The Beijer Institute of Ecological Economics

DISCUSSION PAPER

Beijer Discussion Paper Series No. 265

The Tipping Game

Florian Wagener and Aart de Zeeuw. 2018.

The Beijer Institute of Ecological Economics
The Royal Swedish Academy of Sciences
Box 50005 SE-104 05 Stockholm Sweden
Phone: +46 8 673 95 00 Fax: +46 8 15 24 64
E-mail: beijer@beijer.kva.se

The Tipping Game*

Florian Wagener[†] and Aart de Zeeuw[‡]

12th December 2018

Abstract

This paper shows that in the presence of a tipping point in a natural system, stable partial cooperation may prevent tipping and if this is not possible, the remaining gains of cooperation are small. This is good news, because the big loss of ecosystem services from tipping can often be avoided. Moreover, the usual grim story that a high level of cooperation is hard to achieve and usually leaves large possible gains of cooperation, does not hold in the presence of a tipping point. These results are shown for a simple tipping game, with constant inputs and piecewise linear dynamics, and for the well-known lake game, with time-dependent inputs and convex-concave dynamics. Tipping back to good conditions can also be induced by stable partial cooperation, but this paper shows that this is more vulnerable to free-riding. Therefore, a natural system that is physically reversible may prove to be socially irreversible.

JEL classification: C70; Q20

Keywords: Tipping points, multiple Nash equilibria, partial cooperation, stability, ecological systems

---

*We are grateful for discussions on this topic with Scott Barrett, Michael Finus and Paolo Zeppini. We have benefitted from presentations and from the comments of the participants at EAERE conferences in 2017 and 2018, at SURED and CESifo conferences in 2018, at workshops in Warsaw in 2017 and in Thessaloniki in 2018, and at seminars in Munich in 2017 and in Paris in 2018.

[†]CeNDEF, ASE, Universiteit van Amsterdam, P.O. Box 15867, 1001 NJ Amsterdam, the Netherlands, and Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, the Netherlands, e-mail: *F.O.O.Wagener@uva.nl*

[‡]CentER, TSC, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, the Netherlands, and The Beijer Institute of Ecological Economics, The Royal Swedish Academy of Sciences, Box 50005, 10405 Stockholm, Sweden, e-mail: *A.J.deZeeuw@uvt.nl*

1 Introduction

Natural systems often show tipping points or regime shifts. At some point, a small change in one of the inputs may have substantial consequences, as the system shifts to a different domain of attraction, with an equilibrium that has very different characteristics (Scheffer et al., 2001, Biggs et al., 2012). Furthermore, if this happens, it is very difficult (hysteresis) or even impossible (irreversibility) to restore the original conditions of the natural system. A well-known example is the lake system, where at some point a small increase in phosphorus loading shifts the lake to bad conditions, with a substantial loss of ecosystem services (Carpenter and Cottingham, 1997, Scheffer, 1997). Another example is the coral-reef system, where at some point a small increase in the temperature of the ocean shifts the coral reef to bad conditions, with a substantial loss of coral and fish (Hughes et al., 2003). It is expected that the climate system also has a tipping point, although this has fortunately not yet been observed (Lenton and Ciscar, 2013).

Economic activities yield benefits but also release emissions into the natural systems. At a tipping point, a marginal increase in economic activities and benefits, and thus in emissions, yields a non-marginal increase in costs, i.e. a sudden big loss of ecosystem services. The presence of tipping points is therefore a threat, but it may also have a positive effect. In a public good game on emission reduction, Barrett (2013) has shown that if it is collectively rational to avoid tipping, it may also be individually rational to avoid tipping. More specifically, a tipping point may change a prisoners' dilemma into a coordination game. If tipping cannot be avoided in a non-cooperative equilibrium, cooperation is needed, but Barrett (2013) does not find many cases where stable cooperation solves the problem. This paper analyses a dynamical tipping game, and then extends the results to the lake system. It shows that if it is not individually rational to avoid tipping, stable partial cooperation often solves the problem and if not, the remaining gains of cooperation are small. Stable partial cooperation is defined as the equilibrium between the incentives to cooperate and the incentives to free-ride. The presence of tipping points increases the incentives to cooperate or decreases the incentives to free-ride, and thus increases the size of the stable coalition. These results are in contrast with the usual grim story that a high level of cooperation cannot be sustained, especially when the gains of cooperation are high. In the presence of tipping points, however, the opposite occurs. Tipping is usually avoided by a non-cooperative equilibrium or by stable partial cooperation and if not, the remaining gains of cooperation are small.

The literature on economic analyses in the presence of tipping points is rapidly increasing (e.g., Brock and Starrett, 2003, Mäler et al., 2003, Wagener, 2003, Crépin, 2007, Kossioris et al., 2008, Kiseleva and Wagener, 2010, Polasky et al., 2011, Heijdra and Heijnen, 2013, Dockner and Wagener, 2014, Lemoine and Traeger, 2014, Cai et al., 2015, van der Ploeg and de Zeeuw, 2018). Mäler et al. (2003) compare cooperation and non-cooperation in the lake system, between the users of the lake. They show the existence of two Nash equilibria, one close to the full-cooperative outcome, with a high level of ecosystem services, and one where a big loss of ecosystem services occurs. Welfare in the good Nash equilibrium is somewhat lower than in the full-cooperative outcome, but welfare drops considerably when the game ends up in the bad Nash equilibrium. Which equilibrium can be played depends on the initial conditions of the lake. An important question is how the users of the lake can avoid the bad Nash equilibrium, or how they can get out of it in case the game happens to end up in the bad Nash equilibrium.

Cooperation is the answer, and full cooperation would also yield the first-best outcome. However, cooperation is usually not stable in the sense that free-rider incentives dominate the incentives to cooperate. The literature on international environmental agreements (e.g., Hoel, 1992, Carraro and Siniscalco, 1993, Barrett, 1994, Finus, 2003, Karp and Simon, 2013) shows that the level of stable cooperation, i.e. the size of the stable coalition, is usually very small. However, in the presence of a tipping point, the incentives to free-ride are lower, and the incentives to cooperate are higher. This paper shows that it is possible to have larger stable coalitions that avoid tipping and increase welfare. However, there is a limit to this. In some cases, a very high level of cooperation is needed and the free-rider incentives dominate again. Fortunately, in these cases the remaining gains of cooperation are low. Similarly, when the game has ended up in the bad Nash equilibrium earlier, it is possible to form larger stable coalitions in order to induce tipping back to the good conditions of the lake, but this has its limits as well.

This paper starts with a simple dynamical tipping-point model in order to analyse the basic questions of optimal management, Nash equilibria, and coalition formation. We will show when avoiding the tip is first-best, and when this is also a Nash equilibrium. If it is not a Nash equilibrium, we will show when coalition formation can avoid tipping. This model has constant inputs and linear dynamics with a downward jump, so that the analysis becomes tractable. In the sequel, the paper analyses the same questions for the lake system, with time-dependent phosphorus loadings as inputs and convex-concave system dynamics. Elaborate numerical methods are needed to solve for the equilibria. The most interesting result is that tipping from the good to the bad conditions of the lake may indeed be avoided by stable partial cooperation and if not, the remaining gains of cooperation are small. In case tipping has occurred earlier, tipping back from the bad to the good conditions of the lake may also be induced by stable partial cooperation, but this requires a higher level of cooperation and is therefore more vulnerable to free-riding.

An important conclusion of this paper is that it is possible that stable partial cooperation avoids tipping, or induces tipping back. However, it is not always possible to sustain the level of cooperation needed, because the incentive to free-ride may become higher than the incentive to cooperate. This implies that even if it is physically possible to tip back, it may not always be socially possible. It follows that the combined socio-ecological system has more complicated hysteresis and irreversibility properties than the underlying ecological system itself. In the ecological system, it is always costly to tip back, because the level of the input has to be reduced substantially, but it may still be possible to do so: the ecological system has hysteresis in this case, but it is not irreversible. However, the socio-ecological system may be irreversible, because the level of cooperation that is needed to induce the system to tip back is not stable. In this case, the incentive to free-ride prevents sustaining the high level of cooperation that is needed to tip back, which makes tipping socially irreversible.

Section 2 presents the simple tipping game, with constant inputs and linear dynamics, and analyses optimal management, the Nash equilibria, and coalition formation in this game, with a note on uncertainty about the location of the tipping point. Section 3 presents and analyses the tipping game with time-dependent inputs and convex-concave dynamics, in the form of the lake game. Section 4 concludes the paper.

2 The simple tipping game

The most simple way of modelling a tipping point or regime shift is the following. Suppose that a natural system is affected by inputs a_i, $i = 1, 2, ..., n$, from n economic agents. The essential dynamics of the system is represented by a differential equation in a stock of pollutants s (for example, the stock of phosphorus in the water of the lake or the stock of greenhouse gases in the atmosphere), given by

$$\dot{s}(t) = A - f(s(t)), \quad A = \sum_{i=1}^{n} a_i, \quad s(0) = s_0, \tag{2.1}$$

$$\begin{aligned} f(s) &= bs, \qquad 0 \leq s \leq s_c, \\ f(s) &= bs - 1, \quad s > s_c, \end{aligned} \tag{2.2}$$

where s_c denotes the level of the stock where tipping occurs. Figure 1 shows the dynamics of the system for an initial stock $s(0) = 0$ and for different levels of the total input A. The tipping point is given by (A_{ch}, s_c). A small increase of the total input above A_{ch} shifts the equilibrium level of the stock s to a much higher level. Figure 2 shows the dynamics of the system for a high initial stock s_0 and, again, for different levels of the total input A. In order to shift back to the favourable conditions of the natural system, the total input has to be reduced to just below A_{cl}: (A_{cl}, s_c) is the other tipping point. If the tipping point (A_{cl}, s_c) happens to lie below the s-axis, it is impossible to restore the original conditions of the natural system: the system is physically irreversible.

