of N. The collection of all TU-games with coalition and graph structure is denoted by $\mathcal{CGG}$. Given a fixed player set N, the collection of all TU-games with coalition and graph structure on N is denoted by $\mathcal{CGG}^N$.

A value f on $\mathcal{CGG}$ assigns a unique payoff vector $f(N, v, L, P) \in \mathbb{R}^N$ to every TU-game with coalition and graph structure $(N, v, L, P) \in \mathcal{CGG}$. As a value for TU-games with coalition and graph structure Vázquez-Brage, García-Jurado and Carreras (1996) proposed the Owen value of the graph restricted game, which we denote by VGC:

$$VGC(N, v, L, P) = Ow(N, v^L, P).$$

Alonso-Meijide, Álvarez-Mozos and Fiestras-Janeiro (2009) have suggested two other values for TU-games with coalition and graph structure. They applied Banzhaf (1965) type modifications of the Owen value to the graph restricted game.

PFF-games

Given a set of players N, let $E = \{(S, P) \mid S \in P \in \mathcal{P}^N\}$ be the set of embedded coalitions. An *embedded coalition* consists of a coalition $S \in P$ and a specification of how the (other) players from N are aligned into coalitions in P. We define $\perp = \{(\emptyset, \{\{i\} \mid i \in N\})\}$ and $E_\perp = E \cup \perp$. Then, a *cooperative game with transferable utility in partition function form*, or *PFF-game*, is a pair (N, w), where N is the set of players and $w : E_\perp \to \mathbb{R}$ is a *partition function* on N such that $w(\perp) = 0$.[6] PFF-games were introduced by Thrall

[6] Given a fixed player set N, we sometimes denote a PFF-game by its partition function w.

and Lucas (1963) to allow for externalities in coalition formation.[7] For any embedded coalition (S, P), $w(S, P)$ displays the worth of that embedded coalition. The worth of an embedded coalition can be interpreted as the wealth, measured in units of transferable utility, which the members of that coalition are able to divide among themselves when all the players in N are aligned into the coalitions (elements) of the partition P. The collection of all PFF-games is denoted by $\mathcal{PG}$. Given a fixed player set N, the collection of all PFF-games on N is denoted by $\mathcal{PG}^N$. A value f on $\mathcal{PG}$ assigns a unique payoff vector $f(N, w) \in \mathbb{R}^N$ to every PFF-game $(N, w) \in \mathcal{PG}$.

2.2 River water allocation problems

As mentioned in the introduction, the problem of sharing water among agents located along a river is the central problem of this dissertation. We primarily concentrate on the case in which the agents located along the river are countries, but most results also apply when the agents are, for instance, states, cities or firms. The models outlined in this section, and those in Chapters 3, 4, 5 and 6, all deal with successive rivers and consumptive use of the river water by the agents. We discuss models of non-consumptive use in Chapter 7.

River water allocation problems have been analyzed using different methods. For two early contributions to the problem see, for instance, Rogers (1969), using a systems analysis technique, and Burness and Quirk (1979), using a stochastic model. Barrett (1994) and Kilgour and Dinar (1995) were some of the first to study the river water sharing problem with the help of cooperative game theory.

Barrett (1994) argues that implementing the principle of equitable utilization does not maximize the total benefits of the countries in a river basin. Using concepts from cooperative game theory, he shows that a basin wide approach to the river water allocation problem could lead to an outcome that is both efficient and equitable. Kilgour and Dinar (1995) focus on the dynamic and stochastic aspects of the river water sharing problem. They investigate whether stable water-sharing agreements in international river basins are possible. The stability concept that Kilgour and Dinar (1995) use is a (cooperative) game-theoretic one.

What characterizes the papers of Barrett (1994) and Kilgour and Dinar (1995) is that in modeling water allocation problems the authors combine principles from international watercourse law, as discussed in Chapter 1, with concepts from cooperative game theory, as discussed in Section 2.1. More specifically, they try to answer the (normative) question of how benefits of cooperation have to be distributed among the agents located along a river by referring to established water law principles. The main advantage of this approach is that one does not have to rely on one's own value judgments about the fairness of a particular division of water (benefit) among countries, but can fall back on consensus

[7]Despite its early introduction, the cooperative game in partition function form turned out to be a methodologically tough game. Only during the last decade some progress on the subject has been reported, see e.g., Funaki and Yamato (1999), Ray and Vohra (1999), Albizuri, Arin and Rubio (2005), Gomez (2005), Macho-Stadler, Pérez-Castrillo and Wettstein (2007), De Clippel and Serrano (2008) and Dutta, Ehlers and Kar (2010).

views from a substantial legal literature.

Ambec and Sprumont (2002) apply the methodology of Barrett (1994) and Kilgour and Dinar (1995) when they introduce a static, deterministic river water allocation model. Ambec and Sprumont (2002) ignore the dynamic and stochastic nature of the problem of water allocation because they want to focus on (welfare) distribution issues. Given their model, the aim of Ambec and Sprumont (2002) is to find an 'efficient' and 'fair' allocation of the river water and the welfare that the consumption of this water creates. In the model an allocation of the river water among the agents is efficient when it maximizes the total utility of the agents, which is equal to the sum of their benefits of water consumption. To sustain an efficient water allocation, the agents can compensate each other by paying monetary transfers. Every water allocation and schedule of monetary compensations between the agents yields a welfare distribution, where the utility of an agent is equal to its benefit from water consumption plus its monetary transfer, which can be negative. By deriving a cooperative game from their model, Ambec and Sprumont (2002) find out how the river water can be allocated efficiently over the agents and propose monetary transfers that can be performed in order to realize a fair welfare distribution. Ambec and Sprumont (2002) base their idea of what is fair on two principles from international watercourse law. By translating these principles into two properties for their cooperative game they come up with the downstream incremental solution that satisfies both core lower bounds as well as aspiration upper bounds.

The model of Ambec and Sprumont (2002) has been extended and generalized by several authors. Since most of the work in this dissertation can also be seen as extensions to the Ambec and Sprumont (2002) model, the aim of this section is to present the Ambec and Sprumont (2002) model in more detail and discuss some of the work that has been done by other authors on this model.

River benefit problems and river games

Consider a single-stream river and let $N = \{1, \ldots, n\}$ be a set of *agents* (e.g., countries) located at different points along the river, numbered successively from upstream to downstream. Let $e_i \geq 0$, $i \in N$, be the *inflow* of water in the river, in the form of tributary streams or precipitation, on the territory of agent i. Every agent, $i \in N$, can extract water from the river for (rival) consumption and has a quasi-linear utility function given by $u^i(x_i, t_i) = b_i(x_i) + t_i$, where $x_i \geq 0$ is the amount of water allocated to (consumed by) agent i, $b_i : \mathbb{R}_+ \to \mathbb{R}$ is a function yielding the benefit $b_i(x_i)$, and t_i is a monetary compensation (transfer) to agent i.[8] We call the triple (N, e, b), where N is the set of agents, $e = (e_i)_{i \in N}$ is the vector of inflows and $b = (b_i)_{i \in N}$ is the vector of benefit functions, a *river benefit problem*.

Because of the unidirectionality of the water flow in the river, from upstream to downstream, every agent in a river benefit problem can be assigned at most the water inflow on the territories of itself and its upstream agents. That is, the water inflow downstream of some agent can not be allocated to this agent. A *water allocation* $x \in \mathbb{R}^N_+$ assigns an

[8]When $t_i > 0$ agent i receives a monetary compensation, when $t_i < 0$ agent i pays a monetary compensation.

amount of water x_i, $i \in N$, to agent i under the constraints

$$\sum_{i=1}^{j} x_i \leq \sum_{i=1}^{j} e_i,\ j \in \{1, \dots, n\},$$

i.e., $x \in \mathbb{R}_+^N$ is a water allocation if, for every agent j, the sum of the water assignments $x_1, \dots, x_j$ is at most equal to the sum of the inflows $e_1, \dots, e_j$. A water allocation x yields *total welfare* $\sum_{i=1}^{n} b_i(x_i)$.

Apart from consuming an amount of water x_i, agents are also able to make monetary transfers to each other. Money is assumed to be available in unbounded quantities to perform such side-payments. Since the agents are not allowed to create welfare by simply transferring money, a *compensation scheme* $t \in \mathbb{R}^N$ gives a monetary compensation t_i to agent i, $i \in N$, under the restriction

$$\sum_{i=1}^{n} t_i \leq 0.$$

So, the sum of all positive compensations (agents that receive money) is at most equal to the absolute value of the sum of all negative compensations (agents that have to pay money).

A *welfare distribution* is a pair (x, t) of a water allocation x and a compensation scheme t yielding utility $u^i(x_i, t_i) = b_i(x_i) + t_i$ to every agent $i \in N$. A welfare distribution (x, t) is *Pareto efficient* if there does not exist another welfare distribution (x', t') such that $u^i(x'_i, t'_i) \geq u^i(x_i, t_i)$ for all $i \in N$ with at least one strict inequality. This is the case if the water is distributed efficiently and no money is wasted. So, welfare distribution (y, t) is Pareto efficient if and only if $y \in \mathbb{R}_+^N$ solves the welfare maximization problem

$$\max_{x_1,\dots,x_n} \sum_{i=1}^{n} b_i(x_i) \text{ s.t. } \sum_{i=1}^{j} x_i \leq \sum_{i=1}^{j} e_i,\ j \in \{1, \dots, n\}, \text{ and } x_i \geq 0,\ i \in N, \tag{2.1}$$

and the compensation scheme is *budget balanced*:

$$\sum_{i=1}^{n} t_i = 0.$$

Let x^* be a solution of problem (2.1). Then a Pareto efficient welfare distribution (x^*, t) yields *payoffs* (utilities)

$$z_i = b_i(x_i^*) + t_i,\ i \in N,$$

with the sum of the payoffs equal to the Pareto efficient total welfare $\sum_{i=1}^{n} b_i(x_i^*)$.

Ambec and Sprumont (2002) make the following assumption on the benefit functions of the agents in a river benefit problem.

Assumption 2.2.1 *Every benefit function $b_i : \mathbb{R}_+ \to \mathbb{R}$, $i \in N$, is a strictly increasing and strictly concave function, which is differentiable for $x_i > 0$ with derivative going to infinity as x_i tends to zero.*

The assumption of strictly increasing benefit functions means that agents always want to consume more water, which creates water scarcity. The assumption of strictly concave benefit functions implies that as agents consume more water, they desire additional water less (the marginal benefits of water consumption decrease as the agents consume more).

Given Assumption 2.2.1, Ambec and Sprumont (2002) show that the efficient water allocation x^* is unique. Note also that under the assumption the presence of money in the model is crucial. If there would be no money, each agent would simply consume all the water from the river that is available at its location (since the benefit functions are strictly increasing). Because there is money though, and because the agents value it, it is possible that an upstream agent does not consume all the water at its location and lets some of it flow to its downstream neighbor in exchange for a monetary transfer. This 'trade' in river water can take place through the signing of contracts between the agents.

Ambec and Sprumont (2002) study the distribution of the Pareto efficient total welfare by deriving a TU-game (N, v) from the river benefit problem (N, e, b). The player set in this TU-game is given by the set of agents along the river $N = \{1, \ldots, n\}$. The characteristic function v is defined as follows. The worth $v(N)$ is equal to the Pareto efficient total welfare, i.e., $v(N) = \sum_{i=1}^{n} b_i(x_i^*)$. Further, for any pair of agents $i, j \in N$ with $j > i$, it holds that the water inflow entering the river before the upstream agent i can only be allocated to the downstream agent j if every agent k between agents i and j cooperates. Otherwise, since every benefit function b_k is strictly increasing in x_k, every agent between i and j can increase its utility by confiscating the water flow from i to j. Hence, a coalition T in the TU-game (N, v) is admissible if and only if T is *consecutive*, i.e., $T = \{i, i+1, \ldots, j\}$ for some $i, j \in N,\ i \leq j$. In the sequel, we denote such a coalition of consecutive agents by $[i, j]$. For any consecutive coalition $[i, j]$ the worth $v([i, j])$ is given by

$$v([i,j]) = \sum_{k=i}^{j} b_k(x_k^{[i,j]}) \quad \text{where} \quad x^{[i,j]} = (x_k^{[i,j]})_{k=i}^{j} \quad \text{solves}$$

$$\max_{x_i,\ldots,x_j} \quad \sum_{k=i}^{j} b_k(x_k) \quad \text{s.t.}$$

$$\sum_{g=i}^{\ell} x_g \leq \sum_{g=i}^{\ell} e_g, \ \ \ell \in [i,j], \ \text{ and } x_g \geq 0, \ \ g \in [i,j]. \tag{2.2}$$

So, the worth of a consecutive coalition is obtained by solving a similar maximization problem as for the grand coalition N in (2.1), but restricted to the water inflows and benefits of agents in the consecutive coalition. For any other (non-consecutive) coalition S, the worth $v(S)$ is equal to the sum of the worths of its *maximal consecutive* subsets, where a subset T of S is maximal consecutive if T is consecutive and $T \cup \{h\}$ is not consecutive for any $h \in S \setminus T$. This concludes the definition of the characteristic function v.

For benefit functions satisfying Assumption 2.2.1, we refer to the TU-game (N, v) described above as the *river game* and denote the collection of all river games on N by $\mathcal{R}^N$. It has been shown in Ambec and Sprumont (2000) that every river game is convex.

To find a distribution of the Pareto efficient total welfare $v(N) = \sum_{i=1}^{n} b_i(x_i^*)$ in a river benefit problem (N, e, b), it is now possible to apply any efficient solution from cooperative game theory to the corresponding river game (N, v). Notice that a single-valued solution f assigns payoff vector $z = f(v) \in \mathbb{R}^N$ to game $v \in \mathcal{R}^N$, which can be implemented by the welfare distribution (x^*, t) with $t_i = z_i - b_i(x_i^*)$, $i \in N$. The 'fairness' of such a distribution depends on the properties of the solution.

Solutions for river games

Solutions for river games have been proposed by Ambec and Sprumont (2002), Herings, van der Laan and Talman (2007), van den Brink, van der Laan and Vasil'ev (2007) and Wang (2011).

Ambec and Sprumont (2000, 2002) propose the so-called *downstream incremental solution* for river games. They introduce and axiomatize this solution by using axioms that are based on the ATS and UTI principles from international watercourse law (see Chapter 1). Ambec and Sprumont (2002) argue that the ATS principle implies that every agent along a river is allowed to use the water that it controls as it pleases. Since this holds for each individual agent along the river, they reason that it also holds for all coalitions of agents along the river. As the water that a coalition of agents controls is determined by its location along the river, the welfare that such a coalition can secure for itself (in a river game) is also determined by this location. Hence, Ambec and Sprumont (2002) require the following property to hold for a solution to river games.

Axiom 2.2.2 Core lower bounds
A solution f on the class of river games $\mathcal{R}^N$ satisfies the core lower bounds if for any $(N, v) \in \mathcal{R}^N$ it holds that $\sum_{i \in S} f_i(v) \geq v(S)$ for all $S \subseteq N$.

While the translation of the ATS principle into an axiom for river games is relatively straightforward, the interpretation of the UTI principle is more difficult. According to Ambec and Sprumont (2002) the UTI principle implies that each agent is allowed to consume all of the water that is originating upstream of its location and therefore has a legitimate claim to the welfare level corresponding to this level of water consumption. It is clear, as also explained in Chapter 1, that such welfare claims can be incompatible because the water originating upstream can only be consumed by one agent. For this reason Ambec and Sprumont (2002) do not consider the UTI welfare levels as lower bounds on the welfare that agents along the river can claim, but as aspiration upper bounds. Thus, in a river benefit problem, every agent along the river is maximally allowed to claim the welfare that it could achieve on it own when it would have the full stream of water originating upstream of its location available. Since this holds for each individual agent along the river, Ambec and Sprumont (2002) reason that it also should hold for all coalitions of agents along the river. Hence, the aspiration upper bound for a coalition of agents S is the welfare level that the agents in the coalition can obtain when they can also use the water inflows of the agents not in S, but upstream to the most downstream member of S. Given $i \in N$, let $UP^i = \{j \in N \mid j \leq i\}$ denote the set of all agents upstream of agent

i, including i itself. Then this discussion can be summarized in the next property for a solution to river games.

Axiom 2.2.3 Aspiration upper bounds
A solution f on the class of river games $\mathcal{R}^N$ satisfies the aspiration upper bounds if for any $(N,v) \in \mathcal{R}^N$ it holds that $\sum_{i\in S} f_i(v) \le \sum_{i\in S} b_i(\tilde{x}_i^S)$ for all $S \subseteq N$, where $\tilde{x}^S = (\tilde{x}_i^S)_{i\in S}$ solves

$$\max_{\{x_i|i\in S\}} \sum_{i\in S} b_i(x_i) \quad s.t. \quad \sum_{k\in UP^j\cap S} x_k \le \sum_{k\in UP^j} e_k \text{ for every } j \in S \text{ and } x_k \ge 0,\ k \in S.$$

While there is nothing wrong with the aspiration upper bound property in itself (as a normative property), we do not agree with Ambec and Sprumont (2002) that it provides an interpretation of the UTI principle. The UTI principle clearly implies lower bound type restrictions on the water consumptions of the agents along a river. These restrictions, in turn, can only imply lower bound type (not upper bound) restrictions on the welfare levels that the agents in a river benefit problem or river game can demand. As an alternative, we therefore provide a number of different properties in Chapters 3 and 4 that are based on the UTI principle and other water distribution principles from international watercourse law.

When applying the aspiration upper bounds axiom to the upstream coalition $[1,j]$ of consecutive agents from 1 to j, it requires that the solution of a river game gives a total payoff to coalition $[1,j]$ that is at most equal to the aspiration upper bound $v([1,j])$. On the other hand, the core lower bounds property requires that coalition $[1,j]$ receives at least $v([1,j])$. Therefore, for every upstream coalition $[1,j]$, $j \in N$, the core lower bounds and aspiration upper bounds together imply that the total payoff that the agents in coalition $[1,j]$ receive should be equal to $v([1,j])$. This uniquely determines the *downstream incremental solution* for river games f^d that is given by

$$f_1^d(v) = v(\{1\}) \text{ and } f_i^d(v) = v([1,i]) - v([1,i-1]),\ i \in [2,n].$$

The downstream incremental solution assigns to every agent along the river its contribution to the welfare when it enters the coalition consisting of its upstream agents. It thus can be seen as the marginal vector of the river game, according to ordering $\pi(i) = i$, $i \in N$ (that is, the ordering in which the agents are ordered from upstream to downstream).

Since under Assumption 2.2.1 every river game is convex, and because every marginal vector of a convex game is in the core of the game (see Shapley, 1971), it follows that the downstream incremental solution is in the core and thus satisfies all core lower bounds, not only the ones for upstream coalitions $[1,j]$, $j \in N$. Ambec and Sprumont (2002) show that the downstream incremental solution also satisfies all aspiration upper bounds.

Theorem 2.2.4 (Ambec and Sprumont, 2002)
A solution f on the class of river games $\mathcal{R}^N$ satisfies the core lower bounds and the aspiration upper bounds if and only if it is the downstream incremental solution f^d.

The downstream incremental solution has the property that for every $i \in N \setminus \{n\}$, the total payoff to the agents in the consecutive coalition $[1,i]$ upstream of i (including i

itself) is equal to $v([1,i])$, while the total payoff to the agents in the downstream coalition $[i+1,n]$ is equal to $v(N) - v([1,i]) \geq v([i+1,n])$. This means that all additional welfare that is created when the two coalitions $[1,i]$ and $[i+1,n]$ merge to form the grand coalition N is attributed to the downstream coalition $[i+1,n]$. However, any upstream coalition $[1,i]$ can prevent that coalition $[i+1,n]$ receives a higher welfare than $v([i+1,n])$ by using all its inflows $e_1, \ldots, e_i$ itself. Herings, van der Laan and Talman (2007) and van den Brink, van der Laan and Vasil'ev (2007) argue that a coalition $[1,i]$ can play some type of ultimatum game by claiming that it will use its total water inflow $\sum_{k=1}^{i} e_k$ by itself, unless the agents of the downstream coalition $[i+1,n]$ agree to make a monetary compensation almost equal to the total welfare gain of cooperation. This results in precisely the opposite of the solution proposed by Ambec and Sprumont (2002), namely the *upstream incremental solution* for river games f^u. The upstream incremental solution is given by

$$f^u_n(v) = v(\{n\}) \text{ and } f^u_i(v) = v([i,n]) - v([i+1,n]),\ i \in [1, n-1].$$

This upstream incremental solution also can be obtained as a marginal vector of the river game, but now according to ordering $\pi(i) = n - i + 1$, $i \in N$ (that is, the ordering in which the players are ordered from downstream to upstream). Although the upstream incremental solution does not satisfy the aspiration upper bounds, it does satisfy the core lower bounds (because it is a marginal vector of the river game and every marginal vector of a convex game is in the core).

Wang (2011) proposes a welfare distribution for river benefit problems based on the idea that only consecutive pairs of agents along the river are allowed to trade water. He shows that, under Assumption 2.2.1, the unique efficient water allocation x^* can be obtained through a process of downstream bilateral trading, and that the payoffs that result from this process are in the core of the corresponding river game. In the solution of Wang (2011) any coalition of upstream agents receives a payoff that is in between that of the downstream incremental solution and the upstream incremental solution.

The downstream bilateral trading process sequentially considers the agents along the river, starting with the most upstream agent 1. As soon as it holds that $b'_i(x_i) < b'_{i+1}(x_{i+1})$ for any $i \in [1, n-1]$ (the welfare distribution at the start of the trading process is $(x,t) = (e, \mathbf{0})$), then agent i and agent $i+1$ are forced to trade the amount of water that will ensure that $b'_i(x_i) = b'_{i+1}(x_{i+1})$. Agent i gives up the required amount of water to agent $i+1$ and receives a monetary compensation paid by agent $i+1$ equal to this amount of water times the marginal benefit $b'_i(x_i) = b'_{i+1}(x_{i+1})$. This results in a new water allocation and compensation scheme. The process continues by following the same procedure, again starting by considering agent 1. In his main theorem Wang (2011) shows that the downstream bilateral trading process converges to the efficient water allocation x^* and a unique compensation scheme t^*. The river game solution corresponding to the resulting Pareto efficient welfare distribution satisfies the core lower bounds.

River games with externalities

Ambec and Ehlers (2008) generalize the single-stream river game of Ambec and Sprumont (2002) by weakening the assumption on the benefit functions of the agents located along the river. Instead of Assumption 2.2.1, Ambec and Ehlers (2008) make the following assumption on the benefit functions of the agents in a river benefit problem.

Assumption 2.2.5 *Every benefit function $b_i : \mathbb{R}_+ \to \mathbb{R}$, $i \in N$, is a strictly concave function, which is differentiable for $x_i > 0$ with derivative going to infinity as x_i tends to zero.*

This assumption implies that either b_i is strictly increasing, or there exists a *satiation point* $\hat{x}_i > 0$ such that $b_i'(\hat{x}_i) = 0$. When the benefit function of an agent $j \in N$ has a satiation point, b_j is strictly increasing on $x_j < \hat{x}_j$ and strictly decreasing on $x_j > \hat{x}_j$. If, in this latter case, agent j consumes more water than $\hat{x}_j$ it incurs a loss from overconsumption (for instance because of flooding). This would mean that $\hat{x}_j$ is agent j's optimal water consumption.

The existence of satiation points has serious consequences for the resulting game. Before, under Assumption 2.2.1, only consecutive coalitions of agents were able to cooperate because any river water transferred from an upstream part of a non-consecutive coalition to a downstream part would fully be consumed by any 'in-between' agents. So, a non-consecutive coalition S consisting of two consecutive subsets of agents, say an upstream consecutive subset S_1 and a downstream consecutive subset S_2, would never transfer water from S_1 to S_2 because the strictly increasing benefit functions of the agents would make that all water transferred from S_1 to S_2 would immediately be taken by the agents in between S_1 and S_2. In contrast, under Assumption 2.2.5 it might be profitable for a non-consecutive coalition of agents to transfer water between its non-consecutive parts. When all agents in between S_1 and S_2 have a satiation point and S_1 transfers water to S_2, some of the flow might be taken by the in-between agents. However, the in-between agents will only confiscate water up to their satiation points. When the flow transferred from S_1 to S_2 is big enough it could be that part of it reaches S_2. Thus, depending on the benefit functions of the agents and the flow transferred from S_1 to S_2, it could be that cooperation between the two non-consecutive parts of the coalition S is profitable. As a result of the change in assumption, the worth of a non-consecutive coalition in a river game can now be higher than the sum of the worths of its maximal consecutive subsets.

In addition, the behavior of the agents under Assumption 2.2.5 might cause positive *externalities* on a connected coalition of agents T. Under Assumption 2.2.1, the worth of coalition T follows from the maximization problem (2.2). Yet, when it is profitable for agents upstream of T to transfer water to agents downstream of T the agents in T can take some of this water. The worth of T therefore depends on the coalition formation of the agents outside T (whether agents upstream of T cooperate with agents downstream of T or not).

Situations in which the worth of a coalition $S \subset N$ can depend on the coalition formation of agents outside S can be modeled by a PFF-game. Recall from Section 2.1 that a PFF-game assigns a worth $w(S, P)$ to every pair (S, P) such that $S \in P$ and $P \in \mathcal{P}$

is a partition of N. This means that the worth of a coalition S in P of N depends on the cooperation structure $P \setminus \{S\}$ of the players in $N \setminus S$.

Given a river benefit problem (N, e, b) and Assumption 2.2.5, Ambec and Ehlers (2008) provide an iterative procedure to define the worths $w(S, P)$, for every $P \in \mathcal{P}^N$ and every $S \in P$, of a *river game with externalities*. We denote the collection of all river games with externalities on N by $\mathcal{RE}^N$.

In a river game with externalities $(N, w) \in \mathcal{RE}^N$, when $S \in P$ and every coalition $T \in P$, $T \neq S$, is a singleton, then the agents outside S do not cooperate. We write $P_S = \{S\} \cup \{\{i\}, i \in N \setminus S\}$ for the partition of N where all agents outside S do not cooperate and act as singletons, and $v_*(S) = w(S, P_S)$ for the worth of S in this situation. Ambec and Ehlers (2008) call $v_*(S)$, $S \subseteq N$, the non-cooperative core lower bounds. As Ambec and Sprumont (2002) do for river games, Ambec and Ehlers (2008) maintain that any solution f for river games with externalities should satisfy these (non-cooperative) core lower bounds as well as aspiration upper bounds (which can be defined similarly as before). They then define the downstream incremental solution for river games with externalities $(N, w) \in \mathcal{RE}^N$ as

$$f_1^d(N, w) = v_*(\{1\}) \text{ and } f_i^d(N, w) = v_*([1, i]) - v_*([1, i-1]),\ \ i \in [2, n],$$

and show that this is the only solution that satisfies the non-cooperative core lower bounds and aspiration upper bounds. The downstream incremental solution of Ambec and Ehlers (2008) does not depend on the externalities in the game (N, w). This should not come as a surprise because the maximum welfare (worth) of an upstream coalition $[1, i]$, $i \in N$, does not depend on the behavior of the agents downstream of this coalition (water from the downstream agents cannot reach the upstream coalition so the actions of the downstream agents are irrelevant for the upstream coalition). The downstream incremental solution assigns to every agent along the river its contribution to the welfare when it enters the coalition consisting of its upstream agents and thus does not depend on externalities.

Chapter 3

Independence axioms for river water allocation

3.1 Introduction

In this chapter we consider the problem of sharing water among agents located along a single-stream river. As in the model of Ambec and Sprumont (2002), discussed in Section 2.2, each agent has quasi-linear preferences over river water and money, but now the benefit of consuming an amount of water is given by a continuous and concave benefit function, which is not necessarily differentiable, strictly increasing and strictly concave. As before, a Pareto efficient solution to the river sharing problem efficiently distributes the river water over the agents and wastes no money. We introduce a number of (independence) axioms to characterize two new and two existing Pareto efficient solutions for river benefit problems (with concave benefit functions). Then we apply the solutions to the particular case that every agent has constant marginal benefit of one up to a satiation point, and marginal benefit of zero thereafter. This special case can be seen as representing a situation where the full benefit functions of the agents are unknown and each agent along the river has only specified a single claim on water from the river. In this case we find that two of the solutions (one existing and one new) can be implemented without monetary transfers between the agents.

It follows from the above that the novelty in this chapter, in comparison to the literature discussed in Section 2.2, is threefold. First, we weaken the assumption of Ambec and Ehlers (2008) (and therefore also the assumption of Ambec and Sprumont (2002)) on the benefit functions of the agents in the river sharing model by only requiring concavity and continuity. Second, we characterize two existing solutions for the single-stream river benefit problem by introducing a number of (independence) axioms. Third, we propose and characterize two new solutions for the single-stream river benefit problem, also by using (independence) axioms.

In contrast to the other chapters in this dissertation, in this chapter we avoid the detour of modeling the river benefit problem as a cooperative game. Instead, we immediately impose axioms on the class of all river benefit problems with concave benefit functions. This has as a main advantage that the axioms we propose can directly be interpreted in terms of water (benefit) allocation. While most axioms used in the literature are derived from water distribution principles, they are ultimately axioms on cooperative games and not on water allocation problems. This can lead to friction when trying to interpret the cooperative game axioms directly in terms of water allocation. The approach in this chapter allows for a more straightforward interpretation of the axioms. In addition, it provides new insights into the (implicit) assumptions underlying the construction of river games.

This chapter is based on van den Brink, Estévez-Fernández, van der Laan and Moes (2011) and is organized as follows. In Section 3.2 we weaken the assumptions on the benefit functions of the agents in the single-stream river benefit problem of Ambec and Sprumont (2002). In Section 3.3 we introduce a number of (basic) axioms on the class of river benefit problems. In Section 3.4 we consider the independence of downstream benefits and independence of upstream benefits axioms, and use each of these axioms in a characterization of a solution for river benefit problems. In Section 3.5 we consider the independence of downstream inflows and independence of upstream inflows axioms, and also use each of these axioms in a characterization of a solution for river benefit problems. This leads to a total of four solutions for river benefit problems that we apply in Section 3.6 to the special case where every agent has constant marginal benefit of one up to a satiation point, and marginal benefit of zero thereafter. We conclude with a comparison of the four solutions in Section 3.7.

3.2 River benefit problems with concave benefit functions

Consider a river benefit problem (N, e, b), where $N = \{1, \dots, n\}$ is the set of agents, $e = (e_i)_{i \in N}$ is the vector of inflows and $b = (b_i)_{i \in N}$ is the vector of benefit functions, as introduced in Section 2.2. In Ambec and Sprumont (2002) the problem to find a 'fair' Pareto efficient welfare distribution (x^*, t) for river benefit problems is modeled by a TU-game. It is, however, also possible to find such Pareto efficient distributions by directly imposing axioms on the class of river benefit problems.

For instance, the core lower bounds axiom of Ambec and Sprumont (2002) (Axiom 2.2 in Section 2.2) implies *stability*. Stability says that the total payoff (utility) of each coalition of consecutive agents $[i, j] = \{i, \dots, j\}$, $1 \leq i \leq j \leq n$, at a Pareto efficient pair (x^*, t) is at least equal to the sum of the benefits that the agents in the coalition can obtain by optimally allocating their own water inflows $e_i, \dots, e_j$ among themselves. In case $i = j$ this stability notion reduces to *individual rationality*, which says that the payoff of an agent is at least equal to the benefit that the agent can obtain by maximally consuming its own water inflow.[1] Taking $i = 1$ and $j \geq i$, stability implies *upstream stability*, meaning that for every upstream coalition of consecutive agents $[1, j]$, $j \in N$, the total payoff $\sum_{i=1}^{j} z_i$ of the first j upstream agents at a Pareto efficient pair (x^*, t) is at least equal to the sum of the benefits that these agents can guarantee themselves by

[1]Notice that this notion of individual rationality only holds under the assumption that an agent is the legal owner of its own inflow, more on this in the final section of this chapter.

solving the welfare maximization problem

$$\max_{x_1,\dots,x_j} \sum_{i=1}^{j} b_i(x_i) \text{ s.t. } \sum_{i=1}^{k} x_i \le \sum_{i=1}^{k} e_i,\ k \in [1,j], \text{ and } x_\ell \ge 0,\ \ell \in [1,j]. \tag{3.1}$$

Under Assumption 2.2.1 of Chapter 2, for each $j \in N$ this maximization problem has a unique solution. We denote this solution by $\widetilde{x}^j = (\widetilde{x}_1^j, \dots, \widetilde{x}_j^j)$ and the corresponding total welfare in the river benefit problem (N, e, b) by $v^j(e,b) = \sum_{i=1}^{j} b_i(\widetilde{x}_i^j)$. Notice that $\widetilde{x}_i^n = x_i^*$, $i \in N$ (see Section 2.2), and that $V(N,e,b) := v^n(e,b) = \sum_{i=1}^{n} b_i(x_i^*)$. It follows that upstream stability requires that $\sum_{i=1}^{j} z_i \ge v^j(e,b)$ for every $j \in N$.

Similarly, the aspiration upper bounds axiom of Ambec and Sprumont (2002) (Axiom 2.2.3 in Section 2.2) implies that for every upstream coalition the total payoff $\sum_{i=1}^{j} z_i$ of the agents in such a coalitions is bounded from above by the maximum welfare they can obtain by distributing their own water inflows optimally among themselves. Thus, the aspiration upper bounds axiom requires that for each $j \in N$ the total payoff $\sum_{i=1}^{j} z_i$ of the first j upstream agents is at most equal to the welfare obtained by solving the welfare maximization problem (3.1), i.e., $\sum_{i=1}^{j} z_i \le v^j(e,b)$ for every $j \in N$.

The upstream stability requirement and the upstream aspiration upper bounds together require that $\sum_{i=1}^{j} z_i = v^j(e,b)$ for every $j \in N$, and thus determine the unique payoff vector $z_i = v^i(e,b) - v^{i-1}(e,b)$, $i \in N$, with $v^0(e,b)$ defined to be equal to zero. This payoff vector can be implemented by the Pareto efficient welfare distribution $(\widetilde{x}^n, t)$ with $t_i = z_i - b_i(\widetilde{x}_i^n)$, $i \in N$. The corresponding downstream incremental solution assigns to every river benefit problem (N,e,b) the payoff vector $f^d(N,e,b) \in \mathbb{R}^N$ given by

$$f_i^d(N,e,b) = v^i(e,b) - v^{i-1}(e,b),\ i \in N.$$

This shows that a 'solution' for river benefit problems (a payoff vector that can be implemented by a Pareto efficient welfare distribution) can be found without first transforming the river benefit problem into a river game.

Given the river benefit problem (N,e,b), we now further weaken the assumption of Ambec and Ehlers (2008) (Assumption 2.2.5 in Chapter 2), and therefore also that of Ambec and Sprumont (2002) (Assumption 2.2.1 in Chapter 2), on the benefit functions of the agents by allowing them to be concave instead of strictly concave and continuous instead of differentiable.

Assumption 3.2.1 *In a river benefit problem (N,e,b), every benefit function $b_i : \mathbb{R}_+ \to \mathbb{R}$, $i \in N$, is concave and continuous for $x_i > 0$.*

This assumption allows for a benefit function b_i, $i \in N$, that has an interval $[C_i, \widetilde{C}_i]$, $C_i \le \widetilde{C}_i$, such that b_i is increasing on $x_i < C_i$, constant on $x_i \in [C_i, \widetilde{C}_i]$, and decreasing when $x_i > C_i$. We call C_i the satiation point of agent i because agent i reaches its highest possible benefit of water consumption at C_i. All water consumption levels between C_i and $\widetilde{C}_i$ also yield this maximal benefit, but water consumptions higher than $\widetilde{C}_i$ yield a lower benefit. Assumption 3.2.1 allows for $C_i = 0$ and for $\widetilde{C}_i = \infty$ (meaning that b_i is constant for $x_i \geq C_i$). In particular this means that for some $i \in N$ it could be that $b_i(x_i) = b_i(0)$ for every $x_i \geq 0$.

Under Assumption 3.2.1 the maximization problems (3.1) do not necessarily have a unique solution, but are still well-defined. Let X^j be the set of solutions of the welfare maximization problem (3.1) for agent $j \in N$ under Assumption 3.2.1. Then, for every solution $x^j \in X^j$ it holds that $v^j(e,b) = \sum_{i=1}^{j} b_i(x_i^j)$ and for every $x^n \in X^n$ the budget balanced pair (x^n, t) yields a welfare distribution

$$z_i = b_i(x_i^n) + t_i,\ i \in N,$$

with sum of payoffs equal to the Pareto efficient total welfare $V(N,e,b) = \sum_{i=1}^{n}\ b_i(x_i^n)$.

Under Assumption 3.2.1, the cooperative game of Ambec and Sprumont (2002) corresponding to a river benefit problem (N,e,b) (the river game of Chapter 2) is not well defined unless additional assumptions on the water consumption of agents that have concave, but not strictly concave, benefit functions are made. Consider, for instance, a non-consecutive coalition consisting of a consecutive upstream part and a consecutive downstream part. If some agent j in between these two parts has a benefit function with a satiation point C_j and a point $\widetilde{C}_j > C_j$, then the cooperative game is not well-defined without an additional assumption on the water consumption of agent j in case the water flow transferred by the upstream part to the downstream part is so large that the amount of water available to agent j exceeds its satiation point C_j. Instead of making such an assumption, in the following sections we will impose axioms directly on river benefit problems (N,e,b). We thus derive unique solutions for the welfare distribution problem without modeling the river situation as a cooperative game. Doing so, assumptions in addition to Assumption 3.2.1 are not necessary.

3.3 Basic axioms for river benefit problems

In this section we first formulate three basic axioms concerning the distribution of welfare in river benefit problems (N,e,b), where the preferences of the agents over water are described by benefit functions satisfying Assumption 3.2.1. Let $\mathcal{RB}^N$ denote the collection of all river benefit problems (N,e,b) on N satisfying Assumption 3.2.1. Then a solution to river benefit problems is a function f assigning to every $(N,e,b) \in \mathcal{RB}^N$ a payoff vector $f(N,e,b) \in \mathbb{R}^N$. In the sequel, the component $f_i(N,e,b)$, $i \in N$, is called the payoff of agent i.

In the first two axioms of this section we weaken the stability requirement of the previous section. Ambec and Sprumont (2002) argue that their (core lower bounds) stability requirement reflects the ATS principle (see Section 2.2). However, in Chapter 1 we argued that the ATS principle is no longer a widely accepted principle of international watercourse law. The ATS principle is considered to be unfair because it violates the water needs of other states, has never been used in treaties and agreements between nations, and has been described as self-contradictory by authors in the field of international watercourse law. In short, the ATS principle violates the generally accepted meta-principle of limited territorial sovereignty, which says that there exist legal restrictions on a state's use of an international watercourse. To be more in line with current thinking in international watercourse law we therefore change the stability requirement to an efficiency axiom and a lower bound property.

Axiom 3.3.1 Efficiency
A solution f to river benefit problems is efficient if for every river benefit problem $(N, e, b) \in \mathcal{RB}^N$ it holds that $\sum_{i \in N} f_i(N, e, b) = V(N, e, b)$.

This (Pareto) efficiency axiom only requires stability for the grand coalition N, and requires that the agents in a river benefit problem do not distribute more welfare (utility) than is available to them. It states that the total sum of the payoffs to the agents in a river benefit problem must equal the total welfare $V(N, e, b)$ in an optimal water allocation. Next, consider the following lower bound property.

Axiom 3.3.2 Lower bound property
A solution f to river benefit problems satisfies the lower bound property if for every river benefit problem $(N, e, b) \in \mathcal{RB}^N$ it holds that $f_i(N, e, b) \geq b_i(0)$ for all $i \in N$.

This lower bound property is clearly weaker (from a normative perspective) than the stability interpretation of the ATS principle. Whereas stability essentially makes each country the legal owner of its own water inflow, the lower bound property only guarantees that an agent receives a payoff at least equal to the benefit of consuming no water.

The aspiration upper bounds axiom is used by Ambec and Sprumont (2002) to put an upper bound on the total payoff to each coalition of agents (see Section 2.2). By this property, it is a priori excluded for an upstream coalition $[1, j]$, $j \in N$, to benefit from cooperation with its downstream coalition $[j+1, n]$ because its total payoff is restricted to the welfare level that it can obtain by optimally allocating its own water inflows among its members. Consequently, when an upstream coalition $[1, j]$ cooperates with a downstream coalition $[j+1, n]$, all additional benefits of cooperation are distributed to the agents in the downstream coalition. One might wonder why the coalition of upstream agents $[1, j]$ would agree to such an allocation if it can obtain the same level of welfare on its own (without the downstream coalition). In the next axiom we therefore weaken the aspiration upper bounds axiom in a way that allows an upstream coalition $[1, j]$ to possibly benefit from allocating some of its water inflow to its downstream coalition $[j+1, n]$. This 'weak aspiration level property' requires that no agent obtains a higher payoff than the utility it can obtain when it would have access to all the water inflow in the river. That is, its own water inflow plus all the upstream and downstream water inflows.

Axiom 3.3.3 Weak aspiration level property
A solution f to river benefit problems satisfies the weak aspiration level property if for every river benefit problem $(N, e, b) \in \mathcal{RB}^N$ it holds that $f_i(N, e, b) \leq \max_{x_i \leq \sum_{j \in N} e_j} b_i(x_i)$ for all $i \in N$.[2]

[2]Note that under increasing benefit functions this inequality can be written as $f_i(N, e, b) \leq b_i(\sum_{i \in N} e_i)$ for all $i \in N$.

3.4 Independence of benefits

In Chapter 1 we argued that in the last decade experts in the field of international watercourse law have shifted their focus from water rights to water responsibilities. In light of this shift, one might wonder whether an agent (country) along a river should be held responsible for the benefit that other agents derive from the consumption of river water. For instance, if the demand for river water of an agent increases because it develops a new irrigation technique (or decides to start large-scale farming activity on its territory), should this have any consequences for the other agents along the river? According to the often invoked principle of prior appropriation (which also plays a role as one of the factors in the principles of equitable utilization and territorial integration of all basin states, see Chapter 1) this should not be the case. Recall that the principle of prior appropriation states that a country that first makes use of some quantity of water from an international watercourse has the right to the continued use of that quantity. This principle is reflected in the independence axioms of this section, which imply that an agent should not be held responsible for (changes in) the benefit (utility) that its downstream or its upstream agents derive from the consumption of water from the river.

Independence of downstream benefits and the downstream incremental solution

The first independence axiom states that the payoff of an agent does not depend on the benefit functions of the agents downstream of it.

Axiom 3.4.1 Independence of downstream benefits
A solution f to river benefit problems is independent of downstream benefits if for every pair of river benefit problems $(N, e, b) \in \mathcal{RB}^N$ and $(N, e, b') \in \mathcal{RB}^N$ such that $b_j = b'_j$ for all $j \leq i$, we have that $f_i(N, e, b) = f_i(N, e, b')$.

This axiom, together with the three basic axioms of the previous section, uniquely characterizes the downstream incremental solution $f^d(N, e, b)$ on the class of river benefit problems $\mathcal{RB}^N$.

Theorem 3.4.2 *A solution f on the class $\mathcal{RB}^N$ of river benefit problems is equal to the downstream incremental solution f^d if and only if f satisfies efficiency, the lower bound property, the weak aspiration level property and independence of downstream benefits.*

Proof. It is straightforward to show that the downstream incremental solution satisfies these four axioms. Hence, it is sufficient to show that the four axioms determine a unique solution.

Let $(N, e, b) \in \mathcal{RB}^N$ be a river benefit problem and suppose that solution f satisfies the four axioms. We prove uniqueness by induction on the labels of the agents, starting with the most upstream agent 1.

We first show that $f_1(N, e, b)$ is uniquely determined by the four axioms. Consider the modified river benefit problem (N, e, b^1) given by benefit functions $(b^1)_1 = b_1$, and $(b^1)_j(x) = 0$ for all $x \in \mathbb{R}_+$ and $j \in [2, n]$. Imposing the lower bound property on

(N,e,b^1) requires that $f_j(N,e,b^1) \geq (b^1)_j(0) = 0$ for all $j \in [2,n]$, while imposing the weak aspiration level property requires that $f_j(N,e,b^1) \leq \max_{x_j \leq \sum_{k\in N} e_k} (b^1)_j(x_j) = 0$ for all $j \in [2,n]$. Hence, $f_j(N,e,b^1) = 0$ for all $j \in [2,n]$. By efficiency it then holds that $f_1(N,e,b^1) = v^n(e,b^1)$, being the welfare level at the solution of the maximization problem (3.1) for (N,e,b^1), $j = n$. Since $(b^1)_j(x) = 0$ for every $x \in \mathbb{R}_+$ and $j \in [2,n]$ it follows that $v^n(e,b^1) = v^1(e,b^1) = \max_{x_1 \leq e_1}(b^1)_1(x_1) = \max_{x_1\leq e_1} b_1(x_1) = v^1(e,b)$. Independence of downstream benefits then implies that $f_1(N,e,b) = f_1(N,e,b^1) = v^1(e,b) = f^d_1(N,e,b)$.

Proceeding by induction, assume that $f_k(N,e,b) = f^d_k(N,e,b)$ for all $k < i \leq n$. Next, consider the modified river benefit problem (N,e,b^i) given by $(b^i)_j = b_j$ for all $j \in [1,i]$ and $(b^i)_j(x) = 0$ for all $x \in \mathbb{R}_+$ and $j \in [i+1,n]$. Similar as above, the lower bound property requires that $f_j(N,e,b^i) \geq 0$ for all $j \in [i+1,n]$, while the weak aspiration level property requires that $f_j(N,e,b^i) \leq 0$ for all $j \in [i+1,n]$. Thus

$$f_j(N,e,b^i) = 0 \text{ for all } j \in [i+1,n]. \tag{3.2}$$

Independence of downstream benefits and the induction hypothesis imply that $f_j(N,e,b^i)$ $= f_j(N,e,b) = f^d_j(N,e,b)$ for all $j \in [1,i-1]$. Hence,

$$\sum_{j=1}^{i-1} f_j(N,e,b^i) = \sum_{j=1}^{i-1} f^d_j(N,e,b) = v^{i-1}(e,b). \tag{3.3}$$

Efficiency, the induction hypothesis, (3.2) and (3.3) then determine that

$$f_i(N,e,b^i) = v^n(e,b^i) - \sum_{j=1}^{i-1} f_j(N,e,b^i) - \sum_{j=i+1}^{n} f_j(N,e,b^i) = v^n(e,b^i) - v^{i-1}(e,b). \tag{3.4}$$

Since $(b^i)_j(x) = 0$ for every $x \in \mathbb{R}_+$ and $j \in [i+1,n]$, similar as above it follows that $v^n(e,b^i) = v^i(e,b^i) = v^i(e,b)$. Therefore, with (3.4) it holds that $f_i(N,e,b^i) = v^n(e,b^i) - v^{i-1}(e,b) = v^i(e,b) - v^{i-1}(e,b) = f^d_i(N,e,b)$. Independence of downstream benefits then implies that $f_i(N,e,b) = f_i(N,e,b^i) = f^d_i(N,e,b)$.

□

Logical independence of the axioms in this theorem is shown by giving four alternative solutions. Each of these solutions only satisfies three of the four axioms.

1. The solution $f_i(N,e,b) = b_i(0)$ for all $i \in N$ satisfies the lower bound property, the weak aspiration level property and independence of downstream benefits. It does not satisfy efficiency.

2. The solution $f_i(N,e,b) = \max_{x_i \leq \sum_{j\in N} e_j} b_i(x_i)$ for all $i \in N \setminus \{n\}$, and $f_n(N,e,b) = v^n(e,b) - \sum_{j=1}^{n-1} f_i(N,e,b)$ assigns to every agent, except the most downstream agent, its maximal benefit when it would have access to all water inflows. The payoff of the most downstream agent is obtained by subtracting all these maximal benefits from the total benefit in an efficient allocation. This solution satisfies efficiency, independence of downstream benefits and the weak aspiration level property. It does not satisfy the lower bound property.

3. The solution $f_n(N,e,b) = v^n(e,b) - \sum_{j=1}^{n-1} b_i(0)$ and $f_i(N,e,b) = b_i(0)$ for all $i \in N \setminus \{n\}$ satisfies efficiency, the lower bound property and independence of downstream benefits. It does not satisfy the weak aspiration level property.

4. The downstream solution, to be introduced below, satisfies efficiency, the lower bound property and the weak aspiration level property. It will also be easy to verify that this solution does not satisfy independence of downstream benefits.

Observe that the downstream incremental solution satisfies the stability requirement of Section 3.2 for every coalition of consecutive agents. To see this, consider a coalition $[i,j]$, $1 \leq i \leq j \leq n$. Then $\sum_{k\in[i,j]} f_k^d(N,e,b) = v^j(e,b) - v^{i-1}(e,b)$. Because $v^j(e,b)$ is the maximum welfare that the agents in $[1,j]$ can obtain by distributing $e_1,\ldots,e_j$ among themselves it follows that $v^j(e,b) \geq v^{i-1}(e,b) + v([i,j])$ (where $v([i,j])$ is as defined in (2.2) of Section 2.2) and thus that f^d satisfies stability for $[i,j]$.

Independence of upstream benefits and the downstream solution

As a counterpart of the independence of downstream benefits axiom, we now consider the independence of upstream benefits axiom, which states that the payoff of an agent does not depend on the benefit functions of the agents upstream of it.

Axiom 3.4.3 Independence of upstream benefits
A solution f to river benefit problems is independent of upstream benefits if for every pair of river benefit problems $(N,e,b) \in \mathcal{RB}^N$ and $(N,e,b') \in \mathcal{RB}^N$ such that $b_j = b'_j$ for all $j \geq i$, we have that $f_i(N,e,b) = f_i(N,e,b')$.

Similarly as the independence of downstream benefits axiom, this independence of upstream benefits axiom, together with the three basic axioms of the previous section, uniquely characterizes a solution on the class of river benefit problems $\mathcal{RB}^N$. This is the *downstream solution* f^{ds}, which assigns to a river benefit problem $(N,e,b) \in \mathcal{RB}^N$ the payoff vector $f^{ds}(N,e,b)$ given by

$$f_i^{ds}(N,e,b) = \bar{v}^i(e,b) - \bar{v}^{i+1}(e,b),\ i \in N,$$

where $\bar{v}^{n+1}(e,b) = 0$ and $\bar{v}^j(e,b) = \sum_{i=j}^{n} b_i(\bar{y}_i^j)$, $j \in N$, with $\bar{y}^j = (\bar{y}_j^j,\ldots,\bar{y}_n^j)$ a solution of the welfare maximization problem

$$\max_{x_j,\ldots,x_n} \sum_{i=j}^{n} b_i(x_i) \text{ s.t. } \sum_{i=j}^{k} x_i \leq \sum_{i=1}^{k} e_i,\quad k \in [j,n],\ \text{and } x_i \geq 0,\ i \in [j,n]. \tag{3.5}$$

The maximization problem (3.5) optimally allocates the water inflows $e_1,\ldots,e_n$ over the agents in the coalition $[j,n]$, $j \in N$, given the unidirectionality of the water flow. This reveals that $\bar{v}^j$, $j \in N$, can be seen as an aspiration welfare level of Ambec and Sprumont (2002), i.e., the upper bound welfare level in Axiom 2.2 of Section 2.2. For $j = 1$ the maximization problem (3.5) is equal to problem (3.1) for $j = n$, so that $\bar{v}^1(e,b) = v^n(e,b) = V(N,e,b)$ is the maximum total benefit that can be obtained when allocating all inflows optimally among all agents. Since $\sum_{i=1}^{n} f_i^{ds}(N,e,b) = \bar{v}^1(e,b) = V(N,e,b)$ the downstream solution provides an efficient payoff vector.

In the welfare maximization problem (3.5), the agents in the downstream coalition $[j,n]$, $j \in N$, are able to use the total water inflow into the river $\sum_{i\in N} e_i$. When some of the water is allocated to other (upstream) agents, the downstream solution fully compensates the downstream agents for their loss of benefit via monetary compensations from the upstream agents. Hence, the downstream solution attributes the rights over the use of water from the river to the downstream coalitions. It is not difficult to see that the downstream solution does not satisfy (upstream) stability and therefore violates the Ambec and Sprumont (2002) interpretation of the ATS principle. In fact, the downstream solution seems to be more in line with the UTI principle from international watercourse law (see Chapter 1).

The downstream solution can be characterized by replacing the independence of downstream benefits axiom in the characterization of the downstream incremental solution (Theorem 3.4.2) with the independence of upstream benefits axiom. The downstream and downstream incremental solutions thus only differ in one axiom.

Theorem 3.4.4 *A solution f on the class $\mathcal{RB}^N$ of river benefit problems is equal to the downstream solution f^{ds} if and only if f satisfies efficiency, the lower bound property, the weak aspiration level property and independence of upstream benefits.*

Proof. It is straightforward to prove that the downstream solution satisfies these four axioms, it therefore only has to be shown that the four axioms determine a unique solution.

Let $(N,e,b) \in \mathcal{RB}^N$ be a river benefit problem and suppose that solution f satisfies the four axioms. We apply induction on the labels of the agents, starting with the most downstream agent n.

Consider the modified river benefit problem (N,e,b^n) given by $(b^n)_n = b_n$, and $(b^n)_j(x) = 0$ for all $x \in \mathbb{R}_+$ and $j \in [1,n-1]$. The lower bound property requires that $f_j(N,e,b^n) \geq (b^n)_j(0) = 0$ for all $j \in [1,n-1]$, while the weak aspiration level property requires that

$$f_j(N,e,b^n) \leq \max_{x_j \leq \sum_{k\in N} e_k} (b^n)_j(x_j) = 0, \quad \text{for all } j \in [1,n-1].$$

Thus, it can be concluded that $f_j(N,e,b^n) = 0$ for all $j \in [1,n-1]$. By efficiency it then holds that

$$f_n(N,e,b^n) = v^n(e,b^n) = \bar{v}^1(e,b^n). \tag{3.6}$$

Since $(b^n)_j(x) = 0$ for every $x \in \mathbb{R}_+$ and $j \in [1,n-1]$, and $(b^n)_n = b^n$, it follows that $\bar{v}^1(e,b^n) = \bar{v}^n(e,b^n) = \bar{v}^n(e,b)$ and thus, with (3.6), that $f_n(N,e,b^n) = \bar{v}^1(e,b^n) = \bar{v}^n(e,b) = f_n^{ds}(N,e,b)$. Independence of upstream benefits then implies that $f_n(N,e,b) = f_n(N,e,b^n) = f_n^{ds}(N,e,b)$.