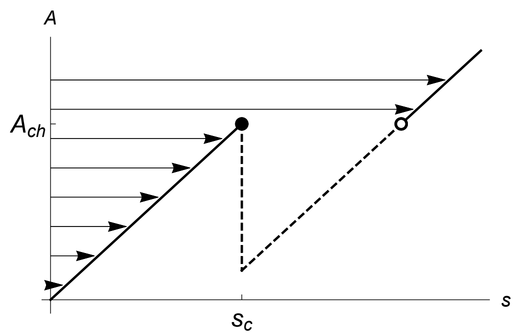

Figure 1: Initial stock $s(0) = 0$

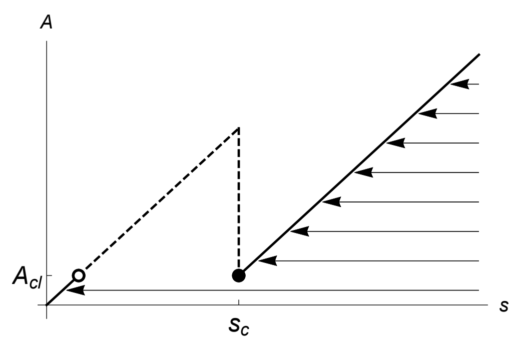

Figure 2: High initial stock

It is clear that bad conditions of the natural system (i.e., a high stock s) imply a loss of welfare, for example in terms of a loss of ecosystem services. On the other hand, the inputs a_i usually represent a benefit, for example agriculture that causes the release of phosphorus on the lake, or the use of cheap fossil fuels that causes the emission of greenhouse gases. We can model this trade-off by maximising the welfare indicators (Mäler et al., 2003), given by

$$\max_{a_i} \left[\ln a_i - cs^2 \right], \quad i = 1, 2, ..., n, \tag{2.3}$$

where the parameter c weighs the benefits and costs. Logarithmic utility is chosen as this functional form will be convenient in the analysis below. The question is how the maximisation of welfare is affected by the existence of a tipping point and how the incentives to cooperate or to free-ride change in the presence of a tipping point. The problem is relatively simple, because we assume constant inputs a_i and we can focus on the steady states that can arise as shown by Figure 1 and Figure 2. We will first derive the results for the tipping point in Figure 1. The analysis for the tipping point in Figure 2 is basically the same.

2.1 Full cooperation

First we solve the full-cooperative case which means that we maximize the sum of the welfare indicators. The Lagrangian becomes

$$L = \sum_{i=1}^{n} \ln a_i - ncs^2 + \lambda(A - f(s)). \tag{2.4}$$

Since $f(s)$ is either part of the line bs or part of the line $bs - 1$, we can characterise the stationary points by considering the first-order condition

$$A = \sum_{i=1}^{n} a_i = -\frac{n}{\lambda} = \frac{nf'(s)}{2ncs} = \frac{b}{2cs}. \tag{2.5}$$

This first-order condition is a hyperbola h in Figure 1 that moves up for decreasing c. It is clear that when h cuts the left part of the possible steady states, this is the optimal outcome. This happens for high values of c, which is intuitively clear because the costs of a high stock s weigh heavily in this case. We will denote this set of possible steady states as the line segment l_1. For a decreasing c, the hyperbola moves up and for some value of c, say $c = \tilde{c}$, it cuts the line segment l_1 in the tipping point (A_{ch}, s_c). Decreasing c further, the hyperbola moves up and at some point it cuts the right part of the possible steady states. We will denote this set of possible steady states as the line segment l_2. At first, welfare in the tipping point is higher than welfare in the intersection point of the hyperbola and the line segment l_2. Decreasing c further implies that welfare in the tipping point increases but welfare in the intersection point increases faster. For some value of c, say $c = \hat{c}$, welfare is the same in the tipping point as in the intersection point corresponding to this $c = \hat{c}$. This means that for $c = \hat{c}$, the highest feasible iso-welfare curve w is tangent to the line segment l_2 and passes through the tipping point (A_{ch}, s_c) (Figure 3). Decreasing c further, i.e. $c < \hat{c}$, implies that the tipping point is not feasible anymore and that the optimal outcome is the intersection point of the hyperbola and the line segment l_2. This is intuitively clear because for a low c a high stock s is not so costly. Note that for intermediate values of c, i.e. $\hat{c} \leq c \leq \tilde{c}$, the maximal welfare is realised in the tipping point (A_{ch}, s_c).

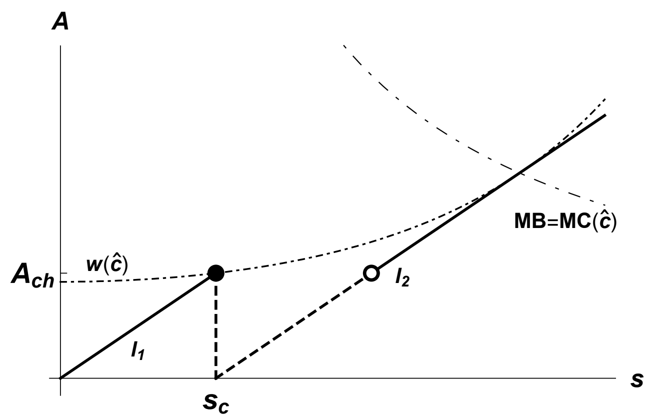

Figure 3: For $c = \hat{c}$ welfare is maximal at the tipping point as well as at the tangency.

2.2 Nash equilibria

The non-cooperative Nash equilibrium can be characterised as follows. The Lagrangians become

$$L_i = \ln a_i - cs^2 + \lambda_i(A - f(s)), \quad i = 1, 2, ..., n. \tag{2.6}$$

Since $f(s)$ is either part of the line bs or part of the line $bs - 1$, we can characterise the candidate symmetric Nash equilibria by considering the first-order condition

$$A = \sum_{i=1}^{n} a_i = -\frac{n}{\lambda_i} = \frac{nf'(s)}{2cs} = \frac{nb}{2cs}. \tag{2.7}$$

This first-order condition is again a hyperbola in Figure 1 that moves up for decreasing c. Note that this hyperbola passes through the tipping point (A_{ch}, s_c) for $c = n\tilde{c}$. It is clear that when this hyperbola cuts the line segment l_1, this is the only Nash equilibrium. However, when c is low enough so that this hyperbola cuts the line segment l_2, there are possibly two Nash equilibria, namely the tipping point (A_{ch}, s_c) and this intersection point. We can investigate whether the tipping point is indeed a Nash equilibrium by checking the Nash property. When the other economic agents stick to their input in the tipping point, the first-order condition for the best response of economic agent i, and the resulting total input A, are given by

$$a_i = -\frac{1}{\lambda_i} = \frac{f'(s)}{2cs} = \frac{b}{2cs} \quad \Longrightarrow \quad A = \frac{b}{2cs} + \frac{n-1}{n} A_{ch}. \tag{2.8}$$

When c is sufficiently low, say $c = \bar{c}$, this expression for total input A will cut the line segment l_2 in a point where welfare of economic agent i is the same as in the tipping point (A_{ch}, s_c) (Figure 4). This implies that for $c < \bar{c}$, economic agent i can improve by deviating, so that the tipping point does not have the Nash property. In that case it is individually rational to tip, because the costs of a higher stock s are not sufficiently high to prevent tipping. The other candidate Nash equilibrium, located on the line segment l_2, becomes the only Nash equilibrium. However, for $\bar{c} \leq c$, economic agent i cannot improve by deviating. It follows that the tipping point (A_{ch}, s_c) is a Nash equilibrium for $\bar{c} \leq c \leq n\tilde{c}$.

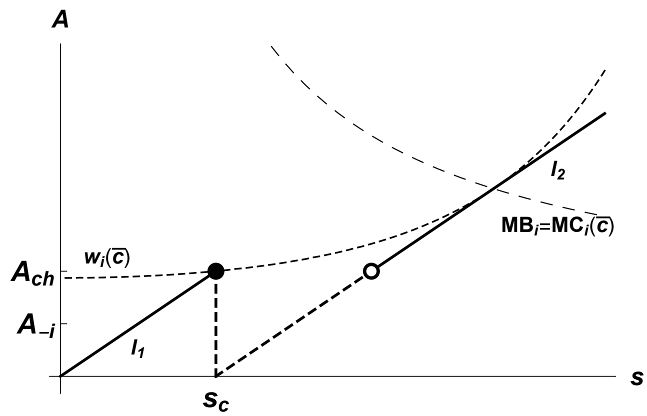

Figure 4: For $c = \bar{c}$ individual welfare of agent i is maximal at the tipping point as well as at the tangency.

The results above for the full-cooperative outcome and the possible Nash equilibria in this simple model with a tipping point have an interesting consequence. It is possible that the tipping point is first-best but also a Nash equilibrium. This means that in this case the full-cooperative outcome is stable in the sense of being a Nash equilibrium. Individual rationality is collectively rational as well in this case. Barrett (2013) also finds, in a static model with a fixed loss of tipping, that a Nash equilibrium prevents tipping in a large part of the parameter space where it is optimal to do so. In our model, the tipping point (A_{ch}, s_c) is first-best and a Nash equilibrium when the areas $\hat{c} \leq c \leq \tilde{c}$, where the tipping point is first-best (see the previous section), and $\bar{c} \leq c \leq n\tilde{c}$, where the tipping point is a Nash equilibrium, overlap. This implies that if $\bar{c}$, the value of c where the tipping point loses the Nash property, lies in the area $\hat{c} \leq c \leq \tilde{c}$, then the tipping point is also a Nash equilibrium for $\bar{c} \leq c \leq \tilde{c}$ but not for $\hat{c} \leq c < \bar{c}$. If $\bar{c} > \tilde{c}$, then the tipping point is first-best for $\hat{c} \leq c \leq \tilde{c}$, but cannot be sustained as a Nash equilibrium.