Proceeding by induction, assume that $f_k(N,e,b) = f_k^{ds}(N,e,b)$ for all $k > i \geq 1$. Next, consider the modified river benefit problem (N,e,b^i) given by $(b^i)_j = b_j$ for all $j \in [i,n]$, and $(b^i)_j(x) = 0$ for all $x \in \mathbb{R}_+$ and $j \in [1,i-1]$. Similar as above, the lower bound property requires that $f_j(N,e,b^i) \geq 0$ for all $j \in [1, i-1]$, while the weak aspiration level property requires that $f_j(N,e,b^i) \leq 0$ for all $j \in [1, i-1]$. Thus,

$$f_j(N,e,b^i) = 0 \text{ for all } j \in [1, i-1]. \tag{3.7}$$

Independence of upstream benefits and the induction hypothesis imply that $f_j(N,e,b^i) = f_j(N,e,b) = f_j^{ds}(N,e,b)$ for all $j \in [i+1, n]$. Therefore,

$$\sum_{j=i+1}^{n} f_j(N,e,b^i) = \sum_{j=i+1}^{n} f_j^{ds}(N,e,b) = \bar{v}^{i+1}(e,b). \tag{3.8}$$

Efficiency, the induction hypothesis, (3.7) and (3.8) then determine that

$$f_i(N,e,b^i) = \bar{v}^1(e,b^i) - \sum_{j=1}^{i-1} f_j(N,e,b^i) - \sum_{j=i+1}^{n} f_j(N,e,b^i) = \bar{v}^1(e,b^i) - \bar{v}^{i+1}(e,b). \tag{3.9}$$

Since $(b^i)_j(x) = 0$ for every $x \in \mathbb{R}_+$ and $j \in [1, i-1]$, and $(b^i)_j = b_j$ for all $j \geq i$, similar as above it follows that $\bar{v}^1(e,b^i) = \bar{v}^i(e,b^i) = \bar{v}^i(e,b)$. Hence, with (3.9) it holds that $f_i(N,e,b^i) = \bar{v}^1(e,b^i) - \bar{v}^{i+1}(e,b) = \bar{v}^i(e,b) - \bar{v}^{i+1}(e,b) = f_i^{ds}(N,e,b)$. Independence of upstream benefits then implies that $f_i(N,e,b) = f_i(N,e,b^i) = f_i^{ds}(N,e,b)$.

□

Also now, logical independence of the axioms is shown by giving four alternative solutions. Each of these solutions only satisfies three of the four axioms.

1. The solution assigning $f_i(N,e,b) = b_i(0)$ to $i \in N$, satisfies the lower bound property, the weak aspiration level property and independence of upstream benefits. It does not satisfy efficiency.

2. The solution $f_i(N,e,b) = \max_{x_i \leq \sum_{j \in N} e_j} b_i(x_i)$ for all $i \in N \setminus \{1\}$, and $f_1(N,e,b) = v^n(e,b) - \sum_{j=2}^{n} f_i(N,e,b)$ assigns to every agent, except the most upstream agent, its maximal benefit when it would have access to all water inflows into the river. The payoff of the most upstream agent is obtained by subtracting all these maximal benefits from the total benefit in an efficient allocation. This solution satisfies efficiency, independence of upstream benefits and the weak aspiration level property. It does not satisfy the lower bound property.

3. The upstream incremental solution, discussed below, satisfies efficiency, the lower bound property and independence of upstream benefits. It does not satisfy the weak aspiration level property.

4. The downstream incremental solution satisfies efficiency, the lower bound property and the weak aspiration level property. It does not satisfy independence of upstream benefits.

3.5 Independence of inflows

In this section we study the dependence of solutions for river benefit problems on the water inflows into the river. The independence of upstream inflows axiom, on the one hand, reflects that a country does not have any right to the water inflows on the territories of its upstream agents. The independence of downstream inflows axiom, on the other hand, restricts the maximal water rights of a country to all the water inflows upstream of its territory. These axioms thus partially respect the ATS principle (see Chapter 1) in the sense that they do not allow a state to infringe the sovereignty of its upstream or its downstream states.

Independence of upstream inflows and the upstream incremental solution

The independence of upstream inflows axiom states that the payoff of an agent does not depend on the water inflows into the river on the territories of its upstream agents.

Axiom 3.5.1 Independence of upstream inflows
A solution f to river benefit problems is independent of upstream inflows if for every pair of river benefit problems $(N, e, b) \in \mathcal{RB}^N$ and $(N, e', b) \in \mathcal{RB}^N$ such that $e_j = e'_j$ for all $j \geq i$, we have that $f_i(N, e, b) = f_i(N, e', b)$.

This independence of upstream inflows axiom is incompatible with the weak aspiration level property of Section 3.3 because the latter property allows an agent to use all the water inflows into the river, also the inflows on the territories of its upstream agents. Since the weak aspiration level property was introduced to weaken the aspiration upper bounds axiom of Ambec and Sprumont (2002) (see Section 2.2), we now weaken the aspiration upper bounds axiom in a different way.

The drought property only requires the aspiration upper bounds axiom to hold for upstream coalitions of which the total water inflow into the river is zero. The drought property therefore says that when there is no water entering the river in the upstream coalition $[1, j]$, $j \in N$, all agents in this coalition receive a payoff that is smaller than, or equal to, their benefit of no water consumption.

Axiom 3.5.2 Drought property
A solution f to river benefit problems satisfies the drought property if for every river benefit problem $(N, e, b) \in \mathcal{RB}^N$ with $e_j = 0$ for all $j \leq i$, it holds that $f_i(N, e, b) \leq b_i(0)$.

This drought property, together with the efficiency axiom, lower bound property and independence of upstream inflows axiom uniquely characterizes the upstream incremental solution of Herings, van der Laan and Talman (2007) and van den Brink, van der Laan and Vasil'ev (2007) (see Section 2.2). Recall that in the upstream incremental solution all gains in benefit that arise when some of the inflows of an upstream coalition $[1, j]$, $j \in N$, are allocated to the agents downstream of it are distributed to $[1, j]$, in the sense that the total payoff to the downstream coalition $[j + 1, n]$ is equal to the total benefit that the agents in this coalition can obtain by allocating their own water inflows optimally among themselves.

To define the upstream incremental solution for river benefit problems $(N, e, b) \in \mathcal{RB}^N$, we consider for every $j \in N$ the welfare maximization problem

$$\max_{x_j,\ldots,x_n} \sum_{i=j}^{n} b_i(x_i) \text{ s.t. } \sum_{i=j}^{k} x_i \leq \sum_{i=j}^{k} e_i, \quad k \in [j, n], \text{ and } x_i \geq 0, \; i \in [j, n], \tag{3.10}$$

i.e., for agent $j \in N$ the maximization problem (3.10) optimally allocates the inflows $e_j, \ldots, e_n$ among the agents in the coalition $[j, n]$, given the unidirectionality of the water flow.[3] Given a solution $y^j = (y^j_j, \ldots, y^j_n)$ of maximization problem (3.10) for agent $j \in N$, denote $\widetilde{v}^j(e, b) = \sum_{i=j}^{n} b_i(y^j_i)$ as the maximum welfare that the agents in $[j, n]$ can obtain by distributing their own inflows among themselves. Notice that for $j = 1$ the maximization problem (3.10) is equal to problem (3.1) for $j = n$, so that $\widetilde{v}^1(e, b) = v^n(e, b) = V(N, e, b)$ is the maximum total benefit that can be obtained when allocating all inflows optimally among all agents. For every solution y^1 of (3.10) for $j = 1$, the budget balanced pair (y^1, t) thus yields a welfare distribution

$$z_i = b_i(y^1_i) + t_i, \quad i \in N,$$

with sum of the payoffs equal to the Pareto efficient total welfare $V(N, e, b) = \sum_{i=1}^{n} b_i(y^1_i)$.

The upstream incremental solution assigns to every agent $i \in N$ the increase in welfare that is created when agent i joins the coalition of its downstream agents $[i+1, n]$. Hence, the upstream incremental solution assigns to every river benefit problem $(N, e, b) \in \mathcal{RB}^N$, the payoff vector $f^u(N, e, b) \in \mathbb{R}^N$ given by

$$f^u_i(N, e, b) = \widetilde{v}^i(e, b) - \widetilde{v}^{i+1}(e, b), \; i \in N,$$

with $\widetilde{v}^{n+1}(e, b) = 0$.

Notice that the solution f^u is fully determined by the welfare levels obtained by solving the welfare maximization problems (3.10). Hence, by definition, the upstream incremental solution satisfies stability for every downstream coalition $[i, n]$, $i \in N$. Moreover, like the downstream incremental solution, the upstream incremental solution also satisfies the stability requirement for every coalition of consecutive agents. To see this consider a coalition $[i, j]$, $1 \leq i \leq j \leq n$. Then $\sum_{k \in [i,j]} f^u_k(N, e, b) = \widetilde{v}^i(e, b) - \widetilde{v}^{j+1}(e, b)$. Because $\widetilde{v}^i(e, b)$ is the maximum welfare that the agents in $[i, n]$ can obtain by distributing $e_i, \ldots, e_n$ among themselves it follows that $\widetilde{v}^i(e, b) \geq v([i, j]) + \widetilde{v}^{j+1}(e, b)$ (where $v([i, j])$ is as defined in (2.2) of Section 2.2) and thus that f^u satisfies stability for $[i, j]$.

The next theorem gives a characterization of the upstream incremental solution on the class of river benefit problems $\mathcal{RB}^N$.

Theorem 3.5.3 *A solution f on the class $\mathcal{RB}^N$ of river benefit problems is equal to the upstream incremental solution f^u if and only if f satisfies efficiency, the lower bound property, the drought property and independence of upstream inflows.*

[3] Notice the difference between the maximization problems (3.5) and (3.10). In (3.10) a downstream coalition $[j, n]$, $j \in N$, can only consume its own water inflow, while in (3.5) it can consume all the water inflows into the river (also the inflows of the agents upstream of the downstream coalition $[j, n]$).

Proof. It is straightforward to show that the upstream incremental solution satisfies these four axioms. So, it is sufficient to prove that the four axioms determine a unique solution.

Let $(N, e, b) \in \mathcal{RB}^N$ be a river benefit problem and suppose that solution f satisfies the four axioms. We apply induction on the labels of the agents, starting with the most downstream agent n.

Consider the modified river benefit problem (N, e^n, b) given by $e^n_n = e_n$, and $e^n_j = 0$ for all $j \in [1, n-1]$. The lower bound property requires that $f_j(N, e^n, b) \geq b_j(0)$ for all $j \in [1, n-1]$, while the drought property requires that $f_j(N, e^n, b) \leq b_j(0)$ for all $j \in [1, n-1]$. Thus, it can be concluded that $f_j(N, e^n, b) = b_j(0)$ for all $j \in [1, n-1]$. By efficiency it then holds that

$$f_n(N, e^n, b) = \widetilde{v}^1(e^n, b) - \sum_{j=1}^{n-1} f_j(N, e, b) = \widetilde{v}^1(e^n, b) - \sum_{j=1}^{n-1} b_j(0). \tag{3.11}$$

Since $e^n_j = 0$ for all $j \in [1, n-1]$ and $e^n_n = e_n$, it follows that $\widetilde{v}^1(e^n, b) = \sum_{j=1}^{n-1} b_j(0) + \widetilde{v}^n(e^n, b) = \sum_{j=1}^{n-1} b_j(0) + \widetilde{v}^n(e, b)$, and thus with (3.11) it holds that $f_n(N, e^n, b) = \widetilde{v}^n(e, b)$. Independence of upstream inflows then implies that $f_n(N, e, b) = f_n(N, e^n, b) = \widetilde{v}^n(e, b) = f^u_n(N, e, b)$.

Proceeding by induction, assume that $f_k(N, e, b) = f^u_k(N, e, b)$ is determined for all $k > i \geq 1$. Next, consider the modified river benefit problem (N, e^i, b) given by $e^i_j = e_j$ for all $j \in [i, n]$, and $e^i_j = 0$ for all $j \in [1, i-1]$. Similar as above, the lower bound property requires that $f_j(N, e^i, b) \geq b_j(0)$ for all $j \in [1, i-1]$, while the drought property requires that $f_j(N, e^i, b) \leq b_j(0)$ for all $j \in [1, i-1]$. Thus,

$$f_j(N, e^i, b) = b_j(0) \text{ for all } j \in [1, i-1]. \tag{3.12}$$

Independence of upstream inflows and the induction hypothesis imply that $f_j(N, e^i, b) = f_j(N, e, b) = f^u_j(N, e, b)$ for all $j \in [i+1, n]$. Hence,

$$\sum_{j=i+1}^{n} f_j(N, e^i, b) = \sum_{j=i+1}^{n} f^u_j(N, e, b) = \widetilde{v}^{i+1}(e, b). \tag{3.13}$$

Efficiency, the induction hypothesis, (3.12) and (3.13) then determine that

$$f_i(N, e^i, b) = \widetilde{v}^1(e^i, b) - \sum_{j=1}^{i-1} f_j(N, e^i, b) - \sum_{j=i+1}^{n} f_j(N, e^i, b) =$$

$$\widetilde{v}^1(e^i, b) - \sum_{j=1}^{i-1} b_j(0) - \widetilde{v}^{i+1}(e, b). \tag{3.14}$$

Since $e^i_j = 0$ for all $j \in [1, i-1]$ and $e^i_j = e_j$ for all $j \in [i, n]$, similar as above it follows that $\widetilde{v}^1(e^i, b) = \sum_{j=1}^{i-1} b_j(0) + \widetilde{v}^i(e^i, b) = \sum_{j=1}^{i-1} b_j(0) + \widetilde{v}^i(e, b)$. Thus, with (3.14) it holds that $f_i(N, e^i, b) = \widetilde{v}^1(e^i, b) - \sum_{j=1}^{i-1} b_j(0) - \widetilde{v}^{i+1}(e, b) = \widetilde{v}^i(e, b) - \widetilde{v}^{i+1}(e, b) = f_i^u(N, e, b)$. Independence of downstream benefits then implies that

$$f_i(N, e, b) = f_i(N, e^i, b) = f_i^u(N, e, b).$$

□

Logical independence of the axioms in Theorem 3.5.3 follows from the following four alternative solutions.

1. The solution $f_i(N, e, b) = b_i(0)$ for all $i \in N$ satisfies the lower bound property, the drought property and independence of upstream inflows. It does not satisfy efficiency.

2. For some $\epsilon > 0$, define the solution f as $f(N, e, b) = f^u(N, e, b)$ if $e_n = 0$. Otherwise, define $f_1(N, e, b) = f_1^u(N, e, b) - \epsilon$, $f_i(N, e, b) = f_i^u(N, e, b)$ for $i \in [2, n-1]$ and $f_n(N, e, b) = f_n^u(N, e, b) + \epsilon$. It is easy to see that f satisfies efficiency, the drought property and independence of upstream inflows since f^u satisfies these properties. The solution f, however, does not satisfy the lower bound property.

3. The solution $f_1(N, e, b) = \widetilde{v}^1(e, b) - \sum_{j=2}^{n} b_j(0)$ and $f_i(N, e, b) = b_i(0)$ for all $i \in N \setminus \{1\}$ satisfies efficiency, the lower bound property and independence of upstream inflows. It does not satisfy the drought property.

4. The downstream incremental solution f^d satisfies efficiency, the lower bound property and the drought property. It does not satisfy independence of upstream inflows.

Independence of downstream inflows and the upstream solution

Finally, we consider the independence of downstream inflows axiom, which states that the payoff of an agent does not depend on the water inflows into the river on the territories of its downstream agents.

Axiom 3.5.4 Independence of downstream inflows
A solution f to river benefit problems is independent of downstream inflows if for every pair of river benefit problems $(N, e, b) \in \mathcal{RB}^N$ and $(N, e', b) \in \mathcal{RB}^N$ such that $e_j = e'_j$ for all $j \leq i$, we have that $f_i(N, e, b) = f_i(N, e', b)$.

Like the independence of upstream inflows axiom, this independence of downstream inflows axiom is not compatible with the weak aspiration level property of Section 3.3. Replacing the weak aspiration level property with the drought property, however, does not do much good because together the efficiency axiom, lower bound property, drought property and independence of upstream inflows axiom do not characterize a unique solution. Hence, to obtain a unique solution for river benefit problems we strengthen the drought property to a property, called the no contribution property, which states that an agent with zero inflow of water on its territory gets, at most, a payoff equal to its benefit of zero water consumption. It thus implies that a country cannot claim any benefit from the water allocation when its own inflow is zero.

Axiom 3.5.5 No contribution property
A solution f to river benefit problems satisfies the no contribution property if for every river benefit problem $(N, e, b) \in \mathcal{RB}^N$ with $e_i = 0$ it holds that $f_i(N, e, b) \leq b_i(0)$.

This no contribution property, together with the efficiency axiom, lower bound property and independence of downstream inflows axiom uniquely characterizes a solution on the class of river benefit problems $\mathcal{RB}^N$. This is the *upstream solution* f^{us}.

To define the upstream solution, we first reconsider the welfare distribution according to the upstream incremental solution. The upstream incremental solution yields a payoff $f_n^u(N, e, b) = \widetilde{v}^n(e, b)$ for the last agent n along the river, where $\widetilde{v}^n(e, b)$ is the highest benefit that agent n can obtain by consuming only its own water inflow. Agent $n-1$ receives $f_{n-1}^u(N, e, b) = \widetilde{v}^{n-1}(e, b) - \widetilde{v}^n(e, b)$, where $\widetilde{v}^{n-1}(e, b)$ is the total benefit that agents $n-1$ and n can jointly obtain by distributing their own water inflows e_{n-1} and e_n optimally among themselves (given the unidirectionality of the water flow). In this way agent $n-1$ receives the marginal contribution to the total benefit of the water inflow e_{n-1} to the water inflow e_n, taking all the upstream inflows equal to zero. In general, in the upstream incremental solution agent $i \in N$ receives the marginal contribution to the total benefit of the water inflow e_i to the downstream inflows e_j, $j > i$, taking all the upstream inflows e_k, $k < i$, equal to zero.

The upstream solution can be defined the other way around, starting with agent 1. When all inflows in a river benefit problem (N, e, b) are put equal to zero and there are no transfers, every agent has payoff $b_i(0)$, $i \in N$. Now, let the most upstream inflow e_1 be non-zero and let it be distributed optimally among all agents. Then in the upstream solution agent 1 receives, in addition to $b_1(0)$, all the benefit that this creates. So, agent 1 receives $b_1(0)$ plus the the marginal contribution to the total benefit when the water inflow e_1 is distributed optimally among all agents along the river, assuming that all other inflows are equal to zero. The upstream solution f^{us} thus distributes to agent 1 the payoff $f_1^{us}(N, e, b) = \widehat{v}^1(e, b)$, where

$$\widehat{v}^1(e, b) = b_1(0) + \sum_{j=1}^{n} (b_j(\hat{y}_j^1) - b_j(0)) = b_1(\hat{y}_1^1) + \sum_{j=2}^{n} (b_j(\hat{y}_j^1) - b_j(0)),$$

with $\hat{y}^1 = (\hat{y}_1^1, \ldots, \hat{y}_n^1)$ a solution of the welfare maximization problem

$$\max_{x_1,\ldots,x_n} \sum_{j=1}^{n} b_j(x_j) \text{ s.t. } \sum_{j=1}^{n} x_j \leq e_1 \text{ and } x_j \geq 0,\ j \in [1, n].$$

Next, given the unidirectionality of the water flow, the inflows e_1 and e_2 are distributed optimally over all agents along the river, assuming that all other inflows are equal to zero. Agent 2 receives its initial payoff $b_2(0)$ plus the additional total benefit that the optimal distribution of e_1 and e_2 adds to the total benefit obtained from optimally distributing e_1 among all agents. Subsequently, for agent $i \in N \setminus \{1, 2\}$ all inflows e_j, $j \leq i$, are distributed optimally over all agents, assuming all inflows of the downstream agents $k > i$ are equal to zero. Agent i then receives its initial payoff $b_i(0)$ plus the additional total benefit that the optimal distribution of the inflows $e_1, \ldots, e_i$ adds to the total benefit already obtained from the optimal distribution of $e_1, \ldots, e_{i-1}$.

In general, the upstream solution f^{us} thus assigns to a river benefit problem $(N,e,b) \in \mathcal{RB}^N$ the payoff vector $f^{us}(N,e,b)$ given by

$$f_i^{us}(N,e,b) = \widehat{v}^i(e,b) - \widehat{v}^{i-1}(e,b), \quad i \in N,$$

where $\widehat{v}^0(e,b) = 0$ and $\widehat{v}^i(e,b) = \sum_{j=1}^{i} b_j(\hat{y}_j^i) + \sum_{j=i+1}^{n} (b_j(\hat{y}_j^i) - b_j(0))$, $i \in N$, with $\hat{y}^i = (\hat{y}_1^i, \ldots, \hat{y}_n^i)$ a solution of the welfare maximization problem

$$\max_{x_1,\ldots,x_n} \sum_{j=1}^{n} b_j(x_j) \text{ s.t. } \begin{cases} \sum_{j=1}^{n} x_j \leq \sum_{j=1}^{i} e_j, \\ \sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} e_j, \ k \in [1, i-1], \\ x_j \geq 0, \ j \in N. \end{cases} \tag{3.15}$$

Observe that this maximization problem optimally distributes the water inflows of the agents in $[1,i]$ over all agents along the river, taking into account that for every agent k the total water consumption of the first k agents is at most equal to the sum of the inflows upstream of k. For $j = n$ the maximization problem (3.15) is again equal to problem (3.1) with $j = n$, so that $\widehat{v}^n(e,b) = v^n(e,b) = V(N,e,b)$ is the maximum total benefit that can be obtained when allocating all inflows optimally among all agents. Since $\sum_{i=1}^{n} f_i^{us}(N,e,b) = \widehat{v}^n(e,b) = V(N,e,b)$ also the upstream solution provides an efficient payoff vector.

The payoffs of the vector $f^{us}(N,e,b)$ can also be written as

$$f_i^{us}(N,e,b) = \widehat{v}^i(e,b) - \widehat{v}^{i-1}(e,b) =$$

$$\sum_{j=1}^{i} b_j(\hat{y}_j^i) + \sum_{j=i+1}^{n} (b_j(\hat{y}_j^i) - b_j(0)) - \left(\sum_{j=1}^{i-1} b_j(\hat{y}_j^{i-1}) + \sum_{j=i}^{n} (b_j(\hat{y}_j^{i-1}) - b_j(0)) \right) =$$

$$b_i(0) + \sum_{j=1}^{n} (b_j(\hat{y}_j^i) - b_j(\hat{y}_j^{i-1})), \ i \in N.$$

According to the upstream solution, every upstream coalition $[1,j]$, $j \in N$, receives the total welfare that can be obtained by allocating the water inflows of such a coalition optimally over all agents along the river. Clearly, for an upstream coalition the welfare at a solution of the welfare maximization problem (3.15) is at least as high as the welfare at a solution of the welfare maximization problem (3.1) (in which the inflows of a coalition $[1,j]$ are distributed optimally over the agents in the coalition). This implies that the upstream solution satisfies stability for the upstream coalitions. However, the upstream solution does not satisfy stability for consecutive coalitions $[i,j]$ in general. For example, agent n receives the marginal benefit $\widehat{v}^n(e,b) - \widehat{v}^{n-1}(e,b)$, which is the difference between the total benefit of the water consumptions $\hat{y}^n$ and $\hat{y}^{n-1}$. Nothing can be said about this difference and the benefit $b_n(e_n)$ that agent n can obtain by consuming its own water inflow. It can happen that $f_n^{us}(N,e,b) < b_n(e_n)$, violating individual rationality and thus stability.

In the next theorem we give a characterization of the upstream solution on the class of river benefit problems $\mathcal{RB}^N$.

Theorem 3.5.6 *A solution f on the class $\mathcal{RB}^N$ of river benefit problems is equal to the upstream solution f^{us} if and only if f satisfies efficiency, the lower bound property, the no contribution property and independence of downstream inflows.*

Proof. It is straightforward to show that the upstream solution satisfies these four axioms. It is therefore sufficient to prove that the four axioms determine a unique solution.

Let $(N, e, b) \in \mathcal{RB}^N$ be a river benefit problem and suppose that solution f satisfies the four axioms. Similar as in the proof of Theorem 3.4.2, we apply induction on the labels of the agents, starting with the most upstream agent 1.

Consider the modified river benefit problem (N, e^1, b) given by $e^1_1 = e_1$, and $e^1_j = 0$ for all $j \in [2, n]$. The lower bound property requires that $f_j(N, e^1, b) \geq b_j(0)$ for all $j \in [2, n]$, while the no contribution property requires that $f_j(N, e^1, b) \leq b_j(0)$ for all $j \in [2, n]$. Thus, it can be concluded that $f_j(N, e^1, b) = b_j(0)$ for all $j \in [2, n]$. By efficiency it then holds that $f_1(N, e^1, b) = v^n(e^1, b) - \sum_{j=2}^n f_j(N, e^1, b) = v^n(e^1, b) - \sum_{j=2}^n b_j(0)$. Since $e^1_j = 0$ for all $j \in [2, n]$ and $e^1_1 = e_1$, it follows that $v^n(e^1, b) = \sum_{j=1}^n b_j(\hat{y}^1_j) = \widehat{v}^1(e, b) + \sum_{j=2}^n b_j(0)$. Independence of downstream inflows then implies that $f_1(N, e, b) = f_1(N, e^1, b) = \widehat{v}^1(e, b) + \sum_{j=2}^n b_j(0) - \sum_{j=2}^n b_j(0) = \widehat{v}^1(e, b) = \widehat{v}^1(e, b) - \widehat{v}^0(e, b) = f^{us}_1(N, e, b)$.

Proceeding by induction, assume that $f_k(N, e, b) = f^{us}_k(N, e, b)$ is determined for all $k < i \leq n$. Next, consider the modified river benefit problem (N, e^i, b) given by $e^i_j = e_j$ for all $j \in [1, i]$ and $e^i_j = 0$ for all $j \in [i+1, n]$. Similar as above, the lower bound property requires that $f_j(N, e^i, b) \geq b_j(0)$ for all $j \in [i+1, n]$, while the no contribution property requires that $f_j(N, e^i, b) \leq b_j(0)$ for all $j \in [i+1, n]$. Thus,

$$f_j(N, e^i, b) = b_j(0) \text{ for all } j \in [i+1, n]. \tag{3.16}$$

Independence of downstream inflows and the induction hypothesis imply that

$$f_j(N, e^i, b) = f_j(N, e, b) = f^{us}_j(N, e, b)$$

for all $j \in [1, i-1]$. Therefore,

$$\sum_{j=1}^{i-1} f_j(N, e^i, b) = \sum_{j=1}^{i-1} f^{us}_j(N, e, b) = \sum_{j=1}^{i-1} [\widehat{v}^j(e, b) - \widehat{v}^{j-1}(e, b)] = \widehat{v}^{i-1}(e, b). \tag{3.17}$$

Efficiency, the induction hypothesis, (3.16) and (3.17) then determine that

$$f_i(N, e^i, b) = v^n(e^i, b) - \sum_{j=1}^{i-1} f_j(N, e^i, b) - \sum_{j=i+1}^{n} f_j(N, e^i, b) =$$

$$v^n(e^i, b) - \widehat{v}^{i-1}(e, b) - \sum_{j=i+1}^{n} b_j(0). \tag{3.18}$$

Since $e^i_j = 0$ for all $j \in [i+1, n]$ and $e^i_j = e_j$ for all $j \in [1, i]$, similar as above it follows that $v^n(e^i, b) = \sum_{j=1}^n b_j(\hat{y}^i_j) = \widehat{v}^i(e, b) + \sum_{j=i+1}^n b_j(0)$. Thus, with (3.18) it holds that

$f_i(N,e^i,b) = \widehat{v}^i(e,b) + \sum_{j=i+1}^{n} b_j(0) - \widehat{v}^{i-1}(e,b) - \sum_{j=i+1}^{n} b_j(0) = \widehat{v}^i(e,b) - \widehat{v}^{i-1}(e,b) = f_i^{us}(N,e,b)$. Independence of downstream inflows then implies that

$$f_i(N,e,b) = f_i(N,e^i,b) = f_i^{us}(N,e,b).$$

□

Logical independence of the axioms in Theorem 3.5.6 follows from the following four alternative solutions.

1. The solution $f_i(N,e,b) = b_i(0)$ for all $i \in N$ satisfies the lower bound property, the no contribution property and independence of downstream inflows. It does not satisfy efficiency.

2. For some $\epsilon > 0$, define the solution f by $f(N,e,b) = f^{us}(N,e,b)$ if $e_1 = 0$. Otherwise, define $f_1(N,e,b) = f_1^{us}(N,e,b) + \epsilon$, $f_i(N,e,b) = f_i^{us}(N,e,b)$ for $i \in [2,n-1]$ and $f_n(N,e,b) = f_n^{us}(N,e,b) - \epsilon$. It is easy to see that f satisfies efficiency, the no contribution property and independence of upstream inflows. It does not satisfy the lower bound property.

3. The solution $f_i(N,e,b) = b_i(0)$ for all $i \in N \setminus \{n\}$ and $f_n(N,e,b) = v^n(e,b) - \sum_{j=1}^{n-1} b_j(0)$ satisfies efficiency, the lower bound property and independence of downstream inflows. It does not satisfy the no contribution property.

4. The upstream incremental solution satisfies efficiency, the lower bound property and the no contribution property. It does not satisfy independence of downstream inflows.

3.6 A special case: the river claim problem

Before comparing the four solutions of this chapter in the next section, in this section we consider river benefit problems in which every agent has constant marginal benefit of one up to its satiation point, and zero marginal benefit thereafter. That is, for every $i \in N$ there exists a $C_i > 0$ such that

$$b_i(x_i) = \begin{cases} x_i & \text{if } x_i \le C_i \\ C_i & \text{if } x_i > C_i. \end{cases} \tag{3.19}$$

This type of benefit functions has (implicitly) been considered by Ansink and Weikard (2012) in river sharing problems in which agents are not allowed, or not able, to make monetary transfers. To see this, consider the single-stream river benefit problem (N,e,b) of Ambec and Sprumont (2002) and assume that, instead of a benefit function b_i, every agent $i \in N$ is endowed with a single *claim* $C_i \ge 0$ on water from the river. We call the triple (N,e,C), where $N = \{1,\ldots,n\}$ is the set of agents, $e = (e_i)_{i \in N}$ is the vector of inflows and $C = (C_i)_{i \in N}$ is the vector of claims, a *river claim problem*. We denote the collection of all river claim problems by $\mathcal{RC}$. As before, a water allocation x assigns an amount of water x_i, $i \in N$, to agent i under the constraints

$$\sum_{i=1}^{j} x_i \leq \sum_{i=1}^{j} e_i, \; j \in N.$$

A solution to a river claim problem is a function g assigning to every $(N, e, C) \in \mathcal{RC}$ a water allocation $g(N, e, C)$ such that $0 \leq g_i(N, e, C) \leq C_i$ for all $i \in N$.

Throughout this section we assume that the following assumption holds (for both river claim problems, as well as river benefit problems).

Assumption 3.6.1 $\sum_{i=j}^{n} C_i \geq \sum_{i=j}^{n} e_i$ *for all* $j \in N$.

This assumption states that the sum of the claims of the $n - j + 1$ most downstream agents is larger than or equal to the sum of their inflows, for all $j \in N$, and can be made without loss of generality. To see this, imagine that the assumption does not hold and consider the most upstream agent $i \in N$ for which it holds that $\sum_{k=i}^{n} e_k > \sum_{k=i}^{n} C_k$. Then all agents $j \geq i$ can be given their claim C_j because there is no shortage of water in the downstream part of the river $i, \ldots, n$. Deleting the agents $i, \ldots, n$ would result in a new claim problem $([1, i-1], (e_1, \ldots, e_{i-1}), (C_1, \ldots, C_{i-1}))$, and benefit problem $([1, i-1], (e_1, \ldots, e_{i-1}), (b_1, \ldots, b_{i-1}))$, in which Assumption 3.6.1 does hold.

Given a river benefit problem (N, e, b) in which every agent has a benefit function of type (3.19), the satiation point C_i can be considered as the claim of agent $i \in N$ on water from the river. We now apply the four solutions of this chapter to this special case in which each agent has a benefit function of type (3.19).

Recall that the downstream incremental solution f^d is given by

$$f_i^d(N, e, b) = v^i(e, b) - v^{i-1}(e, b), \; i \in N, \tag{3.20}$$

where $v^0(e, b) = 0$ and $v^i(e, b)$ is the welfare level at a solution of the welfare maximization problem (3.1) for agent i. For benefit functions of type (3.19) it follows straightforwardly that

$$v^1(e, b) = \min[C_1, e_1],$$

and successively

$$v^j(e, b) = v^{j-1}(e, b) + \min[C_j, \sum_{i=1}^{j} e_i - v^{j-1}(e, b)], \; j \in [2, n].$$

Substituting this in the equations (3.20) it can be seen that

$$f_1^d(N, e, b) = \min[C_1, e_1]$$

and, using the fact that by definition $\sum_{i=1}^{j-1} f_i^d(N, e, b) = v^{j-1}(e, b)$ for all $j \in [2, n]$, recursively

$$f_j^d(N, e, b) = \min\left[C_j, e_j + \sum_{i=1}^{j-1}(e_i - f_i^d(N, e, b))\right],$$

for the agents $j \in [2, n]$.

It can be concluded that in this case the downstream incremental solution can be implemented by assigning to each upstream coalition $[1, j]$, $j \in N$, as much water as possible, given the unidirectionality of the water flows and under the constraint that no agent receives water above its satiation point. When each agent has a benefit function of type (3.19), monetary compensations are thus not needed to implement the downstream incremental solution.

The downstream solution f^{ds} is given by

$$f_i^{ds}(N, e, b) = \bar{v}^i(e, b) - \bar{v}^{i+1}(e, b), \; i \in N, \tag{3.21}$$

where $\bar{v}^{n+1}(e, b) = 0$ and $\bar{v}^i(e, b)$ is the welfare level at a solution of the welfare maximization problem (3.5) for agent i. Under Assumption 3.6.1 it follows that for benefit functions of type (3.19)

$$\bar{v}^n(e, b) = \min[C_n, \sum_{i=1}^{n} e_i].$$

Since, by the same assumption, $\bar{v}^n(e, b) \geq e_n$, it holds that

$$\bar{v}^{n-1}(e, b) = \bar{v}^n(e, b) + \min[C_{n-1}, \sum_{i=1}^{n} e_i - \bar{v}^n(e, b)],$$

and that $\bar{v}^{n-1}(e, b) \geq e_{n-1} + e_n$. Continuing, it follows successively from $j = n - 2$ to $j = 1$ that

$$\bar{v}^j(e, b) = \bar{v}^{j+1}(e, b) + \min[C_j, \sum_{i=1}^{n} e_i - \bar{v}^{j+1}(e, b)].$$

Substituting this in the equations (3.21) one obtains

$$f_n^{ds}(N, e, b) = \min[C_n, \sum_{i=1}^{n} e_i]$$

and, using the fact that by definition $\sum_{i=j+1}^{n} f_i^{ds}(N, e, b) = \bar{v}^{j+1}(e, b)$ for all $j \in [1, n-1]$, recursively from $j = n - 1$ to $j = 1$,

$$f_j^{ds}(N, e, b) = \min[C_j, \sum_{j=1}^{n} e_j - \sum_{i=j+1}^{n} f_i^{ds}(N, e, b)].$$

From these expressions it follows that the downstream solution can be implemented in this case by assigning to each downstream coalition $[j, n]$, $j \in N$, as much water as possible, given the unidirectionality of the water flows and under the constraint that no agent receives water above its satiation point. Hence, when each agent has a benefit function of type (3.19) monetary compensations are not needed to implement the downstream solution.

Recall that the upstream incremental solution f^u is given by

$$f_i^u(N,e,b) = \widetilde{v}^i(e,b) - \widetilde{v}^{i+1}(e,b),\ i \in N, \tag{3.22}$$

where $\widetilde{v}^{n+1}(e,b) = 0$ and $\widetilde{v}^i(e,b)$ is the welfare level at a solution of the welfare maximization problem (3.10) for agent i. Again, without loss of generality one can assume that Assumption 3.6.1 holds. For benefit functions of type (3.19) it then follows straightforwardly that

$$\widetilde{v}^j(e,b) = \sum_{i=j}^{n} e_i,\ j \in N.$$

Substituting this in the equations (3.22) one obtains

$$f_j^u(N,e,b) = \sum_{i=j}^{n} e_i - \sum_{i=j+1}^{n} e_i = e_j,\ j \in N.$$

The upstream incremental solution thus gives precisely payoff e_j to each agent $j \in N$. Observe that when $e_i \leq C_i$ for all $i \in N$ the downstream incremental solution reduces to $f_i^d(N,e,b) = e_i$ for all $i \in N$ so that the downstream incremental and upstream incremental solutions coincide.

It is not difficult to see that it could be impossible to implement the upstream incremental solution without monetary transfers. For example, take $n = 2$, $C_1 < e_1 < C_1 + C_2$ and $e_2 = 0$. The total welfare e_1 is obtained for every solution x^n of the welfare maximization problem (3.1) for $j = n$, so for every x^n with $\max[0, e_1 - C_2] \leq x_1^n \leq C_1$ and $x_2^n = e_1 - x_1^n$. To implement the welfare distribution $f_1(N,e,b) = e_1$, $f_2(N,e,b) = 0$, it is thus required that agent 2 pays a monetary compensation $t_1 = x_2^n$ to agent 1 (since by only consuming water agent 1 cannot reach a higher payoff than $C_1 < e_1$). It can be concluded that when each agent has a benefit function of type (3.19) the upstream incremental solution, in general, cannot be implemented without monetary transfers between the agents.

Finally, consider the upstream solution f^{us} given by

$$f_i^{us}(N,e,b) = \widehat{v}^i(e,b) - \widehat{v}^{i-1}(e,b),\ i \in N, \tag{3.23}$$

where $\widehat{v}^0(e,b) = 0$ and $\widehat{v}^i(e,b)$ is the welfare level at a solution of the welfare maximization problem (3.15) for agent i. Under Assumption 3.6.1 it follows that

$$\widehat{v}^j(e,b) = \sum_{i=1}^{j} e_i,\ j \in N,$$

and substituting this in the equations (3.23) one obtains

$$f_j^{us}(N,e,b) = \sum_{i=1}^{j} e_i - \sum_{i=1}^{j-1} e_i = e_j,\ j \in N.$$

This last equation implies that the upstream solution and the upstream incremental solution coincide when each agent has a benefit function of type (3.19). In general the upstream solution therefore cannot be applied when monetary compensations between the agents are not allowed or are not possible.

We can conclude that when applying the four solutions of this chapter to the particular case that every agent has constant marginal benefit of one up to a satiation point and marginal benefit of zero thereafter, only the downstream and downstream incremental solutions can be implemented without monetary transfers between the agents. This means that when countries along an international river only state a claim on the river water, and are not willing to transfer money to each other, out of the four solutions presented in this chapter only these two solutions are viable.

Since the type of benefit functions considered in this section is also (implicitly) considered in river claim problems, where monetary transfers are not possible, it follows that the downstream and downstream incremental solutions can be seen as solutions to river claim problems. The downstream incremental solution for river claim problems assigns as much water as possible to each upstream coalition $[1,i]$, $i \in N$, given the unidirectionality of the water flows and under the constraint that no agent receives water above its claim. The downstream solution for river claim problems assigns as much water as possible to each downstream coalition $[i,n]$, $i \in N$, given the unidirectionality of the water flows and under the constraint that no agent receives water above its claim.[4] This can also be seen in the next example.

Example 3.6.2 As in case 1 of Ansink and Weikard (2012) let (N,e,C) be a river claim problem such that $N = \{1,2,3,4\}$, $e = (80,10,10,10)$ and $C = (50,10,20,90)$. The downstream incremental solution then results in the water allocation $(50,10,20,30)$ and the downstream solution in the water allocation $(0,0,20,90)$.

□

Clearly, the downstream incremental solution and the downstream solution are very extreme solutions for river claim problems. Ansink and Weikard (2012) suggest sequential sharing rules that may result in more equitable distributions of the water in a river claim problem.

A *sequential sharing rule* is a solution to river claim problems that distributes to each agent the allocation that it receives from repeatedly applying a bankruptcy rule to a sequence of two-player bankruptcy problems. Given a set of players $N \subset \mathbb{N}$, let $ES \geq 0$ be an amount of a certain good (e.g., water) to be distributed among the players. We

[4]The downstream and downstream incremental solutions for river claim problems resemble certain priority rules for rationing problems of Moulin (2000).

refer to the amount ES as the *estate*. Suppose that each player $i \in N$ has a *claim* $C_i \geq 0$ on the estate such that $\sum_{i \in N} C_i \geq ES$. A *bankruptcy problem*, or *rationing problem*, can then be defined as a triple (N, ES, C), where $C = (C_i)_{i \in N}$ is the collection of claims. The *aggregate claim* is written as $AC = \sum_{i \in N} C_i$. The collection of all bankruptcy problems is denoted by $\mathcal{B}$. Foundational work on bankruptcy problems was done by O'Neill (1982) and Aumann and Maschler (1985). For a general introduction to bankruptcy problems see, for instance, Thomson (2003).

A *bankruptcy rule* f on $\mathcal{B}$ assigns to every bankruptcy problem $(N, ES, C) \in \mathcal{B}$ a unique payoff vector $f(N, ES, C) \in \mathbb{R}^N$ that is *efficient*, $\sum_{i \in N} f_i(N, ES, C) = ES$, *individually non-negative*, $f_i(N, ES, C) \geq 0$ for all $i \in N$, and *claim bounded*, $f_i(N, ES, C) \leq C_i$ for all $i \in N$. Three of the best-known bankruptcy rules are the proportional rule, the constrained equal awards rule and the constrained equal losses rule. The *proportional rule* is defined as $p_i(N, ES, C) = \frac{C_i}{AC} ES$, $i \in N$. Let λ^* be the solution to $\sum_{i \in N} \min[C_i, \lambda] = ES$. Then the *constrained equal awards rule* is defined as $cea_i(N, ES, C) = \min[C_i, \lambda^*]$, $i \in N$. Next, let $\widetilde{\lambda}$ be the solution to $\sum_{i \in N} \max[0, C_i - \lambda] = ES$. Then the *constrained equal losses rule* is defined as $cel_i(N, ES, C) = \max[0, C_i - \widetilde{\lambda}]$, $i \in N$.

The sequential sharing rules for river claim problems of Ansink and Weikard (2012) can now be written as follows:

$$\theta_i^f(N, e, C) = f_i(\{i, do_i\}, E_i, (C_i, C_{do_i})), \ i \in N,$$

where $E_i = e_i + \sum_{k \in [1, i-1]} (e_k - \theta_k^f(N, e, C))$, $C_{do_i} = \sum_{k \in [i+1, n]} (C_k - e_k)$ and f_i represents the payoff assigned to player i by any bankruptcy rule f on $\mathcal{B}$ for the bankruptcy problem $(\{i, do_i\}, E_i, (C_i, C_{do_i}))$ with player set $\{i, do_i\}$, estate E_i and vector of claims (C_i, C_{do_i}).[5]

As an alternative to the sequential sharing rules of Ansink and Weikard (2012), we now propose the following (class of) solution(s) for river claim problems:

$$\gamma_i^f(N, e, C) = \sum_{j=1}^{i} f_i([j, n], e_j, C^j), \ i \in N,$$

where $f_i([j, n], e_j, C^j)$ represents the payoff assigned to player i by any bankruptcy rule f on $\mathcal{B}$ for the bankruptcy problem $([j, n], e_j, C^j)$ with player set $[j, n]$, estate e_j and $n-j+1$ dimensional vector of claims $C^j = (C_k^j)_{k \in [j, n]}$ with, for all $j \in N$,

$$C_k^j = C_k - \sum_{m=j+1}^{k} f_k([m, n], e_m, C^m), \ k \in [j, n].$$

Each bankruptcy rule f on $\mathcal{B}$ provides a solution γ^f on $\mathcal{RC}$ for river claim problems.

The solution γ^f decomposes the river claim problem into a set of n bankruptcy problems and applies the same bankruptcy rule to each of them. First, the claim C_n of the most downstream agent along the river is fulfilled as much as possible given its own inflow e_n. That is, the inflow e_n serves as the estate in the 1-player bankruptcy problem $(\{n\}, e_n, (C_n))$. Since any bankruptcy rule f on $\mathcal{B}$ requires efficiency its must be that

[5] Note that do_i is a single player in the bankruptcy problem $(\{i, do_i\}, E_i, (C_i, C_{do_i}))$.

agent n is given its own inflow in the 1-player bankruptcy problem. Because it also holds that any bankruptcy rule is individually non-negative it follows that agent n receives at least its own inflow in the river claim problem. Next, the inflow e_{n-1} serves as the estate in the 2-player bankruptcy problem $(\{n-1,n\}, e_{n-1}, (C_{n-1}^{n-1}, C_n^{n-1}))$, where the claim of agent $n-1$ is equal to its claim in the river claim problem, $C_{n-1}^{n-1} = C_{n-1}$, and the claim of agent n is adjusted for the fact that it has already received its own inflow from the 1-player bankruptcy problem, $C_n^{n-1} = C_n - e_n$. The solution γ^f then applies the bankruptcy rule f to the 2-player bankruptcy problem and adds the payoffs of this to the payoffs already received in the 1-player bankruptcy problem. Since any bankruptcy rule requires efficiency its must be that e_{n-1} is fully distributed among agents $n-1$ and n.

It is not difficult to see that γ^f continues the above procedure until all the inflows e_i, $i \in N$, are fully distributed among the sets of downstream agents $[i, n]$, so that an efficient water allocation is obtained. To gain further insight into how the solution γ^f works, consider the following example.

Example 3.6.3 Let (N, e, C) be as in Example 3.6.2. Given the proportional bankruptcy rule p on $\mathcal{B}$, the solution γ^p can be found by considering the sequence of bankruptcy problems $([j,n], e_j, C^j)$, $j \in N$, starting with the most downstream agent $n = 4$.

When $j = 4$, then $([j,n], e_j, C^j) = (\{4\}, 10, (90))$ so that $p_4(\{4\}, 10, (90)) = 10$.

When $j = 3$, then $([j,n], e_j, C^j) = (\{3,4\}, 10, C^3)$ with $C_3^3 = 20$ and $C_4^3 = 90 - p_4(\{4\}, 10, (90)) = 80$. Hence, $p_3(\{3,4\}, 10, (20, 80)) = 2$ and $p_4(\{3,4\}, 10, (20, 80)) = 8$.

When $j = 2$, then $([j,n], e_j, C^j) = (\{2,3,4\}, 10, C^2)$ with $C^2 = (10, 18, 72)$. Thus, $p_2(\{2,3,4\}, 10, (10, 18, 72)) = 1$, $p_3(\{2,3,4\}, 10, (10, 18, 72)) = 1\frac{4}{5}$ and $p_4(\{2,3,4\}, 10, (10, 18, 72)) = 7\frac{1}{5}$.

Finally, when $j = 1$, then $([j,n], e_j, C^j) = (\{1,2,3,4\}, 80, C^1)$ with $C^1 = (50, 9, 16\frac{1}{5}, 64\frac{4}{5})$. So, $p_1(\{1,2,3,4\}, 80, (50, 9, 16\frac{1}{5}, 64\frac{4}{5})) = 28\frac{4}{7}$, $p_2(\{1,2,3,4\}, 80, (50, 9, 16\frac{1}{5}, 64\frac{4}{5})) = 5\frac{1}{7}$, $p_3(\{1,2,3,4\}, 80, (50, 9, 16\frac{1}{5}, 64\frac{4}{5})) = 9\frac{9}{35}$ and $p_4(\{1,2,3,4\}, 80, (50, 9, 16\frac{1}{5}, 64\frac{4}{5})) = 37\frac{1}{35}$.

This results in $\gamma_1^p = 28\frac{4}{7}$, $\gamma_2^p = 5\frac{1}{7} + 1 = 6\frac{1}{7}$, $\gamma_3^p = 2 + 1\frac{4}{5} + 9\frac{9}{35} = 13\frac{2}{35}$ and $\gamma_4^p = 10 + 8 + 7\frac{1}{5} + 37\frac{1}{35} = 62\frac{8}{35}$, which is the same as the sequential sharing rule of Ansink and Weikard (2012) for $f = p$. However, when $f = cea$ then $\gamma^{cea} = (30\frac{5}{6}, 10, 20, 49\frac{1}{6})$, which is not equal to the outcome $(40, 10, 20, 40)$ of the sequential sharing rule of Ansink and Weikard (2012) for $f = cea$.

□

It is an open question whether, and how, the class of solutions for river claim problems, introduced here, can be axiomatized. What can be said is that it would satisfy appropriate adaptations of the efficiency axiom, lower bound property, drought property and weak aspiration level property to the class of river claim problems. It would, however, not satisfy an appropriate adaptation of the no contribution property (for instance, $\gamma_2^p = 5$

	downstream incremental solution	upstream incremental solution	downstream solution	upstream solution
efficiency	++	++	++	++
lower bound property	++	++	++	++
drought property	+	++	+	+
weak aspiration level property	++		++	
no contribution property		+		++
independence of downstream benefits	++			
independence of upstream benefits		+	++	
independence of upstream inflows		++		
independence of downstream inflows	+			++

Table 3.1: Axioms satisfied by the four solutions of Chapter 3.

if $N = \{1, 2\}$, $e = (10, 0)$ and $C = (10, 10)$) and also not an appropriate adaptation of the independence of (upstream or downstream) inflows axiom (for instance, $\gamma^p = (8, 12)$ if $N = \{1, 2\}$, $e = (10, 10)$ and $C = (40, 20)$; if e changes to $(20, 10)$ then γ_2^p changes to 14, while if e changes to $(10, 15)$ then γ_1^p changes to $8\frac{8}{9}$). Finally, it does not satisfy independence of (upstream or downstream) claims (consider again $N = \{1, 2\}$, $e = (10, 10)$ and $C = (40, 20)$; if C changes to $(50, 20)$ then γ_2^p changes to $11\frac{2}{3}$, while if C changes to $(40, 30)$ then γ_1^p changes to $6\frac{2}{3}$).

3.7 Comparison of the four solutions

Consider Table 3.1. The symbol + or ++ in this table indicates that a particular solution satisfies a particular axiom, with ++ indicating the axiom being used in the characterization of a solution. Observe that the characterization of the downstream incremental solution can be obtained from the characterization of the downstream solution by replacing the independence of upstream benefits axiom by the independence of downstream benefits axiom. The characterization of the upstream incremental solution can be obtained from the characterization of the upstream solution by replacing the independence of downstream inflows axiom by the independence of upstream inflows axiom and weakening the no contribution property to the drought property. This shows, as can also be seen in the proofs of this chapter, that the characterizations of the downstream solutions are based on the independence of certain benefit functions of the agents, while those of the upstream solutions are based on the independence of certain water inflows into the river. Finally, note from the table that among the four solutions considered in this chapter, independence of downstream benefits is only satisfied by the downstream incremental solution and independence of upstream inflows only by the upstream incremental solution. This means that if countries along an international river agree to impose one of these two properties, they immediately select a unique solution out of the four solutions presented in this chapter.

Recall from Section 3.2 that a solution $f(N, e, b)$ for river benefit problems with concave benefit functions satisfies stability if it holds that $\sum_{k\in[i,j]} f_k(N, e, b) \geq v([i, j])$ for all $1 \leq i \leq j \leq n$, where $v([i, j])$ is as defined in (2.2) of Section 2.2. In Sections 3.4 and 3.5 we showed that the downstream incremental solution and the upstream incremental solution satisfy stability, whereas the downstream solution and the upstream solution do not. It is not too difficult to see that the downstream incremental solution favors coalitions of downstream agents $[i, n]$, $i \in N$, as much as possible given the stability requirement (and unidirectionality of the water flow). Likewise, the upstream incremental solution favors coalitions of upstream agents $[1, i]$, $i \in N$, as much as possible given the stability requirement. When ignoring the stability requirement it becomes clear that there are solutions to river benefit problems that favor the coalitions of downstream, or upstream, agents more than the downstream and upstream incremental solutions. In fact, the downstream solution favors coalitions of downstream agents as much as possible and the upstream solution coalitions of upstream agents. If one wants to order the four solutions based on how favorable the solution is for the coalition of upstream agents $[1, i]$, $i \in N$, one would have the following: (1) upstream solution, (2) upstream incremental solution, (3) downstream incremental solution and (4) downstream solution.[6]

Recall from Chapter 2 that the stability requirement was proposed by Ambec and Sprumont (2002) as an axiom to be satisfied by a solution for river games.[7] The core lower bounds axiom is based on the ATS principle from international watercourse law in the sense that it provides a lower bound on the welfare that a coalition of agents can secure for itself. The ATS principle implies that every agent along a river can use the water that it controls as it pleases. Since this also holds for coalitions of agents along the river, and because the water that a coalition controls is determined by its location along the river, the welfare that such a coalition can secure for itself in a river game is similarly determined by its location.

Ambec and Sprumont (2002) define, in the river game (N, v), the welfare that the coalition of consecutive agents $[i, j]$ can secure for itself by $v([i, j])$. In the definition of $v([i, j])$ it is assumed that the agents in $[i, j]$ are able/allowed to use (only) their own water inflows $e_i, \dots, e_j$ in maximizing their own total benefits $\sum_{k\in[i,j]} b_k(x_k)$. The idea that agents are allowed to use their own water inflow, however, is exactly what is set out in the ATS principle. It thus can be argued that the ATS principle does not only appear in the core lower bounds property of Ambec and Sprumont (2002), but also in the definition of the river games $(N, v) \in \mathcal{R}^N$. To illustrate this point further suppose that Assumption 2.2.1 of Chapter 2 holds and consider the following two alternative 'river games'.

The 'UTI river game' $(N, \bar{v})$ induced by (N, e, b) is the TU-game with player set N and characteristic function $\bar{v}$ defined as follows. For any coalition $S \subseteq N$ that has invoked the UTI principle let the worth $\bar{v}(S)$ be given by

$$\bar{v}(S) = \sum_{i\in S} b_i(\bar{y}_i^S) \quad \text{where} \quad \bar{y}^S = (\bar{y}_j^S)_{j\in S} \quad \text{solves}$$

[6] For the downstream agents $[i, n]$, $i \in N$, it is the reverse order.

[7] Ambec and Sprumont (2002) require stability for all coalitions of agents, not only consecutive coalitions, but this is not relevant in this discussion.

$$\max_{\{x_j|j\in S\}} \sum_{j\in S} b_j(x_j) \ \text{ s.t. } \sum_{k\in[1,\ell]\cap S} x_k \leq \sum_{k\in[1,\ell]} e_k,\ \ell \in S,\ \text{and } x_j \geq 0,\ j \in S. \tag{3.24}$$

So, the worth of coalition S is obtained by distributing all the water inflows upstream of the most downstream agent q of S, $e_1, \ldots, e_q$, optimally over the agents in the coalition S, given the unidirectionality of the water flow. Notice that $e_1, \ldots, e_q$ is exactly what the coalition S can claim according to the UTI principle. Also notice that, as in the standard (ATS) river game, in this UTI river game the worth of the grand coalition $\bar{v}(N)$ is equal to the Pareto efficient total welfare, i.e., $\bar{v}(N) = \sum_{i=1}^n b_i(\bar{y}_i^N) = \sum_{i=1}^n b_i(x_i^*)$. Any single-valued solution for TU-games, assigning payoff vector $z = f(\bar{v}) \in \mathbb{R}^N$, can therefore be implemented by the welfare distribution (x^*, t) with $t_i = z_i - b_i(x_i^*)$, $i \in N$. The essential difference between the standard river game and this UTI river game is that, while in the standard river game the worth of a coalition S is derived under the assumption that the agents outside S consume their own water inflows, in the UTI river game the worth of coalition S is derived under the assumption that the agents outside S, and upstream of the most downstream agents of S, are not allowed to consume any water (and therefore also do not do so).[8]

Next, recall from Chapter 1 that the no substantial harm principle says that a riparian state is free to use the water of an international watercourse, provided that this use does not cause substantial harm to other riparian states. Taking the most extreme interpretation of this principle one obtains the 'no harm principle', which implies that a country along a river is not allowed to use water from the river if this causes any harm to any of the other countries along the river. In river benefit problems under Assumption 2.2.1 of Chapter 2 upstream agents can harm downstream agents by consuming any water from the river, but downstream agents are not able to harm upstream agent.[9] This discussion leads to the following definition of the 'no harm river game': for any coalition S such that $n \in S$ let $v'(S) = \sum_{i\in S} b_i(\bar{y}_i^S)$, where $\bar{y}^S$ is as defined in (3.24); otherwise let $v'(S) = \sum_{i\in S} b_i(0)$.[10] Thus, if agent n is not in S, then the agents in S are not allowed (by the no harm principle) to consume any water because this causes harm to agent n. The worth of S in that case is equal to the sum of the benefits of the agents in S when these agents do not consume any water. If agent n is in S, then the agents in S are allowed to consume all the water inflows into the river because the agents outside S are not allowed to consume any water (this would harm agent n which is in S). The worth of S in that case is equal to the sum of the benefits of the agents in S, when these agents optimally distribute all the water inflows among themselves, given the unidirectionality of the water flow. Again, the worth of the grand coalition $v'(N)$ is equal to the Pareto efficient total welfare, i.e., $v'(N) = \sum_{i=1}^n b_i(\bar{y}_i^N) = \sum_{i=1}^n b_i(x_i^*)$ so that any single-valued solution for TU-games, assigning payoff vector $z = f(v') \in \mathbb{R}^N$, can be implemented by the welfare distribution (x^*, t) with $t_i = z_i - b_i(x_i^*)$, $i \in N$.

The above 'critique' of the standard river game and the alternative definitions given with the help of the UTI and no harm principles also apply to the river games with

[8] Unlike the standard river game, the 'UTI river game' is not necessarily convex. Actually, it is not even necessarily superadditive.

[9] Here we consider only physical harm, not legal harm, see Chapter 1.

[10] The 'no harm river game' is also not necessarily superadditive.

externalities of Chapter 2. Nevertheless, in the remaining chapters of this dissertation (except in Chapter 7) we work only with the river game and river game with externalities as defined in Chapter 2. The reason for this is that we want our (solutions to) river games to take voluntary participation into account.

In the maximin tradition of von Neumann and Morgenstern (1944) the worth of a coalition S can be interpreted as the wealth that the members of S can secure for themselves, given that the agents outside S do everything they can to keep this wealth as low as possible. A milder version of this idea would be that the worth of a coalition S can be interpreted as the wealth that the members of S can secure for themselves, given that the agents outside S do everything they can to maximize their own wealth. In defining the worth of a coalition in a game, one thus has to determine what the level of wealth is that the members of the coalition can secure for themselves. In river games this level of wealth depends crucially on whether one considers voluntary participation of countries into water allocation agreements, or not.

We argued in the Introduction and Chapter 1 that when there are no rules of international law that govern, countries are free to do as they please when it comes to withdrawing water from an international river. This is reflected in the ATS principle, that guarantees the sovereignty of a state over the natural resources on its territory. When a country cannot be forced (by an international court of law, (economic) sanctions or military force) to enter a water allocation agreement (give up (part of) its water inflow), it will only do so through voluntary participation. The standard river game takes account of this, because in it each coalition of agents is maximally able (allowed) to consume the water inflows that enter the river on its territory. The standard river game therefore seems to be appropriate under the constraint of voluntary participation.

In the UTI river game the worth of a coalition is derived under the assumption that the agents outside the coalition, and upstream of the most downstream agent of the coalition, are not allowed to consume any water. In the no harm river game the agents in a coalition are not allowed to consume any water if the most downstream agent is not in the coalition. This implies that the level of wealth that certain coalitions of agents can secure for themselves in these games depends on whether upstream agents can be forced to give up (part of) their water inflows (enter a water allocation agreement). The UTI and no harm river games thus seem to be more appropriate when there is an international court of law that declares one of these principles as binding.

Since in the upstream and downstream solutions of this chapter it is possible that an agent receives a lower payoff than the level of wealth it can secure for itself under voluntary participation, these solutions for river benefit problems seem more appropriate in a setting in which there is an international court of law that can implement these solutions than in a setting in which this is not the case. The downstream incremental solution and upstream incremental solution do satisfy the voluntary participation requirement that any agent along the river receives a payoff that is at least equal to the level of wealth it can secure for itself without force or agreement. For this reason we consider in the next chapter an extension of these two solutions, based on the TIBS principle from international watercourse law. In Chapter 7 we return to the issue of voluntary participation when we discuss solutions for river pollution problems that focus on property rights over the river.

Chapter 4

River games with multiple springs and the TIBS principle

4.1 Introduction

In this chapter we introduce and analyze solutions for river games with multiple springs. That is, we consider rivers that possibly have several tributaries merging into one main stream. Examples of river basins with multiple springs, in which there is (potential) conflict over the distribution of water, include some of the world's largest basins: the Amazon basin, the Euphrates and Tigris basin, the Ganges-Brahmaputra-Meghna basin, the Mekong basin and the Nile basin. In the Amazon basin water shortage is currently no real issue, but in 2005 parts of the basin experienced severe droughts. Similar droughts in the future could lead to friction between states sharing the basin. In the Euphrates and Tigris basin, Turkey has more than once been accused by Syria and Iraq of depriving them of water. In the Ganges-Brahmaputra-Meghna basin, the most heavily populated river basin in the world, water shortages in some places are getting worse, with sections of the river running dry for parts of the year. The building of dams by upstream states in the Mekong basin (China and Thailand) is causing low flows in downstream states (Cambodia and Vietnam), that are completely dependent on the river for food and the majority of their developing economies. Yet, the best example of disagreement in an international river basin with multiple springs is probably that of the Nile river. This river, generally regarded as the longest river in the world, has two main tributaries, the White Nile and the Blue Nile, and runs through the eleven territories of Tanzania, Burundi, Rwanda, Democratic Republic of Congo, Uganda, Kenya, South Sudan, Sudan, Ethiopia, Eritrea and Egypt. Historically, Egypt (the most downstream country) has claimed most of the rights over the use of water from the Nile. In recent decades, however, water sharing disputes with upstream countries, including Uganda, Sudan and Ethiopia, have erupted about the Egyptian domination of Nile resources. In 1999 the Nile Basin Initiative was launched with the goals of developing the river in a cooperative manner, sharing substantial socioeconomic benefits, and promoting regional peace and security.

The above examples show that there are many important applications of a water distribution model for rivers with multiple springs. Because multiple springs affect the positions of the agents in negotiations over river water, they also influence the game-theoretic analysis of the river water distribution problem. To see this, consider the following example. In a single-stream river basin with one upstream agent and one downstream agent, the two agents are completely dependent on each other when it comes to the trade of river water. The upstream agent holds water the downstream agent might want, and the downstream agent holds money the upstream agent might want. In a river basin with two springs, located on the territories of two different agents, that merge together to one stream, on the territory of a single downstream agent, the position of the latter agent in negotiations is completely different. While the downstream agent still holds money the upstream agents might want, there are now two upstream agents that hold water the downstream agent might want. Hence, there are now (possibly) two suppliers of the good 'water' and (possibly) only one agent that demands it. This implies that in negotiations over river water the downstream agent is in a much better position when it has two upstream neighbors than when it has only one. Ignoring multiple springs could thus be a serious shortcoming of the analysis.

To capture the consequences of multiple springs in a river basin, in this chapter we change the single-stream river benefit problem and river games of Chapter 2 to a river benefit problem and river games with multiple springs. As before, a distribution of the Pareto efficient total welfare in the benefit problem can be found by applying any efficient solution from cooperative game theory to the corresponding game. The 'fairness' of such a distribution again depends on the properties of the solution.

In the analysis of this chapter we base the concept of 'fairness' on the TIBS principle from international watercourse law. Recall from Chapter 1 that, although the ATS and UTI principles are still used in state practice, the principle of equitable utilization and its efficient extension, the TIBS principle, are considered far more 'fair' principles by a majority of the experts in the field. The TIBS principle states that the water of an international watercourse belongs to all basin states combined, no matter where it enters the watercourse; each riparian state is entitled to a reasonable and equitable share in the optimal use of the waters of the international watercourse (see Chapter 1 and e.g., Lipper (1967) or McCaffrey (2001)). According to this description, the principle requires that (1) the water of an international river is allocated in such a way that the total (combined) welfare of all countries is maximized (optimal use) and (2) each country receives a (reasonable and equitable) share in the total welfare resulting from an optimal allocation of the water.

We apply the TIBS principle to a river basin with multiple springs in the following way. Suppose that, for one reason or another, the agents along a river with multiple springs are cooperating in two separate coalitions. For some agent, say i, one coalition consists of agent i and all its upstream agents, the other coalition is its complement (i.e., it consists of all other agents). This occurs, for instance, when agent i is not willing to cooperate with its unique downstream neighbor. The question that can now be asked is the following: how should the gain in total welfare, that is created when the two coalitions combine into one, be divided among the agents? Evidently, this question can be asked for every single agent along the river (except the unique most downstream agent).

The TIBS principle provides an answer to the above question. Let there be for each agent a nonnegative number, its weight, with sum over all agents equal to one. These weights can be seen as the reasonable and equitable shares mentioned in the TIBS principle. We now interpret the TIBS principle as follows: for each agent i, the gain in welfare that is created by merging i's upstream coalition and its complement is divided among these two coalitions proportional to the sum of the weights of the agents in the coalitions. We show that, for every specific vector of weights, this requirement characterizes a particular distribution of the total welfare resulting from a welfare maximizing water allocation of the river water. The extreme case in which the most downstream agent has weight equal to one yields a payoff vector that can be seen as a generalization of the downstream incremental solution of Ambec and Sprumont (2002). The special case that all weights are taken to be equal yields a payoff vector that can be seen as the average tree solution of Herings, van der Laan and Talman (2008).

This chapter is based on based on van den Brink, van der Laan and Moes (2012a). It provides three novelties in comparison to the literature discussed in Section 2.2. In Section 4.2 we extend the single-stream river model of Ambec and Sprumont (2002) to rivers with multiple springs. In Section 4.3 we introduce and characterize a new class of solutions for river games (with multiple springs) based on the TIBS principle from international watercourse law (the class of weighted hierarchical solutions). And, in Section 4.4 we extend the single-stream river model of Ambec and Ehlers (2008) to rivers with multiple springs and show that the weighted hierarchical solutions for the resulting games are externality-free (do not depend on worths of coalitions that experience externalities).

4.2 River benefit problems and river games with multiple springs

To describe a river system with several tributaries, originating at different springs and merging together into one main river, let $N = \{1, \ldots, n\}$ be the set of players representing the agents, in the sequel also called countries, located along the river. Further, identify each spring by an agent, i.e., consider the most upstream agent along a tributary as its spring. In a river system with multiple springs every agent has precisely one downstream neighbor (except the most downstream agent which has none), but agents can have multiple upstream neighbors, namely in case two (or more) tributaries merge on the territory of an agent.[1] We denote the number of springs by s and let $O = \{1, \ldots, s\}$ be the set of agents located at some spring, i.e., agent j, $j \in \{1, \ldots, s\}$, is located at spring j. Note that $n > s$ is the total number of agents. We index the (unique) most downstream agent by n. For agent $k \in N$, let U^k denote the set of upstream neighbors of k. Every agent $i \in N \setminus \{n\}$ is in exactly one U^k for some $k \in N$. Notice that the structure of the river system is fully determined by the n-tuple of sets $(U^k)_{k \in N}$, with $U^k = \emptyset$ if and only if $k \in O$. We denote the collection of sets U^k, $k \in N$, by $\mathcal{U} = \{U^k | k \in N\}$. Formally, for a set N of agents, a *river structure* is a collection $\mathcal{U} = \{U^k | k \in N\}$, such that (1) for some $1 \leq s < n$, $U^k \neq \emptyset$ if and only if $k > s$, (2) $U^k \cap U^h = \emptyset$ for all $k, h \in N,\ k \neq h$, and (3) $\bigcup_{k \in N} U^k = N \setminus \{n\}$. A pair $(N, \mathcal{U})$ consisting of a set of agents N and a river structure

[1] We assume that on the territory of an agent there is at most one spring.

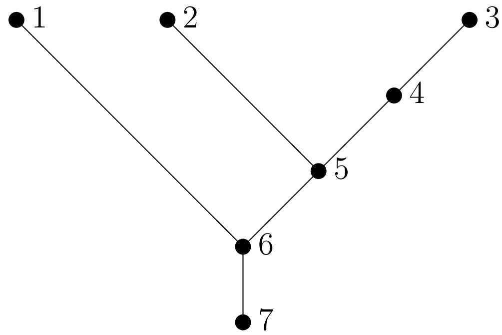

Figure 4.1: River structure from Example 4.2.1.

$\mathcal{U}$ is then called a *river system.* Notice that for a river with a single spring, and agents numbered successively from upstream to downstream by 1 to n, the river system $(N,\mathcal{U})$ is given by $U^1 = \emptyset$ and $U^k = \{k-1\}$ for $k \in \{2,\ldots,n\}$.

Example 4.2.1 Let $(N,\mathcal{U})$ represent a river system with $N = \{1,2,3,4,5,6,7\}$ and $U^1 = U^2 = U^3 = \emptyset$, $U^4 = \{3\}$, $U^5 = \{2,4\}$, $U^6 = \{1,5\}$ and $U^7 = \{6\}$, see Figure 4.1. So, $O = \{1,2,3\}$ is the set of springs and $n = 7$ is the (unique) most downstream agent. The two rivers originating at 2 and 3 merge together at agent 5 and then this stream merges together at agent 6 with the tributary originating at agent 1.

□

For $k \in N$, let UP^k denote the set of all agents upstream of k, including agent k itself. Clearly, (1) $U^k \subseteq UP^k \setminus \{k\}$ for every $k \in N$, (2) if $k \in O$ then $UP^k = \{k\}$, and (3) $UP^n = N$. Further, let $N_k = N \setminus UP^k$, i.e., N_k is the complement of the set UP^k consisting of the set of agents not in UP^k. Thus, N_k contains all agents downstream of agent k, all springs $j \in O \setminus UP^k$ that are not upstream of k, and all agents downstream of these springs. Notice that for every agent $k \in N$, both the $|UP^k|$-tuple $(U^i \cap UP^k)_{i\in UP^k} = (U^i)_{i\in UP^k}$ and the $|N_k|$-tuple $(U^i \cap N_k)_{i\in N_k}$ also provide river structures. So, the pairs $(UP^k, \{U^i \mid i \in UP^k\})$ and $(N_k, \{U^i \cap N_k \mid i \in N_k\})$ are sub-river systems on the sets UP^k and N_k respectively . Finally, let DO_k denote the set of all agents downstream of agent k, including k itself, and for $k \neq n$, let d_k be the unique downstream neighbor of k. Taking $k = 5$ in Example 4.2.1, $UP^5 = \{2,3,4,5\}$, $N_5 = \{1,6,7\}$, $DO_5 = \{5,6,7\}$ and $d_5 = 6$. Further, the sub-river system $(N_5, (U^i \cap N_5)_{i\in N_5})$ is given by the agent set N_5 and the river structure $U^1 \cap N_5 = \emptyset$, $U^6 \cap N_5 = \{1\}$ and $U^7 \cap N_5 = \{6\}$.

To complete the description of a river benefit problem with multiple springs, let $e_i \geq 0$ be the inflow of water on the territory of agent $i \in N$ and assume that each agent has a quasi-linear utility function given by $u^i(x_i,t_i) = b_i(x_i) + t_i$, where $t_i \in \mathbb{R}$ is a monetary compensation to agent i, $x_i \in \mathbb{R}_+$ is the amount of water allocated to (consumed by) agent i, and $b_i : \mathbb{R}_+ \to \mathbb{R}$ is a function yielding the benefit $b_i(x_i)$ to agent i of the consumption x_i of water. We call the quadruple $(N,\mathcal{U},e,b)$, where $(N,\mathcal{U})$ is a river system, $e = (e_i)_{i\in N}$ and $b = (b_i)_{i\in N}$ a *river benefit problem with multiple springs.* Following Ambec and Sprumont (2002) we, for now, assume that Assumption 2.2.1 of Chapter 2 holds. In Section 4.4 we generalize this to Assumption 2.2.5 of Chapter 2.

Because of the unidirectionality of the water flow from upstream to downstream, in a river benefit problem with multiple springs the water inflow downstream of some agent can not be allocated to this agent. As before, every agent can be assigned at most the water inflow on the territories of itself and its upstream agents. Hence, given a river system with multiple springs, a vector $x \in \mathbb{R}^N_+$ is a water allocation only if it satisfies the feasibility restrictions

$$\sum_{i \in UP^j} x_i \leq \sum_{i \in UP^j} e_i, \; j \in N,$$

i.e., for every agent $j \in N$, the sum of the water allocations to agent j and all its upstream agents is at most equal to the sum of the inflows at j and all its upstream agents.

A compensation scheme and welfare distribution are as defined in Section 2.2. Remember that a welfare distribution is Pareto efficient if no water and no money is wasted. So, welfare distribution (y, t) is Pareto efficient if and only if $y \in \mathbb{R}^N_+$ solves the welfare maximization problem

$$\max_{x_1,\ldots,x_n} \sum_{i=1}^{n} b_i(x_i) \;\text{ s.t. } \sum_{i \in UP^j} x_i \leq \sum_{i \in UP^j} e_i, \; j \in N, \;\text{and } x_i \geq 0, \; i \in N \tag{4.1}$$

and the compensation scheme is *budget balanced*: $\sum_{i=1}^{n} t_i = 0$. By Assumption 2.2.1 of Chapter 2, the maximization problem (4.1) has at least one solution and every solution x^* yields the same maximum attainable (social) welfare $\sum_{i=1}^{n} b_i(x_i^*)$. In this chapter, let $V(N, \mathcal{U}, e, b)$ denote the maximum (social) welfare for river system $(N, \mathcal{U})$, vector of inflows $e \in \mathbb{R}^N_+$ and benefit functions $b = (b_i)_{i \in N}$. For a solution x^*, a Pareto efficient welfare distribution (x^*, t) yields payoffs (utilities)

$$z_i = b_i(x^*) + t_i, \; i \in N,$$

with sum of payoffs equal to the total welfare $V(N, \mathcal{U}, e, b)$.

As in Ambec and Sprumont (2002) the problem to find a 'fair' distribution of the Pareto efficient total welfare can be modeled by a TU-game (N, v). In this game, the worth $v(N)$ is given by $v(N) = V(N, \mathcal{U}, e, b)$. To define the worth of any other coalition $S \subseteq N$, consider the river system with multiple springs $(N, \mathcal{U})$. Because of Assumption 2.2.1 of Chapter 2, in such a system a coalition of agents S can only cooperate when (1) there exists a $k \in S$ such that $S \subseteq UP^k$, and (2) for every $i \in S \setminus \{k\}$, every agent in between i and k on a branch of the river is also in S. Condition (1) means that agents located on two different branches of the river can not cooperate if they do not have a common most downstream agent in their coalition. For instance, in Example 4.2.1 the two upstream branches $\{1\}$ and $\{3, 4\}$ cannot benefit from cooperation in the coalition $\{1, 3, 4\}$. Condition (2) generalizes the notion of a coalition of consecutive agents to the case of multiple springs: it implies that when agent j cooperates with an upstream agent i, every agent on the branch between i and j also cooperates. In Example 4.2.1, agents 2 and 6 can only cooperate when 5 also agrees. In this chapter, we say that a coalition S is *connected* when S satisfies (1) and (2). The worth $v(S)$ of a connected coalition S is given by

$$v(S) = \sum_{k \in S} b_k(x_k^S) \quad \text{where} \quad x^S = (x_k^S)_{k \in S} \quad \text{solves}$$

$$\max_{\{x_k | k \in S\}} \sum_{k \in S} b_k(x_k) \;\; \text{s.t.} \sum_{i \in UP^j \cap S} x_i \leq \sum_{i \in UP^j \cap S} e_i, \; j \in S, \; \text{and} \; x_i \geq 0, \; i \in S. \tag{4.2}$$

For any other (non-connected) coalition S, the worth $v(S)$ is equal to the sum of the worths of its maximally connected subsets, where a subset T of S is maximal connected in S if T is connected in the sub-river system $(S, (U^k \cap S)_{k \in S})$ and $T \cup \{\ell\}$ is not connected in this sub-river system for any $\ell \in S \setminus T$.

By definition the set UP^i is connected for every $i \in N$. But, for an agent $k \in N$ with at least two upstream neighbors (i.e., k is an agent at which at least two tributaries merge together) the set $UP^k \setminus \{k\}$ is not connected (in Example 4.2.1 the set $UP^5 \setminus \{5\} = \{2, 3, 4\}$ is not connected, but it contains two maximal connected subsets: $\{2\}$ and $\{3, 4\}$). For every $k \in N$ it holds that

$$v(UP^k \setminus \{k\}) = \sum_{j \in U^k} v(UP^j).$$

Taking the river system $(N, \mathcal{U})$ as given, we refer to the TU-game (N, v) derived above (with benefit functions satisfying Assumption 2.2.1 of Chapter 2) as the *river game with multiple springs*. We denote the collection of all river games with multiple springs on $(N, \mathcal{U})$ by $\mathcal{R}^{(N, \mathcal{U})}$.

To find a distribution of the Pareto efficient total welfare $v(N)$ it is now possible to apply any efficient solution from cooperative game theory. Notice that a (single-valued) solution f assigns payoff vector $z = f(v) \in \mathbb{R}^N$ to game $v \in \mathcal{R}^{(N, \mathcal{U})}$. When $\sum_{i=1}^n z_i = v(N)$, then z can be implemented by any Pareto efficient welfare distribution (x^*, t), with $t_i = z_i - b_i(x_i^*)$, $i \in N$.

4.3 Solutions for river games with multiple springs

In this section we introduce a class of efficient solutions for river games with multiple springs, based on the TIBS principle from international river law. Recall that the TIBS principle does not make any country along a river the legal owner of the river water. Instead, it states that the river water belongs to all the countries combined, no matter where it enters the river, and that each country has the right to 'a reasonable and equitable share' in the optimal (efficient) use of the water.

The TIBS principle explicitly requires 'optimal use of the water' from a river and thus requires that river water is allocated in such a way that the total welfare of the countries located along the river is maximized. Within the game-theoretic framework, outlined in the previous section, this straightforwardly leads to the following axiom to be satisfied by a solution f on $\mathcal{R}^{(N, \mathcal{U})}$.

Axiom 4.3.1 Efficiency
A solution f on the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$ is efficient if $\sum_{i\in N} f_i(v) = v(N)$ for every $v \in \mathcal{R}^{(N,\mathcal{U})}$.

In addition to the efficiency requirement, the TIBS principle stipulates that each country has the right to a 'reasonable and equitable share' in the optimal (efficient) use of the water. It is also possible to translate this aspect of the TIBS principle into an axiom for a solution f on $\mathcal{R}^{(N,\mathcal{U})}$. To do so, suppose that the agents along the river are cooperating according to the partition $P = \{UP^i,\ N_i\}$ for some agent $i \in N \setminus \{n\}$. So, agent i is cooperating with all its upstream agents in coalition UP^i, and all other agents are cooperating in the complement coalition N_i. This situation can occur, for instance, because agent i is not willing to transfer water to its unique downstream neighbor d_i. In this case, the agents in UP^i distribute their total welfare $v(UP^i)$ among themselves and the agents in N_i distribute $v(N_i)$ among themselves. The question that can then be asked is: how should the gain in welfare $v(N) - v(UP^i) - v(N_i)$ that is created by cooperation between the two coalitions UP^i and N_i be divided among the agents? Evidently, this question can be asked for each agent $i \in N \setminus \{n\}$. The TIBS principle provides an answer to the question. Let $\alpha \in \mathbb{R}^N_+$ with $\sum_{k\in N} \alpha_k = 1$ be a vector of weights, where $\alpha_k \geq 0$ is the weight of agent $k \in N$. Then the TIBS principle can be interpreted by saying that the gain in welfare, that is created by joining the two coalitions UP^i and N_i, should be divided among the two coalitions proportional to the sum of the weights of the agents in the coalitions. Denoting $\alpha_S = \sum_{i\in S} \alpha_i$ for every $S \subseteq N$, we thus require that:

$$\frac{\sum_{k\in UP^i} f_k(v) - v(UP^i)}{\sum_{k\in N_i} f_k(v) - v(N_i)} = \frac{\alpha_{UP^i}}{\alpha_{N_i}},$$

assuming that both α_{UP^i} and α_{N_i} are nonzero. This leads to the following fairness axiom (which also is valid in case some weights are zero) for efficient solutions on the class of river games with multiple springs.

Axiom 4.3.2 α-TIBS fairness
Let $\alpha \in \mathbb{R}^N_+$ be such that $\sum_{i\in N} \alpha_i = 1$. An efficient solution f on the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$ satisfies α-TIBS fairness if, for every $v \in \mathcal{R}^{(N,\mathcal{U})}$ and every $i \in N \setminus \{n\}$, it holds that

$$\alpha_{N_i}\left(\sum_{k\in UP^i} f_k(v) - v(UP^i)\right) = \alpha_{UP^i}\left(\sum_{k\in N_i} f_k(v) - v(N_i)\right). \tag{4.3}$$

In this interpretation of the TIBS principle the 'reasonable and equitable shares' in the principle are represented by the weights $\alpha \in \mathbb{R}^N_+$, $\sum_{i\in N} \alpha_i = 1$, in the axiom. Notice that although the TIBS principle speaks about countries having the right to 'a reasonable and equitable share' in the optimal use of the river water, it does not specify the shares and does not require the shares to be equal. We therefore allow any nonnegative weight vector α of which the components add up to one. Later we give special attention to the specific case in which all weights are equal.

Weighted hierarchical solutions

The introduction of the efficiency and α-TIBS fairness axioms allows us to find a class of 'fair' (according to the TIBS principle) solutions for river games with multiple springs $v \in \mathcal{R}^{(N,\mathcal{U})}$. We call this class of solutions the class of *weighted hierarchical solutions*. Each solution from the class of weighted hierarchical solutions assigns to every river game $v \in \mathcal{R}^{(N,\mathcal{U})}$ a weighted *hierarchical outcome*. The notion of hierarchical outcomes has been introduced by Demange (2004) in the context of games with (tree) graph structure (see Section 2.1). In this same context Herings, van der Laan and Talman (2008) propose the average of all hierarchical outcomes as the so-called *average tree solution* (see also Section 2.1). Béal, Rémila and Solal (2009, 2010) extend the average tree solution to the class of all weighted averages of the hierarchical outcomes, but this is done in a different way than in this section. In this chapter we further examine the class of weighted hierarchical solutions from the perspective of river games, without referring to the underlying graph-theoretical concepts.

Given a river game with multiple springs $v \in \mathcal{R}^{(N,\mathcal{U})}$, for any agent $i \in N$ and any agent j downstream of i, let j^i be the last agent before agent j on the river branch from i to j (with $j^i = i$ when $j = d_i$). Now, take some agent $i \in N$ and for every $k \in N$ consider the following payoff $t^i_k(v)$ (recall from the previous section that DO_i is the set of agents downstream of i, including i itself) given by

$$t^i_k(v) = \begin{cases} v(UP^k) - v(UP^k \setminus \{k\}) & \text{if } k \in N \setminus DO_i, \\ v(N_{k^i}) - v(N_{k^i} \setminus \{k\}) & \text{if } k \in DO_i \setminus \{i\}, \\ v(N) - v(UP^k \setminus \{k\}) - v(N_k) & \text{if } k = i. \end{cases} \tag{4.4}$$

The payoff vector $t^i(v)$ gives the hierarchical outcome of Demange (2004) when i is taken to be the 'top' agent in the hierarchy.[2] The set of agents $N \setminus DO_i$ consists of all agents upstream of i and all agents that are neither upstream nor downstream of i. For instance, in Example 4.2.1 the set $N \setminus DO_5$ consists of the agents 2, 3 and 4, upstream of 5, and agent 1, which is neither upstream nor downstream of 5. In the payoff vector $t^i(v)$ each agent k not in DO_i receives its marginal contribution to the coalition of agents UP^k consisting of this agent k and all its upstream agents. Since for an agent $k \in DO_i \setminus \{i\}$, downstream of agent i, k^i denotes the upstream neighbor on the branch from i to k, the agent k receives its marginal contribution to the set N_{k^i}. That is, its marginal contribution to the set of agents who can be reached from k by walking along the river without visiting its upstream neighbor k^i on the branch from i to k. Finally, top agent i receives the surplus $v(N) - v(UP^i \setminus \{i\}) - v(N_i)$. Notice that the sets $UP^i \setminus \{i\}$ of all agents upstream of i and the set N_i of all agents not in UP^i can not cooperate without i and thus $v(N \setminus \{i\}) = v(UP^i \setminus \{i\}) + v(N_i)$. This implies that $t^i_i(v) = v(N) - v(N \setminus \{i\})$, i.e., top agent i receives its marginal contribution to the grand coalition N, which is equal to the additional welfare that it generates by joining together the two coalitions $UP^i \setminus \{i\}$ and N_i.

It is also possible to consider agent i as the top agent in a hierarchy on the set of all agents as follows: for an agent $\ell \neq i$, let i_ℓ be the distance from i to ℓ in the river system

[2] Note that the top agent does not have to be a spring.

$(N,\mathcal{U})$, with the distance defined as the number of agents (including ℓ itself, but not i) that has to be visited when traveling along the river from i to ℓ. For example, taking $i = 4$ in the river system of Figure 4.1, the distance to agent 3 is one, the distance to agent 2 is two and the distance to agents 1 or 7 is equal to three. Now, let π be an ordering on N such that an agent $k \neq i$ is ordered before an agent $\ell \neq i$ if $i_k > i_\ell$ (and with agent i as the last agent). E.g., if $i = 4$ in Figure 4.1 the agents 1 and 7, with distance three, are ordered first and second (in arbitrary order); then the agents 2 and 6, with distance two, are ordered third and fourth (in arbitrary order); next the agents 3 and 5, with distance one, are ordered fifth and sixth (in arbitrary order) and finally agent 4 is ordered seventh and last. It then follows from the fact that for every S the worth $v(S)$ is equal to the sum of the worths of its maximally connected subsets, that $t^i(v)$ is the marginal vector of the game v with respect to such an order π. So, considering the river system as a hierarchy with agent i as the top agent, the agents are ordered successively according to their distance to the top and receive their marginal contribution to the coalition of preceding agents. Since the top agent receives everything that is not assigned to the other agents, every hierarchical outcome provides an efficient payoff vector.

Given a river game with multiple springs $v \in \mathcal{R}^{(N,\mathcal{U})}$, each agent $i \in N$ induces a hierarchical outcome $t^i(v)$. The total number of hierarchical outcomes is therefore equal to n. For every nonnegative vector $\alpha \in \mathbb{R}^N_+$, with $\sum_{i\in N} \alpha_i = 1$, the payoff vector $h^\alpha(v) \in \mathbb{R}^N$ given by

$$h^\alpha(v) = \sum_{i\in N} \alpha_i t^i(v) \tag{4.5}$$

is a *weighted hierarchical outcome* of $v \in \mathcal{R}^{(N,\mathcal{U})}$. This leads to the next definition of a weighted hierarchical solution on the class $\mathcal{R}^{(N,\mathcal{U})}$.

Definition 4.3.3 *A solution f on the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$ is a weighted hierarchical solution if there exists an $\alpha \in \mathbb{R}^N_+$ with $\sum_{i\in N} \alpha_i = 1$, such that for every $v \in \mathcal{R}^{(N,\mathcal{U})}$*

$$f_i(v) = h^\alpha_i(v), \text{ for all } i \in N.$$

We now show that every weighted hierarchical solution is the unique efficient solution that satisfies the corresponding α-TIBS fairness axiom. For this consider the following lemmas.

Lemma 4.3.4 *Given $v \in \mathcal{R}^{(N,\mathcal{U})}$ and an agent $j \in N$, consider the hierarchical outcome $t^j(v)$ and an agent $i \in N$. Then*
(1) $\sum_{k\in UP^i} t^j_k(v) = v(N) - v(N_i)$ if $j \in UP^i$, and
(2) $\sum_{k\in UP^i} t^j_k(v) = v(UP^i)$ if $j \in N_i$.

Proof.
(1) If $j \in UP^i$, then for every $\ell \in N_i$, agent i has to be passed when one travels along the river from j to ℓ. So, for every $\ell \in N_i$ the distance from j to ℓ is bigger than the distance

from j to i, and thus every agent $\ell \in N_i$ is ordered before i in every ordering π that results in the marginal vector $t^j(v)$. Hence, all agents of the set N_i are ordered before i and the total payoff $\sum_{\ell \in N_i} t^j_\ell(v)$ to these agents is equal to $v(N_i)$. From efficiency it then follows that the total payoff to the agents in the complementary set UP^i is equal to $v(N) - v(N_i)$.
(2) If $j \in N_i$ then $k \in N \setminus DO_j$ for every $k \in UP^i$. So every $k \in UP^i$ receives $t^j_k(v) = v(UP^k) - v(UP^k \setminus \{k\})$. Summing over all $k \in UP^i$ gives $\sum_{k \in UP^i} t^j_k(v) = v(UP^i)$. □

Lemma 4.3.5 *Let $\alpha \in \mathbb{R}^N_+$ be so that $\sum_{i\in N} \alpha_i = 1$. Then the solution h^α on the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$ satisfies α-TIBS fairness.*

Proof. Distinguish the following three cases:

Case 1. Consider an agent $i \in N$ such that $\alpha_{UP^i} > 0$ and $\alpha_{N_i} > 0$. Then

$$\begin{aligned}
\sum_{k\in UP^i} h^\alpha_k(v) - v(UP^i) &= \sum_{k\in UP^i}\sum_{j\in N} \alpha_j t^j_k(v) - v(UP^i)\\
&= \sum_{k\in UP^i}\sum_{j\in UP^i} \alpha_j t^j_k(v) + \sum_{k\in UP^i}\sum_{j\in N_i} \alpha_j t^j_k(v) - v(UP^i)\\
&= \sum_{j\in UP^i}\sum_{k\in UP^i} \alpha_j t^j_k(v) + \sum_{j\in N_i}\sum_{k\in UP^i} \alpha_j t^j_k(v) - v(UP^i)\\
&= \sum_{j\in UP^i} \alpha_j \sum_{k\in UP^i} t^j_k(v) + \sum_{j\in N_i} \alpha_j \sum_{k\in UP^i} t^j_k(v) - v(UP^i)\\
&= \alpha_{UP^i}(v(N) - v(N_i)) + \alpha_{N_i} v(UP^i) - v(UP^i)\\
&= \alpha_{UP^i}(v(N) - v(N_i) - v(UP^i)), \qquad (4.6)
\end{aligned}$$

where the first equality follows by definition of h^α, the fifth equality follows from (1) and (2) of Lemma 4.3.4, and the last equality follows since $\alpha_{UP^i} + \alpha_{N_i} = 1$.

In a similar way it holds that

$$\begin{aligned}
\sum_{k\in N_i} h^\alpha_k(v) - v(N_i) &= \sum_{k\in N_i}\sum_{j\in N} \alpha_j t^j_k(v) - v(N_i)\\
&= \sum_{k\in N_i}\sum_{j\in N_i} \alpha_j t^j_k(v) + \sum_{k\in N_i}\sum_{j\in UP^i} \alpha_j t^j_k(v) - v(N_i)\\
&= \sum_{j\in N_i}\sum_{k\in N_i} \alpha_j t^j_k(v) + \sum_{j\in UP^i}\sum_{k\in N_i} \alpha_j t^j_k(v) - v(N_i)\\
&= \sum_{j\in N_i} \alpha_j \sum_{k\in N_i} t^j_k(v) + \sum_{j\in UP^i} \alpha_j \sum_{k\in N_i} t^j_k(v) - v(N_i)\\
&= \alpha_{N_i}(v(N) - v(UP^i)) + \alpha_{UP^i} v(N_i) - v(N_i)\\
&= \alpha_{N_i}(v(N) - v(UP^i) - v(N_i)). \qquad (4.7)
\end{aligned}$$

From (4.6) and (4.7) it follows that

$$\frac{1}{\alpha_{UP^i}}\left(\sum_{k\in UP^i} h_k^\alpha(v) - v(UP^i)\right) = v(N) - v(N_i) - v(UP^i) = \frac{1}{\alpha_{N_i}}\left(\sum_{k\in N^i} h_k^\alpha(v) - v(N_i)\right),$$

which shows that the α-TIBS fairness condition (4.3) is satisfied in this case.

Case 2. Consider an agent $i \in N$ such that $\alpha_{UP^i} = 0$. Then $\alpha_{N_i} = 1$ and, since $\alpha_j > 0$ only if $j \in N_i$, it holds that

$$\sum_{k\in UP^i} h_k^\alpha(v) = v(UP^i),$$

showing that the α-TIBS fairness condition (4.3) is satisfied in this case.

Case 3. Consider an agent $i \in N$ such that $\alpha_{N_i} = 0$. Then $\alpha_{UP^i} = 1$ and, since $\alpha_j > 0$ only if $j \in UP^i$, it holds that

$$\sum_{k\in N_i} h_k^\alpha(v) = v(N_i),$$

showing that the α-TIBS fairness condition (4.3) is also satisfied in this case.

□

Given Lemmas 4.3.4 and 4.3.5 it is possible to state the following characterization result.

Theorem 4.3.6 *Let $\alpha \in \mathbb{R}_+^N$ be so that $\sum_{i\in N} \alpha_i = 1$. A solution f on the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$ satisfies efficiency and α-TIBS fairness if and only if it is the weighted hierarchical solution h^α.*

Proof. Since any hierarchical outcome provides an efficient payoff vector, every weighted hierarchical solution is efficient. Further, it follows from Lemma 4.3.5 that h^α satisfies α-TIBS fairness. It remains to show that the two axioms uniquely determine a solution.

Suppose that solution f satisfies the two axioms and let $v \in \mathcal{R}^{(N,\mathcal{U})}$ be a river game with multiple springs. Since equation (4.3) in Axiom 4.3.2 has to hold for every $i \neq n$, the α-TIBS fairness yields $n-1$ linear independent equations. Thus, together with the efficiency condition that $\sum_{i\in N} f_i(v) = v(N)$ there are n linear independent equations in the n unknown payoffs $f_i(v)$, $i \in N$. Hence, the payoffs are uniquely determined and it must hold that $f(v) = h^\alpha(v)$, for every $v \in \mathcal{R}^{(N,\mathcal{U})}$.

□

We conclude this subsection by considering core-stability of the weighted hierarchical solutions. As mentioned before, in case of a single spring it has been shown by Ambec and Sprumont (2000) that under Assumption 2.2.1 of Chapter 2 the river game is convex, and thus every marginal vector of a river game v belongs to the core of the game. Since, as argued above, every hierarchical outcome is a marginal vector of the game v, this means that every hierarchical outcome is in the core. Because the core is a convex set this, in turn, implies that every weighted hierarchical solution, assigning a weighted hierarchical outcome $h^\alpha(v)$ to every single spring river game v, is core-stable.

The next example shows that a river game with multiple springs does not have to be convex.

Example 4.3.7 Consider $(N, \mathcal{U}, e, b)$ with $N = \{1,2,3\}$, $U^1 = U^2 = \emptyset$, $U^3 = \{1,2\}$, vector of inflows $e = (30, 30, 0)$ and vector of benefit functions b such that $b_1(x_1) = 100x_1 - x_1^2$ for all $x_1 \in [1, 49]$, $b_2(x_2) = 100x_2 - x_2^2$ for all $x_2 \in [1, 49]$ and $b_3(x_3) = 200x_3 - 4x_3^2$ for all $x_3 \in [1, 24]$. Straightforward calculations give $v(\{1\}) = v(\{2\}) = 2100$, $v(\{3\}) = 0$, $v(\{1,2\}) = 4200$, $v(\{1,3\}) = v(\{2,3\}) = 3380$ and $v(\{1,2,3\}) = 5622\frac{2}{9}$. Since $v(\{1,3\}) + v(\{2,3\}) = 6760 > 5622\frac{2}{9} = v(\{1,2,3\}) + v(\{3\})$ it follows that v is not convex.

□

Although a river game with multiple springs does not have to be convex, it is superadditive, as can directly be seen from the maximization problem that determines $v(S)$, $S \subseteq N$. It also holds for every river game with multiple springs $v \in \mathcal{R}^{(N,\mathcal{U})}$ that $v(S \cup T) = v(S) + v(T)$ when $S, T \subset N$, $S \cap T = \emptyset$ and $S \cup T$ is not connected. It follows from these two facts, as has been shown by Demange (2004) within the framework of games on (tree) graph structures, that every hierarchical outcome is in the core. Since the core is a convex set, it follows that also every weighted hierarchical outcome is in the core of the game. This yields the following corollary.

Corollary 4.3.8 *Given a river game with multiple springs $v \in \mathcal{R}^{(N,\mathcal{U})}$, every weighted hierarchical solution of v is in the core of the game v.*

Following the interpretation of Ambec and Sprumont (2002), this corollary implies that every weighted hierarchical solution satisfies the ATS principle. To some extent, this might come as a surprise because the weighted hierarchical solutions were derived with the help of the TIBS principle which, in general, does not encompass the ATS principle (see Chapter 1). Looking closely at the setup of river games with multiple springs, however, reveals that this result is the consequence of the definition of the characteristic function v (as discussed in the previous chapter). Although stylized, the outcome that every weighted hierarchical solution is core-stable is encouraging because it shows that countries in an international river water allocation problem could be made better off by cooperating, while (some) countries still adhere to the ATS principle.

The vector of weights α

What can be concluded from the above is that it is possible to make the TIBS principle from international watercourse law 'operational' in the sense that it provides a water allocation and monetary compensation scheme for the countries along a river. We have done this by introducing the efficiency and α-TIBS fairness axioms for river games with multiple springs, which provide a precise formulation of the TIBS principle. If one accepts this formulation, only the weights of the countries in the vector α remain to be determined.

A simple solution would be to take equal weights for all countries (leading to the average of all hierarchical outcomes, see below), but it is also possible to consider the weights to be exogenously given (for example by existing power structures among countries or by the factors mentioned in Article V of the Helsinki rules and Article 6 of the UN convention, see Chapter 1). When one does not want to impose weights directly, it is even possible to make them the subject of negotiation between countries.

We now consider some specific vectors of weights $\alpha \in \mathbb{R}^N_+$, with $\sum_{i \in N} \alpha_i = 1$. Taking $\alpha_n = 1$, and thus $\alpha_i = 0$ for every $i \in N \setminus \{i\}$, gives $h^\alpha(v) = t^n(v)$. Since $j \in UP^n$ for every $j \neq n$ and $N_n = \emptyset$, the payoffs of this outcome as given in formula (4.4) reduce to

$$t^n_k(v) = v(UP^k) - v(UP^k \setminus \{k\}),\ k \in N.$$

Since, in case of a single spring, this is the downstream incremental solution of Ambec and Sprumont (2002), we call the weighted hierarchical solution h^α with $\alpha_n = 1$ the *generalized downstream incremental solution*. On the class $\mathcal{R}^{(N,\mathcal{U})}$ of river games with multiple springs the generalized downstream incremental solution can thus be characterized by efficiency and α-TIBS fairness with $\alpha_n = 1$. The gain in welfare that is created by joining the two coalitions UP^i and N_i, $i \in N$, is in the generalized downstream incremental solution fully allocated to coalition N_i. In other words, joining the coalitions UP^i and N_i has no effect on the (average) payoff of the agents in the upstream coalition UP^i.

According to Corollary 4.3.8, the generalized downstream incremental solution satisfies the core lower bounds for every river game with multiple springs. In fact, it is straightforward to verify that also in this case it is the unique solution that satisfies the core lower bounds and the aspiration upper bounds of Ambec and Sprumont (2002). However, as for the single spring case, the generalized downstream incremental solution has the undesirable property that all gains from cooperation between an upstream and a downstream coalition are distributed to the agents in the downstream coalition, while the agents in the upstream coalition control the water flows from upstream to downstream.

In case of a single spring the upstream incremental solution is the weighted hierarchical solution h^α with $\alpha_1 = 1$ (thus assigning all weight to the most upstream agent 1). This solution, in which all gains from cooperation between an upstream and a downstream coalition are distributed to the agents in the upstream coalition, was proposed as a counterpart of the downstream incremental solution. There is, however, no straightforward unique generalization of the upstream incremental solution to the case of multiple springs, because in general there is no unique most upstream agent in a river system with multiple springs $(N,\mathcal{U})$. One way to get a generalization of the upstream incremental solution could be to take the average of all hierarchical outcomes corresponding to the agents located at one of the springs (the average over all $t^j(v)$, $j \in O$), but there are many other possibilities.

Instead of generalizing the upstream incremental solution, we now consider the average of all hierarchical outcomes. For the specific case that all weights are equal, i.e., $\alpha_i = \frac{1}{n}$ for all $i \in N$, the α-TIBS fairness axiom yields the following equal weights axiom.

Axiom 4.3.9 Equal weights TIBS fairness
An efficient solution f on the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$ satisfies equal weights TIBS fairness if, for every $v \in \mathcal{R}^{(N,\mathcal{U})}$ and any $i \in N \setminus \{n\}$, it holds that

$$\frac{1}{|UP^i|}\left(\sum_{k \in UP^i} f_k(v) - v(UP^i)\right) = \frac{1}{|N_i|}\left(\sum_{k \in N_i} f_k(v) - v(N_i)\right). \tag{4.8}$$

The unique solution that satisfies efficiency and equal weights TIBS fairness is the hierarchical solution corresponding to $\alpha_i = \frac{1}{n}$ for all $i \in N$. In the following, we denote this average hierarchical solution by h^A. Notice that in the context of games on cycle-free graph structures the average of the hierarchical outcomes was introduced in Herings, van der Laan and Talman (2008) as the average tree solution (see Section 2.1) and characterized by the so-called component efficiency and component fairness axioms. A minor adjustment of these two axioms to the river games setting gives the efficiency and equal weights TIBS fairness axioms.[3]

We feel that the average hierarchical solution h^A is a good alternative for the downstream and upstream incremental solutions, for single stream river games, and the generalized downstream incremental solution, for river games with multiple springs. Like these solutions, the average hierarchical solution is in the core of the game $v \in \mathcal{R}^{(N,\mathcal{U})}$ (see Corollary 4.3.8). It thus satisfies the core lower bounds of Ambec and Sprumont (2002), reflecting the ATS principle. Furthermore, when for some $i \in N \setminus \{n\}$, the upstream coalition UP^i is going to cooperate with its complement N_i, then the equal weights TIBS fairness axiom implies that

$$\frac{1}{|UP^i|}\Big(\sum_{k \in UP^i} h^A_k(v) - v(UP^i)\Big) = \frac{1}{|N_i|}\Big(\sum_{k \in N_i} h^A_k(v) - v(N_i)\Big).$$

Thus, the welfare distribution according to the average hierarchical solution has the property that for every $i \in N \setminus \{n\}$, the welfare gain of cooperation between the upstream coalition UP^i and its complement coalition N_i is divided among the coalitions UP^i and N_i proportional to the number of agents in these two coalitions. Hence, in the average hierarchical solution the average welfare gain of an agent in UP^i is equal to the average welfare gain of an agent in N_i when UP^i and N_i decide to cooperate.

As a special case, consider the implications of this for single spring river games. Recall that when the agents are indexed from upstream to downstream $[i, j]$, $1 \leq i \leq j \leq n$, denotes the coalition of consecutive agents $\{i, \dots, j\}$. The equal weights TIBS fairness axiom then implies that the average hierarchical solution h^A satisfies for every $k \in N \setminus \{n\}$

$$\frac{1}{k}\Big(\sum_{\ell \in [1,k]} h^A_\ell(v) - v([1,k])\Big) = \frac{1}{n-k}\Big(\sum_{\ell \in [k+1,n]} h^A_\ell(v) - v([k+1,n])\Big).$$

Thus, the welfare distribution according to the average hierarchical solution has the property that for every $k \in N \setminus \{n\}$, the welfare gain of the cooperation of the upstream

[3] In Béal, Rémila and Solal (2009) the component fairness axiom of Herings, van der Laan and Talman (2008) is generalized to weighted component fairness for forest games. This generalization differs from α-TIBS fairness in that it assigns weights to so-called cones (of a tree) instead of individual players.

coalition $[1, k]$ and the downstream coalition $[k + 1, n]$ is divided among the coalitions $[1, k]$ and $[k + 1, n]$ proportional to the number of agents in these two coalitions. So, the average welfare gain of an agent in $[1, k]$ is equal to the average welfare gain of an agent in $[k + 1, n]$ when $[1, k]$ and $[k + 1, n]$ decide to cooperate. This respects the TIBS principle in the sense that for each $k \in N$, every agent in the coalition $[k, n]$ is entitled to a share in the optimal use of the water inflow at agent k.

For two-agent river games, equal weights TIBS fairness implies that the welfare gain of cooperation between the upstream agent and the downstream agent is equally divided between the two agents.

4.4 Weighted hierarchical solutions for river games with multiple springs and externalities

As explained in Chapter 2, Ambec and Ehlers (2008) have generalized the single spring river game of Ambec and Sprumont (2002) by replacing Assumption 2.2.1 by Assumption 2.2.5 (recall that Assumption 2.2.5 allows for satiable agents). It is possible to generalize the river games with multiple springs in the same way. For this, assume, without loss of generality (see Ambec and Ehlers, 2008), that $e_i \leq \hat{x}_i$ for all $i \in N$, where $\hat{x}_i$ is the satiation point of agent i (with $\hat{x}_i = \infty$ when b_i is strictly increasing).[4]

In Ambec and Ehlers (2008) the presence of satiable agents in the model could cause positive externalities for coalitions of consecutive agents $[i, j]$, $2 \leq i \leq j \leq n$. It is not difficult to see that in the context of rivers with multiple springs, the same is true for connected coalitions T: under Assumption 2.2.1, the worth of T follows from the maximization problem (4.2); yet, when all agents in T have satiation points, and it is profitable for agents upstream of T to transfer water to agents downstream of T, the worth of T may depend on the coalition formation of the agents outside T.

Recall from Section 2.1 that situations in which the worth of a coalition $S \subset N$ can depend on the coalition formation of agents outside S can be modeled by a PFF-game. Given a partition $P \in \mathcal{P}^N$ of the player set N, PFF-games assign a worth $w(S, P)$ to every pair (S, P) such that $S \in P$. For $S \in P$, the worth $w(S, P)$ denotes the maximum welfare that the agents in coalition S can guarantee themselves by cooperating, when the agents outside S form coalitions T, $T \in P \setminus \{S\}$.

For a river with one spring, Ambec and Ehlers (2008) provide an iterative procedure to find the worths $w(S, P)$, for every $P \in \mathcal{P}^N$ and every $S \in P$, of a river game with externalities. For rivers with multiple springs, it is an open question whether $w(S, P)$ is uniquely determined and if so, how to derive the worth. Nevertheless, we will show that the (weighted hierarchical) solutions of this chapter only depend on the worths $w(S, P)$ for $S = UP^k$ or $S = N_k$, $k \in N$. For these coalitions S the worths do not depend on the partition of the complement of S (see Theorem 4.4.2 below) and follow from solving the maximization problems (4.2). Given the river benefit problem with multiple springs

[4]If $e_i > \hat{x}_i$ then in any welfare maximization problem $e_i - \hat{x}_i$ can be considered as additional inflow on the territory of the downstream neighbor of i, d_i.

$(N,\mathcal{U},e,b)$, we denote the collection of all partition function form games w on $(N,\mathcal{U})$ with benefit functions satisfying Assumption 2.2.5 by $\mathcal{RE}^{(N,\mathcal{U})}$.

Recall from Section 2.1 that P_S is the partition of N where all agents outside S act as singletons, and that $v_*(S)$ is the non-cooperative core lower bound of S. For a connected coalition it holds that $v_*(S) = v(S)$, i.e., $v_*(S)$ is precisely the worth that S can obtain by solving the welfare maximizing problem (4.2). Now, let $v^*(S) = w(S, P^S)$, where $P^S = \{S,\ N \setminus S\}$ is the partition of N in which all agents outside S cooperate. The worth $v^*(S)$ is called the *cooperative core lower bound* of S. It is the amount that the agents in S can guarantee themselves when the agents outside S cooperate together in the coalition $N \setminus S$. Notice that $v_*(S)$ and $v^*(S)$ are defined for every S, not only for connected coalitions. Also notice that $v_*(N) = v^*(N) = w(N, \{N\}) = v(N)$ is the worth of the grand coalition N when all agents cooperate together.

The following results have been stated in Ambec and Ehlers (2008) for rivers with one spring and generalize straightforwardly to rivers with multiple springs.

Lemma 4.4.1 *Let $w \in \mathcal{RE}^{(N,\mathcal{U})}$. Then for any partition P of N:*
(1) $v_*(S) \leq w(S, P)$, for all $S \in P$,
(2) for any two different $S, T \in P$, $w(S, P) + w(T, P) \leq w(S \cup T, P')$ with $P' = (P \setminus \{S, T\}) \cup \{S \cup T\}$.

Notice that (1) also implies that $v_*(S) \leq v^*(S)$ and that (2) implies for every disjoint S and T that

$$v_*(S) + v_*(T) \leq v_*(S \cup T).$$

Hence, the worths $v_*(S)$, $S \subseteq N$, induce a superadditive TU-game (N, v_*). Note that in the case without externalities the above inequality holds with equality when S and T are two disjoint connected coalitions and $S \cup T$ is not connected. This does not need to be true in the river game with externalities. For instance when S is upstream of T, the union $S \cup T$ may benefit from transferring water from S to T.

For rivers with a single spring it is also stated in Ambec and Ehlers (2008) that for coalitions $S = [1, i]$, $i \in N$,

$$w([1, i], P) = v_*([1, i]) \text{ for every } P \text{ with } [1, i] \in P,$$

i.e., when coalition S consists of some agent i and all its upstream agents, then the worth of S does not depend on the partition of the agents outside S. Indeed, by definition, the worth of such an upstream coalition S does not depend on the behavior of the agents downstream of S. In case of a river with multiple springs, this result generalizes to every upstream coalition UP^k and its complement $N_k = N \setminus UP^k$, $k \in N$.

Theorem 4.4.2 *Let $w \in \mathcal{RE}^{(N,\mathcal{U})}$. Then, for every partition $P \in \mathcal{P}^N$ and $S \in P$,*

$$w(S, P) = v_*(S) \text{ if } S = UP^k \text{ or } S = N_k \text{ for some } k \in N.$$

Proof. Let $S = UP^k$ for some $k \in N$. Since UP^k consists of agent k and all its upstream agents, its worth $w(UP^k, P)$ does not depend on the partition $P \setminus \{UP^k\}$ of the agents outside UP^k, because the water inflows of any $T \in P \setminus \{UP^k\}$ cannot reach UP^k anyway. Hence $w(UP^k, P) = v_*(UP^k)$ for all P with $UP^k \in P$.