The levels of $\tilde{c}$, $\hat{c}$ and $\bar{c}$ depend, of course, on the values of the parameters b and s_c, and on the number of economic agents n. Take, for example, $b = 1$, $s_c = 1$ and $n = 2$. The tipping point becomes $(A_{ch}, s_c) = (1, 1)$, with the inputs $a_1 = a_2 = 0.5$. It follows that the largest value of c for which the tipping point is first best is $\tilde{c} = 0.5$, which yields total welfare -2.3863. It is easy to show that for $\hat{c} = 0.0867$, the same level of total welfare -1.5597 is realised in the tipping point $(A_{ch}, s_c) = (1, 1)$ and in the point on the line segment l_2 with the inputs $a_1 = a_2 = 0.97646$ and the stock $s = 2.9529$. This implies that the tipping point is the optimal solution for $0.0867 \leq c \leq 0.5$. Furthermore, the tipping point (A_{ch}, s_c) is the only Nash equilibrium for $c = n\tilde{c} = 1$, with individual welfare -1.6931. The first-order condition for the best response of economic agent 2 yields the condition $A = 1/(2cs) + 0.5$. It is easy to show that for $\bar{c} = 0.13971$, the same level of individual welfare -0.83286 for the economic agent 2 is realised in the tipping point $(A_{ch}, s_c) = (1, 1)$ as in the point on the line segment l_2 with the input $a_2 = 1.285$ and the stock $s = 2.785$. This implies that the tipping point is a Nash equilibrium in the range $0.13971 \leq c \leq 1$, because an individual economic agent only has the incentive to deviate for $c < 0.13971$. It follows that the tipping point is first-best but also a Nash equilibrium for $0.13971 \leq c \leq 0.5$. For $0.5 < c \leq 1$, the tipping point is a Nash equilibrium, but it is optimal to move below the tipping point. For $0.0867 \leq c < 0.13971$, the tipping point is first-best, but it is not a Nash equilibrium anymore.

This most simple way of modelling a tipping point or regime shift already allows for an interesting conclusion. It is intuitively clear, of course, that it is optimal to prevent tipping, unless the weight on the damage in the welfare indicator becomes really small. More interestingly, however, for a range of weights on damage in the welfare indicator, a Nash equilibrium exists that does not tip either. The threat of tipping already induces sufficient individual discipline. However, for smaller weights on damage in the welfare indicator, it may still be optimal to prevent tipping, but this cannot be sustained as a Nash equilibrium anymore. In this case, the question arises whether stable partial cooperation can prevent tipping. This is the subject of the next section.

2.3 Coalition formation

In the previous section, it was found that the tipping point (A_{ch}, s_c) is a Nash equilibrium for $\bar{c} \leq c \leq n\tilde{c}$, but loses the Nash property for $c < \bar{c}$, so that the system will tip. The question is if tipping can be prevented by (partial) cooperation. We assume that one coalition of k cooperating economic agents can be formed and that the other $(n - k)$ economic agents continue to play individually. The respective welfare indicators become

$$\max_{a_1,...,a_k} \left[\sum_{i=1}^{k} \ln a_i - kcs^2 \right],$$

and

$$\max_{a_i} \left[\ln a_i - cs^2 \right], \quad i = k + 1, ..., n.$$

The Lagrangians become

$$L = \sum_{i=1}^{k} \ln a_i - kcs^2 + \lambda(A - f(s)), \tag{2.9}$$

$$L_i = \ln a_i - cs^2 + \lambda_i(A - f(s)), \quad i = k+1, ..., n. \tag{2.10}$$

Since $f(s)$ is either part of the line bs or part of the line $bs - 1$, we can characterise the candidate Nash equilibria between the coalition of size k and the individual economic agents by considering the first-order condition

$$A = \sum_{i=1}^{n} a_i = -\frac{k}{\lambda} - \frac{n-k}{\lambda_i} = \frac{kf'(s)}{2kcs} + \frac{(n-k)f'(s)}{2cs} = \frac{(n-k+1)b}{2cs}. \tag{2.11}$$

Note that this is the same condition as for the Nash equilibrium with $(n-k+1)$ economic agents. The coalition effectively operates as one individual economic agent. This is convenient for the analysis, and it is the result of choosing the logarithm as the functional form for the benefits.

In the example in the previous section, it is immediately clear what happens in the case of two economic agents. For $c < 0.13971$, the tipping point $(A_{ch}, s_c) = (1, 1)$ is not a Nash equilibrium, and the system ends up in the other candidate Nash equilibrium on the line segment l_2. However, if the two economic agents form a coalition, the tipping point is first-best for $0.0867 \leq c \leq 0.5$. It follows that for $0.0867 \leq c < 0.13971$, the system will tip if the economic agents do not form a coalition, but the system will stay at the tipping point if the economic agents cooperate and form a coalition. In the case of two economic agents, a coalition between them allows to reap substantial welfare benefits. For example, if $c = 0.1$, individual welfare at the tipping point is -0.79315. However, if the two economic agents do not form a coalition, the system tips and ends up in the Nash equilibrium on the line segment l_2, with the inputs $a_1 = a_2 = 1.3508$ and the stock $s = 3.7016$. Individual welfare in this Nash equilibrium is -1.0695, a loss of 35% as compared to individual welfare in the tipping point.

In the case of n economic agents, partial cooperation effectively means that the number of players is reduced. If a coalition of size k is formed, the number of players is reduced from n to $n-k+1$. Furthermore, the number of players affects the range of values of c for which the tipping point $(A_{ch}, s_c) = (1, 1)$ is a Nash equilibrium. This implies that it may happen that the system tips in case of $n-k+2$ players, but does not tip in case of $n-k+1$ players. More specifically, in the previous section we have seen that the tipping point $(A_{ch}, s_c) = (1, 1)$ is a Nash equilibrium for the range $\bar{c} \leq c \leq n\tilde{c}$, where $\bar{c}$ depends on the number of players n. This means that for $\bar{c}(n-k+1) \leq c < \bar{c}(n-k+2)$, the system tips in the case of $n-k+2$ players, but does not tip in the case of $n-k+1$

players. This has an important consequence for coalition formation. If an economic agent considers to leave a coalition of size k, so that the number of players effectively increases from $n-k+1$ to $n-k+2$, this agent has to realize that besides the usual effects on welfare of leaving the coalition, the system will tip with an overall negative effect on welfare.

In general, we have to analyse what the incentives are to join a coalition of size k, and what the incentives are to leave a coalition of size k. The literature on international environmental agreements (Hoel, 1992, Carraro and Siniscalco, 1993, Barrett, 1994, Finus, 2003, Karp and Simon, 2013), which is based on the literature on cartel theory (d'Aspremont et al., 1983), usually concludes that cooperation is hard to sustain, because free-rider incentives dominate the incentives to cooperate. However, in the presence of a tipping point there are additional incentives to cooperate, as we have just seen. A coalition of size k is called internally stable, if an economic agent has higher welfare as a member of the coalition than as an outsider to the coalition of size $k-1$, so that this economic agent does not have an incentive to leave the coalition. A coalition of size k is called externally stable, if an economic agent has higher welfare as an outsider to this coalition than as a member of the coalition of size $k+1$, so that this economic agent does not have an incentive to join the coalition. The previous literature on international environmental agreements shows that the size of the stable coalition (i.e., both internally and externally stable) is usually small: $k = 2$ or $k = 3$. The question is what happens in the presence of a tipping point.

The results for our model above are presented in Figure 5. We have fixed the number of economic agents ($n = 10$) and the critical stock ($s_c = 1$), but we have varied the parameters b ($0 < b \leq 1$) and c ($0 < c \leq 0.5$). The dashed line $\hat{c}(b)$ indicates for each b the level of c below which it is not optimal anymore to prevent tipping, from a full-cooperative perspective. Furthermore, note that $\tilde{c} = 1/(2s_c^2) = 0.5$, so that the tipping point $(A_{ch}, s_c) = (b, 1)$ is not optimal anymore for $c > 0.5$ (see the section on full cooperation above). Therefore we restrict ourselves to the range $0 < c \leq 0.5$.

The area S_1 indicates the area in the parameter space where the tipping point $(A_{ch}, s_c) = (b, 1)$ is a Nash equilibrium. The lower border of the area S_1 represents $\bar{c}(n, b)$, where the tipping point loses the Nash property (see the section on Nash equilibria above). This curve is upward sloping because for larger b, the slopes of the lines in Figure 1 are steeper, so that the shift in the stock s is smaller and the incentive to deviate is therefore larger. Below the lower border of the area S_1, the system will tip unless some economic agents cooperate. The areas S_k indicate the areas in the parameter space where the coalition of size k is stable, and where the tipping point is a Nash equilibrium for the game between the coalition and the remaining $(n-k)$ economic agents. Two mechanisms are now at work here. One is that the economic agents have to consider the usual trade-off between the free-rider benefits and the benefits of cooperation. The other one is that the economic agents have to consider the possibility that the system tips if the level of cooperation is not sufficiently high. Figure 5 shows that more cooperation is needed to stay in the tipping point when c becomes smaller and b becomes larger. The reason is that the cost of tipping becomes lower, so that more cooperation is needed to prevent tipping. Figure 5 also shows that the coalitions of size $k = 10$ and size $k = 9$ are not stable. Apparently the incentive to free-ride in these large coalitions is higher than the incentive to cooperate, even though the system will tip when an economic agent leaves the coalition. However, coalitions up to size $k = 8$ are stable, and prevent tipping in the corresponding areas of the parameter space. It also shows that, in the presence of a tipping point, the size of the stable coalition can be larger than $k = 2$ or $k = 3$, which is the usual grim result for the size of the stable coalition.