Next, consider $S = N_k$ for some $k \in N$. By definition of UP^k, agent k is the only agent in UP^k that is connected to an agent in N_k, namely to its unique downstream neighbor d_k. Further, by definition of UP^k, there are no agents in UP^k downstream of k. By the assumption that $e_\ell \leq \hat{x}_\ell$ for every $\ell \in N$, it follows that agent d_k never receives any water from k, independent of the partition of $UP^k = N \setminus N_k$. So, $w(N_k, P) = v_*(N_k)$ for all P with $N_k \in P$.

□

The worth of every coalition of type UP^k or N_k is called *externality-free*. That is, the worth of a coalition of such a type does not depend on the partition of the agents outside the coalition.

Next, we consider the application of weighted hierarchical solutions to the class $\mathcal{RE}^{(N,\mathcal{U})}$ of river games with multiple springs and externalities. Similar as in the case without externalities, a solution is efficient if it always fully distributes the worth of the grand coalition N (the worth $v(N) = w(N, \{N\})$, with $v(N)$ the solution of maximization problem (4.2) for $S = N$).

Axiom 4.4.3 Efficiency for river games with externalities
A solution f on the class of river games with multiple springs and externalities $\mathcal{RE}^{(N,\mathcal{U})}$ is efficient if it holds for any game $w \in \mathcal{RE}^{(N,\mathcal{U})}$ that $\sum_{i \in N} f_i(w) = w(N, \{N\})$.

Given $i \in N \setminus \{n\}$, let $P(i)$ denote the partition $\{UP^i, N_i\}$ and recall that the α-TIBS fairness axiom for games without externalities was obtained by considering the situation in which the agents along the river are cooperating according to the partition $P(i)$. That is, agent i is cooperating with all its upstream agents in coalition UP^i, while all other agents are cooperating in its complement coalition N_i. Given externalities, the agents in UP^i can obtain total welfare $w(UP^i, P(i))$ and the agents in N_i can obtain $w(N_i, P(i))$. Since, by Theorem 4.4.2, the worths of coalitions of type UP^i and N_i are externality-free, it follows that $w(UP^i, P(i)) = v_*(UP^i)$ and $w(N_i, P(i)) = v_*(N_i)$. This leads to the next α-TIBS fairness axiom for the class of river games with multiple springs and externalities.

Axiom 4.4.4 α-TIBS fairness for river games with externalities
Let $\alpha \in \mathbb{R}^N_+$ be such that $\sum_{i \in N} \alpha_i = 1$. An efficient solution f on the class of river games with multiple springs and externalities $\mathcal{RE}^{(N,\mathcal{U})}$ satisfies α-TIBS fairness if, for any $w \in \mathcal{RE}^{(N,\mathcal{U})}$ and any $i \in N \setminus \{n\}$, it holds that

$$\alpha_{N_i} \left(\sum_{k \in UP^i} f_k(v) - v_*(UP^i) \right) = \alpha_{UP^i} \left(\sum_{k \in N_i} f_k(v) - v_*(N_i) \right). \tag{4.9}$$

Notice that this is the same as for the no externality case, only the worths of the coalitions UP^i and N_i are replaced by their non-cooperative core lower bounds v_* in the partition function form game w. So, irrespective of externalities, this axiom states that for any agent $i \in N$ the gain in welfare, that is created by joining the two coalitions UP^i and N_i, should be divided among the two coalitions proportional to the sum of the weights of the agents in these two coalitions.

Similar as in the proof of Theorem 4.3.6 it follows that there is a unique solution that satisfies efficiency for river games with externalities and α-TIBS fairness for river games with externalities. Moreover, similar as in the proof of Lemma 4.3.5 it follows that, given a river game with multiple springs and externalities w, the weighted hierarchical solution of the associated TU-game (N, v_*) satisfies both these axioms. This shows the following theorem.

Theorem 4.4.5 *Let $\alpha \in \mathbb{R}^N_+$ be such that $\sum_{i \in N} \alpha_i = 1$. A solution f on the class of river games with multiple springs and externalities $\mathcal{RE}^{(N,\mathcal{U})}$ satisfies efficiency for river games with externalities and α-TIBS fairness for river games with externalities if and only if $f(w) = h^\alpha(v_*)$ for every $w \in \mathcal{RE}^{(N,\mathcal{U})}$.*

As before, we call the solutions as characterized in this theorem, weighted hierarchical solutions. Given weight vector α, with components adding up to one, the weighted hierarchical solution of a river game with multiple springs and externalities is equal to the weighted hierarchical solution of the associated TU-game (N, v_*).

We now discuss a number of properties that is satisfied by the weighted hierarchical solutions. For this, first consider the following definition.

Definition 4.4.6 *A solution f on the class of river games with multiple springs and externalities $\mathcal{RE}^{(N,\mathcal{U})}$ is externality-free if the payoffs in the solution only depend on the worths of the externality-free coalitions.*

It is not difficult to show that the weighted hierarchical solutions are externality-free. Given a river game w with externalities and only one spring, consider the hierarchical outcome $t^i(v_*)$. Then formula (4.4) reduces to

$$t^i_k(v_*) = \begin{cases} v_*([1,k]) - v_*([1,k-1]) & \text{if } k < i, \\ v_*(N) - v_*([1,k-1]) - v_*([k+1,n]) & \text{if } k = i, \\ v_*([k,n]) - v_*([k+1,n]) & \text{if } k > i. \end{cases} \tag{4.10}$$

Given agent $i \in N$, an agent upstream of agent i receives its marginal contribution in the TU-game (N, v_*) to the coalition consisting of this agent and all agents upstream of it. An agent downstream of agent i receives its marginal contribution to the coalition consisting of this agent and all agents downstream of it. Finally, agent i receives its marginal contribution to the grand coalition N, i.e., agent i receives the gain in welfare that is created by connecting the upstream coalition $[1, i-1]$ and the downstream coalition $[i+1, n]$. Formula (4.10) shows that in every hierarchical outcome the payoffs are fully determined by the worths $v_*(S)$ with S of either type $[1, j]$ or type $[j, n]$ for some $1 \leq j \leq n$, i.e., the payoffs are fully determined by the worths of the upstream coalitions $[1, j]$ and the downstream coalitions $[j, n]$, $j \in N$. The worths of all other coalitions don't affect the payoffs. Since $UP^i = [1, i]$ and $N_i = [i+1, n]$, for $i \in N \setminus \{n\}$, it follows from Theorem

4.4.2 that every coalition that appears in formula (4.10) is externality-free. So, in the case of a river with only one spring, each hierarchical outcome only depends on the worths of externality-free coalitions. This implies that every weighted hierarchical solution for river games with externalities but only one spring is externality-free.

Given a river game with externalities and multiple springs, for some $i \in N$, consider an agent $k \in N \setminus DO_i$. Then, according to formula (4.4) the payoff to k in $t^i(v_*)$ is given by

$$t^i_k(v_*) = v_*(UP^k) - v_*(UP^k \setminus \{k\}).$$

Since $v_*(UP^k \setminus \{k\}) = \sum_{\ell \in U^k} v_*(UP^\ell)$ it follows that the payoff $t^i_k(v_*)$ only depends on the worths of coalitions of type UP^j, $j \in N$. Now, consider an agent $k \in DO_i \setminus \{i\}$. Then, according to formula (4.4) the payoff to k in $t^i(v_*)$ is given by

$$t^i_k(v_*) = v_*(N_{k^i}) - v_*(N_{k^i} \setminus \{k\}).$$

Inspecting $v_*(N_{k^i} \setminus \{k\})$ reveals that $v_*(N_{k^i} \setminus \{k\}) = v_*(N_k) + \sum_{\ell \in U^k \setminus \{k^i\}} v_*(UP^\ell)$. Hence, every term in the above expression is either of type $v_*(UP^j)$ or of type $v_*(N_j)$. Finally, consider agent i. According to formula (4.4) $t^i_i(v_*)$ also only depends on the worths of coalitions of type UP^j and N_j. Since, by Theorem 4.4.2, every coalition of type UP^j or N_j is externality-free, it follows that every hierarchical outcome $t^i(v_*)$, $i \in N$, is externality-free. This shows that the following two corollaries hold.

Corollary 4.4.7 *On the class $\mathcal{RE}^{(N,\mathcal{U})}$ of river games with multiple springs and externalities, every weighted hierarchical solution h^α, assigning payoff vector $h^\alpha(v_*)$ to every $w \in \mathcal{RE}^{(N,\mathcal{U})}$, is externality-free.*

Corollary 4.4.8 *On the class $\mathcal{RE}^{(N,\mathcal{U})}$ of river games with multiple springs and externalities, the axioms of efficiency for river games with externalities and α-TIBS fairness for river games with externalities imply externality-freeness.*

These two corollaries also follow directly by observing that efficiency for river games with externalities and α-TIBS fairness for river games with externalities give n linear independent equations that only depend on the worths of coalitions of the form UP^j and N_j, $j \in N$.

Next, consider core stability of the weighted hierarchical solutions for river games with multiple springs and externalities. Every weighted hierarchical solution is a convex combination of the hierarchical outcomes $t^i(v_*)$, $i \in N$. As seen before, for a river game $v \in \mathcal{R}^{(N,\mathcal{U})}$ without externalities, every hierarchical outcome is in the core of the game v. But, as observed above, for two connected, disjoint coalitions S and T the worth $v_*(S \cup T)$ can be bigger than the sum of the two worths $v_*(S)$ and $v_*(T)$. This means that a hierarchical outcome $t^i(v_*)$ doesn't need to satisfy the non-cooperative core lower bound $v_*(R)$ for every coalition $R \subseteq N$. This, in turn, implies that the weighted hierarchical solution h^α might not be core stable.

Ambec and Ehlers (2008) argue that in river games it is natural to restrict blocking (of an agreement) to connected coalitions, because coordination among agents becomes difficult when agents are not neighboring. Clearly, every hierarchical outcome satisfies the non-cooperative core lower bound for every connected coalition S. Hence, the following corollary holds.

Corollary 4.4.9 *For a river game with multiple springs and externalities $w \in \mathcal{RE}^{(N,\mathcal{U})}$, the weighted hierarchical solution $h^{\alpha}(v_*)$ satisfies the non-cooperative core lower bounds when blocking is restricted to connected coalitions.*

Ambec and Ehlers (2008) show that for the river game with a single spring and externalities, the downstream incremental solution (i.e., the hierarchical outcome $t^n(v_*)$) satisfies all non-cooperative core lower bounds. It is an open question whether this also holds for the generalized downstream incremental solution for river games with multiple springs and externalities.

As a final remark, note that also on the class of river games with multiple springs and externalities the average hierarchical solution satisfies equal weights TIBS fairness. When the coalition UP^i does not cooperate with its complement N_i for some $i \in N$, then the total payoff to the agents in UP^i is equal to $v_*(UP^i)$ and the total payoff to the agents in N_i is equal to $v_*(N_i)$. When the two coalitions, UP^i and N_i, then decide to cooperate and distribute the worth of the grand coalition according to the average hierarchical solution, the average welfare gain of the agents in UP^i is equal to the average welfare gain of the agents in N_i.

Chapter 5

A strategic implementation of the weighted hierarchical solutions

5.1 Introduction

In the previous chapter we introduced the class of weighted hierarchical solutions for river games with multiple springs (and externalities). The derivation of this class of solutions was based entirely on cooperative arguments. The goal was to model the outcome of negotiations between countries sharing a river; or, stated more normatively, to find agreements on river water distribution that have a good chance of being accepted by all the countries located along a river. As in Ambec and Sprumont (2002), we argued that a sustainable (stable and fair) welfare distribution can be reached through an efficient water allocation and appropriate compensation scheme. The compensation scheme that we suggested followed from a fairness axiom based on the TIBS principle from international watercourse law.

While introducing the weighted hierarchical solutions, we did not consider how the countries along a river might reach an agreement that implements a weighted hierarchical solution. The aim of this chapter is to do precisely this. So, we want to find out whether it is possible for the agents along a river with multiple springs to reach the weighted hierarchical solution h^α through non-cooperative behavior. This chapter is thus part of the Nash program in the sense that we try to establish a non-cooperative foundation for a cooperative solution concept.

Stated differently, the question we would like to answer in this chapter is whether it is possible to find a non-cooperative framework that, as the result of equilibrium behavior by the agents, leads to the weighted hierarchical solution h^α. In answering this question we will make use of the non-cooperative, procedural, arguments from (game-theoretic) implementation theory. The paper of Pérez-Castrillo and Wettstein (2001) is one of the central contributions to this theory. In it the authors propose a strategic implementation of the Shapley value for zero-monotonic transferable utility games in characteristic function form. That is, they propose a non-cooperative game (mechanism) of which the unique subgame perfect Nash equilibrium (SPNE) payoffs correspond to the Shapley value payoffs of a zero-monotonic cooperative game.

In the mechanism of Pérez-Castrillo and Wettstein (2001) each player, in a first stage, makes a bid to all other players to become the proposer in the second stage of the mechanism. The player with the highest net bid (the sum of the bids made by this player to the other players minus the sum of the bids made by the other players to this player) pays its bids and becomes the proposer in the second stage. The proposer then makes a proposal to each of the other players about a division of the surplus of cooperation. If the proposal is accepted by all players, the proposer pays the proposed amount to each of the other players and keeps the remains of the surplus of cooperation itself. If the proposal is rejected by a least one of the players, the proposer leaves the mechanism and the mechanism restarts with the remaining players.

This mechanism of Pérez-Castrillo and Wettstein (2001) has been adapted by many authors to implement other solution concepts from cooperative game theory. Pérez-Castrillo and Wettstein (2002) construct a mechanism to obtain efficient outcomes in a general social choice network; Vidal-Puga and Bergantiños (2003) implement the Owen value for games with coalition structures; Mutuswami, Pérez-Castrillo and Wettstein (2004) implement efficient outcomes in local public goods environments; Pérez-Castrillo and Wettstein (2005) construct a mechanism to form efficient networks; Macho-Stadler, Pérez-Castrillo and Wettstein (2006) implement a Shapley-type outcome in games with externalities; Slikker (2007) implements the position value, Myerson value and the component-wise egalitarian value for games with graph structure; Ju and Wettstein (2009) discuss a generalized bidding approach for several values; van den Brink and Funaki (2010) implement discounted Shapley values; and van den Brink, Funaki and Ju (2011) implement egalitarian Shapley values.

In this chapter we adapt the mechanism of Pérez-Castrillo and Wettstein (2001) to implement the weighted hierarchical solutions on the class of zero-monotonic tree games. Since a river game with multiple springs can be seen as a tree game, and because each river game with multiple springs is superadditive (and therefore also zero-monotonic), it follows that this also gives an implementation of the weighted hierarchical solutions on the class of river games with multiple springs.

Given a vector of weights $\alpha \in \mathbb{R}^N$, with $\sum_{i \in N} \alpha_i = 1$, adapting the mechanism of Pérez-Castrillo and Wettstein (2001) results in a non-cooperative mechanism of which the unique SPNE payoffs correspond to the payoffs in the weighted hierarchical solution h^α of a zero-monotonic tree game. Similar as in Pérez-Castrillo and Wettstein (2001), in a first stage of this mechanism each player makes a bid to all other players to become the proposer in the second stage of the mechanism and the player with the highest (weighted) net bid pays its bids and becomes the proposer in the second stage. The proposer then makes a proposal to each of the other players about a division of the surplus of cooperation. If the proposal is accepted by all players, the proposer pays the proposed amount to each of the other players and keeps the remains of the surplus of cooperation itself. If the proposal is rejected, the proposer leaves the mechanism. Next, the graph structure is taken into account. Different from the mechanism in Pérez-Castrillo and Wettstein (2001), after leaving, each of the neighbors of the original proposer in the graph automatically becomes a proposer in the next round of the mechanism. In this next round, each proposer makes a proposal about the division of the surplus of cooperation to the players in its component (the graph now possibly has multiple components because the original proposer has left the mechanism). If the proposal of a proposer in a particular component is accepted by all the players in the component, the proposer pays the proposed amount to each of the other players in the component and keeps the remains of the surplus of cooperation of the component; all players in the component leave the mechanism. If the proposal of a proposer in a particular component is rejected by at least one of the players in the component, the proposer in the component leaves the mechanism and the neighbors of this proposer become the proposer in their own (sub)components in the next round. The mechanism continues until either all players in all components have accepted their proposals or if there are no players left that the remaining proposers can propose a division of a surplus of cooperation to.

What should be observed is that in this mechanism a player is only able to communicate with other players (i.e., to make proposals about a division of the surplus of cooperation) when they are connected in the graph. This immediately highlights the main difference between the mechanism of this chapter and that of Pérez-Castrillo and Wettstein (2001). In their mechanism there is bidding in every round on who will be the proposer, whereas in the mechanism of this chapter, once a first proposer has been established in the bidding stage, the order in which the players are allowed to make proposals is fixed and given by the structure of the graph. Also, in the mechanism of this chapter it is possible that there are multiple proposers in one round (one for each component) while in the mechanism of Pérez-Castrillo and Wettstein (2001) there is a unique proposer in each round. Nevertheless, because the mechanism of this chapter is an adaptation of the mechanism of Pérez-Castrillo and Wettstein (2001) it has the same desirable features. Namely, (1) the players receive the weighted hierarchical solution payoff in every equilibrium outcome, not only in expectation, (2) there is no a priori randomization that imposes an order on the moves of the players (only the graph (river) imposes such an order), (3) the game is finite and (4) the equilibrium strategies are simple, moreover they are unique when the cooperative graph game is strictly zero-monotonic.

This chapter is based on van den Brink, van der Laan and Moes (2012b) and is organized as follows. In Section 5.2 we give an implementation of the hierarchical outcomes of Demange (2004) (see Section 2.1) on the class of zero-monotonic tree games. In Section 5.3 we use this implementation of the hierarchical outcomes to obtain an implementation of the weighted hierarchical solution on the class of zero-monotonic tree games. Finally, in Section 5.4 we discuss how the implementation of the weighted hierarchical solution on the class of zero-monotonic tree games can be used to say something about an implementation of the same solution on the class of river games with multiple springs (and externalities).

5.2 An implementation of the hierarchical outcomes

In this section we give an implementation of the hierarchical outcomes of Demange (2004) on the class of zero-monotonic tree games. Recall from Section 2.1 that we denote the hierarchical outcomes of Demange (2004) for a tree game (N, v, L) by $h^i(N, v, L)$, $i \in N$. Given a vector of weights $\alpha \in \mathbb{R}^N_+$ with $\sum_{i \in N} \alpha_i = 1$, this implies that we can define the weighted hierarchical solution h^α on the class of tree games as:

$$h^\alpha(N,v,L) = \sum_{i\in N} \alpha_i h^i(N,v,L).$$

Observe that, in particular, $h^\alpha(N,v,L) = AT(N,v,L)$ if $\alpha_i = \frac{1}{|N|}$ for all $i \in N$.

To give a strategic implementation of the weighted hierarchical solution h^α on the class of zero-monotonic tree games, we first need to consider an implementation of the hierarchical outcomes $h^i(N,v,L)$, $i \in N$, on the same class. In this section we thus propose for the class of zero-monotonic tree games a non-cooperative game, called a *mechanism*, of which the SPNE payoffs correspond to the payoffs of the hierarchical outcome h^i, $i \in N$. Hence, for a graph game (N,v,L) with (N,v) zero-monotonic and (N,L) a tree, given a player $i \in N$, the mechanism implements the hierarchical outcome $h^i(N,v,L)$ in SPNE.

As pointed out in Section 2.1, the hierarchical outcome $h^i(N,v,L)$ is equal to the marginal vector of the graph restricted game (N,v^L) for any ordering that is consistent with the rooted tree (N,L^i). We therefore first propose for the class of zero-monotonic TU-games a mechanism of which the SPNE payoffs correspond to the payoffs of a marginal vector. For a zero-monotonic game (N,v) and given ordering $\pi \in \Pi^N$, this mechanism implements the marginal vector $m^\pi(N,v)$ in SPNE. We call this Mechanism A.

Unlike in the mechanism of Pérez-Castrillo and Wettstein (2001) that implements the Shapley value, in Mechanism A there is no bidding procedure. Instead, the order in which the players make proposals about a division of the surplus of cooperation is fixed and given by the reverse order of the rank numbers of the players in ordering π. For given ordering $\pi : N \to \{1,\dots,n\}$, let $\rho : \{1,\dots,n\} \to N$ be defined by $\rho(k) = \pi^{-1}(n+1-k)$. So, $\rho(k) \in N$ is the player with rank number $n+1-k$.

Given a zero-monotonic TU-game (N,v) and an ordering $\pi \in \Pi^N$, in each *round* of Mechanism A the proposal procedure consists of three *stages*. In stage 1 of round 1, player $\rho(1)$ is assigned to be the *proposer* and makes a *proposal* about the division of $v(N)$, i.e., the player proposes a payoff to every other player in the game. In stage 2 of round 1, all players, except the proposer $\rho(1)$, sequentially either accept or reject the proposal.[1] The proposal is accepted if all players accept, otherwise it is rejected. In stage 3 of round 1, if the proposal is accepted the accepting players receive the proposed payoffs, the proposer receives $v(N)$ minus the proposed payoffs to the other players, and the mechanism ends. If the proposal is rejected, the proposer leaves the mechanism and receives its singleton worth $v(\{\rho(1)\})$, while the other players go to the next round to bargain over the worth $v(N \setminus \{\rho(1)\})$. This second round has the same three stages as the first round (but with $n-1$ players) and starts with player $\rho(2)$ with rank $n-1$ (the highest rank under the remaining players) as the proposer. In general, in round $k \in \{1,\dots,n-1\}$ the mechanism proceeds with player $\rho(k)$ as the proposer, until either at some round all remaining players accept the offer by the proposer or round $n = |N|$ is reached in which only player $\rho(n)$ is left, who just receives its singleton worth $v(\{\rho(n)\})$.

To describe Mechanism A formally, let $N^t = N \setminus \bigcup_{k=1}^{t-1}\{\rho(k)\} = \bigcup_{k=t}^{n}\{\rho(k)\}$ be the player set at the start of each round t, $t \in \{1,\dots,n\}$. Observe that $N^{t+1} = N^t \setminus \{\rho(t)\}$.

[1] For instance, in (reverse) order of the ordering, but this is not essential.

Mechanism A to implement the marginal vector $m^{\pi}(N,v)$ of TU-game (N,v)

Initiation: Let $\rho(k) = \pi^{-1}(n+1-k)$, $k \in N$. Set $t = 1$ and go to Stage 1 of round 1.

Stage 1: If $t = n$, the mechanism ends and player $\rho(n)$ receives its stand-alone worth $v(\{\rho(n)\})$. If $t < n$, player $\rho(t)$ proposes an offer $y_j^{\rho(t)} \in \mathbb{R}$ to every player $j \in N^{t+1}$. Go to Stage 2.

Stage 2: The players in N^{t+1}, sequentially (according to ρ), either accept or reject the offer. If all players accept, then the proposal is accepted. If at least one player rejects, the proposal is rejected. Go to Stage 3.

Stage 3: If the proposal is accepted, then each player $j \in N^{t+1}$ receives $y_j^{\rho(t)}$, player $\rho(t)$ obtains the remainder $v(N^t) - \sum_{j \in N^{t+1}} y_j^{\rho(t)}$ and the mechanism ends.
If the proposal is rejected, then player $\rho(t)$ leaves the mechanism and obtains its stand-alone worth $v(\{\rho(t)\})$. In the next round N^{t+1} is the set of players, who bargain over $v(N^{t+1})$. Set t equal to $t+1$ and return to Stage 1.

The next theorem states that, given $\pi \in \Pi^N$, Mechanism A implements the marginal vector $m^{\pi}(N,v)$ as a SPNE payoff vector if the TU-game (N,v) is zero-monotonic. The proof is similar to the proof of the SPNE payoffs in the proposal subgame of the mechanism given in Pérez-Castrillo and Wettstein (2001).

Theorem 5.2.1 *Let $(N,v) \in \mathcal{G}$ be a zero-monotonic TU-game and let $\pi \in \Pi^N$ be given. The payoff vector in any SPNE of Mechanism A coincides with the payoff vector $m^{\pi}(N,v)$.*

Proof. We first introduce some notation. For $t \in \{1,\ldots,n\}$ we denote $\pi_t : N^t \to \{1,\ldots,n+1-t\}$ as $\pi_t(k) = \pi(k)$ for all $k \in N^t$, $\pi_t^i = \{j \in N^t \mid \pi_t(j) \leq \pi_t(i)\}$, $i \in N^t$, and $m_i^{\pi_t}(N^t, v_{N^t}) = v_{N^t}(\pi_t^i) - v_{N^t}(\pi_t^i \setminus \{i\})$, $i \in N^t$. That is, $m_i^{\pi_t}(N^t, v_{N^t})$ is the marginal vector on the subgame (N^t, v_{N^t}) with respect to the ordering π_t. Notice that for $i \in N^t$, $m_i^{\pi_t}(N^t, v_{N^t}) = v_{N^t}(\pi_t^i) - v_{N^t}(\pi_t^i \setminus \{i\}) = v(\pi_t^i) - v(\pi_t^i \setminus \{i\}) = v(\pi^i) - v(\pi^i \setminus \{i\}) = m_i^{\pi}(N,v)$. Hence, for every $i \in N^t$, the payoff of i in $m^{\pi_t}(N^t, v_{N^t})$ is equal to the payoff of i in the marginal vector $m^{\pi}(N,v)$ on the TU-game (N,v).

We now show that the marginal vector payoffs are indeed equilibrium payoffs of Mechanism A. Consider the strategies in which in any round $t < n$ in Stage 1, the proposer $\rho(t)$ offers $y_j^{\rho(t)} = m_j^{\pi}(N,v)$ to every player $j \in N^t \setminus \{\rho(t)\}$ and in Stage 2 the players $j \in N^t \setminus \{\rho(t)\}$ accept any offer greater than or equal to $m_j^{\pi}(N,v)$ and reject any offer strictly smaller than $m_j^{\pi}(N,v)$. These strategies constitute a SPNE. Clearly, the strategy at Stage 1 is a best response for the proposer as long as $v(N^t) - \sum_{j \in N^t \setminus \{\rho(t)\}} m_j^{\pi}(N,v) \geq v(\{\rho(t)\})$. Since $\sum_{j \in N^t \setminus \{\rho(t)\}} m_j^{\pi}(N,v) = v(N^T \setminus \{\rho(t)\})$, this holds by the definition of zero-monotonicity. Also, at Stage 2 the strategy of every player $j \in N^t \setminus \{\rho(t)\}$ is a best response as long as $m_j^{\pi}(N,v) = v(\pi^j) - v(\pi^j \setminus \{j\}) \geq v(\{j\})$, which again follows directly from the definition of zero-monotonicity.

Next, we show by induction on the number of players that the payoffs in any SPNE of Mechanism A are equal to the marginal vector payoffs $m_i^{\pi}(N,v)$, $i \in N$. For a game

(N,v) with $|N| = n = 1$ it holds that in the first round $t = 1 = n$ and so the single player $i \in N$ immediately receives its stand-alone worth $v(\{i\})$, which is equal to its marginal vector payoff $m_i^\pi(N,v)$. Hence, any SPNE yields the marginal vector $m^\pi(N,v)$. Now, suppose that for every (N^k, v_{N^k}) with $k \geq t+1$, (thus $|N^k| = n+1-k \leq n-t$) and given ordering π, every SPNE of Mechanism A implements the marginal vector $m_k^\pi(N^k, v_{N^k})$. Then it can be shown that at Stage 2 of round t all players $j \in N \setminus \{\rho(t)\}$ accept the offer if $y_j^{\rho(t)} > m_j^\pi(N,v)$, while the offer is rejected if $y_j^{\rho(t)} < m_j^\pi(N,v)$ for at least one $j \in N^t \setminus \{\rho(t)\} = N^{t+1}$. In case of rejection, player $\rho(t)$ leaves with its singleton worth and by the induction argument, the payoff to a player $j \in N^{t+1}$ is $m_j^{\pi_{t+1}}(N^{t+1}, v_{N^{t+1}}) = m_j^\pi(N,v)$. If in Stage 2 of round t the (last) player $\rho(n)$ is reached, its optimal strategy is to accept any offer higher than $m_{\rho(n)}^{\pi_{t+1}}(N^{t+1}, v_{N^{t+1}}) = m_{\rho(n)}^\pi(N,v)$ and to reject any offer lower than $m_{\rho(n)}^{\pi_{t+1}}(N^{t+1}, v_{N^{t+1}}) = m_{\rho(n)}^\pi(N,v)$. The second to last player $\rho(n-1)$ anticipates the reaction of player $\rho(n)$. So, if $y_{\rho(n)}^{\rho(t)} > m_{\rho(n)}^\pi(N,v)$ and in Stage 2 of round t player $\rho(n-1)$ is reached, then this player accepts the offer if $y_{\rho(n-1)}^{\rho(t)} > m_{\rho(n-1)}^{\pi_{t+1}}(N^{t+1}, v_{N^{t+1}}) = m_{\rho(n-1)}^\pi(N,v)$ and rejects the offer if $y_{\rho(n-1)}^{\rho(t)} < m_{\rho(n-1)}^{\pi_{t+1}}(N^{t+1}, v_{N^{t+1}}) = m_{\rho(n-1)}^\pi(N,v)$. If $y_{\rho(n)}^{\rho(t)} < m_{\rho(n)}^\pi(N,v)$, then player $\rho(n-1)$ is indifferent between accepting or rejecting any offer $y_{\rho(n-1)}^{\rho(t)}$, because player $\rho(n)$ is going to reject the offer $y_{\rho(n)}^{\rho(t)}$ anyway. By backwards induction it follows that, for every $t \in \{1,\ldots,n\}$ and every SPNE, at Stage 2 of round t all players $j \in N^{t+1}$ accept the offer of player $\rho(t)$ if $y_j^{\rho(t)} > m_j^\pi(N,v)$ and that the offer is rejected if $y_j^{\rho(t)} < m_j^\pi(N,v)$ for at least one $j \in N^{t+1}$.

For a round t, we now consider two cases. First, if $v(N^t) > v(N^{t+1}) + v(\{\rho(t)\})$, it follows that the strategies described in the first part of the proof are the only SPNE strategies in the subgame that starts at Stage 1 of round t. To see this notice that in this case rejection of the proposal made by player $\rho(t)$ can not be part of an SPNE because then player $\rho(t)$ would receive $v(\{\rho(t)\})$. Player $\rho(t)$ can improve on this payoff by taking $0 < \epsilon < v(N^t) - v(N^{t+1}) - v(\{\rho(t)\})$ and offering $m_j^{\pi_t}(N^t, v_{N^t}) + \frac{\epsilon}{|N^t|-1} > m_j^\pi(N,v)$ to every $j \in N^{t+1}$. Hence, an SPNE requires acceptance of the offers in Stage 2. Since any proposal with $y_j^{\rho(t)} < m_j^\pi(N,v)$ for some $j \in N^{t+1}$ is rejected, an SPNE also requires that $y_j^{\rho(t)} \geq m_j^\pi(N,v)$ for all $j \in N^{t+1}$. On the other hand, any proposal such that $y_\ell^{\rho(t)} > m_\ell^\pi(N,v)$ for some $\ell \in N^{t+1}$ can not be part of an SPNE, because then $\rho(t)$ could improve on its payoff by taking $0 < \epsilon < y_\ell^{\rho(t)} - m_\ell^\pi(N,v)$ and offering $m_j^\pi(N,v) + \frac{\epsilon}{|N^t|-1} > m_j^\pi(N,v)$ to every $j \in N^{t+1}$. It can be concluded that in any SPNE in the subgame that starts in round t it must hold that $y_j^{\rho(t)} = m_j^\pi(N,v)$ for all $j \in N^{t+1}$ and that these offers are accepted.

Second, we consider the case that $v(N^t) = v(N^{t+1}) + v(\{\rho(t)\})$ and thus $m_{\rho(t)}^\pi = v(N^t) - v(N^{t+1}) = v(\{\rho(t)\})$. As in the previous case, the strategies described in the first part of the proof are SPNE strategies in the subgame that starts at Stage 1 of round t. In addition, also any strategy profile in which at stage 1 of round t, player $\rho(t)$ offers $y_j^{\rho(t)} \leq m_j^\pi(N,v)$ to some players $j \in N^{t+1}$ and, at stage 2, these players j reject any offer $y_j^{\rho(t)} \leq m_j^\pi(N,v)$, constitutes an SPNE. In this SPNE the proposer receives its own worth $v(\{\rho(t)\})$.

In both cases it follows that every player $j \in N^t$ receives $m^{\pi}_j(N,v)$ in every SPNE of the subgame that starts in round t. So, given $(N,v) \in \mathcal{G}$ zero-monotonic and $\pi \in \Pi^N$, every SPNE of Mechanism A yields payoff vector $m^{\pi}(N,v)$.

□

Given this theorem, we are able to propose a mechanism for the class of zero-monotonic tree games of which the SPNE payoffs correspond to the payoffs of the hierarchical outcome h^i, $i \in N$. Mechanism B, below, is a modification of Mechanism A to implement the hierarchical outcome $h^i(N,v,L)$ on the class of zero-monotonic tree games. The order in which the players in Mechanism B are allowed to make proposals, and accept or reject proposals, is given by the directed graph (N,L^i) (instead of the ordering π). Recall that, for $i \in N$, (N,L^i) is the rooted tree with root i induced by (N,L), S^i_j is the set of successors of $j \in N$ in (N,L^i) and $\widehat{S}^i_j$ is the set containing j itself and all its subordinates in (N,L^i).

Given a player $i \in N$, in the first round of Mechanism B the root i in the rooted tree (N,L^i) is the proposer in Stage 1. If its proposal about a division of $v(N)$ is accepted by all the players in $N \setminus \{i\}$, the mechanism ends immediately. Each player $k \in N \setminus \{i\}$ receives the proposed payoff y^i_k and player i receives the remaining surplus of cooperation $v(N) - \sum_{k \in N\setminus\{i\}} y^i_k$. If the proposal of player i is rejected, player i leaves the mechanism with its stand-alone worth $v(\{i\})$.

For $S \subseteq N$, let $(S, D(S))$ with $D(S) = \{(i,j) \in D \mid i,j \in S\}$ be the *directed subgraph* of a directed graph (N,D) on S. For each of the successors of player i in (N,L^i), $j \in S^i_i$, the set $\widehat{S}^i_j$ is a component in the subgraph $(N \setminus \{i\}, L(N \setminus \{i\}))$. In the rooted tree (N,L^i) each of these components $\widehat{S}^i_j$ induces a rooted subtree $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$ with player j as its root. If $\widehat{S}^i_j = \{j\}$ (so that player j has no subordinates in (N,L^i)) then player j also leaves the mechanism with its stand-alone worth $v(\{j\})$. If $j \in S^i_i$ has at least one subordinate in (N,L^i), then the players in $\widehat{S}^i_j$ go to the next round, in which the root j of the rooted subtree $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$ proposes a division of the worth $v(\widehat{S}^i_j)$ to its subordinates in $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$.

In general, if in some round t a proposal of a proposer j to its subordinates in $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$ is accepted, then all players in $\widehat{S}^i_j$ leave the mechanism. Each player $\ell \in \widehat{S}^i_j \setminus \{j\}$ receives the proposed payoff y^j_ℓ and player j receives the remaining surplus of cooperation $v(\widehat{S}^i_j) - \sum_{\ell \in \widehat{S}^i_j \setminus \{j\}} y^j_\ell$. If a proposal of a proposer j is not accepted, then the proposer j leaves the mechanism with its stand-alone worth $v(\{j\})$ and each of its successors in $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$ becomes the root of the directed subtree on the set of players consisting of this successor and its subordinates in (N,L^i). If such a successor has no subordinates, it also leaves the mechanism with its own stand-alone worth, otherwise such a successor becomes the proposer to its subordinates in the next round of the mechanism.

Note that (1) in a round of Mechanism B there can be multiple proposers and (2) in a round of Mechanism B, there can be at the same time several branches $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$ of the rooted tree (N,L^i) in which the proposal of a player j is accepted, and other branches in which the proposal of another player is rejected. In a branch in which the proposal is rejected the players, except the top of the branch, go to the next round and bargain in (possibly) several new subbranches over the division of the surplus of cooperation of that subbranch. Since in each round at least one player leaves the mechanism, and because the number of players is finite, Mechanism B ends within at most $n-1$ rounds.

To describe Mechanism B formally, consider the tree game (N, v, L) with $(N, v) \in \mathcal{G}$ zero-monotonic. For given $i \in N$, we state the mechanism that implements in SPNE the hierarchical outcome $h^i(N, v, L)$. Notice that in each round t, PR^t is the set of proposers.

Mechanism B to implement the hierarchical outcome $h^i(N, v, L)$ of tree game (N, v, L)

Initiation: Set $PR^1 = \{i\}$, set $t = 1$ and go to Stage 1 of round 1.

Stage 1: Every $j \in PR^t$ proposes an offer $y_k^j \in \mathbb{R}$ to every subordinate $k \in \widehat{S}_j^i \setminus \{j\}$ in (N, L^i). Go to Stage 2.

Stage 2: For every $j \in PR^t$, the subordinates of player j sequentially either accept or reject the offer y_k^j. If all subordinates accept, then the proposal of j is accepted; otherwise the proposal of j is rejected. Go to Stage 3.

Stage 3: Consider every $j \in PR^t$. If the proposal of j is accepted, then the players in $\widehat{S}_j^i$ leave the mechanism, each subordinate $k \in \widehat{S}_j^i \setminus \{j\}$ receives y_k^j and player j receives $v(\widehat{S}_j^i) - \sum_{k \in \widehat{S}_j^i \setminus \{j\}} y_k^j$. If the proposal of j is rejected, then player j leaves the mechanism and obtains its stand-alone worth $v(\{j\})$. Set $RE^t = \{j \in PR^t \mid \text{proposal of } j \text{ is rejected}\}$, $H^t = \bigcup_{j \in RE^t} S_j^i$ and $OU^t = \{h \in H^t | S_h^i = \emptyset\}$. If $h \in OU^t$, then h receives its stand-alone worth $v(\{h\})$ and leaves the mechanism. Set $PR^{t+1} = H^t \setminus OU^t$ as the set of proposers in the next round. If $PR^{t+1} = \emptyset$, no players are left and the mechanism stops. Otherwise, set t equal to $t+1$ and return to Stage 1.

The next theorem states that, given $i \in N$, for a tree game (N, v, L) with (N, v) zero-monotonic, Mechanism B implements the hierarchical outcome $h^i(N, v, L)$ as an SPNE payoff vector. The proof follows from the facts that a hierarchical outcome is a marginal vector of the graph restricted game (N, v^L) and that Mechanism B is a modification of Mechanism A.

Theorem 5.2.2 *Let (N, v, L) be a tree game with (N, v) zero-monotonic and let $i \in N$ be given. The payoff vector in any SPNE of Mechanism B coincides with the hierarchical outcome $h^i(N, v, L)$.*

Proof. First, consider the following two observations. One, for every tree $(N, L) \in \mathcal{L}_T^N$ the graph restricted game (N, v^L) is zero-monotonic if (N, v) is zero-monotonic. And two, as discussed in Section 2.1, for a given player $i \in N$ the hierarchical outcome $h^i(N, v, L)$ is equal to the marginal vector $m^\pi(N, v^L)$ of the graph restricted game (N, v^L) for every ordering π that is consistent with the rooted tree (N, L^i). Notice that there can be multiple orderings $\pi \in \Pi^N$ for which this holds, but that for every such π it holds that $h^i(N,v,L) = m^\pi(N,v^L)$.

Next, consider the following strategies. In every round $t \in \{1,\ldots,n-1\}$, every proposer $j \in PR^t$ proposes in Stage 1 of round t the hierarchical outcome payoff $y^j_k = h^i_k(N,v,L)$ to every k in its set of subordinates $\widehat{S}^i_j \setminus \{j\}$. In Stage 2 of round t every subordinate $k \in \widehat{S}^i_j \setminus \{j\}$ of a proposer j accepts any offer at least equal to $h^i_k(N,v,L)$ and rejects any offer strictly smaller than $h^i_k(N,v,L)$.

It is not difficult to see that if the players in Mechanism B follow these strategies, the mechanism ends in round 1 and every player $\ell \in N$ receives payoff $h^i_\ell(N,v,L)$. This payoff is equal to the marginal vector payoff $m^\pi_\ell(N,v^L)$ for an ordering π consistent with (N,L^i).

That the above strategies constitute an SPNE, and that the payoff vector in any SPNE of Mechanism B is equal to the payoff vector resulting from these strategies $h^i(N,v,L)$, follows in a similar way as in Theorem 5.2.1. In fact, Mechanism B is identical to Mechanism A for every π consistent with (N,L^i), except that in a round $t \in \{2,\ldots,n-1\}$ there can be multiple proposers.[2] To see this, let N^t be the set of players in Mechanism B at the start of round t. If $t \in \{2,\ldots,n-1\}$, then the subgraph $(N^t, L(N^t))$ (possibly) consists of several components $\widehat{S}^i_j$ and on each of these components the rooted tree (N,L^i) induces a rooted subtree $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$. Following the above strategies, in each component the root j of the subtree $(\widehat{S}^i_j, L^i(\widehat{S}^i_j))$ proposes the hierarchical outcome $h^i_k(N,v,L)$ to each of its subordinates $k \in \widehat{S}^i_j \setminus \{j\}$. This results in payoffs that are equal to the payoffs in every marginal vector $m^\pi(N,v^L)$, where π is consistent with (N,L^i). Since, given $\pi \in \Pi^N$, Mechanism A implements the marginal vector $m^\pi(N,v)$ in SPNE for a zero-monotonic TU-game (N,v), Mechanism B thus implements the marginal vector $m^\pi(N,v^L)$, for every π consistent with (N,L^i), in SPNE.

□

5.3 An implementation of the weighted hierarchical solution

In this section Theorem 5.2.2 is used to introduce a mechanism of which the payoffs in any SPNE coincide with the payoffs of the weighted hierarchical solution h^α of a zero-monotonic tree game (N,v,L). The mechanism therefore implements the weighted hierarchical solution $h^\alpha(N,v,L)$ in SPNE.

To obtain the mechanism we have to add a bidding procedure to Mechanism B. In the mechanism of Pérez-Castrillo and Wettstein (2001) in every round the players begin with a bidding procedure to determine the proposer in that round. In this procedure every player i, that is still in the mechanism, makes a *bid* b^i_j to every other player j, that is still in the mechanism, to become the proposer in that round. The bid b^i_j is the amount

[2]This is similar as in Slikker (2007), who modifies the mechanism of Pérez-Castrillo and Wettstein (2001) to implement the Myerson value of a graph game.

that player i will pay to each player j when player i becomes the proposer in that round. Given the bids b_j^i, the proposer is chosen with equal probability among the players that have the highest *net bid*, which is equal to the sum of the bids made by a player to the other players minus the sum of the bids made by the other players to this player. So, the net bid of player $i \in N$ in round 1 is equal to $B^i = \sum_{j \in N \setminus \{i\}} \left(b_j^i - b_i^j\right)$.

The proposal procedure in the mechanism of Pérez-Castrillo and Wettstein (2001) is based on the following recursive formula for the Shapley value, stated in Hart and Mas-Colell (1989) and Maschler and Owen (1989):

$$Sh_j(N,v) = \frac{1}{|N|}\Big(v(N) - v(N \setminus \{j\})\Big) + \frac{1}{|N|}\sum_{i \in N \setminus \{j\}} Sh_j(N \setminus \{i\}, v_{N \setminus \{i\}}),\ j \in N.$$

Hence, the Shapley value payoff of a player $j \in N$ is the average of its marginal contribution to the grand coalition and its Shapley value payoffs in the $|N| - 1$ subgames $(N \setminus \{i\}, v_{N\setminus\{i\}})$, $i \in N \setminus \{j\}$. Using this formula, Pérez-Castrillo and Wettstein (2001) prove that the SPNE bids in the bidding procedure of their mechanism are given by $b_j^i = Sh_j(N,v) - Sh_j(N \setminus \{i\}, v_{N\setminus\{i\}})$, $i,j \in N$, $i \neq j$.

Since, for a given nonnegative vector α with $\sum_{i \in N} \alpha_i = 1$, the weighted hierarchical solution payoff for a player $j \in N$ in the tree game (N,v,L) is given by $h_j^\alpha(N,v,L) = \sum_{i \in N} \alpha_i h_j^i(N,v,L)$, and because $h_j^j(N,v,L) = v(N) - \sum_{k \in S_j^j} v(\widehat{S}_k^j) = v^L(N) - v^L(N \setminus \{j\})$, it follows that

$$h_j^\alpha(N,v,L) = \alpha_j\left(v^L(N) - v^L(N \setminus \{j\})\right) + \sum_{i \in N \setminus \{j\}} \alpha_i h_j^i(N,v,L). \tag{5.1}$$

So, the weighted hierarchical solution payoff of a player $j \in N$ is a weighted average (with weights α_k, $k \in N$) of its marginal contribution to the grand coalition in the restricted game (N, v^L) and its payoffs in the $|N| - 1$ hierarchical outcome vectors $h^i(N,v,L)$, $i \in N \setminus \{j\}$. Although there are similarities between formula (5.1) and the recursive formula for the Shapley value, formula (5.1) is not recursive. This explains why it is sufficient to have a single bidding procedure at the start of the mechanism that implements the weighted hierarchical solution h^α, instead of a bidding procedure at the start of each round, as in Pérez-Castrillo and Wettstein (2001).

Hence, consider the following bidding procedure: every player $i \in N$ makes a bid b_j^i to every other player $j \in N \setminus \{i\}$ to become the proposer in the mechanism. The bid b_j^i is the amount that player i will pay to each player j when player i becomes the proposer. Given the bids b_j^i, $i \in N$, the proposer is chosen with equal probability among the players that have the highest *weighted net bid*, which, given the vector α, is equal to $B_\alpha^i = \sum_{j \in N \setminus \{i\}} (\alpha_i b_j^i - \alpha_j b_i^j)$ for each $i \in N$. If we add this bidding procedure to Mechanism B, we obtain a mechanism that implements the weighted hierarchical solution h^α. Using formula (5.1) we will show that in every SPNE of this mechanism the equilibrium bids are given by $b_j^i = h_j^\alpha(N,v,L) - h_j^i(N,v,L)$, $i,j \in N$, $i \neq j$.

To summarize: when, in the mechanism that implements the weighted hierarchical solution h^α, player i wins the bidding procedure, it pays its bids to the players $j \in N \setminus \{i\}$ and becomes the proposer in the first round of the proposal procedure of the mechanism (which is equal to Mechanism B). When the proposal of player i is rejected in this first round, player i leaves the mechanism and the remaining players continue to implement the payoffs of the hierarchical outcome $h^i(N, v, L)$ on the subtrees of (N, L^i). Similar as in Mechanism B, on each subtree the successor j of i becomes a proposer in the next round, without there being a new bidding procedure. So, it follows that, given a vector of weights $\alpha \in \mathbb{R}^N_+$ with $\sum_{i \in N} \alpha_i = 1$, a mechanism to implement the weighted hierarchical solution h^α in SPNE is obtained by starting with a single bidding procedure and then applying Mechanism B.

Mechanism C to implement the weighted hierarchical solution $h^\alpha(N, v, L)$ of tree game (N, v, L)

Bidding procedure: Each player $k \in N$ makes bids $b^k_j \in \mathbb{R}$ to every player $j \in N \setminus \{k\}$. For each $k \in N$, let $B^k_\alpha = \sum_{j \in N \setminus \{k\}} (\alpha_k b^k_j - \alpha_j b^j_k)$ be the weighted net bid of player k. Select the player with the highest weighted net bid and call this player i. In case of a non-unique maximizer, randomly choose, with equal probability, any of the bidders with the highest weighted net bid as player i. Player i is the winner of the bidding procedure and pays every other player $j \in N \setminus \{i\}$ its bid b^i_j. So, each player $j \in N \setminus \{i\}$ receives b^i_j and player i receives $-\sum_{j \in N \setminus \{i\}} b^i_j$. Go to the Proposal procedure.

Proposal procedure: Set $P^1 = \{i\}$ in Stage 1 of round 1 of Mechanism B and follow Mechanism B.

Notice that adding the bidding procedure to Mechanism B changes the final payoffs of the players. The payoff of a player is now equal to the sum of its payoff in the bidding procedure and its payoff in the proposal procedure (Mechanism B).

Suppose that player i is the winner of the bidding procedure. If in round 1 of the proposal procedure the proposal y^i_k, $k \in N \setminus \{i\}$, of player i is accepted, then the total payoff of a player $k \in N \setminus \{i\}$ is equal to $y^i_k + b^i_k$ and the total payoff of player i is equal to $v(N) - \sum_{k \in N \setminus \{i\}} (y^i_k + b^i_k)$. If in round 1 of the proposal procedure the proposal of player i is rejected, then the total payoff of player i is $v(\{i\}) - \sum_{k \in N \setminus \{i\}} b^i_k$.

Next, suppose that in some round $t \in \{2, \ldots, n-1\}$ of the proposal procedure the offers y^j_k of some player j to each subordinate k in $\widehat{S}^i_j \setminus \{j\}$ are accepted, then each subordinate k receives the total payoff $b^i_k + y^j_k$ and player j's total payoff is equal to $b^i_j + v(\widehat{S}^i_j) - \sum_{k \in \widehat{S}^i_j \setminus \{j\}} y^j_k$. If the proposal of player j is rejected, then each successor ℓ of j in (N, L^i) for which it holds that $S^i_\ell = \emptyset$ receives total payoff $b^i_\ell + v(\{\ell\})$ and player j receives total payoff $b^i_j + v(\{j\})$.

Theorem 5.3.2 below states that Mechanism C implements in SPNE the weighted hierarchical solution h^α for zero-monotonic tree games. In proving this theorem we use the following lemma.

Lemma 5.3.1 *Let (N, v, L) be a tree game with (N, v) zero-monotonic. Given the vector of weights $\alpha \in \mathbb{R}^N_+$ with $\sum_{i \in N} \alpha_i = 1$ and given $i, j \in N$, $i \neq j$, let $b^i_j = h^\alpha_j(N, v, L) - h^i_j(N, v, L)$. Then, for every $i \in N$, $B^i_\alpha = 0$.*

Proof. For every $i \in N$ it holds that

$$B_\alpha^i = \sum_{j \in N \setminus \{i\}} \big(\alpha_i h_j^\alpha(N,v,L) - \alpha_i h_j^i(N,v,L) - \alpha_j h_i^\alpha(N,v,L) + \alpha_j h_i^j(N,v,L)\big) =$$

$$\alpha_i \sum_{j \in N \setminus \{i\}} h_j^\alpha(N,v,L) - (1-\alpha_i) h_i^\alpha(N,v,L) + \sum_{j \in N \setminus \{i\}} \big(\alpha_j h_i^j(N,v,L) - \alpha_i h_j^i(N,v,L)\big) =$$

$$\alpha_i \sum_{j \in N} h_j^\alpha(N,v,L) - h_i^\alpha(N,v,L) + \sum_{j \in N \setminus \{i\}} \big(\alpha_j h_i^j(N,v,L) - \alpha_i h_j^i(N,v,L)\big) =$$

$$\alpha_i v(N) - \sum_{k \in N} \alpha_k h_i^k(N,v,L) + \sum_{j \in N \setminus \{i\}} \big(\alpha_j h_i^j(N,v,L) - \alpha_i h_j^i(N,v,L)\big) =$$

$$\alpha_i v(N) - \alpha_i \sum_{k \in N} h_k^i(N,v,L) = 0,$$

where the second equality follows from the fact that $\alpha_i + \sum_{j \in N \setminus \{i\}} \alpha_j = 1$, the fourth because every weighted hierarchical solution provides an efficient payoff vector and the last because every hierarchical outcome provides an efficient payoff vector.

□

Theorem 5.3.2 *Let (N,v,L) be a tree game with (N,v) zero-monotonic and let $\alpha \in \mathbb{R}_+^N$ with $\sum_{i \in N} \alpha_i = 1$ be a vector of weights. The payoff vector in any SPNE of Mechanism C coincides with the payoff vector $h^\alpha(N,v,L)$.*

Proof. We first show that the components of the payoff vector $h^\alpha(N,v,L)$ are indeed SPNE payoffs of Mechanism C. For this, consider the following strategies. In the bidding procedure, player $g \in N$ bids $b_\ell^g = h_\ell^\alpha(N,v,L) - h_\ell^g(N,v,L)$ to every player $\ell \in N \setminus \{g\}$. When player i is the winner of the bidding procedure and player $j \in PR^t$ is a proposer in some round $t \in \{1,\dots,n-1\}$ of the proposal procedure (Mechanism B), then j proposes in Stage 1 the hierarchical outcome payoff $y_k^j = h_k^i(N,v,L)$ to every player k in its set of subordinates $\widehat{S}_j^i \setminus \{j\}$. In Stage 2 every subordinate k of a proposer j accepts any offer at least equal to $h_k^i(N,v,L)$ and rejects any offer strictly smaller than $h_k^i(N,v,L)$.

Following these strategies, the mechanism ends in round 1 of the proposal procedure. The total payoff to the winner i of the bidding procedure is equal to

$$v(N) - \sum_{k \in N \setminus \{i\}} h_k^i(N,v,L) - \sum_{k \in N \setminus \{i\}} b_k^i = h_i^i(N,v,L) - \sum_{k \in N \setminus \{i\}} b_k^i =$$

$$h_i^i(N,v,L) - \sum_{k \in N \setminus \{i\}} \big(h_k^\alpha(N,v,L) - h_k^i(N,v,L)\big) =$$

$$\sum_{k \in N} h_k^i(N,v,L) - \sum_{k \in N \setminus \{i\}} h_k^\alpha(N,v,L) =$$

$$v(N) - \sum_{k \in N \setminus \{i\}} h_k^\alpha(N,v,L) = h_i^\alpha(N,v,L),$$

where the first and fourth equality follow because every hierarchical outcome provides an efficient payoff vector and the fifth equality follows because every weighted hierarchical solution provides an efficient payoff vector. The total payoff to a player $k \in N \setminus \{i\}$ is equal to

$$h_k^i(N,v,L) + b_k^i = h_k^i(N,v,L) + h_k^\alpha(N,v,L) - h_k^i(N,v,L) = h_k^\alpha(N,v,L).$$

The strategy profile outlined above therefore results in the payoff vector $h^\alpha(N,v,L)$.

Next, we show that this strategy profile is indeed an SPNE profile. From Theorem 5.2.2 it follows that the strategies are SPNE strategies in the subgame that starts in Stage 1 of round 1 of the proposal procedure (and in all subsequent subgames). We therefore only have to consider the (sub)game that starts in the bidding procedure, thus the complete Mechanism C.