Figure 5 is quite similar to Figure 2 in Barrett (2013). Both figures show a large area in the parameter space where tipping is avoided in a Nash equilibrium, so that coordination on this Nash equilibrium suffices to avoid a big loss. Both figures also show a smaller area where it is not optimal to avoid tipping, and an area in between where cooperation is needed to avoid tipping. The difference is that this paper shows that stable partial cooperation solves the problem to a large extent, whereas the previous paper does not have very positive results on the possibility of cooperation. Barrett (2013) uses a static public good game on emission reduction, with a fixed loss in case a certain threshold is not reached, but this is not the reason for the difference. The previous paper models cooperation as a treaty in which the remaining coalition of $(n - 1)$ economic agents can commit to some level of emission reduction in case an agent withdraws from the treaty. If optimal behaviour of this coalition reaches the threshold but forces the outsider to reduce more than in the treaty, the treaty will be stable. This happens precisely in the area where coordination suffices, so that the treaty just serves as a coordinating device. If optimal behaviour of this coalition reaches the threshold but allows the outsider to reduce less than in the treaty, Barrett (2013) requires that this outcome is a Nash equilibrium, which can only happen in that same area again. It follows that outcomes in the middle area of Figure 2 in Barrett (2013) only occur, if the optimal behaviour of the remaining coalition of $(n-1)$ economic agents does not reach the threshold. In this case, it proves that the old grim story applies again in the sense that very little can be achieved by cooperation. The conclusion is that coordination is the only relevant answer to the problem. However, this paper shows that stable partial cooperation, without future commitments, can respect the threshold in a substantial part of that middle area, in a game on emission. This implies that a combination of partial cooperation and coordination between the coalition and the outsiders can solve the problem in a very large area of the parameter space.

In general, we can conclude from Figure 5 that the area where the Nash equilibrium prevents tipping is large. However, when c becomes small and b becomes large, cooperation is needed to prevent tipping. Cooperation needs to be stable in the sense that a member of the coalition does not have an incentive to leave the coalition. Figure 5 shows that this stability requirement restricts the possibility to prevent tipping by cooperation. Coalitions can only be stable up to size $k = 8$, and this result leaves an area in the parameter space where tipping cannot be prevented, although welfare could be improved. Fortunately, in this area b is large, so that the shift in the stock s is small, and c is small, so that the cost of tipping is low. Large costs of tipping are controlled by individual rationality or by stable partial cooperation. This shows that tipping points are not only bad news.

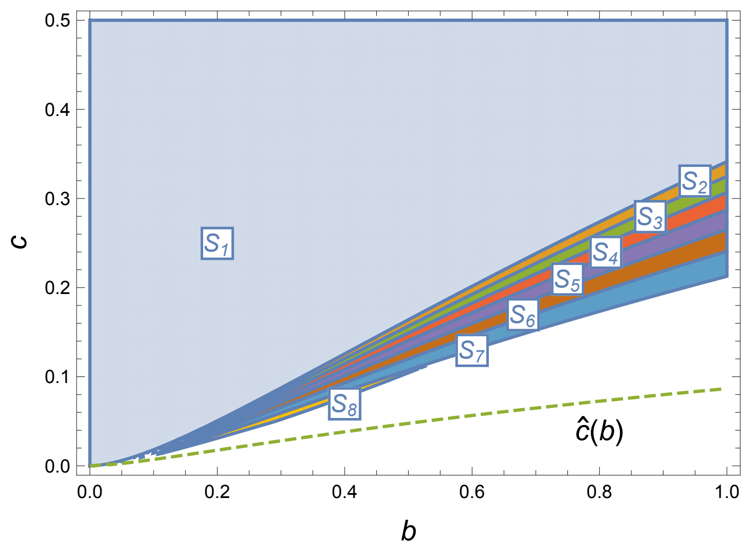

Figure 5: Tipping game.

In conclusion, in case the tipping point is not a Nash equilibrium, tipping can be prevented by stable partial cooperation, but not always and the first-best can usually not be achieved. If the regime shift is modest and if the damage does not weigh much in welfare, tipping will occur but at the same time the consequences are not severe.

2.4 The inverse tipping game

We will now consider what we call the inverse tipping game. This is the situation where the conditions of the system are bad, but where it is possible to tip back to the good conditions.

The inverse tipping game is presented in Figure 2. The initial stock of the system s_0 is high, for example $s_0 = 2$, but a small decrease of the total input below A_{cl} shifts the equilibrium level of the stock s to a much lower level. The situation differs from the tipping game above in the sense that it is now favourable to tip. Therefore, we assume that in this case tipping actually occurs in the tipping point (A_{cl}, s_c), so that the system ends up at $(A_{cl}, s_c - 1/b)$. The situation is now basically the same as the tipping game above and the whole analysis runs parallel, with tipping point $(A_{cl}, s_c - 1/b) = (b-1, 1-1/b)$. Note that b cannot be smaller than 1, because then the lower tipping point $(A_{cl}, s_c) = (b-1, 1)$ disappears below the s-axis and tipping back is not possible. In this case, tipping to the bad conditions is physically irreversible.

The results for this game are presented in Figure 6. The dashed line $\hat{c}(b)$ indicates for each b the level of c below which it is not optimal anymore to tip back in the tipping point $(A_{cl}, s_c) = (b-1, 1)$, from a full-cooperative perspective. For values of b close to 1, the window for tipping back becomes very small and costly to reach, so that in this case tipping back will not be realised. The area S_1 indicates the area in the parameter space where the point $(A_{ch}, s_c - 1/b)$ that results from tipping back is a Nash equilibrium. The areas S_k indicate the areas in the parameter space where the coalition of size k is stable, and where the Nash equilibrium for the game between the coalition and the remaining $(n-k)$ economic agents induces tipping back. Similarly to the tipping game above, Figure 6 shows that the coalitions of size $k = 10$ and size $k = 9$ are not stable. Apparently it is not worthwhile for an economic agent to give up the free-rider benefits and to stay in the coalition in order to reap the benefits of cooperation and to induce tipping back, for the values of c and b below the lower border of the areas S_8 and S_7. A low value of c means that the costs of a large stock s are not sufficiently high to give up the free-rider benefits and to join the coalition in order to induce tipping back. Although it is physically possible in this model to reverse tipping by sufficiently lowering the total input A, it will not happen in this case because the level of cooperation that is needed to tip back is not stable. The system is physically reversible but socially irreversible. However, coalitions up to size $k = 8$ are stable, and induce tipping back in the corresponding areas of the parameter space. It also shows again that the size of the stable coalition can be larger than $k = 2$ or $k = 3$, in the presence of a tipping point.

As in the tipping game above, the stability requirement does not allow to have full cooperation and tip back whenever it is optimal to do so. Stable cooperation can in certain situations help to get out of the bad conditions, but not always. The reason can be that the damage does not weigh much in welfare, so that the consequences are not severe, but the reason can also be that the window for tipping back is very small. The last point means that a system that is close to being physically irreversible may prove to be socially irreversible. In the sections below we will show how the insights from this simple tipping game extend to the lake system, a fully dynamical tipping game with convex-concave dynamics. First, however, we will attend briefly to the issue of uncertainty.

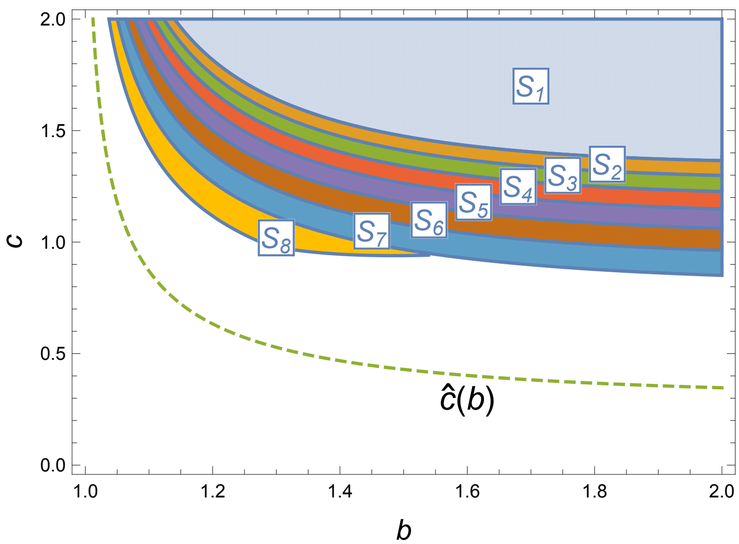

Figure 6: Inverse tipping game.

2.5 Uncertain tipping points

In this paper, we assume that the location of the tipping point (A_c, s_c) (we omit the subscript h here) is known but it is, of course, usually surrounded by uncertainty. For example, Rockström et al. (2009) indicate nine so-called planetary boundaries, like the climate tipping point, but each of those have a zone of uncertainty. This implies that we have to consider the situation that the tipping point (A_c, s_c) is located between a lower bound (A_{c1}, s_{c1}) and an upper bound (A_{c2}, s_{c2}), where this zone of uncertainty has a probability distribution. If the probability of the location between A_{c1} and A_{c2} is given by the distribution function $F(A)$, the response in that zone becomes $s = A/b$, with probability $1 - F(A)$, and $s = (A + 1)/b$, with probability $F(A)$. For example, if the uncertainty is uniform in that zone, the distribution function $F(A)$ and the expected response are given by

$$F(A) = \frac{A - A_{c1}}{A_{c2} - A_{c1}}, \tag{2.12}$$

$$\mathbb{E}s = \frac{A_{c2} - A}{A_{c2} - A_{c1}} \frac{A}{b} + \frac{A - A_{c1}}{A_{c2} - A_{c1}} \frac{A + 1}{b}, \quad A_{c1} \leq A \leq A_{c2}. \tag{2.13}$$

If we want to follow the same analysis as in the previous section, we cannot work with the expected response function, because the expected costs in the objective functional, in the range of uncertainty, are given by

$$c\,\mathbb{E}s^2 = \frac{A_{c2} - A}{A_{c2} - A_{c1}} c \left(\frac{A}{b}\right)^2 + \frac{A - A_{c1}}{A_{c2} - A_{c1}} c \left(\frac{A+1}{b}\right)^2, \; A_{c1} \leq A \leq A_{c2}, \quad (2.14)$$

and $\mathbb{E}s^2$ is not equal to $(\mathbb{E}s)^2$. However, if we replace $\mathbb{E}s^2$ by s^2 in equation (2.14), and solve for A as a function g of s, we create a certainty-equivalent response function in the range of uncertainty. This allows us to solve the problem with an uncertain tipping point in the same way as the problem in the previous sections. The certainty-equivalent response function $f(s)$ becomes

$$\begin{aligned} f(s) &= bs, && 0 \leq s \leq s_{c1}, \\ f(s) &= g(s), && s_{c1} < s \leq s_{c2} + \frac{1}{b}, \\ f(s) &= bs - 1, && s > s_{c2} + \frac{1}{b}, \end{aligned} \quad (2.15)$$

where $g(s)$ is the solution of equation (2.14) with $\mathbb{E}s^2$ replaced by s^2. The uncertainty about the location of the tipping point transforms the function f into a concave-convex response function, and the jump disappears. The resulting situation is depicted in Figure 7. The width of the zone of uncertainty, i.e. the distance between the lower bound (A_{c1}, s_{c1}) and the upper bound (A_{c2}, s_{c2}) of the tipping point, determines the shape of this certainty-equivalent response function.