Suppose that every player $i \in N \setminus \{k\}$ makes bids $b_j^i = h_j^\alpha(N,v,L) - h_j^i(N,v,L)$, $j \in N \setminus \{i\}$, and that some player k deviates and makes bids $(\overline{b}_j^k)_{j\in N\setminus\{k\}}$ with $\overline{b}_j^k \neq h_j^\alpha(N,v,L) - h_j^k(N,v,L)$ for at least one $j \in N \setminus \{k\}$. When, after this deviation by player k, some other player $i \in N \setminus \{k\}$ wins the bidding procedure, then player i pays to $j \in N \setminus \{i\}$ its bids b_j^i in the bidding procedure and the offer $h_j^i(N,v,L)$ in the proposal procedure. The deviation of player k in this case does not change the payoffs. We therefore only have to consider the case that the deviating player k wins the bidding procedure. By Lemma 5.3.1 it holds that in the strategy profile discussed above the weighted net bid $B_\alpha^i = \sum_{j\in N\setminus\{i\}}(\alpha_i b_j^i - \alpha_j b_i^j) = 0$ for every $i \in N$. When it holds for the deviating player k that the weighted net bid $\overline{B}_\alpha^k = \sum_{j\in N\setminus\{k\}}(\alpha_k \overline{b}_j^k - \alpha_j b_k^j) < B_\alpha^k$, then some other player $i \in N \setminus \{k\}$ wins the bidding procedure. Hence, it must be that $\overline{B}_\alpha^k > B_\alpha^k = 0$ and thus $\sum_{j\in N\setminus\{k\}} \overline{b}_j^k > \sum_{j\in N\setminus\{k\}} b_j^k$ (note that $\overline{B}_\alpha^k = B_\alpha^k = 0$ is impossible because then agent k would not be deviating). In the proposal procedure player k makes offers $h_j^k(N,v,L)$ (its SPNE offers in the proposal subgame) and these offers are accepted. Player k's final payoff after deviating and winning the bidding procedure is therefore equal to $v(N) - \sum_{j\in N\setminus\{k\}} h_j^k(N,v,L) - \sum_{j\in N\setminus\{k\}} \overline{b}_j^k = h_k^k(N,v,L) - \sum_{j\in N\setminus\{k\}} \overline{b}_j^k$. Since $h_k^k(N,v,L) - \sum_{j\in N\setminus\{k\}} \overline{b}_j^k < h_k^k(N,v,L) - \sum_{j\in N\setminus\{k\}} b_j^k = h_k^\alpha(N,v,L)$ this means that player k cannot improve its payoff by deviating. It can be concluded that the strategy profile described above is indeed an SPNE profile and that the $h_k^\alpha(N,v,L)$, $k \in N$, are indeed SPNE payoffs of Mechanism C.

It remains to prove that any SPNE of Mechanism C yields the weighted hierarchical solution payoffs. In any SPNE it must be that the weighted net bid $B_\alpha^i = 0$ for all $i \in N$. To see this consider the following. Similar as in Pérez-Castrillo and Wettstein (2001), let $\Omega = \{i \in N \mid B_\alpha^i = \max_{j\in N} B_\alpha^j\}$. Suppose that $\Omega \neq N$ and take some player $i \in \Omega$. When $\alpha_i = 0$ the bids of player i do not affect the weighted net bids $B_\alpha^k = 0$ of all players $k \in N$. This means that player i can change its bids so as to decrease the sum of its payments in the bidding procedure without altering the set Ω (player i maintains the same probability of winning the bidding procedure, but obtains a higher expected payoff). So, Ω must be equal to N after all which implies that $B_\alpha^k = 0$ for all $k \in N$. When $\alpha_i > 0$ take some player $j \in N \setminus \Omega$, $\delta > 0$ and suppose player i changes its bids to $\widehat{b}_k^i = b_k^i + \frac{\delta}{\alpha_i}$ for all $k \in \Omega \setminus \{i\}$, $\widehat{b}_j^i = b_j^i - \frac{|\Omega|\delta}{\alpha_i}$ and $\widehat{b}_\ell^i = b_\ell^i$ for all $\ell \in N \setminus (\Omega \cup \{j\})$. The weighted net bids are then equal to $\widehat{B}^i_\alpha = B^i_\alpha - \delta$, $\widehat{B}^k_\alpha = B^k_\alpha - \delta$ for all $k \in \Omega \setminus \{i\}$, $\widehat{B}^j_\alpha = B^j_\alpha + |\Omega|\delta$ and $\widehat{B}^\ell_\alpha = B^\ell_\alpha$ for all $\ell \in N \setminus (\Omega \cup \{j\})$. If player i now chooses δ so that $B^j_\alpha + |\Omega|\delta < B^i_\alpha - \delta$, then it holds that $\widehat{B}^\ell_\alpha < \widehat{B}^i_\alpha = \widehat{B}^k_\alpha$ for all $\ell \in N \setminus \Omega$ and all $k \in \Omega \setminus \{i\}$. This means that Ω does not change, but $\sum_{g \in N\setminus\{i\}} b^i_g - \frac{\delta}{\alpha_i} < \sum_{g \in N\setminus\{i\}} b^i_g$. Again it follows that Ω must be equal to N after all, which implies that $B^k_\alpha = 0$ for all $k \in N$.

Since all weighted net bids are zero in any SPNE, it follows that in any SPNE each player's payoff must be the same, regardless of the winner of the bidding procedure. If this would not be the case and some player $j \in N$ would prefer to be, or not to be, the proposer at the start of the the proposal procedure, this player j could slightly increase, or decrease, one of its bids b^j_k, $k \in N \setminus \{j\}$, to increase its final payoff. Since none of the players does this in equilibrium it must mean that all players are indifferent to the identity of the winner of the bidding procedure and thus that the payoff of each player $i \in N$ is independent of the identity of the winner of the bidding procedure.

When player ℓ is the winner of the bidding procedure it holds by Theorem 5.2.2 that in any SPNE of the proposal procedure the payoff vector is equal to $h^\ell(N, v, L)$. When player i is the winner of the bidding procedure it therefore follows that the final payoff of player i in the entire Mechanism C is equal to

$$z^i_i = h^i_i(N, v, L) - \sum_{j \in N\setminus\{i\}} b^i_j.$$

This implies that

$$\alpha_i z^i_i = \alpha_i h^i_i(N, v, L) - \sum_{j \in N\setminus\{i\}} \alpha_i b^i_j.$$

When some player $j \in N \setminus \{i\}$ is the winner of the bidding procedure, the final payoff of player i in the entire Mechanism C is equal to

$$z^j_i = h^j_i(N, v, L) + b^j_i.$$

This implies that

$$\alpha_j z^j_i = \alpha_j h^j_i(N, v, L) + \alpha_j b^j_i.$$

Summing over all players $k \in N$ gives

$$\sum_{k \in N} \alpha_k z^k_i = \sum_{k \in N} \alpha_k h^k_i(N, v, L) - B^i_\alpha = \sum_{k \in N} \alpha_k h^k_i(N, v, L) = h^\alpha_i(N, v, L).$$

Since player i is indifferent about the identity of the winner of the bidding procedure, it must be that $z^k_i = z^\ell_i$ for all $k, \ell \in N$ and thus that $z^k_i = h^\alpha_i(N, v, L)$ for all $k \in N$. This shows that any SPNE of Mechanism C yields the weighted hierarchical solution payoffs.

□

Note that, as in the implementation of the Shapley value of Pérez-Castrillo and Wettstein (2001), in any SPNE of Mechanism C all weighted net bids are equal and therefore the choice of a random proposer is the outcome of the strategic bidding process. Further, as for Mechanism A and Mechanism B, it also holds that Mechanism C has a unique SPNE when the game $(N, v) \in \mathcal{G}$ is strictly zero-monotonic. In that case each player $k \in N$ makes bids $b_j^k = h_j^\alpha(N, v, L) - h_j^k(N, v, L)$ to the players $j \in N \setminus \{k\}$ in the bidding procedure, so that some player i is randomly chosen to be the winner and is the proposer in round 1 of the proposal procedure. This player proposes $y_j^i = h_j^i(N, v, L)$ to every player $j \in N \setminus \{i\}$ and every player $j \in N \setminus \{i\}$ accepts the proposal.

5.4 Application to river games

In Theorem 5.3.2 we gave an implementation of the weighted hierarchical solution h^α on the class of zero-monotonic tree games. It is not too difficult to see that this implementation also holds for the class of river games with multiple springs $\mathcal{R}^{(N,\mathcal{U})}$. Thus, for river games with multiple springs under Assumption 2.2.1 of Chapter 2.

Given a river system $(N,\mathcal{U})$ let (N, L) be the tree with $L = \{\{i, j\} \mid i, j \in N$ and $i \in U^j\}$. Because any river game with multiple springs $v \in \mathcal{R}^{(N,\mathcal{U})}$ is superadditive it thus induces a zero-monotonic tree game (N, v, L). This leads to the following corollary.

Corollary 5.4.1 *Let $v \in \mathcal{R}^{(N,\mathcal{U})}$ be a river game with multiple springs and let $\alpha \in \mathbb{R}_+^N$ with $\sum_{i \in N} \alpha_i = 1$ be a vector of weights. The payoff vector in any SPNE of Mechanism C then coincides with the payoff vector $h^\alpha(v)$.*

When interpreting Mechanism C in light of river games with multiple springs, an agent along the river can only make proposals about a division of the surplus of cooperation to other agents along the river when (1) they have a common most downstream agent that also receives a proposal and (2) every agent in between the proposer, or any agent receiving a proposal, and this common most downstream agent also receives a proposal. Hence, in Mechanism C agents are only able to negotiate about a division of the surplus of cooperation when they are ‘connected’ along the river. In light of Assumption 2.2.1 of Chapter 2 this condition makes sure that agents do not divide more welfare than is available to them because some of the water that is transferred to create the welfare is intercepted by some agent that does not cooperate (does not receive a proposal).

At the end of Section 4.4 we discussed the assumption that in river games with multiple springs and externalities blocking (of an agreement) is restricted to connected coalitions because coordination among agents along the river becomes difficult when agents are not neighboring. If we now, by the same argument, make the explicit assumption that agents are only able to negotiate about a division of the surplus of cooperation when they are ‘connected’ along the river, the implementation of the hierarchical solution h^α in Mechanism C also holds for the class of river games with multiple springs and externalities $\mathcal{RE}^{(N,\mathcal{U})}$.

Chapter 6

River basin games

6.1 Introduction

In this chapter we introduce river basin games, in which countries along the river are allowed to be composed of different water users and the river is allowed to have multiple springs (tributaries), multiple sinks (distributaries), or both.

A river basin is the area of land drained by a river and its tributaries. It normally consists of streams and creeks that flow downhill into one another, eventually into one main stream, and can include a delta at the end where the main stream branches off into several smaller streams before it reaches its final destination. The final destination of a river basin is usually an ocean, sea, estuary, lake, flat arid area or man-made reservoir. A river basin can have multiple springs and/or multiple sinks. It can even have anabranches.[1] In short, a river basin represents the most general river configuration one can imagine.

As in Khmelnitskaya (2010) (which models rivers with multiple springs or multiple sinks) and Khmelnitskaya and Talman (2010) (which model rivers with multiple springs, multiple sinks and anabranches), in this chapter we model river basins by making use of directed graphs (see Section 2.1). That is, we model the river basin as a directed graph, where the set of agents in the river basin corresponds to the set of nodes of the graph and the flow of water between the different agents is represented by the directed links of the graph.

In the models of the previous chapters any country located along a river acted as a single entity (agent). In reality though, countries usually consist of heterogeneous groups of water users, e.g., states, cities or individual water users like domestic, industrial and agricultural users. As a consequence, the water allocation problem can be seen as a multiple-level problem. On a first level, water (benefit) has to be allocated among the countries located along the river, and on a second level water (benefit) allocated to a country has to be distributed between the different water users within the country. To capture this situation of conflicting interests of countries along a river, as well as conflicting interests of water users within a country, we make use of TU-games with coalition and graph structure (see Section 2.1).

[1] An anabranch is a branch of the river that splits off from the main stream at some point, but merges again with it further downstream.

Hence, the model of this chapter has a set of agents (let's take cities) N located in a river basin. The cities that are located in the same country are in the same union of the partition P (of the set of cities N). Each city derives benefit by consuming water from the river basin, that picks up volume along its course. The water consumption of a city is restricted by the water inflow in the river basin and the consumption of the cities further upstream in the river basin. The cities value consumption of the river water differently in the sense that some cities have higher needs (marginal benefit) for the consumption of river water than others. The heterogeneous valuations of the cities are introduced by endowing each city with its own benefit function that, together with possible monetary transfers by other cities located in the basin, determine a utility function for each city. We assume that Assumption 2.2.1 of Chapter 2 holds and that cities in the river basin are only able to cooperate ('trade' water for monetary transfers) in coalitions within their own country. In addition, countries are also able to cooperate. When two (or more) countries decide to cooperate it implies that all cities within these countries cooperate in one big coalition (when countries cooperate they force all their constituent cities to cooperate).

A 'fair' division of the maximum total welfare in this model can be found by considering 'fair' solution concepts for a corresponding TU-game with coalition and graph structure. Vázquez-Brage, García-Jurado and Carreras (1996) and Alonso-Meijide, Álvarez-Mozos and Fiestras-Janeiro (2009) suggest three different single-valued solution concepts (values) for TU-games with coalition and graph structure. But, since these values are all based on the Owen value (Owen, 1977) for games with coalition structure, they all implicitly assume that cities (agents) in a country (union) are able to cooperate with other countries (unions), even if the countries (unions) themselves are not (fully) cooperating. Since we assume that when countries cooperate they force all their constituent cities to cooperate, we also want the solution concept to reflect this.

In Section 2.1 we mentioned that recently Kamijo (2011) introduced a value for TU-games with coalition structure in which individual players can cooperate within their union and complete unions can cooperate (when complete unions cooperate, they force all their constituent players to cooperate), but proper subsets of unions cannot cooperate. In this chapter we will show that the ideas of Kamijo (2011) can be combined with the ideas of Myerson (1977), about TU-games with graph structure, to obtain two new Shapley-type values for TU-games with coalition and graph structure. In addition, combining the partition restricted game of Kamijo (2011) with the average tree solution of Herings, van der Laan, Talman and Yang (2010) gives a value for TU-games with coalition and graph structure that, based on the findings of the Chapter 4, seems to be particularly appropriate for river basin games.

This chapter is based on van den Brink, van der Laan and Moes (2011) and is organized as follows. In Section 6.2 we introduce river basin benefit problems and river basin games. In Section 6.3 we propose the graph-partition restricted game and the partition-graph restricted game of a TU-game with coalition and graph structure. In Section 6.4 we apply the Shapley value to both the graph-partition and partition-graph restricted game to obtain two values for TU-games with coalition and graph structure. Finally, in Section 6.5 we show how the values of Section 6.4 can be applied to river basin games, and shortly discuss the average tree solution for river basin games.

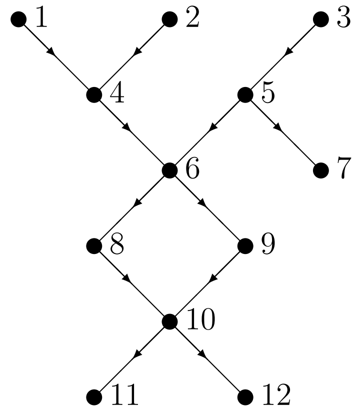

Figure 6.1: An example of a river basin (N, D).

6.2 River basin benefit problems and games

In this chapter we model a river basin as a directed graph (N, D), where the set of nodes N corresponds to the set of agents (cities) in the basin, and D is the collection of directed links that represents the flow of water between the agents. We say that $j \neq i$ is *downstream* of i (and i is a *upstream* of j) if there is a sequence of directed links $(i_\ell, i_{\ell+1}) \in D$, $\ell \in \{1, \dots, k-1\}$, such that $i_1 = i$ and $i_k = j$. A directed link (i, j) is thus in the set D if and only if agent j is a downstream neighbor (successor) of agent i along the river (and agent i is an upstream neighbor (predecessor) of agent j). Each spring and each sink of the river basin is identified by an agent, i.e., we consider the most upstream agent along a tributary as its spring and the most downstream agent along a branch of the river (delta) as its sink. More formally, an agent $i \in N$ is a spring when it has no predecessors in (N, D) and a sink when it has no successors in (N, D). Since river basins are necessarily connected, we only consider directed graphs (N, D) that induce connected undirected graphs $(N, \widehat{D})$ (see Section 2.1). Moreover, for (N, D) to represent a river basin, we require that it is acyclic (see also Section 2.1). Figure 6.1 displays an example of a (connected, acyclic) river basin (N, D) with $N = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ and

$$D = \{(1,4), (2,4), (3,5), (4,6), (5,6), (5,7), (6,8), (6,9), (8,10), (9,10), (10,11), (10,12)\}.$$

For all $i \in N$, let $e_i \geq 0$ be the inflow of water in the river basin on the territory of agent (city) i. As before, agents derive benefit by consuming water from the river basin. Note that an agent can maximally consume all the water that enters the river basin on its territory plus all the water that is not consumed by its upstream agents.[2] A water allocation is therefore given by any vector of water consumption levels $x \in \mathbb{R}^N_+$ that

[2]For $k \in N$, UP^k denotes the set of all agents upstream of agent k in (N, D), including k itself.

satisfies

$$\sum_{i\in S} x_i \leq \sum_{i\in S} e_i + \sum_{j\in \bigcup\limits_{i\in S}(UP^i\setminus S)} (e_j - x_j) \text{ for all } S \subseteq N. \tag{6.1}$$

Condition 6.1 implies that water in the river basin cannot be transferred upstream, between separate branches of a river or that the same water is consumed by multiple agents simultaneously.

As in Khmelnitskaya (2010) and Khmelnitskaya and Talman (2010), we assume that each agent has full control over the water flow at its location. This assumption implies that when a river branches off into several distributaries, the agent located at the split can determine the amounts of water flowing into the separate branches (when it does not consume all the water at its location itself) and will result in a cooperative game in characteristic function form.[3] An alternative assumption could be that the river follows its natural flow and water flows into the separate branches in fixed proportions. This would lead to a situation in which externalities appear and would result in a cooperative game in partition function form, which is beyond the scope of this chapter.

The benefit that agent $i \in N$ derives from consuming the amount of water x_i is, as before, given by its benefit function $b_i : \mathbb{R}_+ \to \mathbb{R}$ which satisfies Assumption 2.2.1 of Chapter 2. We allow for situations in which the countries along the river are composed of different water users. We do this by collecting the agents (cities) that are located in the same country, in the same union of the partition $P \in \mathcal{P}^N$. Cities that are located in the same country k are thus elements of the same union P_k in the partition $P = \{P_j \mid j \in M\}$ with $M = \{1, \ldots, m\}$. So, the agents in N represent the cities in the river basin and the partition (coalition structure) P represents the countries that share the river basin.

We call the quintuple (N, D, P, e, b), where (N, D) is a connected and acyclic directed graph, P is a partition of the set of agents N, e is a vector of inflows, and b is a vector of benefit functions, a *river basin benefit problem.*

Given quasi-linear utility functions of the agents, $u^i(x_i, t_i) = b_i(x_i) + t_i$ for all $i \in N$, and assumption on the monetary compensations, $\sum_{i\in N} t_i \leq 0$, a welfare distribution for a river basin benefit problem (x, t) is given by a water allocation x and a compensation scheme t. A welfare distribution in a river basin benefit problem (y, t) is Pareto efficient if and only if $y \in \mathbb{R}^N_+$ solves the welfare maximization problem

$$\max_{\{x_i \mid i\in N\}} \sum_{i\in N} b_i(x_i) \text{ s.t. inequalities (6.1) and } x_i \geq 0 \text{ for all } i \in N, \tag{6.2}$$

and the compensation scheme is budget balanced: $\sum_{i=1}^n t_i = 0$. Under Assumption 2.2.1, the maximization problem (6.2) has at least one solution and every solution x^* yields the same maximum total welfare $\sum_{i=1}^n b_i(x_i^*)$. The Pareto efficient welfare distribution (x^*, t) again yields payoffs (utilities)

$$z_i = b_i(x^*) + t_i, \quad i \in N,$$

[3] An agent is located at a split of (N, D) when it has multiple successors in (N, D).

with the sum of the payoffs equal to the Pareto efficient total welfare $\sum_{i=1}^{n} b_i(x_i^*)$.

Now, consider the characteristic function v in which the worth $v(N)$ is equal to the Pareto efficient total welfare, i.e., $v(N) = \sum_{i=1}^{n} b_i(x_i^*)$, and the worth $v(S)$ of a coalition $S \subset N$ is given by

$$v(S) = \sum_{k \in S} b_k(x_k^S) \quad \text{where} \quad x^S = (x_k^S)_{k \in S} \quad \text{solves}$$

$$\max_{\{x_k | k \in S\}} \sum_{k \in S} b_k(x_k) \quad \text{s.t.} \quad \sum_{i \in R} x_i \leq \sum_{i \in R} e_i + \sum_{j \in (\bigcup\limits_{i \in R} (UP^i \backslash R)) \cap S} (e_j - x_j) \quad \text{for all } R \subseteq S$$

$$\text{and} \quad x_k \geq 0, \; k \in S. \tag{6.3}$$

The set of agents N and the characteristic function v constitutes a TU-game (N, v). This TU game, however, does not provide a complete picture of the river basin benefit problem under Assumption 2.2.1.

Consider, for instance, a non-connected coalition S consisting of two connected subsets of agents; say an upstream connected subset S_1 and a downstream connected subset S_2.[4] Similar as discussed in Section 2.2 for consecutive coalitions, in maximizing its joint benefit, the coalition S cannot transfer water from S_1 to S_2 because the strictly increasing benefit functions of the agents imply that all water sent from S_1 to S_2 is immediately intercepted by the agents in between S_1 and S_2. In the river basin benefit problem under Assumption 2.2.1, a coalition is therefore admissible if and only if it is connected. For instance, in Figure 6.1 agents 1 and 6 can only cooperate when agent 4 also agrees. The TU-game (N, v) does not take this into account. To solve this problem, we add the (undirected) graph $\widehat{D}$, representing the river basin, to the TU-game (N, v) to obtain a TU-game with graph structure $(N, v, \widehat{D})$ (see Section 2.1) and assume, following Myerson (1977), that the agents in a coalition are only able to cooperate when they are connected in the graph $(N, \widehat{D})$.

Because we also want to take into account that agents (cities) in the river basin are only able to cooperate in coalitions within their own country, and that complete countries can cooperate, we further add the coalition structure P (giving the partition of agents (cities) into unions (countries)) to the TU-game with graph structure $(N, v, \widehat{D})$ to obtain a TU-game with coalition and graph structure $(N, v, \widehat{D}, P)$.

Given a river basin benefit problem (N, D, P, e, b), we call the corresponding TU-game with coalition and graph structure $(N, v, \widehat{D}, P)$ a *river basin game*. So, the player set N in the game corresponds to the set of agents N in the problem, the characteristic function v is as defined in (6.3), the graph L in the game is equal to the graph $\widehat{D}$, and the coalition structure P in the game and model are equal.

[4]Given a river basin (N, D), we say that the set of agents S_1 is a connected subset of N when the (undirected) subgraph $(S_1, \widehat{D}(S_1))$ is connected.

6.3 The graph-partition and partition-graph restricted games

In this chapter we now first propose and characterize two new values for (general) TU-games with coalition and graph structure, which we later apply to river basin games. They are obtained by applying the Shapley value to two restricted games associated with a TU-game with coalition and graph structure. The two restricted games combine the ideas of Myerson (1977) and Kamijo (2011). The first is called the *graph-partition restricted game* and is the partition restricted game of the graph restricted game. That is, first the graph structure is taken into account to obtain the graph restricted game (see Section 2.1), and then the partition structure is taken into account by taking the partition restricted game (also see Section 2.1) of the graph restricted game. The second is called the *partition-graph restricted game* and is obtained the other way around: it is the graph restricted game of the partition restricted game.

It follows from Owen (1986) that for a partition-graph restricted game the (Harsanyi) dividend of any coalition that is not connected in the graph is zero.[5] For a graph-partition restricted game we show that the dividend of every coalition that is neither a subset of a union, nor a union of unions is zero. This implies that, in general, the graph-partition restricted game is not equal to the partition-graph restricted game.

In Section 2.1 we defined a TU-game with coalition and graph structure as a quadruple (N, v, L, P) with $(N, v) \in \mathcal{G}$ a TU-game, $L \in \mathcal{L}^N$ a graph and $P \in \mathcal{P}^N$ a partition of N. We denoted the collection of all TU-games with coalition and graph structure by $\mathcal{CGG}$. Given a TU-game with graph structure (N, v, L), the corresponding graph restricted game is given by (N, v^L) with player set N and characteristic function $v^L(S) = \sum_{T \in C^L(S)} v(T)$ for all $S \subseteq N$; and, given a TU-game with coalition structure (N, v, P), the corresponding partition restricted game is given by $(N, v|_P)$ with player set N and characteristic function $v|_P(S) = \sum_{T \in S/P} v(T), \quad \text{for all} \quad S \subseteq N$.

As stated above, with each TU-game with coalition and graph structure we associate two restricted TU-games. First, we define the *graph-partition restricted game* induced by L and P. This game associates with every $(N, v, L, P) \in \mathcal{CGG}$ the corresponding TU-game $(N, v^L|_P)$. So, given $(N, v, L, P) \in \mathcal{CGG}$, the graph-partition restricted game is obtained by first taking the graph restricted game v^L of (N, v, L) and then the partition restricted game $v^L|_P$ of (N, v^L, P). Second, the *partition-graph restricted game* is defined the other way around and associates with every $(N, v, L, P) \in \mathcal{CGG}$ the corresponding TU-game $(N, (v|_P)^L)$. So, given $(N, v, L, P) \in \mathcal{CGG}$, the partition-graph restricted game is obtained by first taking the partition restricted game $v|_P$ of (N, v, P) and then the graph restricted game $(v|_P)^L$ of $(N, v|_P, L)$.

In general, the game $(N, v^L|_P)$ is not equal to the game $(N, (v|_P)^L)$. This implies that the order in which the cooperation restrictions are applied matters. We illustrate this in the next example.

[5]Recall that the dividend of a coalition is the additional contribution of cooperation among the players in a coalition, that they did not already realize by cooperating in smaller coalitions, see Harsanyi (1963).

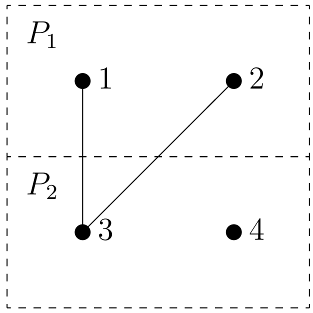

Figure 6.2: $N = \{1,2,3,4\}$, $L = \{\{1,3\},\{2,3\}\}$, $P = \{\{1,2\},\{3,4\}\}$

Example 6.3.1 Let $(N, v, L, P) \in \mathcal{CGG}$ be so that $N = \{1,2,3,4\}$, $L = \{\{1,3\},\{2,3\}\}$ and $P = \{\{1,2\},\{3,4\}\}$, as displayed in Figure 6.2. Then $v^L|_P(S)$ and $(v|_P)^L(S)$, $S \subseteq N$, are as given in Table 6.1 and Table 6.2 respectively. In the last column of both tables the dividends are given. For readability, in the tables we write $v(\{...\})$ as $v(...)$. □

Given a TU-game with graph structure $(N, v, L) \in \mathcal{GG}$, Owen (1986) has shown that for the corresponding graph restricted game (N, v^L) the dividend $\Delta^S(v^L)$ is equal to zero for any coalition S that is not connected in (N, L). Since the partition-graph restricted game is defined as the graph restricted game of the partition restricted game, it holds that in a partition-graph restricted game the dividend of any coalition that is not connected in the graph is zero.

Corollary 6.3.2 *For every $(N, v, L, P) \in \mathcal{CGG}$ and every $S \in 2^N \setminus \{\emptyset\}$, if S is not connected in (N, L) then $\Delta^S((v|_P)^L) = 0$.*

The direct analog of Corollary 6.3.2 does not hold for the graph-partition restricted game $(N, v^L|_P)$. For instance, in Example 6.3.1 it holds that $S = N$ is not connected, but $\Delta^N(v^L|_P) = v(\{1,2,3\}) - v(\{1\}) - v(\{2\}) - v(\{3\})$.

To find the counterpart of Corollary 6.3.2 for $(N, v^L|_P)$, we first consider TU-games with coalition structure $(N, v, P) \in \mathcal{CG}$. Recall that for a fixed player set N, $\mathcal{G}^N$ denotes the collection of all TU-games on N. Then, for $P = \{P_j \mid j \in M\} \in \mathcal{P}^N$, define the mapping $Z_P \colon \mathcal{G}^N \to \mathcal{G}^N$ by

$$Z_P(v) = v|_P.$$

So, Z_P maps each characteristic function $v \in \mathcal{G}^N$ to the characteristic function $v|_P \in \mathcal{G}^N$. Because the elements of the collection S/P are fixed (see Section 2.1), Z_P is a linear mapping. In order to investigate the behavior of the mapping Z_P, we consider the images

S	$v^L(S)$	$v^L\|_P(S)$	$\Delta^S(v^L\|_P)$
$\emptyset$	$v(\emptyset)$	$v^L(\emptyset) = v(\emptyset)$	0
$\{1\}$	$v(1)$	$v^L(1) = v(1)$	$v(1)$
$\{2\}$	$v(2)$	$v^L(2) = v(2)$	$v(2)$
$\{3\}$	$v(3)$	$v^L(3) = v(3)$	$v(3)$
$\{4\}$	$v(4)$	$v^L(4) = v(4)$	$v(4)$
$\{1,2\}$	$v(1)+v(2)$	$v^L(1,2) = v(1)+v(2)$	0
$\{1,3\}$	$v(1,3)$	$v^L(1)+v^L(3) = v(1)+v(3)$	0
$\{1,4\}$	$v(1)+v(4)$	$v^L(1)+v^L(4) = v(1)+v(4)$	0
$\{2,3\}$	$v(2,3)$	$v^L(2)+v^L(3) = v(2)+v(3)$	0
$\{2,4\}$	$v(2)+v(4)$	$v^L(2)+v^L(4) = v(2)+v(4)$	0
$\{3,4\}$	$v(3)+v(4)$	$v^L(3,4) = v(3)+v(4)$	0
$\{1,2,3\}$	$v(1,2,3)$	$v^L(1,2)+v^L(3) = v(1)+v(2)+v(3)$	0
$\{1,2,4\}$	$v(1)+v(2)+v(4)$	$v^L(1,2)+v^L(4) = v(1)+v(2)+v(4)$	0
$\{1,3,4\}$	$v(1,3)+v(4)$	$v^L(1)+v^L(3,4) = v(1)+v(3)+v(4)$	0
$\{2,3,4\}$	$v(2,3)+v(4)$	$v^L(2)+v^L(3,4) = v(2)+v(3)+v(4)$	0
N	$v(1,2,3)+v(4)$	$v^L(N) = v(1,2,3)+v(4)$	$v(1,2,3)-v(1)$ $-v(2)-v(3)$

Table 6.1: Characteristic function and dividends of $(N, v^L|_P)$.

S	$v\|_P(S)$	$(v\|_P)^L(S)$	$\Delta^S((v\|_P)^L)$
$\emptyset$	$v(\emptyset)$	$v\|_P(\emptyset) = v(\emptyset)$	0
$\{1\}$	$v(1)$	$v\|_P(1) = v(1)$	$v(1)$
$\{2\}$	$v(2)$	$v\|_P(2) = v(2)$	$v(2)$
$\{3\}$	$v(3)$	$v\|_P(3) = v(3)$	$v(3)$
$\{4\}$	$v(4)$	$v\|_P(4) = v(4)$	$v(4)$
$\{1,2\}$	$v(1,2)$	$v\|_P(1)+v\|_P(2) = v(1)+v(2)$	0
$\{1,3\}$	$v(1)+v(3)$	$v\|_P(1,3) = v(1)+v(3)$	0
$\{1,4\}$	$v(1)+v(4)$	$v\|_P(1)+v\|_P(4) = v(1)+v(4)$	0
$\{2,3\}$	$v(2)+v(3)$	$v\|_P(2,3) = v(2)+v(3)$	0
$\{2,4\}$	$v(2)+v(4)$	$v\|_P(2)+v\|_P(4) = v(2)+v(4)$	0
$\{3,4\}$	$v(3,4)$	$v\|_P(3)+v\|_P(4) = v(3)+v(4)$	0
$\{1,2,3\}$	$v(1,2)+v(3)$	$v\|_P(1,2,3) = v(1,2)+v(3)$	$v(1,2)$ $-v(1)-v(2)$
$\{1,2,4\}$	$v(1,2)+v(4)$	$v\|_P(1)+v\|_P(2)+v\|_P(4) = v(1)+v(2)+v(4)$	0
$\{1,3,4\}$	$v(1)+v(3,4)$	$v\|_P(1,3)+v\|_P(4) = v(1)+v(3)+v(4)$	0
$\{2,3,4\}$	$v(2)+v(3,4)$	$v\|_P(2,3)+v\|_P(4) = v(2)+v(3)+v(4)$	0
N	$v(N)$	$v\|_P(1,2,3)+v\|_P(4) = v(1,2)+v(3)+v(4)$	0

Table 6.2: Characteristic function and dividends of $(N, (v|_P)^L)$.

of the unanimity games (N, u^T). It is not hard to see that if there is a $j \in M$ with $T \subseteq P_j$, or there is a $Q \subseteq M$ such that $T = \cup_{q \in Q} P_q$, then $Z_P(u^T) = u^T$. But, if T is not of this form, then it holds that $u^T|_P(S) = \sum_{R \in S/P} u^T(R) = \sum_{\{R \in S/P | T \subseteq R\}} u^T(R)$ for all $S \in 2^N \setminus \{\emptyset\}$. Hence, $Z_P(u^T) = d^T$, where d^T is the game given by

$$d^T(S) = \begin{cases} 1 & \text{if there is an } R \in S/P \text{ such that } T \subseteq R \subseteq S, \\ 0 & \text{otherwise.} \end{cases}$$

Since $v = \sum_{S \in 2^N \setminus \{\emptyset\}} \Delta^S(v)\, u^S$ for any game $(N, v) \in \mathcal{G}$, it holds that

$$d^T = \sum_{S \in 2^N \setminus \{\emptyset\}} \Delta^S(d^T) u^S.$$

Note that $\Delta^S(d^T) = 0$ unless $T \subseteq S$. In addition, the next proposition holds.

Proposition 6.3.3 *Let $P = \{P_j | j \in M\}$ be a partition of N and $S \subseteq N$. If there is no $j \in M$ with $S \subseteq P_j$, and there is no $Q \subseteq M$ such that $S = \bigcup_{q \in Q} P_q$, then $\Delta^S(d^T) = 0$ for all $T \in 2^N \setminus \{\emptyset\}$.*

Proof. The proof proceeds along the same lines as the proof of Theorem 2 in Owen (1986). Let $S \in 2^N \setminus \{\emptyset\}$ be such that there is no $j \in M$ with $S \subseteq P_j$, and there is no $Q \subseteq M$ with $S = \bigcup_{q \in Q} P_q$. Let $T \in 2^N \setminus \{\emptyset\}$ be arbitrary.

If $T \not\subseteq S$, then $d^T(H) = 0$ for all $H \subseteq S$. So, $\Delta^S(d^T) = 0$. Also, if $T \subseteq S$ but there is no $R \in S/P$ with $T \subseteq R \subseteq S$, then $d^T(H) = 0$ for all $H \subseteq S$. Again, $\Delta^S(d^T) = 0$.

Next, suppose that $T \subseteq S$ and there is $R \in S/P$ with $T \subseteq R \subseteq S$. For $H \subseteq S$, write $H = H_1 \cup H_2$ with $H_1 \subseteq R$ and $H_2 \subseteq S \setminus R$. It is not difficult to see that $d^T(H) = d^T(H_1)$. Then,

$$\begin{aligned} \Delta^S(d^T) &= \sum_{H \subseteq S} (-1)^{|S|-|H|} d^T(H) \\ &= \sum_{H_1 \subseteq R} \sum_{H_2 \subseteq S \setminus R} (-1)^{|R|-|H_1|} (-1)^{|S|-|R|-|H_2|} d^T(H_1) \\ &= \sum_{H_1 \subseteq R} (-1)^{|R|-|H_1|} d^T(H_1) \Big[\sum_{H_2 \subseteq S \setminus R} (-1)^{|S|-|R|-|H_2|} \Big]. \end{aligned}$$

With $h_2 = |H_2|$ it follows that

$$\Delta^S(d^T) = \sum_{H_1 \subseteq R} (-1)^{|R|-|H_1|} d^T(H_1) \Big[\sum_{h_2=0}^{|S|-|R|} (-1)^{|S|-|R|-h_2} \binom{|S|-|R|}{h_2} \Big], \tag{6.4}$$

because $S \setminus R$ has $\binom{|S|-|R|}{|H_2|}$ subsets of cardinality $|H_2|$. A lemma in Owen (1986), that follows directly from the binomial expansion of $(-1+1)^n$, states that for any integer $n \geq 0$,

$$\sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} = \begin{cases} 0 & \text{if } n \geq 1, \\ 1 & \text{if } n = 0. \end{cases} \tag{6.5}$$

Because there is no $j \in M$ with $S \subseteq P_j$, and there is no $Q \subseteq M$ with $S = \bigcup_{q \in Q} P_q$, and $R \in S/P$ (so that there is a $j \in M$ with $R \subseteq P_j$, or there is a $Q \subseteq M$ with $R = \bigcup_{q \in Q} P_q$) it holds that $S \setminus R \neq \emptyset$, and thus $|S| - |R| \geq 1$. It then follows from (6.5) that the last bracket in equation (6.4) is zero. It can be concluded that $\Delta^S(d^T) = 0$.

$\square$

Let $\mathcal{S}$ be defined by

$$\mathcal{S} = \{S \subseteq N \mid S \subseteq P_j \text{ for some } j \in M\} \cup \{S \subseteq N \mid S = \cup_{q \in Q} P_q \text{ for some } Q \subseteq M\}.$$

Then Proposition 6.3.3 leads to the next theorem.

Theorem 6.3.4 *Let $P = \{P_j | j \in M\}$ be a partition of N. Then the unanimity games u^S, $S \in \mathcal{S}$, form a basis for the image of Z_P.*

Proof. It follows from Proposition 6.3.3 that the image $Z_P(u^T)$ of any unanimity game u^T is a linear combination of unanimity games u^S, $S \in \mathcal{S}$. Additionally, if $S \in \mathcal{S}$, then u^S is its own image. This implies that the u^S, $S \in \mathcal{S}$, span the image space and, because they are independent, form a basis for it.

$\square$

It follows from Theorem 6.3.4 that $\Delta^S(v|_P) = 0$ for any coalition S such that there is no $j \in M$ with $S \subseteq P_j$, and there is no $Q \subseteq M$ such that $S = \bigcup_{q \in Q} P_q$. That is, any coalition that is neither a subset of a union, nor a union of unions has a zero dividend in the partition-restricted TU-game. Since the graph-partition restricted game is defined as the partition restricted game of the graph restricted game, this leads to the next corollary.

Corollary 6.3.5 *For every $(N, v, L, P) \in \mathcal{CGG}$ and every $S \in 2^N \setminus \{\emptyset\}$, if there is no $j \in M$ with $S \subseteq P_j$, and there is no $Q \subseteq M$ such that $S = \bigcup_{q \in Q} P_q$, then $\Delta^S(v^L|_P) = 0$.*

The direct analog of this corollary does not hold for the partition-graph restricted game $(N, (v|_P)^L)$. For instance, in Example 6.3.1 it holds that $S = \{1, 2, 3\}$ is not a subset of any P_j and $S \neq \bigcup_{q \in Q} P_q$ for all $Q \subseteq M$. However, $\Delta^{\{1,2,3\}}((v|_P)^L) = v(\{1, 2\}) - v(\{1\}) - v(\{2\})$.

Above, we represented the restrictions on the cooperation possibilities of the players in a TU-game $v \in \mathcal{G}^N$, given by the partition $P \in \mathcal{P}^N$, by the linear mapping Z_P. In general, it is possible to define a *restriction* on the cooperation possibilities of the players in a TU-game $v \in \mathcal{G}^N$ as a linear mapping $Z : \mathcal{G}^N \to \mathcal{G}^N$. Because every linear mapping from $\mathcal{G}^N$ to $\mathcal{G}^N$ can be represented as a matrix it follows that any restriction can be represented as a matrix. For instance, because the TU-game $v \in \mathcal{G}^N$ can be seen as a vector in the $2^{|N|}$-dimensional Euclidean space $\mathbb{R}^{|N|}$, the partition $P = \{\{1, 2\}, \{3\}\}$ of $N = \{1, 2, 3\}$ can be represented by the $2^3 \times 2^3$-dimensional matrix A in the following matrix equation:

$$\begin{pmatrix} v|_P(\emptyset) \\ v|_P(\{1\}) \\ v|_P(\{2\}) \\ v|_P(\{3\}) \\ v|_P(\{1,2\}) \\ v|_P(\{1,3\}) \\ v|_P(\{2,3\}) \\ v|_P(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&1&0&1&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&1 \end{pmatrix} \begin{pmatrix} v(\emptyset) \\ v(\{1\}) \\ v(\{2\}) \\ v(\{3\}) \\ v(\{1,2\}) \\ v(\{1,3\}) \\ v(\{2,3\}) \\ v(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} v(\emptyset) \\ v(\{1\}) \\ v(\{2\}) \\ v(\{3\}) \\ v(\{1,2\}) \\ v(\{1\})+v(\{3\}) \\ v(\{2\})+v(\{3\}) \\ v(\{1,2,3\}) \end{pmatrix}.$$

Similarly, the graph $L = \{\{1,3\}\}$ on $N = \{1,2,3\}$ can be represented by the $2^3 \times 2^3$-dimensional matrix B in the next matrix equation:

$$\begin{pmatrix} v^L(\emptyset) \\ v^L(\{1\}) \\ v^L(\{2\}) \\ v^L(\{3\}) \\ v^L(\{1,2\}) \\ v^L(\{1,3\}) \\ v^L(\{2,3\}) \\ v^L(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&1&1&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&1&0&0&1&0&0 \end{pmatrix} \begin{pmatrix} v(\emptyset) \\ v(\{1\}) \\ v(\{2\}) \\ v(\{3\}) \\ v(\{1,2\}) \\ v(\{1,3\}) \\ v(\{2,3\}) \\ v(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} v(\emptyset) \\ v(\{1\}) \\ v(\{2\}) \\ v(\{3\}) \\ v(\{1\})+v(\{2\}) \\ v(\{1,3\}) \\ v(\{2\})+v(\{3\}) \\ v(\{1,3\})+v(\{2\}) \end{pmatrix}.$$

Since $AB \neq BA$ this immediately reveals that the graph-partition restricted TU-game $(N, v^L|_P)$ is not equal to the partition-graph restricted TU-game $(N, (v|_P)^L)$ in this example. This gives another illustration of the fact that the order in which cooperation restrictions are applied to a TU-game $v \in \mathcal{G}^N$ matters.

Conversely, if Y is a real $|N| \times |N|$-dimensional matrix, then it describes a linear mapping $Y : \mathbb{R}^{|N|} \to \mathbb{R}^{|N|}$. This implies that any real $|N| \times |N|$-dimensional matrix can serve as a restriction on the cooperation possibilities of the players in a TU-game (as long as the worth of the empty coalition remains $v(\emptyset) = 0$). To see this, consider, for instance, the restricted game (N, v^r) implied by the $2^3 \times 2^3$-dimensional matrix in the next matrix equation:

$$\begin{pmatrix} v^r(\emptyset) \\ v^r(\{1\}) \\ v^r(\{2\}) \\ v^r(\{3\}) \\ v^r(\{1,2\}) \\ v^r(\{1,3\}) \\ v^r(\{2,3\}) \\ v^r(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&2&0&0&0&0&0 \\ 0&1&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&1&0&0&5&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1 \end{pmatrix} \begin{pmatrix} v(\emptyset) \\ v(\{1\}) \\ v(\{2\}) \\ v(\{3\}) \\ v(\{1,2\}) \\ v(\{1,3\}) \\ v(\{2,3\}) \\ v(\{1,2,3\}) \end{pmatrix} = \begin{pmatrix} v(\emptyset) \\ v(\{1\}) \\ 2v(\{2\}) \\ v(\{1\})+v(\{3\}) \\ v(\{1,2\}) \\ 5v(\{1,3\})+v(\{2\}) \\ v(\{1\}) \\ v(\{1,2,3\}) \end{pmatrix}.$$

The cooperation restrictions implied by this matrix cannot be represented by a partition P or by a graph L. It might, however, be possible to find classes of matrices that represent only partition restrictions, or only graph restrictions.

6.4 Two Shapley-type values for TU-games with coalition and graph structure

In this section we define two new values for TU-games with coalition and graph structure as the Shapley values of the two types of restricted game introduced in the previous section. We show that the Shapley value of the graph-partition restricted game can be characterized by the axioms of 'graph efficiency', 'balanced contributions' and 'collective balanced contributions' and that the Shapley value of the partition-graph restricted game can be characterized by the axioms of 'partition component efficiency' and 'fairness'. Recall from Section 2.1 that a value f on $\mathcal{CGG}$ assigns a unique payoff vector $f(N, v, L, P) \in \mathbb{R}^N$ to every TU-game with coalition and graph structure $(N, v, L, P) \in \mathcal{CGG}$.

Definition 6.4.1
(1) The graph-partition value on the class of TU-games with coalition and graph structure is the value ϕ assigning to every $(N, v, L, P) \in \mathcal{CGG}$ the payoff vector $\phi(N, v, L, P) = Sh(N, v^L|_P)$.
(2) The partition-graph value on the class of TU-games with coalition and graph structure is the value ψ assigning to every $(N, v, L, P) \in \mathcal{CGG}$ the payoff vector $\psi(N, v, L, P) = Sh(N, (v|_P)^L)$.

Note that $\phi(N, v, L, P) = Ka(N, v^L, P)$, where Ka represents the collective value for TU-games with coalition structure (as discussed in Section 2.1) and that $\psi(N, v, L, P) = My(N, v|_P, L)$, where My represents the Myerson value for TU-games with graph structure (also discussed in Section 2.1). Because $v^L|_P$ does not have to be equal to $(v|_P)^L$, in general $\phi(N, v, L, P)$ is not equal to $\psi(N, v, L, P)$.

Example 6.4.2 Let $(N, v, L, P) \in \mathcal{CGG}$ be as in Example 6.3.1. From the dividends derived in that example it follows straightforwardly that the graph-partition value is given by

$$\phi_i(N, v, L, P) = v(\{i\}) + \frac{1}{4}[v(\{1, 2, 3\}) - v(\{1\}) - v(\{2\}) - v(\{3\})],\ i = 1, 2, 3, 4$$

and that the partition-graph value is given by

$$\psi_j(N, v, L, P) = v(\{j\}) + \frac{1}{3}[v(\{1, 2\}) - v(\{1\}) - v(\{2\})],\ j = 1, 2, 3,$$

and $\psi_4(N, v, L, P) = v(\{4\})$.

□

It is obvious that for special structures the graph-partition value and partition-graph value are equal to the Myerson value and Kamijo's collective value, respectively. By definition, the Myerson value only takes into account the graph structure and ignores the coalition structure. The graph-partition value is therefore equal to the Myerson value when $P = \{N\}$.

Proposition 6.4.3 *Let* $(N, v, L, P) \in \mathcal{CGG}$*. If* $P = \{N\}$ *then*

$$\phi(N, v, L, P) = My(N, v, L).$$

The collective value of Kamijo (2011) only takes into account the coalition structure and ignores the graph structure. The partition-graph value is therefore equal to the collective value when L is the complete graph.

Proposition 6.4.4 *Let* $(N, v, L, P) \in \mathcal{CGG}$*. If* $L = \{\{i, j\} \mid i, j \in N,\ i \neq j\}$ *then*

$$\psi(N, v, L, P) = Ka(N, v, P).$$

To characterize the graph-partition and partition-graph value for TU-games with coalition and graph structure we generalize the axiomatizations of the collective value for TU-games with coalition structure (Kamijo, 2011) and the Myerson value for TU-games with graph structure (Myerson, 1977) to the class of TU-games with coalition and graph structure.

A characterization of the graph-partition value

The efficiency property as implicitly used in Kamijo (2011) states that the players in N distribute the worth $v(N)$ among themselves. Here we formulate this axiom in the context of games with coalition and graph structure.

Axiom 6.4.5 Efficiency
A value f *on the class of TU-games with coalition and graph structure* $\mathcal{CGG}$ *is efficient if for any* $(N, v, L, P) \in \mathcal{CGG}$ *it holds that* $\sum_{i \in N} f_i(N, v, L, P) = v(N)$*.*

The graph-partition value does not satisfy this axiom in general. However, it does satisfy an alternative version stating that the players in N distribute the sum of the worths of the connected components of (N, L) among themselves. This takes into account that in a game with graph structure players can only cooperate when they are connected in the graph.

Axiom 6.4.6 Graph efficiency
A value f *on the class of TU-games with coalition and graph structure* $\mathcal{CGG}$ *is graph efficient if for any* $(N, v, L, P) \in \mathcal{CGG}$ *it holds that* $\sum_{i \in N} f_i(N, v, L, P) = \sum_{K \in C^L(N)} v(K)$*.*

Clearly, when N is connected in (N, L), then for every solution f satisfying graph efficiency it holds that $\sum_{i \in N} f_i(N, v, L, P) = v(N)$.

Next, we generalize the balanced contributions axiom for TU-games with coalition structure, used in Kamijo (2011), to the setting of TU-games with coalition and graph structure. This axiom states that, given the coalition structure $P = \{N\}$, the loss in value that player $i \in N$ experiences when player $j \in N$ leaves the game is equal to the loss in value that player j experiences when player i leaves the game. For convenience, for every $j \in N$, we denote $N_{-j} = N \setminus \{j\}$, $v_{-j} = v_{N \setminus \{j\}}$ (which is the characteristic function of the subgame on $N \setminus \{j\}$, see Section 2.1) and $L_{-j} = L(N \setminus \{j\})$.

Axiom 6.4.7 Balanced contributions
A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ satisfies balanced contributions if for any $(N, v, L, \{N\}) \in \mathcal{CGG}$ it holds that

$$f_i(N, v, L, \{N\}) - f_i(N_{-j}, v_{-j}, L_{-j}, \{N_{-j}\}) = f_j(N, v, L, \{N\}) - f_j(N_{-i}, v_{-i}, L_{-i}, \{N_{-i}\}),$$

for all $i, j \in N$.

The following collective balanced contributions axiom is a generalization to the setting of TU-games with coalition and graph structure of the collective balanced contributions axiom for TU-games with coalition structure of Kamijo (2011). It states that, given two different unions P_k and P_ℓ in P, for every $i \in P_k$ and $j \in P_\ell$, the loss in value that player i experiences when union $P_\ell \in P$ leaves the game is equal to the loss in value that player j experiences when union $P_k \in P$ leaves the game. Again for convenience, for every $k \in M$, we denote $N_{-P_k} = N \setminus P_k$, $v_{-P_k} = v_{N \setminus P_k}$ (which is the characteristic function of the subgame on $N \setminus P_k$), $L_{-P_k} = L(N \setminus P_k)$ and $P_{-P_k} = P \setminus \{P_k\}$.

Axiom 6.4.8 Collective balanced contributions
A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ satisfies collective balanced contributions if for any $(N, v, L, P) \in \mathcal{CGG}$ with $|P| \geq 2$ it holds that

$$f_i(N, v, L, P) - f_i(N_{-P_\ell}, v_{-P_\ell}, L_{-P_\ell}, P_{-P_\ell}) = f_j(N, v, L, P) - f_j(N_{-P_k}, v_{-P_k}, L_{-P_k}, P_{-P_k})$$

for every two different unions P_k and P_ℓ in P, all $i \in P_k \in P$ and all $j \in P_\ell \in P$.

The axioms 6.4.6-6.4.8 characterize the graph-partition value.

Theorem 6.4.9 *A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ satisfies graph efficiency, balanced contributions and collective balanced contributions if and only if $f(N, v, L, P) = \phi(N, v, L, P)$ for every $(N, v, L, P) \in \mathcal{CGG}$.*

Proof. First, we show that ϕ satisfies graph efficiency, balanced contributions and collective balanced contributions. Graph efficiency follows from

$$\sum_{i \in N} \phi_i(N, v, L, P) = \sum_{i \in N} Sh_i(N, v^L|_P) = \sum_{i \in N} \sum_{\{S \subseteq N | i \in S\}} \frac{\Delta^S(v^L|_P)}{|S|} =$$

$$\sum_{S \subseteq N} \Delta^S(v^L|_P) = v^L|_P(N) = \sum_{S \in N/P} v^L(S) = v^L(N) = \sum_{K \in C^L(N)} v(K),$$

where the first, second, fifth and seventh equalities follow by definition, the third by rearranging terms, the fourth by the expression for the dividends and the sixth because $N/P = \{N\}$.

Next, for every pair $i, j \in N$ it holds that

$$\phi_i(N, v, L, \{N\}) - \phi_i(N_{-j}, v_{-j}, L_{-j}, \{N_{-j}\}) =$$

$$Sh_i(N, v^L|_{\{N\}}) - Sh_i(N_{-j}, (v_{-j})^{L_{-j}}|_{\{N_{-j}\}}) = Sh_i(N, v^L) - Sh_i(N_{-j}, (v_{-j})^{L_{-j}}) =$$

$$My_i(N,v,L) - My_i(N_{-j},v_{-j},L_{-j}) = My_j(N,v,L) - My_j(N_{-i},v_{-i},L_{-i}) =$$
$$Sh_j(N,v^L) - Sh_j(N_{-i},(v_{-i})^{L_{-i}}) = Sh_j(N,v^L|_{\{N\}}) - Sh_j(N_{-i},(v_{-i})^{L_{-i}}|_{\{N_{-i}\}}) =$$
$$\phi_j(N,v,L,\{N\}) - \phi_j(N_{-i},v_{-i},L_{-i},\{N_{-i}\}),$$

where all the equalities follow by definition, except the fourth equality, which follows because the value of Myerson (1977) satisfies balanced contributions for TU-games with graph structure.[6] Hence, ϕ satisfies balanced contributions.

Now, given $L \in \mathcal{L}^N$ and $P_j \in P \in \mathcal{P}^N$ consider the TU-games $(N_{-P_j},(v_{-P_j})^{L_{-P_j}})$ and $(N_{-P_j},(v^L)_{-P_j})$. Because for all $S \subseteq N \setminus P_j$

$$(v_{-P_j})^{L_{-P_j}}(S) = \sum_{T \in C^{L_{-P_j}}(S)} v_{-P_j}(T) = \sum_{T \in C^{L_{-P_j}}(S)} v(T) = v^{L_{-P_j}}(S) = v^L(S) = (v^L)_{-P_j}(S),$$

it holds that these TU-games are equal. Next, consider any $(N,v,L,P) \in \mathcal{CGG}$ with $|P| \geq 2$ and take any $i \in P_k \in P$ and any $j \in P_\ell \in P$, $P_k \neq P_\ell$. Then,

$$\phi_i(N,v,L,P) - \phi_i(N_{-P_\ell},v_{-P_\ell},L_{-P_\ell},P_{-P_\ell}) = Ka_i(N,v^L,P) - Ka_i(N_{-P_\ell},(v_{-P_\ell})^{L_{-P_\ell}},P_{-P_\ell})$$
$$= Ka_i(N,v^L,P) - Ka_i(N_{-P_\ell},(v^L)_{-P_\ell},P_{-P_\ell}) = Ka_j(N,v^L,P) - Ka_j(N_{-P_k},(v^L)_{-P_k},P_{-P_k})$$
$$= Ka_j(N,v^L,P) - Ka_j(N_{-P_k},(v_{-P_k})^{L_{-P_k}},P_{-P_k})$$
$$= \phi_j(N,v,L,P) - \phi_j(N_{-P_k},v_{-P_k},L_{-P_k},P_{-P_k}),$$

where the first and last equality follow by definition, the second and fourth because the TU-games in the expressions are equal and the third because the value of Kamijo (2011) satisfies collective balanced contributions for TU-games with coalition structure (see Kamijo (2011)). Hence, ϕ satisfies collective balanced contributions.

Second, we show that there can be at most one value that satisfies graph efficiency, balanced contributions and collective balanced contributions.

Suppose that f satisfies these axioms and consider first all games $(N,v,L,P) \in \mathcal{CGG}$ with $P = \{N\}$. We uniquely determine $f(N,v,L,P)$ for these games by induction on the number of players n. When $n = 1$ it follows directly from graph efficiency that $f_i(\{i\},v,L,\{\{i\}\}) = v(\{i\}) = \phi_i(\{i\},v,L,\{\{i\}\})$, $i \in N$. Next, suppose that f has been uniquely determined for all games $(K,v,L,P) \in \mathcal{CGG}$ with $P = \{K\}$ and $|K| \leq n-1$. Then applying the balanced contributions property to f for $(N,v,L,P) \in \mathcal{CGG}$ with $P = \{N\}$ and $|N| = n$ gives

$$f_i(N,v,L,\{N\}) - f_i(N_{-j},v_{-j},L_{-j},\{N_{-j}\}) = f_j(N,v,L,\{N\}) - f_j(N_{-i},v_{-i},L_{-i},\{N_{-i}\})$$

for all $i,j \in N$. Notice that for all $i,j \in N$, the values $f_i(N_{-j},v_{-j},L_{-j},\{N_{-j}\})$ and $f_j(N_{-i},v_{-i},L_{-i},\{N_{-i}\})$ are known by the induction hypothesis. For some particular $i \in N$, say $i = i_0$, there now are $n-1$ equations of the above type with $i = i_0$. Together with the graph efficiency equation this gives a system of $(n-1)+1 = n$ linearly independent equations in n unknowns. Hence, this system uniquely determines $f_i(N,v,L,\{N\})$, $i \in N$.

[6] This follows in a similar way as is shown in Myerson (1980) for a fixed player set.

Finally, consider all games $(N, v, L, P) \in \mathcal{CGG}$ with $|P| \geq 2$. Suppose that there are two different values f^1 and f^2 that both satisfy graph efficiency, balanced contributions and collective balanced contributions. Let P be a partition with a minimum number of unions (elements of P) such that $f^1(N, v, L, P) \neq f^2(N, v, L, P)$. It follows by the minimality of P that if P_g is any element of P, then $f^1(N_{-P_g}, v_{-P_g}, L_{-P_g}, P_{-P_g}) = f^2(N_{-P_g}, v_{-P_g}, L_{-P_g}, P_{-P_g})$. Now, by collective balanced contributions (and rearranging terms) it holds for all $i \in P_k \in P$ and all $j \in P_\ell \in P$, $P_k \neq P_\ell$, that

$$f_i^1(N, v, L, P) - f_j^1(N, v, L, P) = f_i^1(N_{-P_\ell}, v_{-P_\ell}, L_{-P_\ell}, P_{-P_\ell}) - f_j^1(N_{-P_k}, v_{-P_k}, L_{-P_k}, P_{-P_k}) =$$

$$f_i^2(N_{-P_\ell}, v_{-P_\ell}, L_{-P_\ell}, P_{-P_\ell}) - f_j^2(N_{-P_k}, v_{-P_k}, L_{-P_k}, P_{-P_k}) = f_i^2(N, v, L, P) - f_j^2(N, v, L, P).$$

So, $f_i^1(N, v, L, P) - f_i^2(N, v, L, P) = f_j^1(N, v, L, P) - f_j^2(N, v, L, P)$ for any $i \in P_k \in P$ and $j \in P_\ell \in P$, $P_k \neq P_\ell$. This, in turn, implies that there exists a $\beta \in \mathbb{R}$ such that $f_i^1(N, v, L, P) - f_i^2(N, v, L, P) = \beta$, for all $i \in N$. It follows from graph efficiency that $\sum_{i \in N} f_i^1(N, v, L, P) = \sum_{K \in C^L(N)} v(K) = \sum_{i \in N} f_i^2(N, v, L, P)$ so that

$$0 = \sum_{i \in N} \Big(f_i^1(N, v, L, P) - f_i^2(N, v, L, P) \Big) = |N| \beta.$$

Since $|N| > 0$ this means that $\beta = 0$, so that $f^1(N, v, L, P) = f^2(N, v, L, P)$, a contradiction.

□

Logical independence of the axioms in this theorem is shown by giving three alternative solutions. Each of these solutions only satisfies two of the three axioms.

1. The value $f(N, v, L, P) = Ka(N, v, P)$ satisfies balanced contributions and collective balanced contributions. It does not satisfy graph efficiency.

2. The value $f(N, v, L, P) = My(N, v, L)$ satisfies graph efficiency and balanced contributions. It does not satisfy collective balanced contributions.

3. The value $f_i(N, v, L, P) = \frac{Sh_j^w(M, (v^L)^P)}{|P_j|}$ for all $i \in P_j$, $j \in M$, where Sh^w is the weighted Shapley value (see Kalai and Samet (1987)) with vector of weights $w = (|P_j|)_{j \in M}$ and the quotient game $(M, (v^L)^P)$ is as described in Section 2.1, satisfies graph efficiency and collective balanced contributions. It does not satisfy balanced contributions.

From the proof of Theorem 6.4.9 it follows directly that the Myerson value for TU-games with graph structure (Myerson, 1977) can be characterized be the following two axioms.

Axiom 6.4.10 Graph efficiency for graph games
A value f on the class of TU-games graph structure $\mathcal{GG}$ is graph efficient if for any $(N, v, L) \in \mathcal{GG}$ it holds that $\sum_{i \in N} f_i(N, v, L) = \sum_{K \in C^L(N)} v(K)$.

Axiom 6.4.11 Balanced contributions for graph games
A value f on the class of TU-games graph structure $\mathcal{GG}$ satisfies balanced contributions if for any $(N,v,L) \in \mathcal{GG}$ it holds that

$$f_i(N,v,L) - f_i(N_{-j}, v_{-j}, L_{-j}) = f_j(N,v,L) - f_j(N_{-i}, v_{-i}, L_{-i}),$$

for all $i, j \in N$.

Thus, the following corollary holds.

Corollary 6.4.12 *A value f on the class of TU-games with graph structure $\mathcal{GG}$ satisfies graph efficiency for graph games and balanced contributions for graph games if and only if $f(N,v,L) = My(N,v,L)$ for every $(N,v,L) \in \mathcal{GG}$.*

In addition, we can generalize the characterization of the collective value by Kamijo (2011) to the class of TU-games with coalition and graph structure. We do this by replacing graph efficiency in Theorem 6.4.9 by efficiency.

Corollary 6.4.13 *A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ satisfies efficiency, balanced contributions and collective balanced contributions if and only if $f(N,v,L,P) = Ka(N,v,P)$ for every $(N,v,L,P) \in \mathcal{CGG}$.*

Since graph efficiency takes into account that the players in N can only realize the sum of the worths of the components of the graph (N,L), the value ϕ can thus be seen as a modification of Kamijo's collective value to the setting of TU-games with coalition and graph structure.

A characterization of the partition-graph value

The partition-graph value can be characterized by axioms similar to those of the original characterization of the Myerson value for graph games (Myerson, 1977). The component efficiency axiom for TU-games with graph structure of Myerson (1977) says that the players in a component of the graph distribute exactly the worth of this component among themselves. Component efficiency for TU-games with coalition and graph structure is stated in Vázquez-Brage, García-Jurado and Carreras (1996) and similarly says that, for every component K in $C^L(N)$, the players of K distribute the worth $v(K)$ among themselves.[7]

Axiom 6.4.14 Component efficiency
A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ is component efficient if for any $(N,v,L,P) \in \mathcal{CGG}$ it holds that $\sum_{i \in K} f_i(N,v,L,P) = v(K)$, for all $K \in C^L(N)$.

[7]Note that any solution for TU-games with coalition and graph structure that satisfies component efficiency, also satisfies graph efficiency.

The partition-graph value does not satisfy this axiom in general, because it takes into account that within a component K of (N, L) the players can only realize the sum of the worths of the coalitions T in K/P. However, it does satisfy that the players in every component K of (N, L) distribute the worths of the coalitions in K/P among themselves, as stated in the next axiom.

Axiom 6.4.15 Partition component efficiency
A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ is partition component efficient if for any $(N, v, L, P) \in \mathcal{CGG}$ it holds that $\sum_{i\in K} f_i(N, v, L, P) = \sum_{T\in K/P} v(T)$, for all $K \in C^L(N)$.

When (N, L) is connected, then N is the unique component of (N, L) and also the unique element of N/P. In this case partition component efficiency implies efficiency.

Next, we translate the fairness axiom of Myerson (1977) to the setting of TU-games with coalition and graph structure. Given two different players $i, j \in N$ that are linked in the graph (N, L) (i.e., $\{i, j\} \in L$), we require that both their values change by the same amount when the link between them is severed. To simplify notation we write $L \setminus \{i, j\}$ instead of $L \setminus \{\{i, j\}\}$.

Axiom 6.4.16 Fairness
A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ is fair if for any $(N, v, L, P) \in \mathcal{CGG}$ it holds that

$$f_i(N, v, L, P) - f_i(N, v, L \setminus \{i, j\}, P) = f_j(N, v, L, P) - f_j(N, v, L \setminus \{i, j\}, P)$$

for all $i, j \in N$ such that $\{i, j\} \in L$.

The axioms 6.4.15 and 6.4.16 characterize the partition-graph value.

Theorem 6.4.17 *A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ satisfies partition component efficiency and fairness if and only if $f(N, v, L, P) = \psi(N, v, L, P)$ for every $(N, v, L, P) \in \mathcal{CGG}$.*

Proof. First, we show that ψ satisfies partition component efficiency and fairness. For every $K \in C^L(N)$ it holds that

$$\sum_{i\in K} \psi_i(N, v, L, P) = \sum_{i\in K} Sh_i(N, (v|_P)^L) = \sum_{i\in K} \sum_{\{S\subseteq N | i\in S\}} \frac{\Delta^S((v|_P)^L)}{|S|} =$$

$$\sum_{S\subseteq K} \Delta^S((v|_P)^L) = (v|_P)^L(K) = \sum_{S\in C^L(K)} v|_P(S) = v|_P(K) = \sum_{T\in K/P} v(T),$$

where the first two equalities follow by definition, the third by rearranging terms, the fourth by the expression for the dividends and the last three again by definition. Hence, ψ satisfies partition component efficiency.

Next, for every pair $i, j \in N$ such that $\{i, j\} \in L$ it holds that

$$\psi_i(N, v, L, P) - \psi_i(N, v, L \setminus \{i, j\}, P) = My_i(N, v|_P, L) - My_i(N, v|_P, L \setminus \{i, j\}) =$$

$$My_j(N, v|_P, L) - My_j(N, v|_P, L \setminus \{i,j\}) = \psi_j(N, v, L, P) - \psi_j(N, v, L \setminus \{i,j\}, P),$$

where the first and the last equality follow by definition and the second because the value of Myerson (1977) satisfies fairness for TU-games with graph structure (see Myerson (1977)). So, ψ satisfies fairness.

Second, we show that there can be at most one value that satisfies partition component efficiency and fairness. This proceeds along the same lines as the first part of the proof of the Theorem in Myerson (1977). If $i \in K \in C^L(N)$ with $|K| = 1$, then partition component efficiency determines that $f_i(N, v, L, P) = \psi_i(N, v, L, P)$.

Next, suppose that there are two different values f^1 and f^2 that both satisfy partition component efficiency and fairness. Let (N, L) be a graph with a minimum number of links (elements of L) such that $f^1(N, v, L, P) \neq f^2(N, v, L, P)$ (note that $|L| > 0$ in this case). If $\{i,j\} \in L$ is a given link of L then it follows by the minimality of L that $f^1(N, v, L \setminus \{i,j\}, P) = f^2(N, v, L \setminus \{i,j\}, P)$. By the fairness axiom (and rearranging terms) it therefore holds that

$$f_i^1(N, v, L, P) - f_j^1(N, v, L, P) = f_i^1(N, v, L \setminus \{i,j\}, P) - f_j^1(N, v, L \setminus \{i,j\}, P) =$$

$$f_i^2(N, v, L \setminus \{i,j\}, P) - f_j^2(N, v, L \setminus \{i,j\}, P) = f_i^2(N, v, L, P) - f_j^2(N, v, L, P).$$

Since this holds for any $\{i,j\} \in L$ it also holds for all $i, j \in K \in C^L(N)$. Hence, there exists a $\beta_K(L) \in \mathbb{R}$ such that

$$f_i^1(N, v, L, P) - f_i^2(N, v, L, P) = \beta_K(L)$$

for all $i \in K \in C^L(N)$. Note that $\beta_K(L)$ depends only on K and L but not on i. It follows from partition component efficiency that $\sum_{i \in K} f_i^1(N, v, L, P) = \sum_{T \in K/P} v(T) = \sum_{i \in K} f_i^2(N, v, L, P)$ so that

$$0 = \sum_{i \in K} \Big(f_i^1(N, v, L, P) - f_i^2(N, v, L, P) \Big) = |K| \beta_K(L).$$

Since $|K| > 1$ this implies that $\beta_K(L) = 0$, so that $f^1(N, v, L, P) = f^2(N, v, L, P)$, a contradiction.

□

Logical independence of the axioms in this theorem is shown by giving two alternative solutions. Each of these solutions only satisfies one of the two axioms.

1. The value $f_i(N, v, L, P) = \frac{v|_P(K)}{|K|}$ for all $i \in K$, $K \in C^L(N)$, satisfies partition component efficiency. It does not satisfy fairness.

2. The value $f(N, v, L, P) = My(N, v, L)$ satisfies fairness. It does not satisfy partition component efficiency.

We can generalize the characterization of the Myerson value (Myerson, 1977) to the class of TU-games with coalition and graph structure. We do this by replacing partition component efficiency in Theorem 6.4.17 by component efficiency.

Corollary 6.4.18 *A value f on the class of TU-games with coalition and graph structure $\mathcal{CGG}$ satisfies component efficiency and fairness if and only if $f(N,v,L,P) = My(N,v,L)$ for every $(N,v,L,P) \in \mathcal{CGG}$.*

Since partition component efficiency takes into account that the players in a component K of (N,L) can only realize the sum of the worths of the elements of K/P, the value ψ can thus be seen as a modification of the Myerson value to the setting of TU-games with coalition and graph structure.

6.5 Solutions for river basin games

As mentioned in Section 6.2, a river basin game is a TU-game with coalition and graph structure. Because the graph $(N,\widehat{D})$ in a river basin game is connected, it follows that

$$v(N) = v^{\widehat{D}}(N) = v|_P(N) = v^{\widehat{D}}|_P(N) = (v|_P)^{\widehat{D}}(N)$$

for any river basin game $(N,v,\widehat{D},P)$. Furthermore, since by Assumption 2.2.1 of Chapter 2 the characteristic function v of a river basin game is superadditive, as can be seen from (6.3) in Section 6.2, it follows that for any river basin game

$$v^{\widehat{D}}|_P(S) = \sum_{T\in S/P} v^{\widehat{D}}(T) \leq v^{\widehat{D}}(S), \text{ for all } S \subseteq N,$$

and

$$(v|_P)^{\widehat{D}}(S) = \sum_{T\in C^{\widehat{D}}(S)} v|_P(T) = \sum_{T\in C^{\widehat{D}}(S)} \sum_{R\in T/P} v(R) \leq \sum_{T\in C^{\widehat{D}}(S)} v(T) = v^{\widehat{D}}(S), \text{ for all } S \subseteq N.$$

Hence, both in the graph-partition restricted game $(N, v^{\widehat{D}}|_P)$ as well as in the partition-graph restricted game $(N,(v|_P)^{\widehat{D}})$ the worth of any coalition $S \subseteq N$ is not larger than the maximum welfare the agents in S can obtain under Assumption 2.2.1 of Chapter 2, i.e., the worth $v^{\widehat{D}}(S)$.

We can apply both the graph-partition value ϕ, as well as the partition-graph value ψ, to river basin games. The graph-partition value ϕ can be implemented by the welfare distribution (x^N,t) with x^N as defined in (6.3) of Section 6.2 for $S = N$ and $t_i = \phi_i(N,v,\widehat{D},P) - b_i(x_i^N)$ for all $i \in N$. The partition-graph value ψ can be implemented by the welfare distribution (x^N,t) with $t_i = \psi_i(N,v,\widehat{D},P) - b_i(x_i^N)$ for all $i \in N$.

Unlike the axioms in Chapters 3 and 4, the axioms that characterize the graph-partition value ϕ and the partition-graph value ψ were not inspired by water distribution principles from international watercourse law. Although this is no problem in itself, it still raises the question whether it is possible to find a value for river basin games that is. A potential candidate is to apply the average tree solution AT, as introduced in Herings, van der Laan and Talman (2008) for cycle-free graph games and Herings, van der Laan, Talman and Yang (2010) for general graph games (and as discussed in Section 2.1), to the TU-game with graph structure $(N, v|_P, L)$. This gives the following generalization of the average tree solution to the class of TU-games with coalition and graph structure:

$$AT(N, v, L, P) = AT(N, v|_P, L)$$

for every $(N, v, L, P) \in \mathcal{CGG}$. Based on the findings of Chapter 4 the average tree solution for TU-games with coalition and graph structure seems to be particularly suitable to apply to river basin games with coalition structure $(N, v, \widehat{D}, P)$. It remains, however, an objective for future research to characterize the average tree solution on the class of river basin games by using axioms that are inspired by water distribution principles.

What we can say, is that it is not too difficult to generalize the characterization of the average tree solution on the class of cycle-free graph games of Herings, van der Laan and Talman (2008) to the class of TU-games with coalition and *cycle-free* graph structure. This only requires a slight modification of two axioms. Axiom 6.4.15 of the previous section can be stated on the class of TU-games with coalition and cycle-free graph structure as follows.

Axiom 6.5.1 Partition component efficiency
A value f on the class of TU-games with coalition and cycle-free graph structure is partition component efficient if for any TU-game with coalition and cycle-free graph structure (N, v, L, P) it holds that $\sum_{i \in K} f_i(N, v, L, P) = \sum_{T \in K/P} v(T)$, for all $K \in C^L(N)$.

A slight modification of the component fairness axiom of Herings, van der Laan and Talman (2008) to the class of TU-games with coalition and cycle-free graph structure gives the following axiom.

Axiom 6.5.2 Component fairness
A value f on the class of TU-games with coalition and cycle-free graph structure is component fair if for any TU-game with coalition and cycle-free graph structure (N, v, L, P) and for any link $\{i, j\} \in L$ it holds that

$$\frac{1}{|K^i|} \sum_{\ell \in K^i} \Big(f_\ell(N, v, L, P) - f_\ell(N, v, L \setminus \{i, j\}, P) \Big) = \frac{1}{|K^j|} \sum_{\ell \in K^j} \Big(f_\ell(N, v, L, P) - f_\ell(N, v, L \setminus \{i, j\}, P) \Big),$$

where K^ℓ, $\ell \in \{i, j\}$, denotes the component in the cycle-free graph $(N, L \setminus \{i, j\})$ that contains player ℓ.

This leads to the next proposition of which we omit the proof because it is similar to the proofs of Theorems 3.4 and 3.7 in Herings, van der Laan and Talman (2008).

Proposition 6.5.3 *A value f on the class of TU-games with coalition and cycle-free graph structure satisfies partition component efficiency and component fairness if and only if*

$$f(N, v, L, P) = AT(N, v, L, P)$$

for every TU-game with coalition and cycle-free graph structure.

We conclude this chapter with an example of a river basin game and a comparison of the three solutions of this chapter for this river basin game.

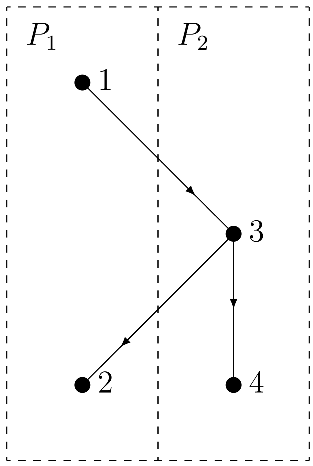

Figure 6.3: River basin from Example 6.5.4.

Example 6.5.4 The tuple (N, D, P, e, b) with $N = \{1, 2, 3, 4\}$, $D = \{(1, 3), (3, 2), (3, 4)\}$, $P = \{\{1, 2\}, \{3, 4\}\}$, $e = (50, 0, 50, 0)$ and $b(x_i) = \sqrt{x_i}$, $i \in N$, constitutes a river basin benefit problem, see Figure 6.3. The corresponding river basin game is equal to $(N, v, \widehat{D}, P)$ with characteristic function v as given in Table 6.3 below, and $\widehat{D} = \{\{1, 3\}, \{3, 2\}, \{3, 4\}\}$. This river basin game provides the following values:

City	$\phi(N, v, \widehat{D}, P)$	$\psi(N, v, \widehat{D}, P)$	$AT(N, v, \widehat{D}, P)$
1	7.80	8.05	7.80
2	0.73	0.98	0.73
3	9.27	9.51	10.73
4	2.20	1.46	0.73.

What can be seen in this table is that the value ψ distributes more of the worth of the grand coalition N to the players 1, 2 and 3 than the value ϕ, at the expense of player 4. The values ϕ and AT are equal for players 1 and 2, but AT distributes more to player 3, and less to player 4, than ϕ. Comparing ψ with AT shows that AT distributes more of the worth of the grand coalition to the 'most central' player in the graph $\widehat{D}$, player 3, and less to the other players.

□

S	$v(S)$	$v^L(S)$	$v\|_P(S)$	$v^L\|_P(S)$	$(v\|_P)^L(S)$
$\emptyset$	0	0	0	0	0
1	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$
2	0	0	0	0	0
3	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$
4	0	0	0	0	0
1,2	10	$5\sqrt{2}$	10	$5\sqrt{2}$	$5\sqrt{2}$
1,3	$10\sqrt{2}$	$10\sqrt{2}$	$10\sqrt{2}$	$10\sqrt{2}$	$10\sqrt{2}$
1,4	10	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$
2,3	10	10	$5\sqrt{2}$	$5\sqrt{2}$	$5\sqrt{2}$
2,4	0	0	0	0	0
3,4	10	10	10	10	10
1,2,3	$10\sqrt{3}$	$10\sqrt{3}$	$10 + 5\sqrt{2}$	$10\sqrt{2}$	$10 + 5\sqrt{2}$
1,2,4	$10\sqrt{\frac{3}{2}}$	$5\sqrt{2}$	10	$5\sqrt{2}$	$5\sqrt{2}$
1,3,4	$10\sqrt{3}$	$10\sqrt{3}$	$5\sqrt{2} + 10$	$5\sqrt{2} + 10$	$5\sqrt{2} + 10$
2,3,4	$10\sqrt{\frac{3}{2}}$	$10\sqrt{\frac{3}{2}}$	10	10	10
N	20	20	20	20	20

Table 6.3: Characteristic functions from Example 6.5.4.

Chapter 7

River pollution models

7.1 Introduction

River water is often not only used directly for consumption (drinking water, irrigation), but also indirectly for the discharge of agricultural, biological and industrial waste products. The discharge of these products in a river can lead to pollution, which, in turn, can cause environmental damage. River pollution provides a classic example of a negative externality: when an upstream agent (e.g., country, state, city or firm) pollutes a river, this can create external costs for the agents downstream of it. Conversely, downstream agents cannot inflict external costs on upstream agents because water in a river, and therefore pollution, is not able to flow upstream. Asymmetric dependence on a water resource, like this, can cause disputes about the use of the resource, especially if property rights over it are not clearly defined. Since upstream agents obtain all the benefits, but only bear part of the social costs, while polluting a river, a situation of over-pollution relative to the social optimum is likely to arise in (international) rivers.

The well-known theorem of Coase (1960) states that when trade in an externality (pollution caused by an upstream agent to a downstream agent) is possible and there are no transaction costs, bargaining leads to an efficient outcome, regardless of the initial allocation of property rights. Because countries are able to bargain over agreements that would reduce pollution in an international river, in practice, we expect to observe similar levels of pollution in intranational and international rivers. Sigman (2002), however, finds that at water quality monitoring stations immediately upstream of international borders the pollution levels are more than 40 percent higher than the average levels at control stations. Sigman concludes that, while rivers would seem to provide a good case for international cooperation (because they involve small numbers of countries and relatively well defined benefits and costs), cooperation on river pollution has not evolved between countries sharing rivers.[1] The reason for this lack of cooperation in international river pollution problems is the absence of clearly defined property rights over the river (water). All countries sharing an international river usually claim property rights over it (at least

[1]Sigman (2002) notes that the countries in the European Union seem to be an exception. See also Barrett (1994) for an example of an agreement between Switzerland, Germany, France and the Netherlands concerning the reduction in salt pollution of the Rhine river by a French potash mine.

that part of the river on their territories) and none are normally willing to reduce their pollution or pay compensation to countries suffering from it.

In this chapter we study how international water law doctrines can be used to solve river pollution problems through cooperation. As mentioned before, a river is considered 'international' if it is shared by two or more sovereign states (Barrett, 1994). International rivers fall into two categories: boundary (or contiguous) rivers and successive rivers. A boundary river flows between the territories of two (or more) states and hence forms the border between the states. A successive river flows from the territory of one state into the territory of another state (Garretson, Hayton and Olmstead (1967)). It is also possible that an international river is (partly) a boundary river and (partly) a successive river. As in the previous chapters, in this chapter we only consider successive international rivers.

We explained in this dissertation that several water resource issues have been modeled using models from (cooperative) game theory. Recently, especially the rival consumption of water from successive international rivers has received attention. Here the main problem is that water consumed by an upstream country can no longer be consumed by a downstream country. It is clear that in water stressed regions this can create tension between countries sharing a river, because the population of a downstream country might (also) depend on the water inflow in the river in an upstream country. Kilgour and Dinar (1995, 2001), Ambec and Sprumont (2002), van den Brink, van der Laan and Vasil'ev (2007), Ambec and Ehlers (2008), Khmelnitskaya (2010), Wang (2011) and Chapters 4 and 6 of this dissertation all use game-theoretic models to investigate the distribution of water among countries sharing an international river. In Ansink and Weikard (2012) and Chapter 3 of this dissertation a closely related axiomatic approach is followed.

The economic literature on the non-rival use of (international) rivers appears to be limited. Apart from the above mentioned paper of Sigman (2002), there exist three empirical papers of Gray and Shadbegian (2004), Sigman (2005) and Lipscomb and Mobarak (2007) that study transboundary river pollution between states and counties in the United States and Brazil. Mäler (1990), Barrett (1994), Fernandez (2002, 2009) and Dinar (2006) all study two-country river pollution problems. Two theoretical papers that model a multi-country setting are that of Ni and Wang (2007) and Gengenbach, Weikard and Ansink (2010).

The model of Gengenbach, Weikard and Ansink (2010) is close to the one we introduce in this chapter in the sense that there is a river with a unidirectional flow of pollution and the agents (countries) along the river are able to choose their own level of pollution abatement (in this chapter agents choose pollution levels instead of pollution abatement levels). Within their model they analyze how voluntary joint action of the agents along the river can increase pollution abatement. The main difference between the paper of Gengenbach, Weikard and Ansink (2010) and the model of this chapter is that their emphasis is on the stability of coalitions of cooperating agents, while we focus on property rights and the distribution of the gain in social welfare that arises when countries along an international river switch from no cooperation on pollution levels to full cooperation. In this respect, it is important to note that some of the solutions to river pollution problems we discuss in this chapter can violate voluntary participation constraints (see the discussion in Section 3.7). We, however, do not consider this a big problem because

(1) voluntary participation constraints will only be violated by some solutions in some instances, and there is always at least one solution that does not violate them and (2) in this chapter we want to focus on the case in which property rights over (international) rivers can be enforced (by an international court of law, (economic) sanctions or military force), so that participation objections can be overruled.

The model of this chapter also differs substantially from the river pollution model of Ni and Wang (2007). In their model pollution levels are not specified. Instead, it is assumed there is a set of agents N along an international river and each agent has exogenously given (environmental) costs caused by the pollution of the agent itself and all agents upstream to it. The problem then is to divide the total costs of pollution among the agents located along the river. For this problem Ni and Wang (2007) provide and characterize two solutions. The Local Responsibility Sharing method holds each agent responsible for the costs on its own territory and therefore requires that each agent pays its own costs. The Upstream Equal Sharing method recognizes that the costs on the territory of an agent are caused by the agent itself and all its upstream agents and thus requires that these costs are divided equally among those agents.

In this chapter we model the pollution problem by assuming that each agent (country) chooses a level of pollution. Several agents are located along a single-stream river from upstream to downstream. Each agent can perform an activity that causes pollution. The higher the level of the activity, the higher the corresponding level of pollution caused by the agent. An agent derives benefits from its activity level, and thus its own level of pollution, but also incurs environmental costs if polluted river water flows through its territory. An agent therefore does not only suffer from its own level of pollution, but also from the pollution levels of all its upstream agents.[2] The agents value pollution of the river water differently in the sense that some agents have higher needs (marginal utility) for the emission of pollutants than others. The heterogeneous valuations of the agents are introduced by endowing each agent with an agent specific benefit and cost function. Together these two functions determine the utility function of the agent. The benefit function of an agent depends only on its own pollution level; its cost function depends on the pollution emissions of the agent itself and of all the agents that are located upstream of it. So, while in the rival consumption river problems of the previous chapters the water consumption of an agent was restricted by the consumption of the agents upstream to it, in this non-rival case of pollution the use of river water by an upstream agent enters the utility functions of all agents downstream to it.

As explained in Chapter 1, in absence of clearly defined property rights in international river situations, typically each country claims to have the right over the river on its own territory, and therefore also the right to choose its own level of pollution. In the model of this chapter, under non-cooperative behavior, each agent chooses a pollution level that maximizes its own utility, given the pollution levels of the others. The resulting non-cooperative Nash equilibrium is usually inefficient, i.e., the sum of all the utilities of the

[2] This is, for instance, the case when river water is used in an industrial process that creates some sort of benefit for the polluter, but at the same time causes environmental damage at the locations of the polluter and all agents downstream to it.

agents along the river (social welfare) can be increased by coordinating the pollution levels among the agents. However, coordination of pollution levels to maximize social welfare will normally result in lower utility for some of the agents along the river; unless the agents are able to reach an agreement on both the optimal pollution levels, as well as a distribution of the total social welfare through monetary transfers. We thus assume that the agents in the model are able to make monetary transfers to each other. Under well-specified property rights, the Coase theorem then implies that the agents are able to reach an agreement and determine appropriate monetary compensations.

Since property rights over international rivers are often not specified, we have to find a way to determine them. We do this, as before, by referring to doctrines from international watercourse law. The doctrines that we consider in this chapter are the principle of Absolute Territorial Sovereignty (ATS), the principle of Unlimited Territorial Integrity (UTI) and the principle of Territorial Integration of all Basin States (TIBS). We find that each of these principles allocates the property rights over the river in a different way, so that each of the principles provides a different answer to the question of what monetary transfers are appropriate, and necessary, to establish cooperation among the agents in international river pollution models. The ATS principle will lead to a solution for river pollution problems that is similar to the downstream incremental solution for river games of Ambec and Sprumont (2002) (see Section 2.2), the UTI principle to a solution that is similar to the downstream solution for river benefit problems (see Section 3.4) and the TIBS principle to a solution that is similar to the weighted hierarchical solution for river games with multiple springs (and externalities) (see Chapter 4).

This chapter is based on van der Laan and Moes (2012) and is organized as follows. In Section 7.2 we introduce a model for (international) river pollution problems and show that the total level of pollution in the model is always lower under cooperation (if agents coordinate their pollution levels) than under individual action. In Section 7.3 we discuss the distribution of cooperative gains in the two-agent case, when the agents switch from their Nash equilibrium to their socially optimal pollution levels. In Section 7.4 we discuss the ATS and UTI values for river pollution problems and give axiomatizations of these solutions. In Section 7.5 we examine solutions for river pollution problems based on the TIBS principle from international watercourse law. Finally, in Section 7.6 we extend the river pollution model to the case of multiple springs and multiple sinks.

7.2 River pollution problems

Consider a successive river flowing through a finite set of agents (countries). The set of agents is denoted by $N \subset \mathbb{N}$.[3] Unless stated otherwise, we assume without loss of generality that $N = \{1, \ldots, n\}$ for some $n \in \mathbb{N}$ and that the agents are labeled from upstream to downstream, i.e., agent 1 is the most upstream agent, followed by agent 2 and so on, until the most downstream agent n. Thus, as before, for two agents $i, j \in N$ it holds that agent i is upstream of agent j (and agent j is downstream of agent i) when $i < j$. For each agent $i \in N$, we write $UP^i = \{1, \ldots, i\}$ as the subset of N containing

[3] In the theorems of this chapter we consider a variable player set.

agent i and all its upstream agents, and $DO_i = \{i, \dots, n\}$ as the subset of N containing i and all its downstream agents.

Each agent $i \in N$ chooses a level $p_i \in \mathbb{R}_+$ of *pollution*.[4] We collect these individual pollution levels in the $|N|$-dimensional *pollution vector* $p \in \mathbb{R}^N_+$. Because the river is transporting the pollution caused by some agent to all its downstream agents, the pollution experienced by agent $i \in N$ depends on the levels of pollution of the agent itself and all its upstream agents. We assume that the pollution experienced by agent i is given by the function $q_i \colon \mathbb{R}^N_+ \to \mathbb{R}_+$ defined by $q_i(p) = \sum_{j=1}^{i} p_j$, i.e., the level of pollution experienced by i is equal to the sum of all pollution levels of the agents in UP^i. This assumption is also made by Gengenbach, Weikard and Ansink (2010) and implies that pollution is not diluted as it flows (further) downstream. Since we consider the situation in which the amount of water flowing in the river is fixed (there are no additional water inflows into the river), it seems to be a reasonable assumption.

We further assume that each agent along the river derives benefit while causing pollution, but also incurs (environmental) costs of experiencing it. The benefit of an agent $i \in N$ only depends on its own pollution level and is given by a function $b_i \colon \mathbb{R}_+ \to \mathbb{R}_+$, yielding benefit $b_i(p_i)$ for every $p_i \geq 0$. The pollution costs of an agent $i \in N$ depend on the total pollution $q_i(p)$ of the agents in UP^i and are given by a function $c_i \colon \mathbb{R}_+ \to \mathbb{R}_+$, yielding costs $c_i(q_i)$ for every $q_i \geq 0$. In the sequel b_i' and b_i'' denote the first and second order derivatives of b_i with respect to p_i, and c_i' and c_i'' denote the first and second order derivatives of c_i with respect to q_i. We make the following assumptions about the benefit and cost functions of the agents.

Assumption 7.2.1
(1) For every $i \in N$: $b_i(0) = 0$ and, for all $p_i > 0$, b_i is twice differentiable with $b_i'(p_i) > 0$ and $b_i''(p_i) < 0$. In addition, $b_i'(p_i) \to \infty$ as $p_i \to 0$ and $b_i'(p_i) \to 0$ as $p_i \to \infty$.
(2) For every $i \in N$: $c_i(0) = 0$ and, for all $q_i > 0$, c_i is twice differentiable with $c_i'(0) > 0$ and $c_i''(q_i) > 0$.

The first assumption states that agents obtain no benefit when there is no pollution and that the marginal benefits of pollution are positive and (strictly) decreasing. Further, the marginal benefits tend to infinity when pollution tends to zero and tend to zero when pollution tends to infinity. The second assumption states that agents incur no costs when there is no pollution and implies that the marginal costs of pollution are positive and (strictly) increasing. Notice that under Assumption 7.2.1, for every $i \in N$ there exists a unique positive real number, say r_i, such that $b_i'(r_i) = c_i'(r_i)$. Let $r \in \mathbb{R}^N_+$ be the vector of these positive real numbers.

Pollution levels $p \in \mathbb{R}^N_+$ result in utilities

$$u_i(p) = b_i(p_i) - c_i(q_i(p)),\ \ i \in N.$$

That is, the utility of agent i is the difference between its *pollution benefit* $b_i(p_i)$ and the *pollution costs* $c_i(q_i(p)) = c_i(\sum_{j=1}^{i} p_j)$. We assume that utility is transferable. This

[4]One could also let the agents choose the level of production in some industrial process that causes river pollution. If it is then assumed that pollution is strictly increasing in the production level and the subsequent assumptions are modified appropriately, this model would lead to similar conclusions.

means that agents are able to transfer utility to each other by making monetary transfers. The monetary transfer to agent $i \in N$ is equal to $t_i \in \mathbb{R}$. When $t_i > 0$ agent i receives a monetary transfer and when $t_i < 0$ agent i pays a monetary transfer. A (monetary) compensation scheme is a vector $t \in \mathbb{R}^N$ that satisfies the restriction

$$\sum_{i=1}^{n} t_i \leq 0, \tag{7.1}$$

i.e., the sum of all monetary transfers is at most equal to zero. A compensation scheme is said to be budget balanced if $\sum_{i=1}^{n} t_i = 0$. Pollution levels p and a compensation scheme t result in payoffs

$$z_i(p,t) = u_i(p) + t_i, \; i \in N.$$

In the rest of this chapter we assume, as before, that the agents in the model are rational utility maximizers and that all benefit and cost functions are common knowledge. The tuple (N, b, c), with $b = \{b_i | i \in N\}$ the collection of benefit functions and $c = \{c_i | i \in N\}$ the collection of cost functions, constitutes a *river pollution problem.* Although the river pollution problem (N, b, c) resembles the river benefit problem (N, e, b), discussed in Section 2.2, observe that there are two important differences: (1) in a river pollution problem there are no water inflows e and (2) in a river pollution problem there are cost functions c that ensure that agents do not pollute an infinite amount. Hence, in a river benefit problem (N, e, b) the water consumption of an agent is restricted by the consumption of the agents upstream to it, while in a river pollution problem (N, b, c) the use of river water by an upstream agent enters the utility functions of all agents downstream to it through the cost functions. The *output* of a river pollution problem (N, b, c) is a pair (p, t) of pollution levels and monetary transfers, yielding payoffs $z_i(p, t)$, $i \in N$. Given a river pollution problem (N, b, c), the aim in this chapter is to make both positive and normative statements about the output (p, t) under the restriction that t satisfies (7.1).

The Nash equilibrium output

We start the analysis of the river pollution problem (N, b, c) by considering the situation in which each agent acts individually. In this situation there (clearly) are no monetary transfers and each agent $i \in N$ maximizes its utility $u_i(p)$ with respect to the variable p_i under its control. So, each agent i chooses its pollution level p_i so as to maximize its own utility, given the pollution levels of the other agents. This behavior results in *Nash equilibrium pollution levels.* The next proposition shows that in the Nash equilibrium each agent $i \in N$ sets the unique (strictly positive) optimal pollution level $\widehat{p}_i$ at which its marginal benefit of pollution is equal to its marginal cost.

Proposition 7.2.2 *For a river pollution problem (N, b, c) that satisfies Assumption 7.2.1, there exists a unique Nash equilibrium pollution vector $\widehat{p} \in \mathbb{R}^N_+$. In $\widehat{p}$ all pollution levels are strictly positive, $\widehat{p}_i > 0$, $i \in N$.*

Proof. When each agent $i \in N$ acts individually, it maximizes its utility $u_i(p) = b_i(p_i) - c_i(q_i(p))$ given the pollution levels p_j, $j < i$, of its upstream agents. We show the uniqueness (and existence) of the Nash equilibrium pollution levels by induction on the labels of the agents.

The utility of the most upstream agent 1 is independent of the pollution levels of all other agents and is given by $u_1(p) = b_1(p_1) - c_1(p_1)$. Maximizing this with respect to $p_1 \geq 0$ gives the first order condition[5]

$$b_1'(p_1) - c_1'(q_1(p))\frac{\partial q_1(p)}{\partial p_1} = b_1'(p_1) - c_1'(p_1) \leq 0 \perp p_1 \geq 0.$$

By Assumption 7.2.1 it follows that there exists a unique solution $\widehat{p}_1 > 0$ (note that $\widehat{p}_1 = r_1$). By the same assumption it holds that $b_1''(p_1) < 0$ and $c_1''(q_1) = c_1''(p_1) > 0$ for every $p_1 > 0$, and thus $\widehat{p}_1$ satisfies the second order condition $b_1''(p_1) - c_1''(p_1) < 0$ for utility maximization.

Proceeding by induction, assume that for some $1 < i \leq n$, $p_j = \widehat{p}_j > 0$ has been uniquely determined for all $j < i$. The utility of agent i is given by $u_i(p) = b_i(p_i) - c_i(q_i(p))$. Maximizing this utility function with respect to $p_i \geq 0$ gives the first order condition

$$b_i'(p_i) - c_i'(q_i(p))\frac{\partial q_i(p)}{\partial p_i} \leq 0 \perp p_i \geq 0.$$

With $q_i(p) = p_i + \sum_{j=1}^{i-1} \widehat{p}_j$ this gives the system

$$\begin{aligned} & b_i'(p_i) - c_i'(q_i) \leq 0 \perp p_i \geq 0, \\ & q_i = p_i + \textstyle\sum_{j=1}^{i-1} \widehat{p}_j. \end{aligned} \tag{7.2}$$

By Assumption 7.2.1.1 b_i' is strictly decreasing in p_i with $b_i'(p_i) \to \infty$ as $p_i \to 0$ and $b_i'(p_i) \to 0$ as $p_i \to \infty$. By Assumption 7.2.1.2 $c_i'(0) > 0$ and c_i' is strictly increasing in q_i (and therefore strictly increasing in p_i). Hence, for the given pollution levels $\widehat{p}_j$, $j < i$, there exists a unique pollution level $\widehat{p}_i > 0$ that satisfies (7.2). Since, by the same assumptions, b_i'' is negative and c_i'' is positive, it follows that $\widehat{p}_i$ also satisfies the second order condition $b_i''(p_i) - c_i''(q_i) < 0$ for utility maximization.

□

Notice that in the Nash equilibrium output all monetary transfers are equal to zero so that the payoffs are given by $z_i(\widehat{p}, \mathbf{0}) = u_i(\widehat{p}) = b_i(\widehat{p}_i) - c_i(\sum_{j=1}^{i} \widehat{p}_j)$, $i \in N$.

Social welfare and Pareto efficiency

In the river pollution problem (N, b, c) the social welfare associated with pollution levels $p \in \mathbb{R}_+^N$ can be measured by the difference between the *total social benefit* $\sum_{i \in N} b_i(p_i)$ and the *total social costs* $\sum_{i \in N} c_i(\sum_{j=1}^{i} p_j)$. The social welfare function $V \colon \mathbb{R}_+^N \to \mathbb{R}$

[5] In this expression, and below, the orthogonality symbol $\perp$ denotes that at least one of the two inequalities must be an equality.

assigns to each vector $p \in \mathbb{R}^N_+$ of pollution levels the *social welfare*

$$V(p) = \sum_{i \in N} b_i(p_i) - \sum_{i \in N} c_i(\sum_{j=1}^{i} p_j).$$

In the next proposition we show that there exist unique and strictly positive pollution levels $\widetilde{p}_i$, $i \in N$, that maximize social welfare.

Proposition 7.2.3 *For a river pollution problem* (N, b, c) *that satisfies Assumption 7.2.1, there exists a unique vector of pollution levels* $\widetilde{p} \in \mathbb{R}^N_+$ *that maximizes social welfare* $V(p)$*. In* $\widetilde{p}$ *all pollution levels are strictly positive.*

Proof. Maximization of $V(p)$ with respect to $p_i \geq 0$, $i \in N$, yields the system of n first order conditions

$$\frac{\partial V(p)}{\partial p_i} = \frac{\partial b_i(p_i)}{\partial p_i} - \sum_{k=i}^{n} \frac{\partial c_k(q_k(p))}{\partial q_k} \frac{\partial q_k(p)}{\partial p_i} \leq 0 \perp p_i \geq 0, \; i \in N.$$

Since $q_j(p) = \sum_{i=1}^{j} p_i$, it holds that $\frac{\partial q_j(p)}{\partial p_i} = 1$ for every $i, j \in N$ with $i \leq j$ and thus the system reduces to

$$\frac{\partial b_i(p_i)}{\partial p_i} - \sum_{k=i}^{n} \frac{\partial c_k(q_k(p))}{\partial q_k} \leq 0 \perp p_i \; \geq 0, \; i \in N. \tag{7.3}$$

First, observe that at a solution to this system $p_i \leq r_i$ for all $i \in N$ because, for every $p_j \geq 0$, $j < i$, it holds that $b_i'(p_i) < c_i'(\sum_{j=1}^{i-1} p_j + p_i)$ if $p_i > r_i$. Second, at a solution it must hold that $p_i > 0$ for all $i \in N$, because $b_i'(p_i) \to \infty$ as $p_i \to 0$ and $c_i'(\sum_{j=1}^{i} p_j)$ is bounded from above by $c_i'(\sum_{j=1}^{i} r_j)$ for all $p_j \in [0, r_j]$, $j \leq i$. So, any solution of the system (7.3) is strictly positive (and bounded from above by the vector r). To maximize the social welfare $V(p)$ we thus have to find a strictly positive solution to the system

$$\frac{\partial b_i(p_i)}{\partial p_i} - \sum_{k=i}^{n} \frac{\partial c_k(q_k(p))}{\partial q_k} = 0, \; i \in N. \tag{7.4}$$

For agent n the system yields

$$\frac{\partial b_n(p_n)}{\partial p_n} - \frac{\partial c_n(q_n(p))}{\partial q_n} = 0. \tag{7.5}$$

For an agent $\ell \in N \setminus \{n\}$ it holds that

$$\frac{\partial b_\ell(p_\ell)}{\partial p_\ell} = \sum_{k=\ell}^{n} \frac{\partial c_k(q_k(p))}{\partial q_k} = \frac{\partial c_\ell(q_\ell(p))}{\partial q_\ell} + \sum_{k=\ell+1}^{n} \frac{\partial c_k(q_k(p))}{\partial q_k}.$$

Using

$$\sum_{k=\ell+1}^{n} \frac{\partial c_k(q_k(p))}{\partial q_k} = \frac{\partial b_{\ell+1}(p_{\ell+1})}{\partial p_{\ell+1}}$$

it follows that

$$\frac{\partial b_{\ell+1}(p_{\ell+1})}{\partial p_{\ell+1}} = \frac{\partial b_\ell(p_\ell)}{\partial p_\ell} - \frac{\partial c_\ell(q_\ell(p))}{\partial q_\ell},\ \ell \in N \setminus \{n\}. \tag{7.6}$$

So, a solution of the welfare maximization problem has to satisfy the system (7.5) and (7.6) of n equations. Now, take some $p_1 > 0$. Since b'_ℓ is strictly decreasing in p_ℓ and c'_ℓ is strictly increasing in q_ℓ for all $\ell \in N$, it follows that for each value of p_1 there exists a unique positive value for p_2 that solves equation (7.6) for $\ell = 1$ (as long as the right hand side of the equation is positive) and that this value of p_2 is increasing in p_1. Continuing in this way, it follows that for each value of $p_1 > 0$ there is a sequence of unique positive values for $p_2, p_3, \ldots, p_n$ that sequentially solves equation (7.6) for $\ell = 1, 2, \ldots, n-1$ (as long as the right hand sides of all equations are positive) and that all these values are increasing in p_1. Hence, there exists a unique value of p_1 such that the value p_n, obtained from sequentially solving the equations (7.6) for $\ell = 1, 2, \ldots, n-1$, solves equation (7.5). It can be concluded that the system (7.5) and (7.6) of n equations has a unique solution $\widetilde{p}_i$, $i \in N$.

It remains to show that $\widetilde{p}$ yields a maximum of the social welfare function $V(p)$. Recall that the components r_i, $i \in N$, of the vector $r \in \mathbb{R}^N_+$ satisfy $b'_i(r_i) = c'_i(r_i)$. Since $\widetilde{p}$ also satisfies the system (7.4), it follows that $\widetilde{p}_i < r_i$, $i \in N$. Since the objective function $V(p)$ is continuous in p, it follows by Weierstrass' (extreme value) theorem that $V(p)$ has a maximum on the compact set

$$\{p \in \mathbb{R}^N \mid 0 \leq p_i \leq r_i,\ i \in N\}.$$

Since $\frac{\partial V(p)}{\partial p_i} > 0$ if $p_i = 0$ and $\frac{\partial V(p)}{\partial p_i} < 0$ if $p_i = r_i$, $i \in N$, it follows that the maximum is achieved in the interior of this set and thus has to satisfy the first order condition (7.4). Hence, the unique solution to this system yields the maximum.

□

The following proposition shows that the total pollution in the outcome that maximizes the social welfare is always lower than the total pollution in the Nash equilibrium output.

Proposition 7.2.4 *For the river pollution problem (N, b, c), $|N| \geq 2$, satisfying Assumption 7.2.1, it holds that $\sum_{i=1}^n \widetilde{p}_i < \sum_{i=1}^n \widehat{p}_i$.*

Proof. We prove this proposition by induction on the number of agents n. We first consider a river model (N, b, c) with $n = 2$. As noticed in the proofs of Proposition 7.2.2 and Proposition 7.2.3, when $n = 2$ it holds both in the Nash equilibrium output and in the output that maximizes $V(p)$ that agent 2 sets its pollution level p_2 so that

$$\frac{\partial b_2(p_2)}{\partial p_2} - \frac{\partial c_2(q_2)}{\partial q_2}\frac{\partial q_2(p)}{\partial p_2} = \frac{\partial b_2(p_2)}{\partial p_2} - \frac{\partial c_2(q_2)}{\partial q_2} = 0.$$

Because $\frac{\partial c_2}{\partial q_2}$ is continuous and strictly increasing it has an inverse $\frac{\partial c_2}{\partial q_2}^{-1}$. It follows that in both cases it must hold that

$$q_2 = p_1 + p_2 = \frac{\partial c_2}{\partial q_2}^{-1}\frac{\partial b_2(p_2)}{\partial p_2}. \tag{7.7}$$

It also follows from the proof of Proposition 7.2.3 that $\widetilde{p}_1$ and $\widetilde{p}_2$ are such that

$$\frac{\partial b_1(\widetilde{p}_1)}{\partial p_1} - \frac{\partial c_1(\widetilde{p}_1)}{\partial q_1} = \frac{\partial b_2(\widetilde{p}_2)}{\partial p_2}.$$

Since $b_2'(p_2) > 0$ at every $p_2 > 0$ and $\widetilde{p}_2 > 0$ (see the proof of Proposition 7.2.3) it must be that

$$\frac{\partial b_1(\widetilde{p}_1)}{\partial p_1} > \frac{\partial c_1(\widetilde{p}_1)}{\partial q_1}.$$

By Assumption 7.2.1 it follows that $\widetilde{p}_1 < \widehat{p}_1$. When agent 1 chooses $p_1 = \widetilde{p}_1 < \widehat{p}_1$ and agent 2 would pollute at (or below) its Nash equilibrium pollution level, $p_2 \leq \widehat{p}_2$, it would hold that

$$u_1((p_1, p_2)) + u_2((p_1, p_2)) < b_1(\widehat{p}_1) - c_1(\widehat{p}_1) + b_2(\widehat{p}_2) - c_2(\widehat{p}_2).$$

This would mean that agent 1 and 2 would obtain a higher social welfare in the Nash equilibrium than in the output that maximizes $V(p)$, a contradiction. It therefore must be that $\widetilde{p}_2 > \widehat{p}_2$. Now, since $\widetilde{p}_2 > \widehat{p}_2$ and b_2' is strictly decreasing in p_2 it follows that $\frac{\partial b_2(\widetilde{p}_2)}{\partial p_2} < \frac{\partial b_2(\widehat{p}_2)}{\partial p_2}$. Because $\frac{\partial c_2}{\partial q_2}^{-1}$ is strictly increasing in its argument it follows from equation (7.7) that

$$\widetilde{p}_1 + \widetilde{p}_2 = \frac{\partial c_2}{\partial q_2}^{-1} \frac{\partial b_2(\widetilde{p}_2)}{\partial p_2} < \frac{\partial c_2}{\partial q_2}^{-1} \frac{\partial b_2(\widehat{p}_2)}{\partial p_2} = \widehat{p}_1 + \widehat{p}_2.$$

Let $K = \{1, ..., k\}$, $k < n$. We now denote the vectors of the unique Nash equilibrium and social welfare maximizing pollution levels for a tuple (K, b, c) with $k = |K|$ agents by $\widehat{p}^k$ and $\widetilde{p}^k$ respectively. Proceeding by induction, assume that

$$\sum_{i=1}^{k} \widetilde{p}_i^k < \sum_{i=1}^{k} \widehat{p}_i^k. \tag{7.8}$$

for every river pollution problem (K, b, c) with $k = |K| < n$. For some (N, b, c) with $|N| = n$, let $(N \setminus \{n\}, b, c)$ denote the model in which the last agent n is deleted. By definition, $\widetilde{p}_i^{n-1}$, $i \in N \setminus \{n\}$, is the solution to the welfare maximization problem

$$\max_{p_1, \ldots, p_{n-1}} \sum_{i=1}^{n-1} b_i(p_i) - \sum_{i=1}^{n-1} c_i\Big(\sum_{j=1}^{i} p_j\Big) \tag{7.9}$$

and $\widetilde{p}_i^n$, $i \in N$, is the solution to the welfare maximization problem

$$\max_{p_1, \ldots, p_n} \sum_{i=1}^{n-1} b_i(p_i) - \sum_{i=1}^{n-1} c_i\Big(\sum_{j=1}^{i} p_j\Big) + \Big[b_n(p_n) - c_n\Big(\sum_{j=1}^{n-1} p_j + p_n\Big)\Big]. \tag{7.10}$$

Since $c_n'(q_n) > 0$ at every $q_n = \sum_{j=1}^{n-1} p_j + p_n$, it follows from comparing problem (7.9) with problem (7.10) that

$$\sum_{i=1}^{n-1} \widetilde{p}_i^n \leq \sum_{i=1}^{n-1} \widetilde{p}_i^{n-1}. \tag{7.11}$$

On the other hand it holds that

$$\sum_{i=1}^{n-1} \widehat{p}_i^{n-1} = \sum_{i=1}^{n-1} \widehat{p}_i^{n}, \tag{7.12}$$

because the unique Nash equilibrium pollution levels of the agents $1, \ldots, n-1$ do not depend on the action (or presence) of agent n. From inequality (7.8) with $k = n-1$, and the (in)equalities (7.11) and (7.12) it follows that

$$\sum_{i=1}^{n-1} \widetilde{p}_i^{n} < \sum_{i=1}^{n-1} \widehat{p}_i^{n}.$$

As noticed in the proofs of Proposition 7.2.2 and Proposition 7.2.3, both in the Nash equilibrium and in the output that maximizes $V(p)$, agent n sets it pollution level p_n so that

$$\frac{\partial b_n(p_n)}{\partial p_n} - \frac{\partial c_n(\sum_{j=1}^{n} p_j)}{\partial q_n} = 0.$$

Because b_n' is strictly decreasing, c_n' is strictly increasing and $\sum_{i=1}^{n-1} \widetilde{p}_i^{n} < \sum_{i=1}^{n-1} \widehat{p}_i^{n}$, it follows that

$$\frac{\partial b_n(p_n)}{\partial p_n} - \frac{\partial c_n(\sum_{i=1}^{n-1} \widetilde{p}_i^{n} + p_n)}{\partial q_n} > 0$$

for any $p_n \leq \widehat{p}_n$. So, it must be that $\widetilde{p}_n^{n} > \widehat{p}_n^{n}$. Further, because c_n' is continuous and strictly increasing it has an inverse $\frac{\partial c_n}{\partial q_n}^{-1}$ that is also strictly increasing in its argument. Analogously as for the case $n = 2$ it now follows that

$$\sum_{i=1}^{n} \widetilde{p}_i^{n} = \frac{\partial c_n}{\partial q_n}^{-1} \frac{\partial b_n(\widetilde{p}_n^{n})}{\partial p_n} < \frac{\partial c_n}{\partial q_n}^{-1} \frac{\partial b_n(\widehat{p}_n^{n})}{\partial p_n} = \sum_{i=1}^{n} \widehat{p}_i^{n}.$$

□

With slight abuse of notation, in the sequel we denote the highest social welfare that can be obtained in the river pollution problem (N, b, c) by $V(N, b, c)$. That is, $V(N, b, c)$ is the social welfare $V(\widetilde{p})$ at the pollution levels $\widetilde{p} \in \mathbb{R}_+^N$ in the river pollution problem (N, b, c). Payoff vector $z(p, t) \in \mathbb{R}^N$ at pollution vector $p \in \mathbb{R}^N$ and compensation scheme $t \in \mathbb{R}^N$ is Pareto efficient if there does not exist another pair (p', t') such that $z_i(p', t') \geq z_i(p, t)$ for all $i \in N$ with at least one strict inequality. Clearly, $z(p, t)$ is Pareto efficient if and only if $p = \widetilde{p}$ and $\sum_{i \in N} t_i = 0$, and thus $\sum_{i \in N} z_i(p, t) = V(\widetilde{p}) = V(N, b, c)$. It therefore follows that any Pareto efficient payoff vector $z \in \mathbb{R}^N$ can be implemented by the vector $\widetilde{p} \in \mathbb{R}_+^N$ of efficient pollution levels and the budget balanced compensation scheme $t_i = z_i - u_i(\widetilde{p})$, $i \in N$. We conclude this section with an example, which also will be used to illustrate the discussion in the subsequent sections.