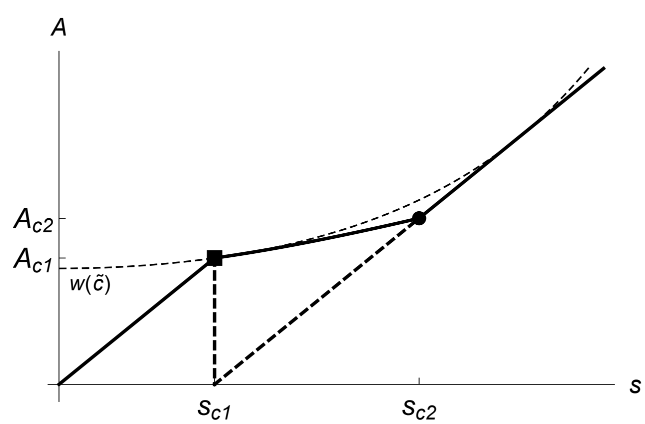

Figure 7: A small amount of uncertainty does not change the optimum.

If the width of the zone of uncertainty is small, the results are the same as in the absence of uncertainty. We can show this by using the example from the previous sections: $n = 2$, $b = 1$ and $s_{c1} = 1$, so that $A_{c1} = 1$. The certainty-equivalent response function $f(s)$ becomes

$$
\begin{aligned}
f(s) &= s, \qquad 0 \leq s \leq 1, && (2.16)\\
f(s) &= \frac{1+\sqrt{1+4(A_{c2}+1)(1+(A_{c2}-1)s^2)}}{2(A_{c2}+1)}, \qquad 1 < s \leq s_{c2}+1, &&\\
f(s) &= s-1, \quad s > s_{c2}+1. &&
\end{aligned}
$$

Similar to the analysis in the previous sections, we can characterise the stationary points on the extended curve through the middle segment of the response function $f(s)$ by

$$A = \frac{f'(s)}{2cs} = f(s). \tag{2.17}$$

It is straightforward to calculate that for $c = (A_{c2}-1)/(2A_{c2}+1)$, the stationary point is the tipping point $(A_{c1}, s_{c1}) = (1, 1)$. It follows that it does not pay to move above the tipping point into the zone of uncertainty for $(A_{c2}-1)/(2A_{c2}+1) \leq c$. In the previous sections, we have seen that in the absence of uncertainty, 0.5 is the largest value of c for which the tipping point $(A_{c1}, s_{c1}) = (1, 1)$ is first best, with the inputs $a_1 = a_2 = 0.5$ and total welfare -2.3863. Furthermore, in the absence of uncertainty, the tipping point is the optimal solution in the range $0.0867 \leq c \leq 0.5$, because for $c = 0.0867$ the same level of total welfare -1.5597 is realised in the tipping point $(A_{c1}, s_{c1}) = (1, 1)$ and in the point on the line segment l_2 (Figure 3), with the inputs $a_1 = a_2 = 0.97646$ and the stock $s = 2.9529$. It follows that in the presence of uncertainty, the conclusion does not change if the upper bound $A_{c2} \leq 1.3147$, so that $(A_{c2}-1)/(2A_{c2}+1) \leq 0.0876$. For $A_{c2} = 1.3147$, the curve through the middle segment of the response function $f(s)$ is tangent to the iso-welfare curve in the tipping point (Figure 7). For $A_{c2} > 1.3147$, the curve through the middle segment of the response function $f(s)$ is steeper, and total welfare increases by moving into the zone of uncertainty (Figure 8).

We can characterise the candidate symmetric Nash equilibria on the curve through the middle segment of the response function $f(s)$ by

$$A = \frac{f'(s)}{cs} = f(s). \tag{2.18}$$

It is easy to see now that for $c = 2(A_{c2}-1)/(2A_{c2}+1)$, the tipping point $(A_{c1}, s_{c1}) = (1, 1)$ is a candidate Nash equilibrium. We have to investigate the incentive to deviate. We fix the input of economic agent 1 at $a_1 = 0.5$, and investigate the best response of economic agent 2. In the previous sections, we have seen that in the absence of uncertainty, the tipping point is a Nash equilibrium in the range $0.13971 \leq c \leq 1$, because for $c = 0.13971$ the same level of individual welfare -0.83286 for economic agent 2 is realised in the tipping point $(A_{c1}, s_{c1}) = (1, 1)$ and in the point on the line segment l_2, with

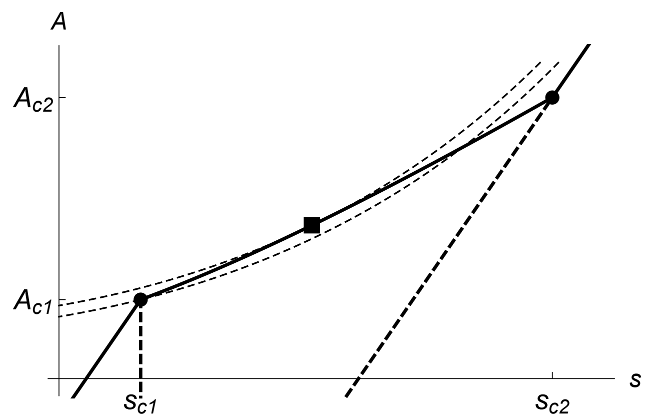

Figure 8: Under intermediate uncertainty, it is optimal to allow the possibility of tipping.

the input $a_2 = 1.285$ and the stock $s = 2.785$, so that an individual economic agent only has the incentive to deviate for $c < 0.13971$. It follows that in the presence of uncertainty, the conclusion does not change if the upper bound $A_{c2} \leq 1.2436$, so that $2(A_{c2} - 1)/(2A_{c2} + 1) \leq 0.13971$. For $A_{c2} = 1.2436$, the curve through the middle segment of the response function $f(s)$ is tangent to the iso-welfare curve for economic agent 2 in the tipping point. The conclusion is that the results remain the same as in the absence of uncertainty, if the width of the zone of uncertainty is small (in the example $A_{c1} = 1$ and $A_{c2} \leq 1.2436$). The tipping point $(A_{c1}, s_{c1}) = (1, 1)$ is first-best but also a Nash equilibrium for $0.13971 \leq c \leq 0.5$.

For a larger upper bound A_{c2}, we do not get the same result, but we can get a similar result. For example, if $A_{c2} = 2$ and $c = 0.12$, the optimal point is $(A^*, s^*) = (1.2648, 1.5919)$ on the middle segment of the response function $f(s)$, with total welfare -1.5247 and a probability of 0.2648 of tipping. The candidate Nash equilibria are the points $(A^{n1}, s^{n1}) = (1.7521, 2.5411)$, with individual welfare -0.9072 and a probability of 0.7521 of tipping, and $(A^{n2}, s^{n2}) = (2.4297, 3.4297)$, with individual welfare -1.2169, on the middle and the right segment of the response function $f(s)$, respectively. An individual economic agent does not have the incentive to deviate from the good Nash equilibrium $(A^{n1}, s^{n1}) = (1.7521, 2.5411)$ for these parameter values, because the best response would be the input $a_i = 1.3085$, which yields the stock $s = 3.1845$ with lower welfare -0.9480. Figure 9 depicts the outcome. In this case, the economic agents can still coordinate on the good Nash equilibrium, although they cannot achieve the first best in a Nash equilibrium. This implies that even in the case of large uncertainty on the location of the tipping point, coordination on a good Nash equilibrium is still possible and worthwhile, which is not in line with proposition 6 in Barrett (2013). However, when decreasing the value of c, at some point the good Nash equilibrium loses the Nash property, but this is the same type of result as in the absence of uncertainty. In the sequel, we will focus on the certainty case but extend the analysis to the lake system, a fully dynamical tipping game with convex-concave dynamics.

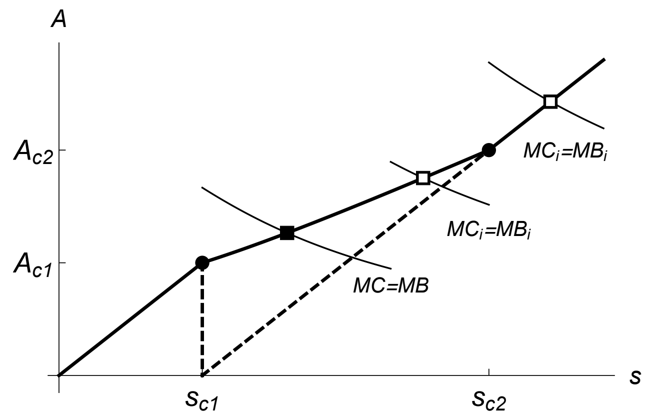

Figure 9: Even under uncertainty, coordination on a good Nash equilibrium is possible.