Example 7.2.5 Let (N, b, c) be a river pollution problem with $N = \{1, 2\}$, $b_i(p_i) = \sqrt{p_i}$ and $c_i(q_i) = q_i^2$, $i = 1, 2$. Then the Nash equilibrium pollution levels are given by $\widehat{p}_1 = 0.3969$ and $\widehat{p}_2 = 0.1847$, yielding utilities $u_1(\widehat{p}) = 0.473$ for the upstream agent 1 and $u_2(\widehat{p}) = 0.092$ for the downstream agent 2. The social welfare in the Nash equilibrium is $V(\widehat{p}) = 0.565$. The Pareto efficient pollution levels are $\widetilde{p}_1 = 0.1621$ and $\widetilde{p}_2 = 0.2968$, yielding utilities $u_1(\widetilde{p}) = 0.376$ and $u_2(\widetilde{p}) = 0.334$. Notice that indeed $\widetilde{p}_1 + \widetilde{p}_2 = 0.4589 < 0.5816 = \widehat{p}_1 + \widehat{p}_2$. The maximal social welfare is equal to $V(\widetilde{p}) = 0.710$.

Observe that $u_1(\widetilde{p}) = 0.376 < 0.473 = u_1(\widehat{p})$, so that without monetary transfers agent 1 prefers the Nash equilibrium output above the Pareto efficient output. When $t_1 = -t_2$ and $0.097 \leq t_1 \leq 0.242$ both agents have at least the same payoff in the Pareto efficient output $(\widetilde{p}, t)$ as at the Nash equilibrium pollution levels $\widehat{p}$ without monetary compensations.

□

7.3 Distribution of the cooperative gains

In the previous section we have seen that the agents in a river pollution problem are able to realize the maximum social welfare $V(N, b, c)$ by choosing the Pareto efficient pollution levels $\widetilde{p}_i$, $i \in N$. In this section we discuss, for the two agent case, what compensation schemes $t = (t_1, t_2)$ would allow the agents to sustain these Pareto efficient pollution levels. In particular, in Example 7.2.5 the Pareto efficient pollution levels $\widetilde{p}_1$ and $\widetilde{p}_2$, together with a monetary compensation scheme $t = (t_1, t_2)$ such that $0.097 \leq t_1 \leq 0.242$ and $t_2 = -t_1$, yield both agents a payoff that is at least equal to its Nash equilibrium payoff. A question that can now be asked is the following: is it reasonable to restrict the value of t_1 between 0.097 and 0.242?

According to Coase (1960) the answer to this question depends on the allocation of property rights. The well-known Coase theorem states that when trade in an externality (pollution caused by the upstream agent to the downstream agent) is possible and there are no transaction costs, bargaining leads to an efficient outcome, regardless of the initial allocation of property rights. It are exactly the property rights that determine how the welfare gain from cooperation is distributed among the agents. For the two-agent river pollution problem the Coase theorem implies that cooperation leads to the Pareto efficient pollution levels $p_i = \widetilde{p}_i$, $i = 1, 2$. The transfers t_1 and t_2 then determine how the maximal social welfare $V(N, b, c)$ is distributed over the two agents.

When the upstream agent 1 has the property rights over the river it can cause as much pollution as it pleases, without taking into account the harmful consequences this might have for the downstream agent 2. It thus can be argued that when agent 1 has the property rights over the river it has a legitimate claim to a payoff that is at least equal to the payoff it obtains in the Nash equilibrium output $z_1(\widehat{p}, \mathbf{0}) = u_1(\widehat{p})$. In this case, agent 1 would only be willing to cooperate with agent 2, and pollute at its Pareto efficient pollution level, if it receives a monetary compensation t_1 that is at least equal to

$$u_1(\widehat{p}) - u_1(\widetilde{p}) = \Big(b_1(\widehat{p}_1) - c_1(\widehat{p}_1)\Big) - \Big(b_1(\widetilde{p}_1) - c_1(\widetilde{p}_1)\Big).$$

On the other hand, when agent 1 has the property rights over the river, agent 2 knows that without cooperation agent 1 would pollute at its Nash equilibrium level. Then the optimal action of agent 2 is also to pollute at its Nash equilibrium level. Hence, agent 2 would not be willing to cooperate with agent 1, and make a monetary transfer, if this would lead to a payoff below its payoff in the Nash equilibrium $z_2(\widehat{p}, \mathbf{0}) = u_2(\widehat{p})$. Thus, the compensation $t_1 = -t_2$ that agent 2 is willing to pay is at most equal to

$$u_2(\widetilde{p}) - u_2(\widehat{p}) = \Big(b_2(\widetilde{p}_2) - c_2(\widetilde{p}_1 + \widetilde{p}_2)\Big) - \Big(b_2(\widehat{p}_2) - c_2(\widehat{p}_1 + \widehat{p}_2)\Big).$$

It can be concluded that when agent 1 has the property rights over the river, the agents are willing to bargain on a transfer t_1 between

$$\Big(b_1(\widehat{p}_1) - c_1(\widehat{p}_1)\Big) - \Big(b_1(\widetilde{p}_1) - c_1(\widetilde{p}_1)\Big)$$

and

$$\Big(b_2(\widetilde{p}_2) - c_2(\widetilde{p}_1 + \widetilde{p}_2)\Big) - \Big(b_2(\widehat{p}_2) - c_2(\widehat{p}_1 + \widehat{p}_2)\Big).$$

In Example 7.2.5 this bargaining interval is $0.097 \leq t_1 \leq 0.242$.

Now consider the case that the downstream agent 2 has the property rights over the river, in the sense that it has the right to claim (and the possibility to enforce) that agent 1 does not cause any pollution (thus $p_1 = 0$). In this case agent 2 can claim a minimal payoff equal to $z_2((0, r_2), \mathbf{0}) = u_2((0, r_2)) = b_2(r_2) - c_2(r_2)$ (recall that r_i, $i \in N$, is the optimal pollution level of agent i when all other pollution levels are zero). Now, agent 2 is only willing to cooperate with agent 1, and set its Pareto efficient pollution level $\widetilde{p}_2$, if agent 1 pays a monetary transfer $t_2 = -t_1$ that is at least equal to

$$u_2((0, r_2)) - u_2(\widetilde{p}) = \Big(b_2(r_2) - c_2(r_2)\Big) - \Big(b_2(\widetilde{p}_2) - c_2(\widetilde{p}_1 + \widetilde{p}_2)\Big).$$

On the other hand, when agent 2 has the property rights over the river, without cooperation agent 1 has a payoff equal to zero, $z_1((0, r_2), \mathbf{0}) = u_1((0, r_2)) = 0$. Agent 1 would therefore not be willing to pay more than $u_1(\widetilde{p}) - u_1((0, r_2)) = b_1(\widetilde{p}_1) - c_1(\widetilde{p}_1)$ to establish cooperation. It can be concluded that when agent 2 has the property rights over the river the agents are willing to bargain on a transfer t_2 between

$$\Big(b_2(r_2) - c_2(r_2)\Big) - \Big(b_2(\widetilde{p}_2) - c_2(\widetilde{p}_1 + \widetilde{p}_2)\Big)$$

and $b_1(\widetilde{p}_1) - c_1(\widetilde{p}_1)$. For Example 7.2.5 it follows straightforwardly that $b_2(r_2) - c_2(r_2) = 0.473$; so the bargaining interval is $0.139 \leq t_2 \leq 0.376$.

When the property rights over the river are unambiguously defined it follows from the above that, at least in the two agent case, a well-defined bargaining problem is obtained. Every solution to such a bargaining problem results in a distribution of the cooperative gains. Yet, the bargaining problem is not so obvious when the property rights over the river are not clearly defined. For instance, what would be the output of the two agent river pollution problem when both agents claim to have the property rights over the river and neither of the agents accepts the claim of the other agent? In this situation each agent $i \in \{1,2\}$ claims a payoff that is at least equal to $b_i(r_i) - c_i(r_i)$. In Example 7.2.5 this would mean that both agents claim at least 0.473. Since the total social welfare (of cooperation) is equal to 0.710, an outcome in which both agents obtain at least their claim is infeasible. This leads to the question how to distribute the deficit that results when each agent claims the property rights over the river. In the following sections we suggest answers to this question by taking into account principles from international watercourse law.

7.4 Two values for the river pollution problem

In this section we propose and characterize two solutions for the welfare distribution problem resulting from the river pollution problem (N, b, c). To do this we use the concept of a value from the theory of cooperative games (see Section 2.1). To apply the notion of a value to polluted rivers, let $\mathcal{RP}^N$ be the class of all river pollution problems (N, b, c) with fixed set of agents N, satisfying Assumption 7.2.1. Further, let $\mathcal{RP} = \cup_{N \subset \mathbb{N}} \mathcal{RP}^N$ be the class of all river pollution problems satisfying Assumption 7.2.1. A value now is a function f that assigns to every $(N, b, c) \in \mathcal{RP}$ a payoff vector $f(N, b, c) \in \mathbb{R}^N$.

Ideally, a value for the river pollution problem would be based directly on international watercourse law. But, since there currently is no binding international law for managing international rivers, the only guidelines that are available are the international watercourse doctrines discussed in Chapter 1. We now first apply the ATS and UTI principles to the class $\mathcal{RP}$ of river pollution problems, which results in two values. The first value, based on the ATS principle, is similar to the downstream incremental solution for river benefit problems of Ambec and Sprumont (2002) (see Section 3.2). The second value, based on the UTI principle, is similar to the downstream solution for river benefit problems (see Section 3.4).

The ATS value

Recall from Chapter 1 that the principle of absolute territorial sovereignty (also known as the Harmon doctrine) states that each country (agent) along an international river has absolute sovereignty over the part of the river on its territory. For river pollution problems the ATS principle favors upstream agents over downstream agents in the sense that it allows an (upstream) agent to choose any pollution level it prefers, without taking into account the consequences for downstream agents. It is not difficult to see that without cooperation, the ATS principle would yield the Nash equilibrium output. It is, however, also possible to apply the ATS principle when the agents along the river do cooperate. As observed in the previous section, the Coase theorem implies that under cooperation all agents pollute at their Pareto efficient pollution level. It are the property rights that determine how the welfare gain from cooperation is distributed among the agents. As in Ambec and Sprumont (2002), we propose that the property rights over an international river are determined by principles from international watercourse law.

When a group of upstream agents UP^i, $i \in N$, decides to cooperate, the ATS principle implies that such a group of agents can pollute as much as it pleases, because it has absolute sovereignty over its territory. So, every upstream set of agents UP^i can claim a total (combined) payoff under full cooperation (of all agents) that is at least equal to the total welfare that it can attain on its own. If it would not receive at least this welfare level, it would be optimal for the group to cease cooperation with the downstream agents. Let p^i_j, $j \in UP^i$, $i \in N$, be a solution to the maximization problem

$$\max_{p_1,\ldots,p_i} \sum_{j=1}^{i} \Big(b_j(p_j) - c_j(\sum_{k=1}^{j} p_k)\Big) \tag{7.13}$$

and denote

$$v^i(N,b,c) = \sum_{j=1}^{i} \Big(b_j(p^i_j) - c_j(\sum_{k=1}^{j} p^i_k)\Big).$$

That is, $v^i(N,b,c)$ is the highest welfare that the set of upstream agents UP^i can obtain without taking into account the consequences of its pollution to the downstream agents. Notice that $v^n(N,b,c) = V(N,b,c)$. The ATS principle thus implies that each group of upstream agents UP^i, $i \in N$, can claim at least a total payoff $v^i(N,b,c)$.

We now define the *ATS value*, denoted by ATS, as the function on the class $\mathcal{RP}$ of river pollution problems that for every $N \subset \mathbb{N}$ and every $(N,b,c) \in \mathcal{RP}^N$ assigns to every agent $j \in N$ the payoff $ATS_j(N,b,c)$ equal to

$$ATS_j(N,b,c) = v^j(N,b,c) - v^{j-1}(N,b,c),$$

with $v^0(N,b,c) = 0$. So, the ATS value distributes to every upstream set of agents UP^i, $i \in N$, a total payoff equal to $\sum_{j=1}^{i} ATS_j(N,b,c) = v^i(N,b,c)$. The ATS value can be implemented by the Pareto efficient pollution levels $\widetilde{p}_i$, $i \in N$, and a budget balanced compensation scheme t such that $t_i = ATS_i(N,b,c) - u_i(\widetilde{p})$, $i \in N$.

If we compare the ATS value for river pollution problems with the downstream incremental solution for river benefit problems (see Section 3.2) we see that both in the ATS value and in the downstream incremental solution every coalition of upstream agents receives precisely the minimum payoff it can claim according to the ATS principle. In the downstream incremental solution for river benefit problems, each coalition of upstream agents UP_i, $i \in N$, receives the maximum welfare that it can obtain by optimally distributing its own water inflows $e_1,\ldots,e_i$ among its agents (not taking into account the agents in DO_{i+1}). In the ATS value for river pollution problems, each UP_i receives the maximum welfare that it can obtain by optimally choosing its own pollution levels $p_1,\ldots,p_i$ (also not taking into account the agents in DO_{i+1}). The ATS value can thus be seen as the downstream incremental solution applied to river pollution problems.

In the sequel, for any river pollution problem $(N,b,c) \in \mathcal{RP}$ and some agent $i \in N$, let $(UP^i, b^{1,i}, c^{1,i})$ denote the river pollution problem restricted to the upstream set of agents UP^i. Hence, $(UP^i, b^{1,i}, c^{1,i})$ is a river pollution problem in $\mathcal{RP}^{UP_i}$ with set of agents UP^i, benefit functions $b^{1,i}_j = b_j$, $j \in UP^i$, and cost functions $c^{1,i}_j = c_j$, $j \in UP^i$. Notice that for every $i \in N$,

$$V(UP^i, b^{1,i}, c^{1,i}) = v^i(N,b,c),$$

i.e., the worth $v^i(N,b,c)$ that the agents in UP^i can guarantee themselves under the ATS principle within the river pollution problem (N,b,c) is equal to the total social welfare that UP^i can attain within the (sub)river problem $(UP^i, b^{1,i}, c^{1,i})$. So, the ATS value satisfies

$$\sum_{j=1}^{i} ATS_j(N,b,c) = V(UP^i, b^{1,i}, c^{1,i}), \text{ for all } i \in N. \tag{7.14}$$

Using this, it follows that the ATS value can be characterized by an *efficiency* and an *upstream autonomy* axiom.

Axiom 7.4.1 Efficiency
A value f on the class of river pollution problems $\mathcal{RP}$ is efficient if it holds for every $(N,b,c) \in \mathcal{RP}$ that $\sum_{i\in N} f_i(N,b,c) = V(N,b,c)$.

In the model of this chapter efficiency follows from the Coase theorem. As stated before, the Coase theorem implies that all agents pollute at their Pareto efficient pollution levels and the property rights determine how the maximum social welfare $V(N,b,c)$ is distributed over the agents.

Axiom 7.4.2 Upstream autonomy
A value f on the class of river pollution problems $\mathcal{RP}$ satisfies upstream autonomy if for every $(N,b,c) \in \mathcal{RP}$ and any $i \in N$ it holds that $f_i(N,b,c) = f_i(UP^i, b^{1,i}, c^{1,i})$.

When all agents downstream of i are not present, upstream autonomy implies that agent i receives the same payoff as it would receive when these agents are present.[6] So, it states that the payoff of an agent does not depend on its downstream agents. We now can state and prove the following simple characterization theorem for the ATS value on the class $\mathcal{RP}$.

Theorem 7.4.3 *A value f on the class of river pollution problems $\mathcal{RP}$ satisfies efficiency and upstream autonomy if and only if f is the ATS value.*

Proof. We first show that the ATS value satisfies the two axioms. Efficiency follows straightforwardly from the definition of ATS, since $\sum_{i\in N} ATS_i(N,b,c) = V(N,b,c)$. Upstream autonomy follows from equation (7.14), because for every $i \in N$,

$$ATS_i(N,b,c) = \sum_{j=1}^{i} ATS_j(N,b,c) - \sum_{j=1}^{i-1} ATS_j(N,b,c) =$$

$$V(UP^i, b^{1,i}, c^{1,i}) - V(UP^{i-1}, b^{1,i-1}, c^{1,i-1}) = ATS_i(UP^i, b^{1,i}, c^{1,i}).$$

Next, take $(N,b,c) \in \mathcal{RP}$ and assume that f satisfies efficiency and upstream autonomy. We prove uniqueness by induction on the labels of the agents, starting with the

[6] The upstream autonomy axiom can be seen as a consistency requirement, discussed in, e.g., Hart and Mas-Colell (1989).

most upstream agent 1. For $i = 1$ it holds by upstream autonomy that $f_1(N, b, c) = f_1(UP^1, b^{1,1}, c^{1,1})$. Thus, agent 1's payoff in the $|N|$-agent river pollution problem (N, b, c) is equal to agent 1's payoff in the 1-agent river pollution problem $(UP^1, b^{1,1}, c^{1,1})$. By efficiency it then holds that $f_1(UP^1, b^{1,1}, c^{1,1}) = V(UP^1, b^{1,1}, c^{1,1})$. So,

$$f_1(N, b, c) = ATS_1(N, b, c).$$

Now, assume by induction that $f_k(N, b, c) = ATS_k(N, b, c)$ for all $k < i \leq n$. Then

$$f_i(N, b, c) = f_i(UP^i, b^{1,i}, c^{1,i}) = V(UP^i, b^{1,i}, c^{1,i}) - \sum_{k=1}^{i-1} f_k(UP^i, b^{1,i}, c^{1,i}),$$

where the first equality follows from upstream autonomy and the second from efficiency. Since, again by upstream autonomy, $f_k(UP^i, b^{1,i}, c^{1,i}) = f_k(N, b, c)$, it follows by the induction hypotheses and equation (7.14) that

$$f_i(N, b, c) = V(UP^i, b^{1,i}, c^{1,i}) - \sum_{k=1}^{i-1} f_k(N, b, c) =$$

$$V(UP^i, b^{1,i}, c^{1,i}) - \sum_{k=1}^{i-1} ATS_k(N, b, c) =$$

$$V(UP^i, b^{1,i}, c^{1,i}) - V(UP^{i-1}, b^{1,i-1}, c^{1,i-1}) = ATS_i(N, b, c).$$

□

For the two agent river pollution problem (N, b, c) with $N = \{1, 2\}$ the ATS value gives the payoffs

$$ATS_1(N, b, c) = V(UP^1, b^{1,1}, c^{1,1}) = b_1(r_1) - c_1(r_1) = u_1(r)$$

and

$$ATS_2(N, b, c) = V(N, b, c) - ATS_1(N, b, c) = V(N, b, c) - u_1(r).$$

So, upstream agent 1 receives a payoff equal to its Nash equilibrium payoff and all gains from cooperation go to the downstream agent 2. The ATS value in this instance corresponds to the outcome discussed in the previous section in which agent 1 has the property rights over the river and agent 2 pays the minimum possible transfer to agent 1 in order for it to be compensated for its loss in utility when switching from the Nash equilibrium to the Pareto efficient pollution level. In Example 7.2.5 the ATS value would mean that agent 2 pays $t_1 = 0.097$ to agent 1.

For the case with more than two agents, the upstream autonomy axiom implies that property rights are assigned subsequently from upstream to downstream along the river. First agent 1 has the right to choose its optimal pollution level, regardless of the other agents. Then agents 1 and 2 cooperate and have the right to choose their joint optimal pollution levels, without considering the other agents, and so on. The ATS value assigns, at each step, the gain of cooperation between the agents in UP^{i-1} and the next agent i to agent i, $i \in \{2, \ldots, n\}$. So, each time an agent i joins its set of upstream agents UP^{i-1} all the gain of cooperation goes to agent i and the upstream agents are just compensated to keep their payoffs equal.

The next theorem states that the ATS value gives each agent $i \in N$ a payoff that is at least equal to the payoff it would receive in the Nash equilibrium output. Each agent in a river pollution problem therefore weakly prefers its payoff according to the ATS value to its payoff in the Nash equilibrium output. This shows that the ATS value satisfies a minimal voluntary participation requirement.

Theorem 7.4.4 *Let $(N, b, c) \in \mathcal{RP}$ be a river pollution problem satisfying Assumption 7.2.1. Then, for any $i \in N$, $ATS_i(N, b, c) \geq z_i(\widehat{p}, \mathbf{0}) = u_i(\widehat{p})$.*

Proof. For agent 1 the theorem is true by definition of the ATS value. Next consider some agent $\ell \geq 2$ and take $i = \ell - 1$. Note that $V(UP^i, b^{1,i}, c^{1,i}) = \sum_{j=1}^{i} \Big(b_j(p_j^i) - c_j(\sum_{k=1}^{j} p_k^i)\Big)$, where p_j^i, $j \in UP^i$, is a solution to the maximization problem (7.13). Let $\bar{p}_\ell$ be the optimal pollution level of agent ℓ, given that all its upstream agents $j \leq \ell - 1$ choose p_j^i. This yields utility $\bar{u}_\ell = b_\ell(\bar{p}_\ell) - c_\ell(\sum_{k=1}^{i} p_k^i + \bar{p}_\ell)$ to agent ℓ. By definition of the ATS value it follows that

$$ATS_\ell(N, b, c) = V(UP^\ell, b^{1,\ell}, c^{1,\ell}) - V(UP^{\ell-1}, b^{1,\ell-1}, c^{1,\ell-1}) \geq \bar{u}_\ell.$$

Further, applying Proposition 7.2.4 to the river pollution problem $(UP^i, b^{1,i}, c^{1,i})$ it follows that $\sum_{j=1}^{i} p_j^i < \sum_{j=1}^{i} \widehat{p}_j$. Hence

$$ATS_\ell(N, b, c) \geq \bar{u}_\ell > b_\ell(\widehat{p}_\ell) - c_\ell(\sum_{k=1}^{\ell} \widehat{p}_k) = u_\ell(\widehat{p}).$$

□

For the single-stream river game discussed in Section 2.2 recall that Herings, van der Laan and Talman (2007) and van den Brink, van der Laan and Vasil'ev (2007) propose the upstream incremental solution as an alternative to the downstream incremental solution. In the upstream incremental solution each coalition of downstream agents DO_i, $i \in \{1, \ldots, n\}$, receives the maximum welfare that it can obtain by optimally distributing its own water inflows $e_i, \ldots, e_n$ among its agents. Note that this maximum welfare does not depend on the behavior of the agents in UP^{i-1}, because in the river game under Assumption 2.2.1 of Chapter 2 the agents in UP^{i-1} are guaranteed to consume $e_1, \ldots, e_{i-1}$. The upstream incremental solution fully distributes the gains of cooperation between a downstream coalition DO_i, $i \in \{2, \ldots, n\}$, and the preceding agent along the river $i - 1$ to the agent $i - 1$.

In a river pollution problem (N, b, c) the maximum welfare that a coalition DO_i, $i \in \{1, \ldots, n\}$, can obtain depends on the pollution levels of the agents in UP^{i-1}. Since these pollution levels do depend on whether the agents in UP^{i-1} are cooperating with each other or not, the welfare that a downstream coalition of agents can obtain without cooperating with its upstream agents is not unambiguously defined (it depends on the behavior of the upstream agents). This implies that, unlike the downstream incremental solution, the upstream incremental solution cannot be directly applied to river pollution problems. In the next subsection, we thus consider the principle of unlimited territorial integrity to define a counterpart of the ATS value. This leads to a value for river pollution problems that is similar to the downstream solution for river benefit problems.

The UTI value

As explained in Chapter 1, the principle of unlimited territorial integrity states that each country (agent) along an international river has the right to demand the natural flow of the river into its territory that is both undiminished in quantity and unchanged in quality by the countries (agents) upstream to it. For river pollution problems the UTI principle favors downstream agents over upstream agents in the sense that an (upstream) agent is only allowed to pollute the river if it has the explicit consent of all agents downstream to it. When the downstream agents in DO_i, $i \in N$, decide to cooperate, the UTI principle implies that such a group of agents can claim a completely clean river. This means that none of the agents upstream of the group DO_i is allowed to cause any pollution. Thus, in a river pollution problem $(N, b, c) \in \mathcal{RP}$ any group of downstream agents DO_i can claim a total (combined) payoff under full cooperation (of all agents) that is at least equal to the total welfare that DO_i can attain under the condition that all upstream agents $j < i$ set pollution level $p_j = 0$. If it would not receive at least this welfare level under full cooperation, it would be optimal for the group of downstream agents to cease cooperation with the upstream agents and invoke the UTI principle. Let s^i_j, $j \in DO_i$, $i \in N$, be a solution to the maximization problem

$$\max_{p_i,\dots,p_n} \sum_{j=i}^{n} \Big(b_j(p_j) - c_j(\sum_{k=i}^{j} p_k)\Big) \tag{7.15}$$

and denote

$$\bar{v}^i(N, b, c) = \sum_{j=i}^{n} \Big(b_j(s^i_j) - c_j(\sum_{k=i}^{j} s^i_k)\Big).$$

That is, $\bar{v}^i(N, b, c)$ is the highest welfare that the downstream group DO_i can obtain under the condition that the pollution levels of all the upstream agents are equal to zero. Notice that $\bar{v}^1(N, b, c) = V(N, b, c)$. The UTI principle implies that each group of downstream agents DO_i, $i \in N$, can claim at least a total payoff $\bar{v}^i(N, b, c)$.

We now define the *UTI value*, denoted by UTI, as the function on the class $\mathcal{RP}$ of river pollution problems that for every $N \subset \mathbb{N}$ and every $(N, b, c) \in \mathcal{RP}^N$ assigns to every agent $j \in N$ payoff $UTI_j(N, b, c)$ equal to

$$UTI_j(N, b, c) = \bar{v}^j(N, b, c) - \bar{v}^{j+1}(N, b, c),$$

with $\bar{v}^{n+1}(N, b, c) = 0$. So, the UTI value distributes to every downstream set of agents DO_i, $i \in N$, a total payoff equal to $\sum_{j=i}^{n} UTI_j(N, b, c) = \bar{v}^i(N, b, c)$. The UTI value can be implemented by the Pareto efficient pollution levels $\widetilde{p}_i$, $i \in N$, and a budget balanced compensations scheme t such that $t_i = UTI_i(N,b,c) - u_i(\widetilde{p})$, $i \in N$.

If we compare the UTI value for river pollution problems with the downstream solution for river benefit problems (see Section 3.4) we see that both in the UTI value and in the downstream solution every coalition of downstream agents receives precisely the minimum payoff it can claim according to the UTI principle. In the downstream solution for river benefit problems, each coalition of downstream agents DO_i, $i \in N$, receives the maximum welfare that it can obtain under the condition that the agents in UP^{i-1} do not consume any water (i.e., each DO_i receives the maximum welfare that it can obtain by optimally distributing all the water inflows $e_1, \ldots, e_n$ among its agents). In the UTI value for river pollution problems, each DO_i receives the maximum welfare that it can obtain under the condition that the pollution levels of all the agents in UP^{i-1} are equal to zero (i.e., each DO_i receives the maximum welfare that it can obtain by optimally choosing its own pollution levels $p_i, \ldots, p_n$, given that $p_j = 0$ for all $j \in UP^{i-1}$). The UTI value can thus be seen as some sort of downstream solution applied to river pollution problems.

In the sequel, for any river pollution problem $(N,b,c) \in \mathcal{RP}$ and some agent $i \in N$, let $(DO_i, b^{i,n}, c^{i,n})$ denote the river pollution problem restricted to the downstream set of agents DO_i. So, $(DO_i, b^{i,n}, c^{i,n})$ is a river problem in $\mathcal{RP}^{DO_i}$ with set of agents DO_i, benefit functions $b_j^{i,n} = b_j$, $j \in DO_i$, and cost functions $c_j^{i,n} = c_j$, $j \in DO_i$.[7] Observe that for every $i \in N$,

$$V(DO_i, b^{i,n}, c^{i,n}) = \bar{v}^i(N,b,c),$$

i.e., the worth $\bar{v}^i(N,b,c)$ that the agents in DO_i can guarantee themselves under the UTI principle within the river pollution problem (N,b,c) is equal to the total social welfare that DO_i can attain within the (sub)river pollution problem $(DO_i, b^{i,n}, c^{i,n})$. Hence, the UTI value satisfies

$$\sum_{j=i}^{n} UTI_j(N,b,c) = V(DO_i, b^{i,n}, c^{i,n}), \text{ for all } i \in N. \tag{7.16}$$

Using this, it follows that the UTI value can be characterized by the efficiency axiom and a *downstream autonomy* axiom.

Axiom 7.4.5 Downstream autonomy
A value f on the class of river pollution problems $\mathcal{RP}$ satisfies downstream autonomy if for every $(N,b,c) \in \mathcal{RP}$ and any $i \in N$ it holds that $f_i(N,b,c) = f_i(DO_i, b^{i,n}, c^{i,n})$.

When all agents upstream of agent $i \in N$ are not present, downstream autonomy implies that agent i receives the same payoff as it would receive when these agents are present.[8] So, downstream autonomy states that the payoff of an agent does not depend on its upstream agents. It is now possible to state and prove the following simple characterization theorem for the UTI value.

[7] In this river pollution problem the agents are numbered from i to n.

[8] Also the downstream autonomy axiom can be seen as a consistency requirement.

Theorem 7.4.6 *A value f on the class of river pollution problems $\mathcal{RP}$ satisfies efficiency and downstream autonomy if and only if f is the UTI value.*

Proof. We first show that the UTI value satisfies the two axioms. Efficiency follows straightforwardly from the definition of UTI, since $\sum_{i \in N} UTI_i(N,b,c) = V(N,b,c)$. Downstream autonomy follows from equation (7.16), because for every $i \in N$,

$$UTI_i(N,b,c) = \sum_{j=i}^{n} UTI_j(N,b,c) - \sum_{j=i+1}^{n} UTI_j(N,b,c) =$$

$$V(DO_i, b^{i,n}, c^{i,n}) - V(DO_{i+1}, b^{i+1,n}, c^{i+1,n}) = UTI_i(DO_i, b^{i,n}, c^{i,n}).$$

Next, take $(N,b,c) \in \mathcal{RP}$ and assume that f satisfies efficiency and downstream autonomy. We prove uniqueness by induction on the labels of the agents, starting with the most downstream agent n. For $i = n$ it holds by downstream autonomy that $f_n(N,b,c) = f_n(DO_n, b^{n,n}, c^{n,n})$. So, the payoff of agent n in the $|N|$-agent river pollution problem (N,b,c) is equal to the payoff of agent n in the 1-agent river pollution problem $(DO_n, b^{n,n}, c^{n,n})$. By efficiency it holds that $f_n(DO_n, b^{n,n}, c^{n,n}) = V(DO_n, b^{n,n}, c^{n,n})$ and thus $f_n(N,b,c) = UTI_n(N,b,c)$. Now, assume by induction that $f_k(N,b,c) = UTI_k(N,b,c)$ for all $k > i \geq 1$. Then

$$f_i(N,b,c) = f_i(DO_i, b^{i,n}, c^{i,n}) = V(DO_i, b^{i,n}, c^{i,n}) - \sum_{k=i+1}^{n} f_k(DO_i, b^{i,n}, c^{i,n}),$$

where the first equality follows from downstream autonomy and the second from efficiency. Since, again by downstream autonomy, $f_k(DO_i, b^{i,n}, c^{i,n}) = f_k(N,b,c)$ it follows by the induction hypothesis and equation (7.16) that

$$f_i(N,b,c) = V(DO_i, b^{i,n}, c^{i,n}) - \sum_{k=i+1}^{n} f_k(N,b,c) = V(DO_i, b^{i,n}, c^{i,n}) - \sum_{k=i+1}^{n} UTI_k(N,b,c)$$

$$= V(DO_i, b^{i,n}, c^{i,n}) - V(DO_{i+1}, b^{i+1,n}, c^{i+1,n}) = UTI_i(N,b,c).$$

□

For the two agent river pollution problem (N,b,c) with $N = \{1,2\}$ the UTI value gives the payoffs

$$UTI_2(N,b,c) = V(DO_2, b^{2,2}, c^{2,2}) = b_2(r_2) - c_2(r_2) = u_2((0, r_2))$$

and

$$UTI_1(N,b,c) = V(N,b,c) - UTI_2(N,b,c) = V(N,b,c) - u_2((0, r_2)).$$

So, downstream agent 2 receives a payoff equal to the minimal payoff it can achieve when it can claim (and enforce) that agent 1 does not cause any pollution. The UTI value in this instance corresponds to the outcome of the previous section in which agent 2 has the property rights over the river and agent 1 pays the minimum possible transfer to agent 2 in order for it to be compensated for its loss in utility when agent 2 gives up its right to a clean river and agrees to cooperate on the Pareto efficient pollution levels. In Example 7.2.5 the UTI value would mean that agent 1 pays $t_2 = 0.139$ to agent 2.

For the case with more than two agents, the downstream autonomy axiom implies that the property rights are assigned subsequently from downstream to upstream along the river. First agent n is given the right to clean water; so, to choose its optimal pollution level under the restriction that all upstream pollution levels are zero. Then the agents n and $n-1$ cooperate and have the right to choose their joint optimal pollution levels under the restriction that all upstream pollution levels are zero, and so on. The UTI value assigns, at each step, the gain in welfare that is created when the agents in DO_{i+1} share their UTI rights with the upstream neighboring agent i, to agent i, $i \in \{1, \ldots, n-1\}$. So, each time an agent i joins its set of downstream agents DO_{i+1} all the gain in total welfare goes to agent i. The downstream agents are just compensated to keep their payoffs equal.

In Theorem 7.4.4 we have seen that the ATS value assigns to each agent a payoff that is at least equal to the utility it would receive in the Nash equilibrium output. This does not hold for the UTI value, as can be seen from Example 7.2.5. The UTI value, however, does satisfy a property that is not satisfied by the ATS value: it guarantees that all agents receive a non-negative payoff. To see that the ATS value does not guarantee non-negative payoffs, consider a two agent river pollution problem and suppose that agent 2 has much higher costs of pollution than agent 1. Then it could be that $V(N, b, c) = [b_1(\widetilde{p}_1) - c_1(\widetilde{p}_1)] + [b_2(\widetilde{p}_2) - c_2(\widetilde{p}_1 + \widetilde{p}_2)]$ is smaller than $V(UP_1, b^{1,1}, c^{1,1}) = b_1(r_1) - c_1(r_1)$ which would mean that $ATS_2 = V(N, b, c) - V(UP_1, b^{1,1}, c^{1,1}) < 0$. Agent 2, however, would still be willing to cooperate with agent 1 because its ATS payoff is at least equal to its Nash equilibrium payoff. The next theorem shows that all UTI payoffs are non-negative.

Theorem 7.4.7 *Let $(N, b, c) \in \mathcal{RP}$ be a river pollution problem satisfying Assumption 7.2.1. Then $UTI_i(N, b, c) \geq 0$ for every $i \in N$.*

Proof. For agent n the theorem is true, because $UTI_n(N, b, c) = b_n(r_n) - c_n(r_n) > 0$. Next, consider some agent $i \leq n-1$. According to the UTI value this agent receives $UTI_i(N, b, c) = V(DO_i, b^{i,n}, c^{i,n}) - V(DO_{i+1}, b^{i+1,n}, c^{i+1,n})$. Since the pollution levels $p_i = 0$ and $p_j = s_j^{i+1}$ for $j > i$ are feasible for the maximization problem (7.15) with respect to agent i, and the levels $p_j = s_j^{i+1}$, $j > i$ are a solution for the maximization problem (7.15) with respect to agent $i+1$, it follows that $V(DO_i, b^{i,n}, c^{i,n}) \geq V(DO_{i+1}, b^{i+1,n}, c^{i+1,n})$, so that $UTI_i(N, b, c) \geq 0$.

□

As already mentioned at the end of the previous section, for the two agent case, the ATS and UTI values are incompatible. In Example 7.2.5 the ATS claim $V(UP_1, b^{1,1}, c^{1,1})$ of agent 1 as well as the UTI claim $V(DO_2, b^{2,2}, c^{2,2})$ of agent 2 is equal to 0.473, while the social welfare is equal to 0.710. This means that it is impossible to satisfy both claims simultaneously. In general, it holds for every $i \in \{1, \ldots, n-1\}$ that the sum of the ATS claim $V(UP_i, b^{1,i}, c^{1,i})$ of the upstream set of agents UP^i and the UTI claim

$V(DO_{i+1}, b^{i+1,n}, c^{i+1,n})$ of its downstream complement DO_{i+1} exceeds the maximal total available welfare $V(N, b, c)$. In the next section we therefore discuss compromise solutions.

7.5 TIBS values for the river pollution problem

In the previous section we have introduced two values for river pollution problems based on the ATS and UTI principles from international watercourse law. Recall from Chapter 1 that there exist three main points of critique against these principles in the legal literature. First, they are considered unfair because they ignore the water needs of other states. Second, they have never been used in treaties and agreements between countries sharing a river. And third, they are considered self-contradictory from a legal point of view. In this section we therefore propose TIBS values for river pollution problems that force both the upstream and the downstream agents along the river to make concessions with respect to their ATS or UTI claims. These TIBS values for river pollution problems are similar to the weighted hierarchical solutions for river games with multiple springs (and externalities) of Chapter 4.

To introduce the TIBS values for river pollution problems, consider a river pollution problem $(N, b, c) \in \mathcal{RP}$ and an agent $j \in N$. Suppose that all agents along the river pollute at their Pareto efficient level $\widetilde{p}_i$, $i \in N$, and that each agent upstream of agent j is given its ATS value payoff, while each agent downstream of agent j is given its UTI value payoff. Since the agents along the river are maximally able to divide the maximum social welfare $V(N, b, c)$ among themselves, if one would like to obtain an efficient payoff vector for the river pollution problem (N, b, c), it must be that agent j receives (pays) the entire surplus (deficit) $V(N, b, c) - \sum_{k \in UP^{j-1}} ATS_k(N, b, c) - \sum_{k \in DO_{j+1}} UTI_k(N, b, c)$. More formally, for all $i, j \in N$ let $t_i^j(N, b, c)$ be defined as

$$t_i^j(N,b,c) = \begin{cases} ATS_i(N,b,c) & \text{if } i < j, \\ V(N,b,c) - \sum_{k \in UP^{i-1}} ATS_k(N,b,c) - \sum_{k \in DO_{i+1}} UTI_k(N,b,c) & \text{if } i = j, \\ UTI_i(N,b,c) & \text{if } i > j. \end{cases}$$

In this way, each agent $j \in N$ induces the value t^j on the class of river pollution problems $\mathcal{RP}$. The value t^j assigns to each $(N, b, c) \in \mathcal{RP}$ the payoff vector $t^j(N, b, c) \in \mathbb{R}^N$. For $j = 1$ it holds that $t^1(N, b, c) = UTI(N, b, c)$ and for $j = n$ that $t^n(N, b, c) = ATS(N, b, c)$.

Notice that the above formula for $t^j(N, b, c)$, $j \in N$, is similar to formula (4.4) of Section 4.3, given a river system $(N, \mathcal{U})$ with a single spring.[9] There is, however, one important difference. Given a river system with a single spring, in formula (4.4) an agent $i > j$ receives a payoff $v(DO_i) - v(DO_i \setminus \{i\})$, where v gives the maximum welfare that the agents in DO_k, $k \in N$, can obtain under the condition that the agents in UP^{k-1} do not take into account the agents in DO_k (the agents in UP^{k-1} consume $e_1, \ldots, e_{k-1}$). In contrast, in the formula for $t^j(N, b, c)$ of this section an agent $i > j$ receives a payoff $\bar{v}^i(N, b, c) - \bar{v}^{i+1}(N, b, c)$, where $\bar{v}^k(N, b, c)$ gives the maximum welfare that the agents in DO_k, $k \in N$, can obtain under the condition that the agents in UP^{k-1} are not allowed

[9] The value $t^j(N, b, c)$ resembles the hierarchical outcome of Demange (2004), in which agent j is the top agent in the hierarchy.

to pollute (they are forced to take into account the agents in DO_k). Thus, for a river system $(N,\mathcal{U})$ with a single spring, formula (4.4) of Section 4.3 gives each agent $i > j$ its upstream incremental solution, whereas the formula for $t^j(N,b,c)$ gives each agent $i > j$ a 'downstream-type solution'.

It is not difficult to see that for river pollution problems the value t^j, $j \in N$, can result in a (large) negative payoff $t^j_j(N,b,c) = V(N,b,c) - \sum_{k\in UP^{j-1}} ATS_k(N,b,c) - \sum_{k\in DO_{j+1}} UTI_k(N,b,c)$ for agent j. Similar to what we did in Chapter 4, in the following we are going to consider weighted averages of the values t^j, $j \in N$. Let $\alpha \in \mathbb{R}^N_+$ with $\sum_{i\in N} \alpha_i = 1$. As before, we call α_j the weight of agent $j \in N$. Given the vector of weights α, we define the *TIBS*$^\alpha$ *value* as the function $TIBS^\alpha$ on the class of river pollution problems $\mathcal{RP}$ that for every $N \subset \mathbb{N}$ and every $(N,b,c) \in \mathcal{RP}^N$ assigns to every agent $i \in N$ the payoff equal to

$$TIBS^\alpha_i(N,b,c) = \sum_{j\in N} \alpha_j t^j_i(N,b,c).$$

We call the α-weighted average of the values t^j, $j \in N$, the TIBS$^\alpha$ value because, as in Chapter 4, it can be seen as reflecting the TIBS principle from international watercourse law.

In Chapter 4 we characterized the weighted hierarchical solutions for river games with multiple springs (and externalities) by using an efficiency and an α-TIBS fairness axiom. For the class of river pollution problems we can state the α-TIBS fairness axiom as follows.

Axiom 7.5.1 **α-TIBS fairness**
Given $\alpha \in \mathbb{R}^N_+$ with $\sum_{i\in N} \alpha_i = 1$, a value f on the class of river pollution problems $\mathcal{RP}$ satisfies α-TIBS fairness if for every $(N,b,c) \in \mathcal{RP}$ and any $i \in N \setminus \{n\}$ it holds that

$$\sum_{j\in DO_{i+1}} \alpha_j \Big[\sum_{j\in UP^i} \Big(f_j(N,b,c) - f_j(UP^i, b^{1,i}, c^{1,i}) \Big) \Big] =$$

$$\sum_{j\in UP^i} \alpha_j \Big[\sum_{j\in DO_{i+1}} \Big(f_j(N,b,c) - f_j(DO_{i+1}, b^{i+1,n}, c^{i+1,n}) \Big) \Big]. \tag{7.17}$$

Apart from the fact that Axiom 4.3.2 of Chapter 4 was defined for rivers with multiple springs, and here the α-TIBS fairness axiom is defined for single-stream rivers, note that the worths of the upstream coalition UP^i and its complement N_i in Axiom 4.3.2 are replaced by the sums of the payoffs of the agents in UP^i and DO_{i+1} respectively, in the sub-river pollution problems $(UP^i, b^{1,i}, c^{1,i})$ and $(DO_{i+1}, b^{i+1,n}, c^{i+1,n})$ respectively. We are able to do this because in this chapter we are working with a variable player set, whereas in Chapter 4 we were working within a river system $(N,\mathcal{U})$ for a fixed player set N.

The efficiency and α-TIBS fairness axioms of this chapter characterize the TIBS$^\alpha$ value for river pollution problems. The proof of this theorem is similar to the proof of Theorem 4.3.6 and its preceding lemmas.

Theorem 7.5.2 *Given $\alpha \in \mathbb{R}^N_+$ with $\sum_{i\in N} \alpha_i = 1$, a value f on the class of river pollution problems $\mathcal{RP}$ satisfies efficiency and α-TIBS fairness if and only if f is the $TIBS^\alpha$ value.*

Proof. We first show that the TIBS$^\alpha$ value satisfies efficiency and α-TIBS fairness. Efficiency follows straightforwardly from the definition of $TIBS^\alpha$ since

$$\sum_{i\in N} TIBS_i^\alpha(N,b,c) = \sum_{i\in N}\sum_{j\in N}\alpha_j t_i^j(N,b,c) =$$

$$\sum_{j\in N}\alpha_j \sum_{i\in N} t_i^j(N,b,c) = \sum_{j\in N}\alpha_j\ V(N,b,c) = V(N,b,c).$$

To show α-TIBS fairness, consider an agent $i \in N\setminus\{n\}$. Then

$$\sum_{j\in UP^i} TIBS_j^\alpha(N,b,c) = \sum_{j\in UP^i}\sum_{k\in N}\alpha_k t_j^k(N,b,c) =$$

$$\sum_{j\in UP^i}\Big(\sum_{k\in UP^i}\alpha_k t_j^k(N,b,c) + \sum_{k\in DO_{i+1}}\alpha_k t_j^k(N,b,c)\Big) =$$

$$\sum_{k\in UP^i}\alpha_k\sum_{j\in UP^i} t_j^k(N,b,c) + \sum_{k\in DO_{i+1}}\alpha_k\sum_{j\in UP^i} t_j^k(N,b,c) =$$

$$\sum_{k\in UP^i}\alpha_k\Big(V(N,b,c) - \sum_{j\in DO_{i+1}} t_j^k(N,b,c)\Big) + \sum_{k\in DO_{i+1}}\alpha_k\sum_{j\in UP^i} t_j^k(N,b,c) =$$

$$\sum_{k\in UP^i}\alpha_k\Big(V(N,b,c) - \sum_{j\in DO_{i+1}} UTI_j(N,b,c)\Big) + \sum_{k\in DO_{i+1}}\alpha_k\sum_{j\in UP^i} ATS_j(N,b,c), \quad (7.18)$$

where the last two equalities follow from the definition of the payoff vectors $t^j(N,b,c)$, $j\in N$. Substituting equations (7.14) and (7.16) into equation (7.18) yields

$$\sum_{j\in UP^i} TIBS_j^\alpha(N,b,c) =$$

$$\sum_{k\in UP^i}\alpha_k\Big(V(N,b,c) - V(DO_{i+1},b^{i+1,n},c^{i+1,n})\Big) + \sum_{k\in DO_{i+1}}\alpha_k V(UP^i,b^{1,i},c^{1,i}). \quad (7.19)$$

By efficiency of $TIBS^\alpha$ in the (sub)river pollution problem $(UP^i,b^{1,i},c^{1,i})$ it holds that

$$\sum_{j\in UP^i} TIBS_j^\alpha(UP^i,b^{1,i},c^{1,i}) = V(UP^i,b^{1,i},c^{1,i}) = \sum_{k\in N}\alpha_k V(UP^i,b^{1,i},c^{1,i}). \quad (7.20)$$

Subtracting equation (7.20) from equation (7.19) results in

$$\sum_{j\in UP^i}\Big(TIBS_j^\alpha(N,b,c) - TIBS_j^\alpha(UP^i,b^{1,i},c^{1,i})\Big) =$$

$$\sum_{k\in UP^i}\alpha_k\Big(V(N,b,c) - V(DO_{i+1},b^{i+1,n},c^{i+1,n})\Big) + \Big(\sum_{k\in DO_{i+1}}\alpha_k - \sum_{k\in N}\alpha_k\Big)V(UP^i,b^{1,i},c^{1,i}) =$$

$$\sum_{k\in UP^i}\alpha_k\Big(V(N,b,c) - V(DO_{i+1},b^{i+1,n},c^{i+1,n}) - V(UP^i,b^{1,i},c^{1,i})\Big). \quad (7.21)$$

Analogously, it follows for the agents in DO_{i+1} that

$$\sum_{j\in DO_{i+1}}\Big(TIBS_j^{\alpha}(N,b,c)-TIBS_j^{\alpha}(DO_{i+1},b^{i+1,n},c^{i+1,n})\Big)=$$

$$\sum_{k\in DO_{i+1}}\alpha_k\Big(V(N,b,c)-V(UP^i,b^{1,i},c^{1,i})-V(DO_{i+1},b^{i+1,n},c^{i+1,n})\Big).\tag{7.22}$$

Multiplying equation (7.21) with $\sum_{k\in DO_{i+1}}\alpha_k$ and equation (7.22) with $\sum_{k\in UP^i}\alpha_k$ shows that the α-TIBS fairness property (7.17) in Axiom 7.5.1 is satisfied.

Next, we prove that there exists a unique value that satisfies efficiency and α-TIBS fairness by induction on the number of agents. Let (K,b,c) be a one-agent river pollution problem with $K=\{k\}$ for some $k\in\mathbb{N}$, i.e., k is the single agent in K. Then by efficiency it holds that $f_k(K,b,c)=V(K,b,c)$, where $V(K,b,c)=b_k(r_k)-c_k(r_k)$ with b_k and c_k the benefit and cost functions of k and r_k the optimal level of pollution.

Now, assume by induction that $f(K,b,c)$ is determined uniquely by efficiency and α-TIBS fairness for every river pollution problem (K,b,c) with number of agents $k=|K|<n$, and let (N,b,c) be a river pollution problem with $n=|N|$ agents. For every $i\in N\backslash\{n\}$, the (sub)river pollution problems $(UP^i,b^{1,i},c^{1,i})$ and $(DO_{i+1},b^{i+1,n},c^{i+1,n})$ have at most $n-1$ agents and so the payoff vectors $f(UP^i,b^{1,i},c^{1,i})\in\mathbb{R}^{UP^i}$ and $f(DO_{i+1},b^{i+1,n},c^{i+1,n})\in\mathbb{R}^{DO_{i+1}}$ have been determined already. Efficiency of f implies on the (sub)river problem $(UP^i,b^{1,i},c^{1,i})$ that

$$\sum_{j\in UP^i} f_j(UP^i,b^{1,i},c^{1,i})=V(UP^i,b^{1,i},c^{1,i})$$

and on the (sub)river problem $(DO_{i+1},b^{i+1,n},c^{i+1,n})$ that

$$\sum_{j\in DO_{i+1}} f_j(DO_{i+1},b^{i+1,n},c^{i+1,n})=V(DO_{i+1},b^{i+1,n},c^{i+1,n}).$$

So, the α-TIBS fairness axiom reduces to

$$\sum_{j\in DO_{i+1}}\alpha_j\Big[\sum_{j\in UP^i} f_j(N,b,c)-V(UP^i,b^{1,i},c^{1,i})\Big]=$$

$$\sum_{j\in UP^i}\alpha_j\Big[\sum_{j\in DO_{i+1}} f_j(N,b,c)-V(DO_{i+1},b^{i+1,n},c^{i+1,n})\Big]\tag{7.23}$$

for all $i\in N\setminus\{n\}$. Since there are $n-1$ equations of type (7.23) and by efficiency it must hold that $\sum_{i\in N} f_i(N,b,c)=V(N,b,c)$, there are n linearly independent equations in n unknowns. Hence, the payoffs $f_i(N,b,c)$, $i\in N$, are uniquely determined.

□

As the class of weighted hierarchical solutions in Chapter 4, the class of TIBS$^{\alpha}$ values encompasses a lot of values. It follows, for instance, directly from the definition of t^j that $TIBS^{\alpha}(N,b,c)=ATS(N,b,c)$ if $\alpha_n=1$ and that $TIBS^{\alpha}(N,b,c)=UTI(N,b,c)$ if

$\alpha_1 = 1$. So, the case that the full weight is given to the most downstream agent along the river reflects the ATS principle. Every upstream coalition in this case receives the payoff that it can obtain when it has the right to pollute the river as it pleases. Conversely, the case that the full weight is given to the most upstream agent along the river reflects the UTI principle. Every downstream coalition then receives the payoff that it can obtain when it has the right to a completely clean river.

Given $i \in N$, when $\alpha_j = 0$ for all $j \leq i$, then $TIBS^{\alpha}(N, b, c)$ is a weighted average of the vectors $t^j(N, b, c)$, $j \geq i + 1$, and every agent in the upstream set UP^i receives its ATS value payoff. Similarly, when $\alpha_j = 0$ for all $j \geq i + 1$, then $TIBS^{\alpha}(N, b, c)$ is a weighted average of the vectors $t^j(N, b, c)$, $j \leq i$, and every agent in the downstream set DO_{i+1} receives its UTI value payoff.