3 The lake game

In the previous sections we assumed constant inputs a_i, $i = 1, 2, ..., n$, and sufficient time for the system to reach the steady state. This is the most simple representation of a tipping point, and it allowed us to focus the analysis on the total input A and the resulting steady state s. In the sequel, we extend the analysis to a fully dynamical model. We use the well-known lake model (Carpenter et al., 1999, Brock and Starrett, 2003, Mäler et al., 2003, Wagener, 2003) where the response function $f(s)$ is not, as in the sections above, linear with a downward jump, but convex-concave instead. The lake model is given by

$$\dot{s}(t) = A(t) - f(s(t)), \quad A = \sum_{i=1}^{n} a_i, \quad s(0) = s_0, \tag{3.1}$$

$$f(s) = bs - \frac{s^2}{1+s^2}, \tag{3.2}$$

where a_i denotes the loading of phosphorus on the lake by the economic agent i, and s denotes the accumulated stock of phosphorus in the water of the lake that causes the loss of ecosystem services. It is easy to show that for $0.5 < b < 3\sqrt{3}/8$, the curve $f(s)$ has the same type of tipping points — at the local maximum and at the local minimum of f — as the simple model in the previous sections. Note, however, that in this model tipping to the bad conditions is physically irreversible for $b < 0.5$, whereas this irreversibility in the simple tipping game above occurred for $b < 1$. Otherwise, the parameter b is an indicator for the steepness of the curve and in that sense it plays a similar role in the two models. The response function is typical in ecological models, and therefore the lake model can be seen as a metaphor for many problems with tipping points. We model the trade-off between the benefits of loading and the loss of ecosystem services by maximising the infinite-horizon discounted welfare indicators

$$\max_{a_i(.)} \int_0^\infty \left[\ln a_i(t) - cs^2(t)\right] e^{-rt} dt, \quad i = 1, 2, ..., n, \tag{3.3}$$

where the parameter c weighs the benefits and costs.

The model has turned into a so-called differential game (Basar and Olsder, 1982). The full-cooperative case is a straightforward optimal-control problem, but in case of a Nash equilibrium we have to distinguish the open-loop Nash equilibrium, where the inputs a_i are only a function of time t, and the feedback Nash equilibrium, where the inputs a_i are also a function of the state s and decisions are taken when the state s is realised. Mäler et al. (2003) derive the open-loop Nash equilibrium for the lake game, and Kossioris et al. (2008) and Dockner and Wagener (2014) investigate the feedback Nash equilibria for the lake game. However, we consider coalition formation and then a choice has to be made regarding the timing. Either the economic agents decide whether or not they want to be a member of the coalition before the game unravels, or the economic agents consider their membership of the coalition at each moment in time. The second approach is consistent with the feedback Nash equilibrium but it is very difficult to solve. We leave this for further research. The first approach is a natural extension of the simple tipping game that we analysed in the previous section, and will be pursued in the sequel. That is, in this paper, we use the open-loop Nash equilibrium concept, preceded by the membership game where the economic agents choose to become a member of the coalition or not.

3.1 Full cooperation and Nash equilibria

In the full-cooperative case we maximize the sum of the welfare indicators. The Hamiltonian becomes

$$H = \sum_{i=1}^{n} \ln a_i - ncs^2 + \lambda(A - f(s)), \tag{3.4}$$

which yields the necessary conditions

$$A = \sum_{i=1}^{n} a_i = -\frac{n}{\lambda}, \tag{3.5}$$

$$\dot{A}(t) = -\left[r + f'(s(t))\right] A(t) + 2cs(t)A^2(t). \tag{3.6}$$

Wagener (2003) shows that there is again a value $\hat{c}$ of c, which depends on the initial state $s(0)$, such that the lake system under optimal total loading of phosphorus converges to a steady state in the low-phosphorus regime of the lake for $c \geq \hat{c}$, but converges to a steady state in the high-phosphorus regime for $c < \hat{c}$.

For the symmetric non-cooperative Nash equilibrium, the Hamiltonians become

$$H_i = \ln a_i - cs^2 + \lambda_i(A - f(s)), \quad i = 1, 2, ..., n, \tag{3.7}$$

which yields the necessary conditions

$$A = \sum_{i=1}^{n} a_i = -\frac{n}{\lambda_i}, \tag{3.8}$$

$$\dot{A}(t) = -\left[r + f'(s(t))\right] A(t) + 2\frac{c}{n}s(t)A^2(t). \tag{3.9}$$

Mäler et al. (2003) show, for $b = 0.6$, $c = 1$, $n = 2$ and $r = 0.03$, that the necessary conditions for the full-cooperative outcome have one steady state in the low-phosphorus regime, but in case of a Nash equilibrium there are two steady states, one in the low- and one in the high-phosphorus regime. Depending on the initial conditions (see Grass et al., 2017), the economic agents can end up in the low- of in the high-phosphorus Nash equilibrium. The last situation can be seen as a pollution trap. When the agents cooperate, they move to the low-phosphorus regime.

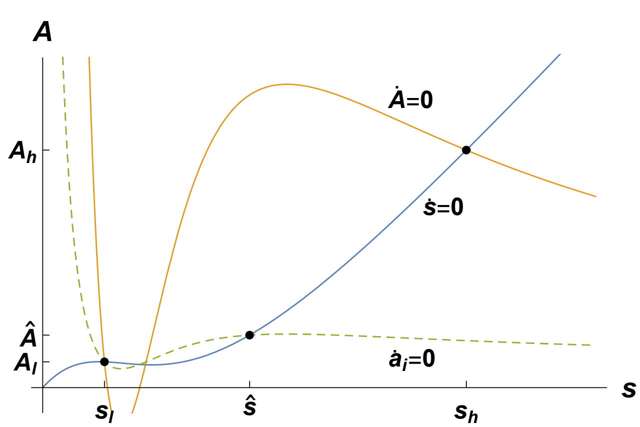

Figure 10: Isoclines.

As for the simple tipping game above, we have to be able to check whether the Nash property holds in a candidate Nash equilibrium. The steady states are the intersection points of the $\dot{s} = 0$ and $\dot{A} = 0$ isoclines. Mäler et al. (2003) show that these isoclines either have one or three intersection points. If there are three intersection points, the middle one is unstable and cannot be a Nash equilibrium, but the other two intersection points are the steady states of the candidate Nash equilibria (Figure 10). We refer to these candidate Nash equilibria by their steady states (A_l, s_l) and (A_h, s_h), where l means low and h means high. The question is again whether an individual economic agent i has an incentive to deviate from the candidate Nash equilibrium (A_l, s_l) in the low-phosphorus regime. Suppose that the other $n-1$ economic agents stick to their phosphorus loadings A_l/n. The necessary conditions for the optimisation of economic agent i lead to

$$A = \sum_{i=1}^{n} a_i = a_i - \frac{n-1}{n} A_l, \tag{3.10}$$

$$\dot{a}_i(t) = -\left[r + f'(s(t))\right] a_i(t) + 2cs(t)a_i^2(t). \tag{3.11}$$

Adding the constant $(n-1)A_l/n$ to the $\dot{a}_i = 0$ isocline yields a curve that has again three intersection points with the $\dot{s} = 0$ isocline. These indicate the candidate optimal trajectories for economic agent i (Figure 10). The middle one is a minimum and the other two are local maxima. (A_l, s_l) denotes one of these local maxima, the second one is denoted by the steady state $(\hat{A}, \hat{s})$. The question is which one is the best for economic agent i. If the trajectory leading to (A_l, s_l) is the best, economic agent i does not have an incentive to deviate, so that (A_l, s_l) is a Nash equilibrium, but if the trajectory leading to $(\hat{A}, \hat{s})$ is the best, economic agent i has an incentive to deviate and (A_l, s_l) is not a Nash equilibrium. The trajectories are presented in Figure 11.

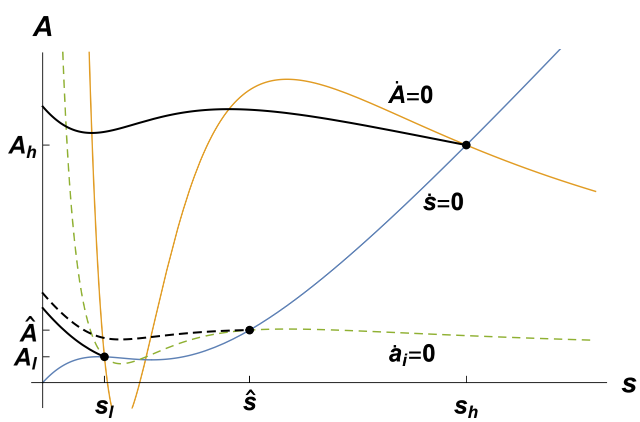

Figure 11: Trajectories.

In case (A_l, s_l) is not a Nash equilibrium, the question is whether stable partial cooperation can keep the lake system in the low-phosphorus regime. In order to analyse this question, we need to solve for coalition formation in this differential game.

3.2 Coalition formation

We assume again that one coalition of k cooperating economic agents can be formed, and that the other $(n-k)$ economic agents continue to play individually. The respective Hamiltonians become

$$H = \sum_{i=1}^{k} \ln a_i - kcs^2 + \lambda(A - f(s)), \tag{3.12}$$

$$H_i = \ln a_i - cs^2 + \lambda_i(A - f(s)), \quad i = k+1, ..., n, \tag{3.13}$$

which yield the necessary conditions

$$a_i = \frac{-1}{\lambda}, \quad i = 1, 2, ..., k, \tag{3.14}$$

$$a_i = \frac{-1}{\lambda_i}, \quad i = k+1, ..., n, \tag{3.15}$$

$$\dot{\lambda}(t) = [r + f'(s(t))]\,\lambda(t) + 2kcs(t), \tag{3.16}$$

$$\dot{\lambda}_i(t) = [r + f'(s(t))]\,\lambda_i(t) + 2cs(t), \quad i = k+1, ..., n, \tag{3.17}$$

so that $\lambda/k = \lambda_i$ and the necessary conditions become

$$A = \sum_{i=1}^{n} a_i = -\frac{k}{\lambda} - \frac{n-k}{\lambda_i} = -\frac{n-k+1}{\lambda_i}, \tag{3.18}$$

$$\dot{A}(t) = -[r + f'(s(t))]\,A(t) + 2\frac{c}{n-k+1}s(t)A^2(t). \tag{3.19}$$

Note that these are the same conditions as for the Nash equilibrium with $(n-k+1)$ economic agents. Again the coalition effectively operates as one individual economic agent.

We use numerical algorithms to calculate the resulting paths and welfare values of the optimal-control problems and the open-loop Nash equilibria, for different values of the parameters. With the optimal-control algorithm, we can cover the cases of full cooperation and a possible deviation from a Nash equilibrium. With the Nash-equilibrium algorithm, we can cover the Nash equilibria, but also the cases with partial cooperation. The analysis requires a substantial numerical effort, but all these cases can be solved for the lake game.

To illustrate the magnitude of the effects, we take the values $b = 0.60$, $r = 0.03$, $n = 2$ from Mäler et al. (2003), with $s_0 = 0$. For $c < 0.6648$, the point (A_l, s_l), which is close to the tipping point, proves not to be a Nash equilibrium, and the system ends up in the high-phosphorus regime at the point (A_h, s_h). However, if the two economic agents form a coalition, the steady state in the low-phosphorus regime is first-best for $c \geq 0.4346$. This implies that the system will tip for $0.4346 \leq c < 0.6648$ in the absence of cooperation, but it will remain in the low-phosphorus regime if the two economic agents form a coalition. Note that in the two-player game the coalition is always stable. The welfare implications are large. For instance, if $c = 0.7$, cooperation is not very beneficial. The point (A_l, s_l) is the steady state of a Nash equilibrium, and individual welfare is -100.3824, just below the welfare of -100.2386 which is obtained if the economic agents cooperate. This is typical for the lake game. Cooperation cannot exactly be sustained by a Nash equilibrium, but the loss of not cooperating is minor. However, if $c = 0.6$, the corresponding point (A_l, s_l) fails to be a Nash equilibrium, and the individual welfare drops to -120.1273 in the Nash equilibrium that ends up in the high-phosphorus regime. This is much lower than the individual welfare of -99.8047 under cooperation, for this value of c: it amounts to a loss of 20%.

Suppose that the values of the parameters are such that full cooperation keeps the lake in the low-phosphorus regime. We can show when a Nash equilibrium in the low-phosphorus regime exists and if it does not exist, whether stable partial cooperation can keep the lake in the low-phosphorus regime. In our model, the coalition of k economic agents behaves as just another player in the interactions with the $n - k$ outsiders. Therefore, there are two conditions that have to be fulfilled, for a given set of parameter values. First, the lake game with $n - k + 1$ economic agents must have a Nash equilibrium in the low-phosphorus regime. Second, a coalition member does not have an incentive to leave the coalition. If a coalition member leaves the coalition, the system may tip to the high-phosphorus regime, because the Nash equilibrium may cease to exist when the lake game is played with a larger number of agents. This may prevent leaving the coalition but if the free-rider benefits are high, the coalition member may leave the coalition anyway, even if the system tips. It is important to note that if a larger coalition is needed in order to have a Nash equilibrium in the low-phosphorus regime, the free-rider benefits increase and this may destabilize the coalition.

As it is very complex to present the results for the dynamical tipping game in the (b, c) parameter space as in Figure 5, we present the results as a cut through the parameter space for a fixed $b = 0.60$. Figure 12 shows the results for $n = 10$, $r = 0.03$, $s_0 = 0$ and $0.01 \leq c \leq 2.0$. As in the tipping game in the previous section, the coalitions of size $k = 10$ and $k = 9$ are not stable, but the coalitions up to size $k = 8$ can be stable.

For $c \geq 1.69$, the steady state in the low-phosphorus regime (A_l, s_l) is a Nash equilibrium of the game between the 10 economic agents, and the

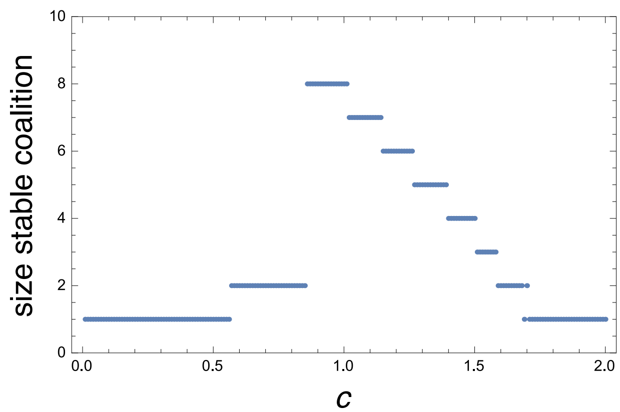

Figure 12: The dynamical tipping game for $b = 0.60$.

welfare of the economic agents in this equilibrium is only just below the first-best outcome. For $1.58 \leq c < 1.69$, the steady state (A_l, s_l) does not have the Nash property, but a Nash equilibrium in the low-phosphorus regime exists for the game between a coalition of size $k = 2$ and the $(n-2)$ outsiders. This provides an incentive for two economic agents to form a coalition. Although their welfare will be lower than that of the other 8 economic agents, it proves to be preferable to the individual welfare in the Nash equilibrium (A_h, s_h).

Figure 12 shows that lowering c further, the level of cooperation increases further in order to prevent shifting to the low-phosphorus regime. This mechanism is effective for $0.88 < c$, leading to a stable coalition up to size $k = 8$. It breaks down for lower values of c: if the costs of tipping are too low, the benefits of not tipping cannot outweigh the free-rider benefits in this stage anymore. It follows that tipping is avoided for $0.88 < c$, either in the Nash equilibrium or by stable partial cooperation. One can say that the "economic resilience" is substantial. Note that it would be beneficial to avoid tipping for some range of lower values of c, from a full-cooperative perspective, but the stability requirement for cooperation prevents this, as in Figure 5. The coalitions of size $k = 10$ and $k = 9$ are not stable.

For $c < 0.88$, the lake system tips, either because the coalitions of size $k = 10$ and $k = 9$ are not stable or because the costs of a high stock s are too low, so that it is simply optimal to tip. In this case, the old grim picture of international environmental agreements applies. The size of the stable coalition is very small, and in this model equal to either 1 or 2. For $0.58 \leq c < 0.88$, there is a still a coalition of size $k = 2$ but for $c < 0.58$, there is no stable partial cooperation.

It is interesting to see what happens if we vary the parameter b. Figure 13 shows the results for a cut through the (b, c) parameter space for a fixed $b = 0.52$. The pattern is very similar to the pattern in Figure 12. The differences are that the lake system does not tip in a larger range of values for c, and that tipping is prevented in the Nash equilibrium in a larger range of values for c. The reason is the same as for the tipping game in the previous section (Figure 5). A smaller value of b means that the curve is less steep, so that the shift in the stock s is larger and the incentive to deviate is smaller. The general conclusion is again that if the regime shift is modest (i.e., b is large) and if the damage does not weigh much in welfare (i.e., c is small), tipping of the lake will occur, but at the same time the consequences are not severe. In the next section we turn to what we call the inverse lake game.

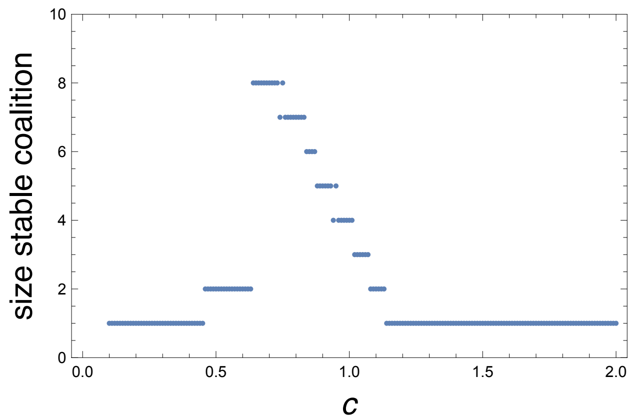

Figure 13: The dynamical tipping game for $b = 0.52$.

3.3 The inverse lake game

When tipping has occurred earlier, the question is whether it is possible to induce tipping back to the low-phosphorus regime of the lake. The Nash equilibrium may achieve this for certain values of the parameters, and otherwise the question is whether stable partial cooperation may induce tipping back to the favourable conditions of the lake.