For the weight vector $\alpha^e \in \mathbb{R}^N_+$ with $\alpha^e_1 = \alpha^e_n = \frac{1}{2}$, the α-TIBS fairness property (7.17) reduces for every $i \in N \setminus \{n\}$ to

$$\sum_{j \in UP^i} \Big(f_j(N, b, c) - f_j(UP^i, b^{1,i}, c^{1,i}) \Big) = \sum_{j \in DO_{i+1}} \Big(f_j(N, b, c) - f_j(DO_{i+1}, b^{i+1,n}, c^{i+1,n}) \Big).$$

Observe that for $TIBS^{\alpha^e}$ it holds that $\sum_{j \in UP^i} TIBS^{\alpha^e}_j(UP^i, b^{1,i}, c^{1,i}) = V(UP^i, b^{1,i}, c^{1,i})$ and $\sum_{j \in DO_{i+1}} TIBS^{\alpha^e}_j(DO_{i+1}, b^{i+1,n}, c^{i+1,n}) = V(DO_{i+1}, b^{i+1,n}, c^{i+1,n})$. The α^e-TIBS fairness axiom thus states that, for every agent $i \in N \setminus \{n\}$, the total (combined) loss that the agents in UP^i experience when they are forced to cooperate with the agents in DO_{i+1} should be equal to the total (combined) loss that the agents in DO_{i+1} experience.[10] Together with the efficiency axiom, the α^e-TIBS fairness axiom characterizes the value

$$TIBS^{\alpha^e}_i(N, b, c) = \frac{ATS_i(N, b, c) + UTI_i(N, b, c)}{2},$$

the average of the ATS and UTI values.[11] More generally, every weight vector $\alpha \in \mathbb{R}^N_+$ with $\alpha_1 + \alpha_n = 1$ results in a weighted average of the ATS and UTI values.

Taking weight vector $\alpha^a \in \mathbb{R}^N_+$ with $\alpha^a_i = \frac{1}{n}$ for all $i \in N$, the α-TIBS fairness property (7.17) reduces for every $i \in N \setminus \{n\}$ to

$$\frac{1}{i} \Big[\sum_{j \in UP^i} \Big(f_j(N, b, c) - f_j(UP^i, b^{1,i}, c^{1,i}) \Big) \Big] =$$

$$\frac{1}{n-i} \Big[\sum_{j \in DO_{i+1}} \Big(f_j(N, b, c) - f_j(DO_{i+1}, b^{i+1,n}, c^{i+1,n}) \Big) \Big].$$

Hence, the α^a-TIBS fairness axiom states that, for every agent $i \in N \setminus \{n\}$, the average loss that the agents in UP^i experience when they are forced to cooperate with the agents

[10] The α^e-TIBS fairness axiom resembles the equal loss property for line-graph games of van den Brink, van der Laan and Vasil'ev (2007).

[11] The value $TIBS^{\alpha^e}$ resembles the equal gain splitting solution for sequencing problems of Curiel (1988).

in DO_{i+1} should be equal to the average loss that the agents in DO_{i+1} experience.[12] Together with the efficiency axiom the α^a-TIBS fairness axiom characterizes the value

$$TIBS_i^{\alpha^a}(N,b,c) = \frac{1}{n}\sum_{j\in N} t_i^j(N,b,c),$$

which is the average of all values $t^j(N,b,c)$, $j \in N$. The $TIBS^{\alpha^a}$ is similar to the average tree solution of Herings, van der Laan and Talman (2008), which we applied to river games with multiple springs in Chapter 4 and to river basin games in Chapter 6.

Above, we argued that the payoff vector $t^j(v)$, $j \in N$, of Section 4.3 gives each agent $i > j$ along a single-stream river its upstream incremental solution (and each agent $i < j$ its downstream incremental solution), while the formula for $t^j(N,b,c)$ of this chapter gives each agent $i > j$ a 'downstream-type solution' (and each agent $i < j$ a 'downstream incremental-type solution'). Since the weighted hierarchical solutions of Chapter 4 and TIBS values of this chapter are weighted averages of these payoff vectors, this immediately reveals the main difference between these two solutions. This difference between the weighted hierarchical solutions for river games with multiple springs (and externalities) and the TIBS values for river pollution problems is also reflected in the role of the vector of weights $\alpha \in \mathbb{R}_+^N$ in both chapters.

In Chapter 4, the vector of weights $\alpha \in \mathbb{R}_+^N$ represents the 'reasonable and equitable' shares mentioned in the TIBS principle (see Chapter 1). That is, it determines how the gain in welfare, relative to a situation of no cooperation between an upstream and its (complement) downstream coalition, is divided between the two coalitions. In this chapter, the vector of weights $\alpha \in \mathbb{R}_+^N$ can still be seen as containing information on the 'reasonable and equitable' shares in the TIBS principle. However, now the weights in the vector can be seen as some sort of counterparts of the property rights: they show how the loss in welfare that results from enforced cooperation between an upstream set UP^i, $i \in N$, and its downstream complement DO_{i+1} is distributed between the two groups, relative to the 'ideal' situations for both groups (where the ideal situation for a downstream coalition would be that the upstream agents do not pollute, and the ideal situation for an upstream coalition would be that it has no responsibility towards the downstream agents and can pollute whatever it likes). When $\alpha_1 = 1$ all the loss is suffered by UP^i, when $\alpha_n = 1$ all the loss is suffered by DO_{i+1}, when $\alpha = \alpha^e$ both groups equally share the loss and when $\alpha = \alpha^a$ the average loss of the agents in both groups is equal. It thus can be argued that $\sum_{j\in UP^i} \alpha_j$ and $\sum_{j\in DO_{i+1}} \alpha_j$ reflect the responsibilities of both groups to prevent river pollution. The higher $\sum_{j\in UP^i} \alpha_j$ is, the larger the loss that the group of agents UP^i has to take relative to its total payoff $\sum_{j\in UP^i} ATS_j(N,b,c)$ in its most ideal situation; and the higher $\sum_{j\in DO_{i+1}} \alpha_j$ is, the larger the loss that the group of agents DO_{i+1} has to take relative to its total payoff $\sum_{j\in DO_{i+1}} UTI_j(N,b,c)$ in its most ideal situation.

Although, as in Chapter 4, in this chapter the weights α_i, $i \in N$, are exogenous, it is also possible to envision a model in which they are the subject of negotiation between the

[12] The α^a-TIBS fairness axiom resembles the component fairness axiom for cycle-free graph games of Herings, van der Laan and Talman (2008).

agents. In that case, agents would bargain over weights α_i, $i \in N$, that in combination with efficiency and α-TIBS fairness would lead to a unique solution for river pollution problems.

7.6 Rivers with multiple springs and multiple sinks

In this section we generalize the river pollution problem (N, b, c) to a river pollution problem with multiple springs and multiple sinks, i.e., the river is allowed to have multiple tributaries and/or multiple distributaries. Given the findings of the previous chapters, this is a relatively straightforward exercise. Recall, however, that in Section 6.2, to incorporate multiple sinks, we had to make the assumption that an agent located at a split of the river had full control over the river flow at its location. In this chapter it is not necessary to make such an assumption. When an upstream agent pollutes the river, all the agents downstream of it are affected, regardless of whether the river splits into several branches or not.

Similar as in the river basin benefit problems of Chapter 6, in this section we describe a river system with multiple springs and sinks by a directed graph (N, D). So, a directed link (i, j), $i, j \in N$, $i \neq j$, is in the set D if and only if j is a downstream neighbor (successor) of i along the river (and thus i is an upstream neighbor (predecessor) of j). For $i \in N$, we let $NE^i \subset N$ be the set of all neighbors (upstream and downstream) of agent i and, as before, let UP^i be the set of all agents upstream of, and including, agent i. Different than in the river basin benefit problems of Chapter 6, we here only consider river systems that are represented by connected cycle-free directed graphs; a directed graph is cycle-free if the undirected graph $(N, \widehat{D})$ induced by (N, D) is cycle-free. A connected cycle-free directed graph (N, D) gives the most general possible definition of a river, except that it does not allow for anabranches (parts of a river where it splits into two or more separate streams that merge again further downstream). Note that for a river pollution problem (N, b, c) with a single spring, single sink and agents numbered successively from upstream to downstream, $N = \{1, \ldots, n\}$ and $D = \{(i, i+1) | i \in N \setminus \{n\}\}$.

Example 7.6.1 Let (N, D) represent a river system with $N = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and $D = \{(1, 4), (2, 4), (3, 5), (4, 5), (4, 6), (5, 7), (5, 8)\}$, see Figure 7.1. The two streams originating at 1 and 2 merge together at agent 4. There the river immediately splits again into two streams, one to agent 5 and one to agent 6. The stream at agent 5 is joined by a stream originating at agent 3. The resulting stream, in turn, splits into two streams, one to agent 7 and one to agent 8. For $i = 5$, it holds that $NE^5 = \{3, 4, 7, 8\}$ where 3 and 4 are upstream neighbors and 7 and 8 are downstream neighbors. Further, $UP^5 = \{1, 2, 3, 4, 5\}$ is the set of agents upstream of agent 5, including agent 5 itself. Notice that agent 6 is not in UP^5, because along the (undirected) path from 6 to 5, one has to travel upstream when going from 6 to 4.

□

A river pollution problem with multiple springs and multiple sinks is now given by (N, D, b, c) with (N, D) the river system, and, as before, $b = \{b_i | i \in N\}$ the collection of

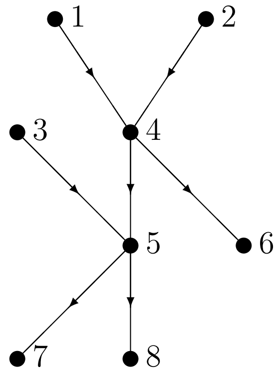

Figure 7.1: River system (N, D) from Example 7.6.1.

benefit functions and $c = \{c_i | i \in N\}$ the collection of cost functions.

Given vector $p \in \mathbb{R}^N_+$ of pollution levels, the total pollution experienced by an agent $i \in N$ is given by $q_i(p) = \sum_{j \in UP^i} p_j$; this is the total pollution of agent i itself and all its upstream agents. Notice that the pollution caused by some agent i affects all its downstream agents. Thus, pollution of, for instance, agent 4 in Example 7.6.1 affects agent 4 itself, but also all its downstream agents 5, 7 and 8.

The output of a river pollution problem (N, D, b, c) is a pair (p, t) of pollution levels and monetary transfers, yielding payoffs

$$z_i(p, t) = u_i(p) + t_i = b_i(p_i) - c_i(\sum_{j \in UP^i} p_j) + t_i, \ i \in N.$$

Let $\widetilde{p}$ be a solution of the welfare maximization problem

$$\max_{p \in \mathbb{R}^N_+} \sum_{i \in N} \Big(b_i(p_i) - c_i(\sum_{j \in UP^i} p_j) \Big) \tag{7.24}$$

and write $V(N, D, b, c) = \sum_{i \in N} u_i(\widetilde{p})$ for the highest social welfare that can be obtained in the river pollution problem (N, D, b, c). The class of all river pollution problems with multiple springs and multiple sinks is denoted by $\mathcal{RPM}$ and a value on this class is a function f that assigns to every $(N, D, b, c) \in \mathcal{RPM}$ a payoff vector $f(N, D, b, c) \in \mathbb{R}^N$.

We now generalize the efficiency axiom, α-TIBS fairness axiom and TIBS$^\alpha$ value to the class $\mathcal{RPM}$.

Axiom 7.6.2 Efficiency on $\mathcal{RPM}$
A value f on the class of river pollution problems $\mathcal{RPM}$ is efficient if it holds for every $(N, D, b, c) \in \mathcal{RPM}$ that $\sum_{i \in N} f_i(N, D, b, c) = V(N, D, b, c)$.

To state the α-TIBS fairness axiom on $\mathcal{RPM}$ consider a connected, cycle-free river system (N, D) and suppose that a directed link $(i, j) \in D$ is deleted from (N, D). That is, there no longer is a water flow from agent i to its downstream neighbor, agent j. This implies that there are two separate connected cycle-free directed graphs that, individually, again represent (part of) a river. For instance, deleting $(4, 5)$ from D in Example 7.6.1 gives two separate river systems, namely $(\{1, 2, 4, 6\}, \{(1, 4), (2, 4), (4, 6)\})$ and

$(\{3,5,7,8\},\{(3,5),(5,7),(5,8)\})$. Now, let (N,D) be a river system and suppose that either $(i,j)\in D$ or $(j,i)\in D$. Then we denote by $(N^{i|j},D^{i|j})$ and $(N^{j|i},D^{j|i})$ the two subriver systems that result when deleting the link (i,j) or (j,i) from D, where $(N^{i|j},D^{i|j})$ represents the river system that contains agent i and $(N^{j|i},D^{j|i})$ represents the river system that contains agent j. Write $b^{N^{i|j}}$ for the set of benefit functions $b_k^{N^{i|j}}=b_k$, $k\in N^{i|j}$, and $c^{N^{i|j}}$ for the set of cost functions $c_k^{N^{i|j}}=c_k$, $k\in N^{i|j}$. Analogously, let $b^{N^{j|i}}$ and $c^{N^{j|i}}$ be the sets of benefit and cost functions for the agents $k\in N^{j|i}$. Now it is possible to state the α-TIBS fairness axiom on $\mathcal{RPM}$.

Axiom 7.6.3 α-TIBS fairness on $\mathcal{RPM}$
Given $\alpha\in\mathbb{R}^N_+$ with $\sum_{i\in N}\alpha_i=1$, a value f on the class of river pollution problems $\mathcal{RPM}$ satisfies α-TIBS fairness if for every $(N,D,b,c)\in\mathcal{RPM}$ and any $(i,j)\in D$ it holds that

$$\sum_{k\in N^{j|i}}\alpha_j\Big[\sum_{k\in N^{i|j}}\Big(f_k(N,D,b,c)-f_k(N^{i|j},D^{i|j},b^{N^{i|j}},c^{N^{i|j}})\Big)\Big]=$$

$$\sum_{k\in N^{i|j}}\alpha_j\Big[\sum_{k\in N^{j|i}}\Big(f_k(N,D,b,c)-f_k(N^{j|i},D^{j|i},b^{N^{j|i}},c^{N^{j|i}})\Big)\Big].$$

Similar as on the class $\mathcal{RP}$ of river pollution problems with a single spring and a single sink, the two axioms characterize a unique value on the class $\mathcal{RPM}$. This value is a generalization of the $TIBS^\alpha$ value to the class $\mathcal{RPM}$. To state the value, let $i,j\in N$ be two different agents and let ℓ^i_j be the second agent on the unique path in $(N,\widehat{D})$ from i to j.[13] For instance, in Example 7.6.1 if $i=4$ and $j=3$ then $\ell^i_j=5$. As in the previous section, we define for each $j\in N$ a payoff vector $t^j(N,D,b,c)$. First, the payoff of agent $j\in N$ itself is defined as

$$t^j_j(N,D,b,c)=V(N,D,b,c)-\sum_{k\in NE^j}V(N^{k|j},D^{k|j},b^{N^{k|j}},c^{N^{k|j}}).$$

For each $i\in N\setminus\{j\}$ the payoff of agent i is defined as

$$t^j_i(N,D,b,c)=V(N^{i|\ell^i_j},D^{N^{i|\ell^i_j}},b^{N^{i|\ell^i_j}},c^{N^{i|\ell^i_j}})-\sum_{k\in NE^i\setminus\{\ell^i_j\}}V(N^{k|i},D^{k|i},b^{N^{k|i}},c^{N^{k|i}}).$$

Observe that, given an agent $j\in N$, for each agent $k\in NE^j$ it holds that

$$\sum_{\ell\in N^{k|j}}t^j_\ell(N,D,b,c)=V(N^{k|j},D^{k|j},b^{N^{k|j}},c^{N^{k|j}}).$$

Thus, each set of agents $N^{k|j}$ that results from deleting (k,j) from D (when k is upstream of j) or deleting (j,k) from D (when k is downstream of j), receives as a total (combined) payoff the welfare that it can attain on its own, ignoring the other agents. When agent k is an upstream neighbor of agent j, this means that the agents in $N^{k|j}$ realize the same

[13]Notice that $\ell^i_j=j$ if j is a neighbor of i in (N,D).

welfare as the welfare they can realize under the ATS principle (not taking into account the effect of their pollution on the downstream agents). When agent k is a downstream neighbor of j it means that the agents in $N^{k|j}$ realize the same welfare as the welfare they can realize under the UTI principle (demanding zero pollution by the upstream agents). Stated differently, the payoff vector $t^j(N, D, b, c)$ assigns to upstream sets $N^{k|j}$, $k \in NE^j$, the ATS claims and to downstream sets $N^{k|j}$, $k \in NE^j$, the UTI claims.

For a given weight vector $\alpha \in \mathbb{R}^N_+$ with $\sum_{i \in N} \alpha_i = 1$, the $TIBS^\alpha$ value assigns to each river pollution problem in $\mathcal{RPM}$ the weighted average of the payoff vectors $t^j(N, D, b, c)$, $j \in N$. So,

$$TIBS^\alpha(N, D, b, c) = \sum_{j \in N} \alpha_j t^j(N, D, b, c).$$

The next theorem states that efficiency and α-TIBS fairness on $\mathcal{RPM}$ characterize the $TIBS^\alpha$ value on $\mathcal{RPM}$.[14] The proof goes along the same lines as the proof of Theorem 7.5.2 and is therefore omitted.

Theorem 7.6.4 *Given $\alpha \in \mathbb{R}^N_+$ with $\sum_{i \in N} \alpha_i = 1$, a value f on the class $\mathcal{RPM}$ of river pollution problems with multiple springs and multiple sinks satisfies efficiency and α-TIBS fairness on $\mathcal{RPM}$ if and only if f is the $TIBS^\alpha$ value.*

Also within the class $\mathcal{RPM}$ the weights $\alpha \in \mathbb{R}^N_+$ can be seen as some sort of counterparts of the property rights over the river. When a link (i, j) (thus j is downstream of i) is deleted from D, the weights determine how the loss of welfare that results from enforced cooperation between the upstream set $N^{i|j}$ and the downstream set $N^{j|i}$ is distributed between the two groups, relative to the most ideal situations for both groups. The higher $\sum_{k \in N^{i|j}} \alpha_k$, the larger the loss that the group of agents $N^{i|j}$ has to take relative to the total welfare it can attain without considering the downstream set $N^{j|i}$; and the higher $\sum_{k \in N^{j|i}} \alpha_k$, the larger the loss that the group of agents $N^{j|i}$ has to take relative to the total welfare it can attain under the condition of zero pollution by its upstream agents.

[14] Observe that in this characterization efficiency is used, while one might expect to need a stronger axiom similar to 'component efficiency' of Herings, van der Laan and Talman (2008). The reason for this is that the axiomatizations of this chapter are on a variable player set. We compare the payoffs $f_k(N, D, b, c)$ with the payoffs $f_k(N^{i|j}, D^{i|j}, b^{N^{i|j}}, c^{N^{i|j}})$ and $f_k(N^{j|i}, D^{j|i}, b^{N^{j|i}}, c^{N^{j|i}})$. The sums of the latter payoffs follow from the efficiency axiom on the river pollution problems $(N^{i|j}, D^{i|j}, b^{N^{i|j}}, c^{N^{i|j}})$ and $(N^{j|i}, D^{j|i}, b^{N^{j|i}}, c^{N^{j|i}})$ respectively.

Conclusion

In this dissertation we modeled (the outcome of) negotiations between water users sharing a river. More specifically, we focused on the problem of distributing the welfare that results from optimally allocating the water that flows in a river among the agents (e.g., countries, cities, firms) located along the river. We did this by first extending a single-stream river water allocation model, originally introduced by Ambec and Sprumont (2002), to situations in which rivers are allowed to have several tributaries (and distributaries). Later, we also allowed the countries in an international river water allocation model to be composed of different water users. Finally, we analyzed the difference between the rival and non-rival use of river water by introducing a river pollution model.

In pursuing these three main goals we employed, and further developed, the methodology pioneered by Barrett (1994) and Kilgour and Dinar (1995). These authors modeled river water allocation problems by combining principles from international watercourse law with concepts from (cooperative) game theory. Hence, they tried to answer the question of how benefits of cooperation have to be distributed among the agents located along a river by referring to established international water law principles. We argued that the main advantage of this approach is that one does not have to rely on one's own value judgment about the fairness of a particular division of water (benefit) among agents, but can fall back on consensus views from a substantial legal literature.

As discussed in Chapter 1 though, the consensus views of legal experts, and countries sharing watercourses, have changed dramatically over the course of the 20th century. They shifted from focusing on absolute water rights (as exemplified by the ATS and UTI principles discussed in Chapter 1) to focusing more on responsibilities towards other water users and cooperation (as represented by the equitable utilization and TIBS principles also discussed in Chapter 1). In this dissertation we have tried to take this into account in our solutions for welfare distribution problems resulting from river water allocation problems. We moved from solutions based on the ordering of the agents along the river (which are primarily inspired by the ATS and UTI principles, see Chapter 3) to solutions that may result in more equitable divisions of the cooperative gain among the agents located along the river (which are inspired by the TIBS principle, see Chapter 4 and further).

Although the models that we have used to derive our solutions are highly stylized, they do seem to suggest that it is possible to make some of the principles of international watercourse law operational in actual river water disputes. Both in theory and in practice coordination of water extraction (or pollution) policies may lead to mutually beneficial outcomes relative to individual action, when monetary transfers between the agents along a river are possible. This last condition, however, still seems more feasible in theory than in practice. In reality, we hardly observe direct monetary compensations between upstream and downstream water users, while theoretically this may lead to Pareto superior outcomes for both of them. A fundamental policy recommendation of this dissertation to countries (or other agents) sharing an international watercourse would therefore be to be more open to the possibility of monetary transfers. This would allow for more direct 'trade' in river water which may benefit all of them.

We believe that especially our solutions (and corresponding transfers) for river sharing problems that are inspired by the TIBS principle could find widespread support among both upstream and downstream water users along a river. We have made the TIBS principle 'operational' for both river water allocation as well as river pollution problems in the sense that it provides a water allocation, or pollution schedule, and a monetary compensation scheme for the agents along a river. We did this by introducing efficiency and α-TIBS fairness axioms. These axioms provide a precise formulation of the TIBS principle. If one accepts this formulation, the weights of the agents in the weight vector α still remain to be determined. We discussed that a simple solution would be to take equal weights for all agents, but it is also possible to consider the weights to be exogenously given (for example by existing power structures among agents or by the factors mentioned in Article V of the Helsinki rules and Article 6 of the UN convention, see Chapter 1). When one does not want to impose weights directly, it is even possible to make them the subject of negotiation between agents.

In this dissertation we argued that international river disputes mostly arise because property rights over international watercourses are not (clearly) defined. While principles from international watercourse law can be used to define them, these principles themselves are often conflicting or also not clearly defined. We demonstrated that the axiomatic approach from cooperative game theory is particularly suitable to make principles from international watercourse law precise so that property rights over international watercourses can be properly established. We translated several principles from international watercourse law into axioms for cooperative river games. Different (combinations of) axioms then led to different solutions for these games. In contrast, explicit negotiations between water users sharing a river can also be modeled using tools from non-cooperative game theory. Since cooperative game-theoretic arguments have so far received the most attention in the river water allocation literature, more (theoretical) research is needed to gain insight into the strategic, non-cooperative, side of the river water allocation problem. The theories of non-cooperative bargaining and the implementation of cooperative solution concepts seem to be especially suited for this and thus provide an interesting area for future research.

We briefly discussed the non-cooperative implementation of the weighted hierarchical solutions (a cooperative solution concept) in Chapter 5. This implementation is, however, mostly of theoretical interest because some essential features of actual river water allocation problems are abstracted away in the underlying model. Two of the main difficulties that one would face when one would want to implement a bargaining framework based on the implementation results of this dissertation, but also when one would want to directly impose the underlying cooperative solution concept, are the dynamic and stochastic nature of actual river water allocation problems. In reality, river water, or river welfare, distribution problems are not one shot problems but occur continuously through time. This continuous aspect could be modeled by looking at repeated water flows in discrete time periods, for instance years, but then one might run into trouble with the stochastic nature of river sharing problems. Suppose that a group of water users along a river divides a certain amount of water (or welfare) in a particular year and bases a water sharing agreement for the following years on this division. If in the next year a drought occurs, and the water inflow into the river is much smaller than in the previous year, it is clear that the stability of the water sharing agreement will be tested because some (coalitions of) water users cannot be given the amount of water (or welfare) agreed upon in the previous year. The stochastic nature of actual river sharing problems thus seems crucial for the stability of water sharing agreements and needs to be considered when one would want to give policy advice on the implementation of theoretical solutions to river sharing problems. In the spirit of von Neumann and Morgenstern (1944), in this dissertation we followed an approach of 'divide the difficulties' by concentrating on generalizations of the static, deterministic, river sharing model of Ambec and Sprumont (2002). The results in this dissertation therefore should not be seen as the complete story on the modeling of river sharing problems, but rather as a first step that can (and should) be extended in several directions. Future research on river sharing problems might focus on dynamic (repeated) games, stochastic (cooperative) games or a combination of the two. Also the theory of differential games and optimal control problems seems applicable to river sharing problems.

In the previous paragraph we mentioned that water sharing agreements might not be stable in case of a drought. Throughout this dissertation when discussing water allocation problems we have focused on shortages of water in a river. In reality we, of course, also observe situations in which a river cannot process a certain amount of water quickly enough and a flood occurs. Several countries along international watercourses have constructed dams on their territories to generate hydroelectric power and control the flow of river water to prevent flooding of the land. This type of use of a river is, in principle, non-rival because river water is not consumed directly and only stored for a specific time period to be released later in generating power. The construction of dams for generating hydroelectric power combined with flood control has not yet been studied from a game theoretic point of view and might be another interesting area for future research.

What should become clear from this short conclusion is that there are plenty of issues related to river sharing problems that still have to be investigated theoretically. We therefore hope that what we have done in this dissertation will become part of a much larger literature studying (cooperative) decision making in river sharing problems.

Bibliography

Albizuri, M., Arin, J. and J. Rubio (2005), An axiom system for a value for games in partition function form, *International Game Theory Review* 72, 7-63.

Alonso-Meijide, J.M., M. Álvarez-Mozos and M.G. Fiestras-Janeiro (2009), Values of games with graph restricted communication and a priori unions, *Mathematical Social Sciences* 58, 202-213.

Ambec, S. and L. Ehlers (2008), Sharing a river among satiable agents, *Games and Economic Behavior* 64, 35-50.

Ambec, S. and Y. Sprumont (2000), Sharing a river, cahier 2000-08, Département de sciences économiques, Université de Montréal, Montréal.

Ambec, S. and Y. Sprumont (2002), Sharing a river, *Journal of Economic Theory* 107, 453-462.

Ansink, E. (2009), *Game-Theoretic Models of Water Allocation in Transboundary River Basins*, Ph.D. Thesis, Wageningen University, Wageningen.

Ansink, E. and H.P. Weikard (2012), Sequential sharing rules for river sharing problems, *Social Choice and Welfare* 38, 187-210.

Aumann, R.J. and J.H. Drèze (1974), Cooperative games with coalition structures, *International Journal of Game Theory* 3, 217-237.

Aumann, R.J. and M. Maschler (1985), Game theoretic analysis of a bankruptcy problem from the Talmud, *Journal of Economic Theory* 36, 195-213.

Banzhaf, J.F. (1965), Weighted voting doesn't work: a mathematical analysis, *Rutgers Law Review* 19, 317-343.

Barrett, S. (1994), Conflict and cooperation in managing international water resources, Policy Research Working Paper 1303, World Bank, Washington D.C.

Barrett, S. (2003), *Environment and Statecraft*, Oxford University Press, Oxford.

Béal, S., E. Rémila and P. Solal (2009), Weighted component fairness for forest games, Munich Personal RePEc Archive Paper No. 17455.

Béal, S., E. Rémila and P. Solal (2010), Rooted-tree solutions for tree games, *European Journal of Operational Research* 203, 404-408.

Boisson de Chazournes, L. and S.M.A. Salman (2005), *Water Resources and International Law*, Martinus Nijhoff Publishers, The Hague.

Bondareva, O. (1963), Some applications of linear programming methods to the theory of cooperative games, *Problemy Kybernetiki* 10, 119-139 (in Russian).

Bourne, C.B. (1996), The International Law Association's contribution to international water resources law, *Natural Resources Journal* 36, 154-216.

Brink, J.R. van den, A. Estévez-Fernández, G. van der Laan and N. Moes (2011), Independence axioms for water allocation, Tinbergen Institute Discussion Paper 2011-128/1, Tinbergen Institute, Amsterdam.

Brink, J.R. van den, and Y. Funaki (2010), Axiomatization and implementation of discounted Shapley values, Tinbergen Institute Discussion Paper 2010-065/1, Tinbergen Institute, Amsterdam.

Brink, J.R. van den, Y. Funaki, and Y. Ju (2011), Reconciling marginalism with egalitarianism: consistency, monotonicity and implementation of egalitarian Shapley values, *Social Choice and Welfare*, doi:10.1007/s00355-011-0634-2.

Brink, J.R. van den, G. van der Laan and N. Moes (2011), Two values for transferable utility games with coalition and graph structure, Tinbergen Institute Discussion Paper 2011-164/1, Tinbergen Institute, Amsterdam.

Brink, J.R. van den, G. van der Laan and N. Moes (2012a), Fair agreements for sharing international rivers with multiple springs and externalities, *Journal of Environmental Economics and Management* 63, 388-403.

Brink, J.R. van den, G. van der Laan and N. Moes (2012b), A strategic implementation of the average tree solution for cycle-free graph games, Tinbergen Institute Discussion Paper 2012-050/1, Tinbergen Institute, Amsterdam.

Brink, J.R. van den, G. van der Laan and V. Vasil'ev (2007), Component efficient solutions in line-graph games with applications, *Economic Theory* 33, 349-364.

Burness, H.S. and J.P. Quirk (1979), Water rights and the efficient allocation of resources, *American Economic Review* 69, 25-37.

Caponera, D.A. (1985), Patterns of cooperation in international water law: principles and institutions, *Natural Resources Journal* 25, 563-588.

Clippel, G. de, and R. Serrano (2008), Marginal contributions and externalities in the value, *Econometrica* 76, 1413-1436.

Coase, R.H. (1960), The problem of social cost, *Journal of Law and Economics* 3, 1-44.

Curiel, I. (1988), *Cooperative Game Theory and Applications*, Ph.D. Thesis, Katholieke Universiteit Nijmegen, Nijmegen.

Demange, G. (2004), On group stability in hierarchies and networks, *Journal of Political Economy* 112, 754-778.

Dinar, S. (2006), Assessing side-payment and cost-sharing patterns in international water agreements: the geographic and economic connection, *Political Geography* 25, 412-437.

Dutta, B., L. Ehlers and A. Kar (2010), Externalities, potential, value and consistency, *Journal of Economic Theory* 145, 2380-2411.

Fernandez, L. (2002), Solving water pollution problems along the US-Mexico Border, *Environment and Development Economics* 7, 715-732.

Fernandez, L. (2009), Wastewater pollution abatement across an international border, *Environment and Development Economics* 14, 67-88.

Funaki, Y. and T. Yamato (1999), The core of an economy with a common pool resource: a partition function form approach, *International Journal of Game Theory* 28, 157-171.

Garretson, A.H., R.D. Hayton and C.J. Olmstead (1967), *The Law of International Drainage Basins*, Oceana Publications, Inc., Dobbs Ferry, New York.

Gengenbach, M.F., H.P. Weikard and E. Ansink (2010), Cleaning a river: an analysis of voluntary joint action, *Natural Resource Modeling* 23, 565-589.

Gillies, D.B. (1959), Solutions to general non-zero-sum games, in *Contributions to the Theory of Games, Vol. IV*, eds. A.W. Tucker and R.D. Luce, Princeton University Press, Princeton, 47-85.

Gomez, A. (2005), Multilateral contracting with externalities, *Econometrica* 73, 1329-1350.

Grabisch, M. (2010), The lattice of embedded subsets, *Discrete Applied Mathematics* 158, 479-488.

Gray, W.B. and R.J. Shadbegian (2004), 'Optimal' pollution abatement - whose benefits matter, and how much?, *Journal of Environmental Economics and Management* 47, 510-534.

Harsanyi, J.C. (1963), A simplified bargaining model for the n-person cooperative game, *International Economic Review* 4, 194-220.

Hart, S. and A. Mas-Colell (1989), Potential, value and consistency, *Econometrica* 57, 589-614.

Herings, P.J.J., G. van der Laan and A.J.J. Talman (2007), The socially stable core in structured transferable utility games, *Games and Economic Behavior* 59, 85-104.

Herings, P.J.J., G. van der Laan and A.J.J. Talman (2008), The average tree solution for cycle-free graph games, *Games and Economic Behavior* 62, 77-92.

Herings, P.J.J, G. van der Laan, A.J.J. Talman and Z. Yang (2010), The average tree solution for cooperative games with communication structure, *Games and Economic Behavior* 68, 626-633.

International Law Association (1966), The Helsinki rules, *Report of the Fifty-Second Conference*, 484-533.

International Law Association (2004), Water resources law, *Report of the Seventy-First Conference*, 334-421.

Jovanovic, D. (1985), Ethiopian interests in the division of the Nile river waters, *Water International* 10, 82-85.

Ju, Y. and D. Wettstein (2009), Implementing cooperative solution concepts: a generalized bidding approach, *Economic Theory* 39, 307-330.

Kalai, E. and D. Samet (1987), On weighted Shapley values, *International Journal of Game Theory* 16, 205-222.

Kamijo, Y. (2011), The collective value: a new solution for games with coalition structures, *TOP*, doi:10.1007/s11750-011-0191-y.

Khmelnitskaya, A.B. (2010), Values for rooted-tree and sink-tree digraph games and sharing a river, *Theory and Decision* 69, 657-669.

Khmelnitskaya, A.B. and A.J.J. Talman (2010), Tree-type values for cycle-free directed graph games, CentER discussion paper 2010-113, Tilburg University, Tilburg.

Kilgour, D.M. and A. Dinar (1995), Are stable agreements for sharing international river waters now possible?, Policy Research Working Paper 1474, World Bank, Washington D.C.

Kilgour, D.M. and A. Dinar (2001), Flexible water sharing within an international river basin, *Environmental and Resource Economics* 18, 43-60.

Laan, G. van der, and N. Moes (2012), Transboundary externalities and property rights: an international river pollution model, Tinbergen Institute Discussion Paper 2012-006/1, Tinbergen Institute, Amsterdam.

Lammers, J.G. (1984), *Pollution of International Watercourses*, Martinus Nijhoff Publishers, The Hague.

LeMarquand, D. (1993), The international joint commission and changing Canada-United States boundary relations, *Natural Resources Journal* 33, 59-92.

Lipper, J. (1967), Equitable utilization, in *The Law of International Drainage Basins*, ed. A.H. Garretson, R.D. Hayton and C.J. Olmstead, Oceana Publications, Inc., Dobbs Ferry, New York, 15-88.

Lipscomb M. and A.M. Mobarak (2007), Decentralization and water quality: evidence from the re-drawing of county boundaries in Brazil, BREAD working paper no. 159, Bureau for Research and Economic Analysis of Development, Durham, North Carolina.

Macho-Stadler, I., D. Pérez-Castrillo and D. Wettstein (2006), Efficient bidding with externalities, *Games and Economic Behavior* 57, 304-320.

Macho-Stadler, I., D. Pérez-Castrillo and D. Wettstein (2007), Sharing the surplus: an extension of the Shapley value for environments with externalities, *Journal of Economic Theory* 135, 339-356.

Mäler, K.G. (1990), International environmental problems, *Oxford Review of Economic Policy* 6, 80-108.

Maschler, M. and G. Owen (1989), The consistent Shapley value for hyperplane games, *International Journal of Game Theory* 18, 389-407.

McCaffrey, S.C. (1996), The Harmon doctrine one hundred years later: buried, not praised, *Natural Resources Journal* 36, 549-590.

McCaffrey, S.C. (2001), *The Law of International Watercourses*, Oxford University Press, Oxford.

McCaffrey, S.C. and M. Sinjela (1998), The 1997 United Nations convention on international watercourses, *American Journal of International Law* 92, 97-107.

Moermond, J.O. and E. Shirley (1987), A survey of the international law of rivers, *Denver Journal of International Law and Policy* 16, 139-159.

Moulin, H. (2000), Priority rules and other asymmetric rationing methods, *Econometrica* 68, 643-684.

Mutuswami, S., D. Pérez-Castrillo and D. Wettstein (2004), Bidding for the surplus: realizing efficient outcomes in economic environments, *Games and Economic Behavior* 48, 111-123.

Myerson, R.B. (1977), Graphs and cooperation in games, *Mathematics of Operations Research* 2, 225-229.

Myerson, R.B. (1980), Conference structures and fair allocation rules, *International Journal of Game Theory* 9, 169-182.

Myerson, R.B. (1991), *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, Massachusetts.

Neumann, J. von, and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.

Ni, D. and Y. Wang (2007), Sharing a polluted river, *Games and Economic Behavior* 60, 176-186.

O'Neill, B. (1982), A problem of rights arbitration from the Talmud, *Mathematical Social Sciences* 2, 345-371.

Owen, G. (1977), Values of games with a priori unions, in *Mathematical Economics and Game Theory*, ed. R. Henn and O. Moeschlin, Springer-Verlag, Berlin, 76-88.

Owen, G. (1986), Values of graph-restricted games, *SIAM Journal on Algebraic and Discrete Methods* 7, 210-220.

Parrachino, I., A. Dinar and F. Patrone (2006), Cooperative game theory and its application to natural, environmental and water resource issues: 3. Application to water resources, World Bank Policy Research Working Paper 4074, World Bank, Washington D.C.

Peleg, B. and P. Sudhölter (2003), *Introduction to the Theory of Cooperative Games*, Kluwer Academic Publishers, Dordrecht.

Pérez-Castrillo, D. and D. Wettstein (2001), Bidding for the surplus: a non-cooperative approach to the Shapley value, *Journal of Economic Theory* 100, 274-294.

Pérez-Castrillo, D. and D. Wettstein (2002), Choosing wisely: a multi-bidding approach, *American Economic Review* 92, 1577-1587.

Pérez-Castrillo, D. and D. Wettstein (2005), Forming efficient networks, *Economic Letters* 87, 83-87.

Ray, D. and R. Vohra (1999), A theory of endogenous coalition structures, *Games and Economic Behavior* 26, 286-336.

Rogers, P. (1969), A game theory approach to the problems of international river basins, *Water Resources Research* 5, 749-760.

Salman, S.M.A. (2007a), The United Nations watercourses convention ten years later: why has its entry into force proven difficult?, *Water International* 32, 1-15.

Salman, S.M.A. (2007b), The Helsinki rules, the UN watercourses convention and the Berlin rules: perspectives on international water law, *Water Resources Development* 23, 625-640.

Schmeidler, D. (1969), The nucleolus of a characteristic function game, *SIAM Journal on Applied Mathematics* 17, 1163-1170.

Shapley, L.S. (1953), A value for n-person games, in *Contributions to the Theory of Games, Vol. II*, ed. H.W. Kuhn and A.W. Tucker, Princeton University Press, Princeton, 307-317.

Shapley, L.S. (1967), On balanced sets and cores, *Naval Research Logistics Quarterly* 14, 453-460.

Shapley, L.S. (1971), Cores of convex games, *International Journal of Game Theory* 1, 11-26.

Sigman, H. (2002), International spillovers and water quality in rivers: do countries free-ride?, *American Economic Review* 92, 1152-1159.

Sigman, H. (2005), Transboundary spillovers and decentralization of environmental policies, *Journal of Environmental Economics and Management* 50, 82-101.

Slikker, M. (2007), Bidding for surplus in network allocation problems, *Journal of Economic Theory* 137, 493-511.

Tanzi, A. and M. Arcari (2001), *The United Nations Convention on the Law of International Watercourses*, Kluwer Law International, The Hague.

Thomson, W. (2003), Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: a survey, *Mathematical Social Sciences* 45, 249-297.

Thrall, R.M. and W.F. Lucas (1963), n-Person games in partition function form, *Naval Research Logistics Quarterly* 10, 281-298.

United Nations (1997), *Convention on the Law of the Non-Navigational Uses of International Watercourses*, New York.

United Nations (2009), *The United Nations World Water Development Report 3: Water in a Changing World*, UNESCO publishing, Paris, and Earthscan, London.

Vázquez-Brage, M., I. García-Jurado and F. Carreras (1996), The Owen value applied to games with graph-restricted communication, *Games and Economic Behavior* 12, 42-53.

Vidal-Puga, J. and G. Bergantiños (2003), An implementation of the Owen value, *Games and Economic Behavior* 44, 412-427.

Wang, Y. (2011), Trading water along a river, *Mathematical Social Sciences* 61, 124-130.

Wolf, A.T. (1999), Criteria for equitable allocations: the heart of international water conflict, *Natural Resources Forum* 23, 3-30.

Wolf, A.T. (2007), Shared waters: conflict and cooperation, *Annual Review of Environment and Resources* 32, 241-269.

Samenvatting (summary in Dutch)

Rivierwater allocatie als een coöperatief beslissingsprobleem

Dit proefschrift presenteert recente ontwikkelingen in het modelleren van de allocatie van water dat stroomt in internationale rivieren. Rivieren, en de watervlaktes waar deze in uitmonden, vormen de belangrijkste bron van zoet water ter wereld. Een rivier wordt ‘internationaal’ genoemd als deze door ten minste twee staten stroomt. Het water uit internationale rivieren wordt op grote schaal gebruikt in de landbouw, industrie en door huishoudens. Bovendien worden rivieren gebruikt als visgrond en voor transport, het opwekken van elektriciteit, recreatie, en het afvoeren van afval. Het is niet moeilijk om voor te stellen dat het gebruik van een rivier voor dit soort activiteiten door een bovenstrooms land negatieve gevolgen kan hebben voor benedenstroomse landen. Vooral in gebieden waar schaarste aan water heerst kan de consumptie van grote hoeveelheden rivierwater door een bovenstrooms land rampzalig zijn voor benedenstroomse landen.

Asymmetrische afhankelijkheid van een waterbron, zoals in bovenstaand voorbeeld, ligt vaak ten grondslag aan geschillen over het gebruik van de bron. Dit geldt met name voor waterbronnen waarvan het eigendomsrecht niet is vastgesteld. In nationale aangelegenheden kunnen problemen rond waterbronnen meestal worden opgelost via het nationale rechtssysteem. In internationale aangelegenheden bestaat er echter geen supranationale autoriteit die haar wil kan opleggen aan de twistende partijen.

Internationale organisaties, zoals de Verenigde Naties, hebben geprobeerd om afspraken over het gebruik van internationale rivieren vast te leggen in verdragen. Hoewel dit, tot op heden, nog niet geresulteerd heeft in globaal bindende afspraken die het gebruik van water uit internationale rivieren reguleren, heeft het wel geholpen bij het tot stand brengen van meer dan vierhonderd (bilaterale) verdragen tussen staten die een internationale rivier delen. Deze verdragen, en de principes waarop de afspraken in deze verdragen zijn gebaseerd, zijn het onderwerp van een omvangrijke literatuur op het gebied van internationaal rivierrecht.

De literatuur op het gebied van internationaal rivierrecht vormt een van de twee belangrijkste pijlers in modellen die tot doel hebben de allocatie van water uit internationale rivieren te beschrijven, voorschrijven en te voorspellen. De andere pijler is de (coöperatieve) speltheorie.

Een speltheoreticus bestudeert conflict en samenwerking tussen een beperkt aantal economische agenten (spelers) in een wiskundig model (spel). In de niet- coöperatieve speltheorie wordt aangenomen dat spelers niet in staat zijn om bindende afspraken te maken. In de coöperatieve speltheorie is dit wel het geval en worden onderliggende strategische procedures die leiden tot de afspraken grotendeels buiten beschouwing gelaten. In plaats daarvan wordt een coöperatief spel gevormd door een verzameling spelers en een karakteristieke functie die voor elke coalitie (groep) van spelers een (maximaal) te bereiken waarde (hoeveelheid nut) aangeeft. Gegeven zo'n coöperatief spel ligt de nadruk op vragen zoals: welke coalitie van spelers wordt gevormd in het spel? En, hoe wordt de waarde van een coalitie verdeeld onder de leden van de coalitie? Een oplossing voor coöperatieve spelen is een functie die voor elk spel een uitbetaling aan elke speler toekent.

Aangezien afspraken over het gebruik van een internationale rivier meestal gemaakt worden tussen een beperkt aantal landen is het niet vreemd dat de (coöperatieve) speltheorie aan de basis ligt van een kleine, maar groeiende, literatuur die de verdeling van water uit internationale rivieren modelleert.

In het uitbreiden van deze literatuur staan in dit proefschrift drie doelen centraal. Het eerste doel is het generaliseren van een bestaand enkel-strooms internationaal rivierwater allocatiemodel naar een model waarin de rivier mogelijk bestaat uit een hoofdrivier met meerdere zijrivieren en/of een rivierdelta. Het tweede doel is het uitbreiden van hetzelfde model naar een model waarin landen mogelijk bestaan uit meerdere water gebruikers (b.v. provincies, steden of individuele gebruikers). Het derde doel is het analyseren van de verschillen tussen het rivaliserend gebruik van water (als het geconsumeerd is, kan het niet nogmaals worden geconsumeerd door anderen) en het niet-rivaliserend gebruik van water (b.v. door vervuiling van het water of het gebruik bij het opwekken van elektriciteit) in een internationaal rivierwater allocatiemodel.

Het onderzoek naar deze drie doelen heeft geleid tot dit proefschrift dat bestaat uit zeven hoofdstukken. De eerste twee daarvan hebben een introducerend karakter. De laatste vijf bevatten de nieuwe bevindingen.

Hoofdstuk 1 geeft een uitgebreide samenvatting van de bestaande literatuur op het gebied van internationaal rivierrecht. Het verschaft zowel een algemene introductie in het water distributie probleem in internationale rivierbekkens als een specifieke discussie van de aan het internationaal rivierrecht onderliggende principes.

In Hoofdstuk 2 wordt een wiskundige inleiding gegeven in de coöperatieve speltheorie door verschillende spelen met overdraagbaar nut te introduceren. De meeste van de speltheoretische concepten die in dit proefschrift worden gebruikt, kunnen worden teruggevonden in dit hoofdstuk. Verder behandelt Hoofdstuk 2 de belangrijkste bevindingen uit de literatuur over coöperatieve beslissingsmodellen in water allocatieproblemen. In deze literatuur wordt het internationaal rivierrecht uit Hoofdstuk 1 gecombineerd met speltheoretische modellen om te komen tot een unieke klasse van modellen. Deze modellen zijn, aan de ene kant, vrij technisch, maar, aan de andere kant, geven ze een duidelijk inzicht over hoe bepaalde principes uit het internationaal rivierrecht kunnen worden toegepast op bestaande rivierwater allocatieproblemen.

In Hoofdstuk 3 wordt een water allocatiemodel bekeken waarin een beperkt aantal agenten (landen) opeenvolgend langs een enkel-strooms rivier is gepositioneerd. Er worden verschillende (onafhankelijkheids)axioma's voor dit model geïntroduceerd en deze axioma's worden gebruikt om twee bestaande en twee nieuwe oplossingen te karakteriseren. Uiteindelijk worden alle vier de oplossingen toegepast op een specifiek geval van het model waarin elke agent een constant marginaal nut van waterconsumptie heeft, tot een verzadigingspunt, en geen marginaal nut daarboven. Hierbij volgt het dat twee van de oplossingen kunnen worden geïmplementeerd zonder monetaire transfers tussen de agenten.

Hoofdstuk 4 richt zich op het eerste van bovengenoemde doelen door het model uit Hoofdstuk 3 te generaliseren naar een model waarin de uiteindelijke hoofdrivier mogelijk ontstaat uit verschillende bronnen. Dit betekend dat er een hoofdrivier is die mogelijk meerdere zijrivieren heeft waarlangs agenten gepositioneerd kunnen zijn. In dit model worden twee verschillende aannamen op de nutsfuncties van de agenten bekeken. De eerste aanname leidt tot een type coöperatief spel waarin agenten altijd meer water willen consumeren, de tweede tot een ander type coöperatief spel waarin agenten mogelijk een verzadigingspunt hebben, wat de mogelijkheid geeft tot samenwerkingsexternaliteiten. Voor beide spelen wordt de klasse van gewogen hiërarchische uitkomsten voorgesteld als klasse van oplossingen die voldoet aan het "TIBS" principe uit het internationaal rivierrecht.

In Hoofdstuk 5 wordt een strategische implementatie van de oplossingen uit de klasse van gewogen hiërarchische uitkomsten gegeven (een niet-coöperatief spel waarvan de uitbetalingen in een evenwicht gelijk zijn aan een gewogen hiërarchische uitkomst in het coöperatieve spel).

Hoofdstuk 6 focust op zowel het eerste als het tweede van bovengenoemde doelen, door het model van Hoofdstuk 3 en Hoofdstuk 4 verder uit te breiden naar een model waarin de rivier mogelijk bestaat uit een hoofdrivier met meerdere zijrivieren en/of een rivierdelta, en waarin de agenten mogelijk bestaan uit meerdere water gebruikers. Om dit te bereiken wordt gebruik gemaakt van spelen met overdraagbaar nut die zowel een graaf structuur (die de rivier weergeeft) als een coalitie structuur (die de verschillende water gebruikers in een land weergeeft) hebben. Er worden twee nieuwe oplossingen geïntroduceerd en gekarakteriseerd voor dit soort spelen die zijn gebaseerd op de Shapley waarde voor gewone spelen met overdraagbaar nut. Verder wordt er nog een extra oplossing voorgesteld die dichter bij de klasse van oplossingen uit Hoofdstuk 4 staat.

Hoofdstuk 7 is het laatste hoofdstuk en behandelt het derde van bovengenoemde doelen. In dit hoofdstuk wordt een model bestudeerd waarin de agenten het water uit de rivier niet direct consumeren (zodat dit niet meer gebruikt kan worden door andere) maar indirect, door het te vervuilen (zodat dit nog wel door andere gebruikt kan worden, maar van lagere kwaliteit is). Uit dit model volgt dat de vervuiling van de rivier afneemt als de agenten langs de rivier besluiten samen te werken. De resulterende toename in welvaart kan tussen de agenten worden verdeeld op basis van de eigendomsrechten over de rivier. Met behulp van enkele principes uit het internationaal rivierrecht worden verschillende manieren voorgesteld om de eigendomsrechten, en daarom ook de toename in welvaart na samenwerking, 'eerlijk' te verdelen tussen de agenten langs de rivier.

Acknowledgments

There is a number of people without whom the publication of this dissertation would not have been possible. I would like to thank my supervisors, Gerard van der Laan and René van den Brink, for between them co-authoring all of my papers and giving me the freedom to explore the field of (cooperative) game theory. The main thing that I have learned from Gerard en René is to be very precise, not only in my mathematics, but in my thinking and writing in general. I would like to thank Arantza Estévez-Fernández for co-authoring one of the papers that has made it into this dissertation and all of the other members of the Department of Econometrics of the Vrije Universiteit for providing a pleasant working environment. Further, I would like to thank Dolf Talman for introducing me into the field of mathematical economics and, in chronological order, Pawel, Warren, Bahar, Chris, Suzanne and Ping for sharing an office with me at the VU and in that way being exposed to my ever changing, often controversial and not always consistent opinions. Finally, I would like to thank my family, especially my father, mother, brother and sister, for supporting me in everything I do.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

493 R.E. VLAHU, *Three Essays on Banking*
494 N.E. VIKANDER, *Essays on Teams and the Social Side of Consumption*
495 E. DEMIREL, *Economic Models for Inland Navigation in the Context of Climate Change*
496 V.A.C. VAN DEN BERG, *Congestion pricing with Heterogeneous travellers*
497 E.R. DE WIT, *Liquidity and Price Discovery in Real Estate Assets*
498 C. LEE, *Psychological Aspects of the Disposition Effect: An Experimental Investigation*
499 MHA. RIDHWAN, *Regional Dimensions of Monetary Policy in Indonesia*
500 J. GARCÍA, *The moral herd: Groups and the Evolution of Altruism and Cooperation*
501 F.H. LAMP, *Essays in Corporate Finance and Accounting*
502 J. SOL, *Incentives and Social Relations in the Workplace*
503 A.I.W. HINDRAYANTO, *Periodic Seasonal Time Series Models with applications to U.S. macroeconomic data*
504 J.J. DE HOOP, *Keeping Kids in School: Cash Transfers and Selective Education in Malawi*
505 O. SOKOLINSKIY, *Essays on Financial Risk: Forecasts and Investor Perceptions*
506 T. KISELEVA, *Structural Analysis of Complex Ecological Economic Optimal Management Problems*
507 U. KILINC, *Essays on Firm Dynamics, Competition and Productivity*
508 M.J.L. DE HEIDE, *R&D, Innovation and the Policy Mix*
509 F. DE VOR, *The Impact and Performance of Industrial Sites: Evidence from the Netherlands*
510 J.A. NON, *Do ut Des: Incentives, Reciprocity, and Organizational Performance*
511 S.J.J. KONIJN, *Empirical Studies on Credit Risk*
512 H. VRIJBURG, *Enhanced Cooperation in Corporate Taxation*
513 P.ZEPPINI, *Behavioural Models of Technological Change*
514 P.H.STEFFENS, *Its Communication, Stupid! Essays on Communication, Reputation and (Committee) Decision-Making*

515 K.C. YU, *Essays on Executive Compensation - Managerial Incentives and Disincentives*
516 P. EXTERKATE, *Of Needles and Haystacks: Novel Techniques for Data-Rich Economic Forecasting*
517 M. TYSZLER, *Political Economics in the Laboratory*
518 Z. WOLF, *Aggregate Productivity Growth under the Microscope*
519 M.K. KIRCHNER, *Fiscal Policy and the Business Cycle The Impact of Government Expenditures, Public Debt, and Sovereign Risk on Macroeconomic Fluctuations*
520 P.R. KOSTER, *The cost of travel time variability for air and car travelers*
521 Y.ZU, *Essays of nonparametric econometrics of stochastic volatility*
522 B.KAYNAR, *Rare Event Simulation Techniques for Stochastic Design Problems in Markovian Setting*
523 P. JANUS, *Developments in Measuring and Modeling Financial Volatility*
524 F.P.W. SCHILDER, *Essays on the Economics of Housing Subsidies*
525 S. M MOGHAYER, *Bifurcations of Indifference Points in Discrete Time Optimal Control Problems*
526 C. ÇAKMAKLI, *Exploiting Common Features in Macroeconomic and Financial Data*
527 J. LINDE, *Experimenting with new combinations of old ideas*
528 D. MASSARO, *Bounded rationality and heterogeneous expectations in macroeconomics*
529 J. GILLET, *Groups in Economics*
530 R. LEGERSTEE, *Evaluating Econometric Models and Expert Intuition*
531 M.R.C. BERSEM, *Essays on the Political Economy of Finance*
532 T. WILLEMS, *Essays on Optimal Experimentation*
533 Z. GAO, *Essays on Empirical Likelihood in Economics*
534 J. SWART, *Natural Resources and the Environment: Implications for Economic Development and International Relations*
535 A. KOTHIYAL, *Subjective Probability and Ambiguity*
536 B. VOOGT, *Essays on Consumer Search and Dynamic Committees*
537 T. DE HAAN, *Strategic Communication: Theory and Experiment*
538 T. BUSER, *Essays in Behavioural Economics*
539 J.A. ROSERO MONCAYO, *On the importance of families and public policies for child development outcomes*
540 E. ERDOGAN CIFTCI, *Health Perceptions and Labor Force Participation of Older Workers*
541 T.WANG, *Essays on Empirical Market Microstructure*
542 T. BAO, *Experiments on Heterogeneous Expectations and Switching Behavior*
543 S.D. LANSDORP, *On Risks and Opportunities in Financial Markets*