As in the lake game above, we present the results as a cut through the (b, c) parameter space for a fixed $b = 0.60$, with $n = 10$, $r = 0.03$ and the high initial condition $s_0 = 1.75$. Figure 14 shows the results for $0.1 \leq c \leq 7.0$. The pattern is again very similar to the patterns above. The coalition of size $k = 10$ is not stable, but the coalitions up to size $k = 9$ can be stable. As compared to the simple tipping game in Figure 6, we get a similar pattern as for $b = 1.3$ in that game. As compared to the lake game in Figure 12, the pivotal value of c, for which partial cooperation starts to be needed in order to induce tipping back, is much larger in the inverse lake game: $c < 6.0$. Furthermore, the value of c below which tipping back cannot be induced anymore is substantially larger: $c < 1.5$. This indicates that more cooperation is needed to induce tipping back

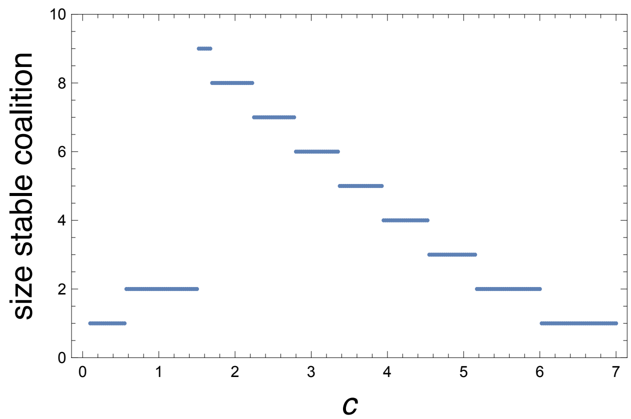

Figure 14: The inverse dynamical tipping game.

than to prevent tipping. Moreover, if tipping occurs, $c < 0.88$, as we have seen in the previous section, and for these values of c tipping back is not possible. It reflects the asymmetry in the tipping game. As the low-phosphorus steady state is close to the tipping point, an economic agent can tip the lake system from the low-phosphorus to the high-phosphorus regime easily, which places a lot of discipline on the economic agents in the Nash equilibrium. On the other hand, the high-phosphorus steady state is far away from the lower tipping point, so that it takes strong coordinated action of a group of economic agents to bring the lake system past the lower tipping point. This also implies that social irreversibility may arise. Since a high level of cooperation is needed to tip back, the incentive to free-ride increases and the stability of the coalition may break down, so that tipping back may be physically but not socially feasible.

In conclusion, Figure 6 shows that for $b = 0.60$ and $c = 2.0$, it requires a stable coalition of size $k = 8$ to lead the users of the lake out of the pollution trap in the high-phosphorus regime. However, when the low-phosphorus regime has been reached, Figure 12 shows that no cooperation is needed to prevent tipping again. The Nash equilibrium will keep the lake system in the low-phosphorus regime. A substantial effort is needed to push the system from "brown" (green in case of a lake) to "green" (blue in case of a lake) environmental conditions, but it is relatively easy to keep the system in a good state.

4 Conclusion

Tipping points are often observed in natural systems. When a natural system tips, it shifts to another domain of attraction, which usually leads to a substantial loss of ecosystem services. Tipping occurs when accumulated emissions from economic activities cross a certain threshold. Full cooperation keeps the natural system in good conditions, unless a very low value is attached to the loss of ecosystem services. Non-cooperative behaviour can keep the natural system in good conditions as well, if the incentive to deviate is suppressed by strong consequences of tipping. In the case non-cooperative behaviour causes tipping anyway, partial cooperation may prevent the shift. However, partial cooperation has to be stable in the sense that a member of the coalition does not have an incentive to leave the coalition and free-ride. Therefore, it is not always possible to prevent tipping by stable partial cooperation.

This paper first presents a simple tipping point model, with constant inputs and linear dynamics with a downward jump. In this model, it is relatively easy to derive all the results. It is shown that stable partial cooperation can indeed prevent tipping, but not in all cases where it is optimal to do so. It is not possible if the value attached to the loss of ecosystem services is low and the shift is small, but this means that the consequences of tipping are fortunately not severe. It is also shown that in case tipping has occurred, stable partial cooperation can induce tipping back to the good conditions, but again not in all cases where it is optimal to do so. This means that it may happen that the system is physically reversible but socially irreversible, because the incentives to free-ride prevent the level of cooperation that is needed to tip back.

Then this paper analyses a tipping point model with time-dependent inputs and convex-concave dynamics, representing the well-known lake system. This system is a metaphor for many natural systems with tipping points. The analysis is more complicated, and requires advanced numerical methods, but it basically leads to the same results. An important policy conclusion is that when the users of the lake system are trapped in a non-cooperative Nash equilibrium with a low level of ecosystem services, stable partial cooperation may lead the users out of this pollution trap. However, this is not always possible, so that social irreversibility of tipping may arise, even if tipping is physically reversible. Furthermore, when a non-cooperative Nash equilibrium with a high level of ecosystem services cannot be sustained, because tipping will occur, stable partial cooperation may prevent the lake from tipping. Generally, a higher level of cooperation is needed to induce tipping back than to prevent tipping.

The usual grim picture in the literature on international environmental agreements is that the stable coalitions are small, especially when the possible gains of cooperation are large. In the presence of a tipping point, this picture is reversed. This paper shows that the stable coalitions can be large, in order to prevent large losses or in order to induce large gains. Furthermore, if the size of the stable coalition cannot be increased anymore, because free-rider benefits become too high, the consequences are not severe, because the losses from tipping are relatively low in this case.

5 References

d'Aspremont, C., A. Jacquemin, J. Gabszewicz, and J. Weymark, On the stability of collusive price leadership, Canadian Journal of Economics 16, 1, 17-25, 1983.

Barrett, S., Self-enforcing international environmental agreements, Oxford Economic Papers 46, 878-894, 1994.

Barrett, S., Climate treaties and approaching catastrophes, Journal of Environmental Economics and Management 66, 2, 235-250, 2013.

Basar, T., and G.J. Olsder, Dynamic Noncooperative Game Theory, Academic Press, New York, 1982.

Biggs, R., Th. Blenckner, C. Folke, L.J. Gordon, A.V. Norström, M. Nyström, and G.D. Peterson, Regime shifts, in A. Hastings and L. Gross (eds.), Encyclopedia in Theoretical Ecology, University of California Press, Berkeley and Los Angeles, 609-616, 2012.

Brock, W.A., and D. Starrett, Managing systems with non-convex positive feedback, Environmental & Resource Economics 26, 4, 575-602, 2003.

Cai, Y., K.L. Judd, and T.S. Lontzek, The social cost of carbon with economic and climate risks, arXiv:1504.06909 preprint, Cornell University, 2015.

Carpenter, S.R., Regime Shifts in Lake Ecosystems: Pattern and Variation, International Ecology Institute, Oldendorf/Luhe, Germany, 2003.

Carpenter, S.R., D. Ludwig, and W.A. Brock, Management of eutrophication for lakes subject to potentially irreversible change, Ecological Applications 9, 3, 751-771, 1999.

Carraro, C., and D. Siniscalco, Strategies for the international protection of the environment, Journal of Public Economics 52, 3, 309-328, 1993.

Crépin, A.-S., Using fast and slow processes to manage resources with thresholds, Environmental & Resource Economics 36, 2, 191-213, 2007.

Dockner, E., and F.O.O. Wagener, Markov perfect Nash equilibria in models with a single capital stock, Economic Theory 56, 3, 585-625, 2014.

Finus, M., Stability and design of international environmental agreements: the case of transboundary pollution, in H. Folmer and T. Tietenberg (eds.), The International Yearbook of Environmental and Resource Economics 2003/2004, Edward Elgar, Cheltenham, UK, 82-158, 2003.

Grass, D., A. Xepapadeas, and A. de Zeeuw, Optimal management of ecosystem services with pollution traps: the lake model revisited, Journal of the Association of Environmental and Resource Economists 4, 4, 1121-1154, 2017.

Heijdra, B., and P. Heijnen, Environmental abatement and the macroeconomy in the presence of ecological thresholds, Environmental & Resource Economics 55, 1, 47-70, 2013.

Hoel, M., International environmental conventions: the case of uniform reduction of emissions, Environmental & Resource Economics 2, 2, 141-159, 1992.

Hughes, T., A. Baird, D. Bellwood, M. Card, S. Connolly, C. Folke, R. Grosberg, O. Hoegh-Guldberg, J. Jackson, J. Kleypas, J. Lough, P. Marshall, M. Nyström, S. Palumbi, J. Panolfi, B. Rosen, and J. Roughgarden, Climate change, human impacts, and the resilience of coral reefs, Science 301, 929-33, 2003.

Karp, L., and L. Simon, Participation games and international environmental agreements: A non-parametric model, Journal of Environmental Economics and Management 65, 2, 326-344, 2013.

Kiseleva, T., and F.O.O. Wagener, Bifurcations of optimal vector fields in the shallow lake model, Journal of Economic Dynamics and Control 34, 5, 825-843, 2010.

Kossioris, G., M. Plexousakis, A. Xepapadeas, A. de Zeeuw, and K.-G. Mäler, Feedback Nash equilibria for non-linear differential games in pollution control, Journal of Economic Dynamics and Control 32, 4, 1312-1331, 2008.

Lemoine, D., and C.P. Traeger, Watch your step: optimal policy in a tipping climate, American Economic Journal: Economic Policy 6, 1, 137-166, 2014.

Lenton, T.M., and J.-C. Ciscar, Integrating tipping points into climate impact assessments, Climate Change 117, 3, 585-597, 2013.

Mäler, K.-G., A. Xepapadeas, and A. de Zeeuw, The economics of shallow lakes, Environmental & Resource Economics 26, 4, 603-624, 2003.

Ploeg, F. van der, and A. de Zeeuw, Climate tipping and economic growth: precautionary capital and the price of carbon, Journal of the European Economic Association 16, 5, 1577–1617, 2018.

Polasky, S., A. de Zeeuw, and F.O.O. Wagener, Optimal management with potential regime shifts, Journal of Environmental Economics and Management 62, 2, 229-240, 2011.

Rockström, J., W. Steffen, K. Noone, Å. Persson, F.S. Chapin III, E.F. Lambin, T.M. Lenton, M. Scheffer, C. Folke, H.J. Schellnhuber, B. Nykvist, C.A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P.K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R.W. Corell, V.J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J.A. Foley, A safe operating space for humanity, Nature 461, 7263, 472–475, 2009.

Scheffer, M., Ecology of Shallow Lakes, Chapman and Hall, London, 1997.

Scheffer, M., S.R. Carpenter, J.A. Foley, C. Folke and B. Walker, Catastrophic shifts in ecosystems, Nature 413, 6856, 591-596, 2001.

Wagener, F.O.O., Skiba points and heteroclinic bifurcations, with applications to the shallow lake system, Journal of Economic Dynamics and Control 27, 9, 1533-1561, 2003.