



HAL
open science

How Do Coalitions Break Down? An Alternative View

Raouf Boucekkine, Carmen Camacho, Weihua Ruan, Benteng Zou

► **To cite this version:**

Raouf Boucekkine, Carmen Camacho, Weihua Ruan, Benteng Zou. How Do Coalitions Break Down? An Alternative View. *Dynamic Games and Applications*, 2024, 14 (1), pp.157-194. 10.1007/s13235-024-00559-6 . hal-04488481

HAL Id: hal-04488481

<https://rennes-sb.hal.science/hal-04488481v1>

Submitted on 8 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

How do coalitions break down? An alternative view *

Raouf Boucekkine

raouf.boucekkine@rennes-sb.com

Rennes School of Business

Carmen Camacho

Paris School of Economics

Weihua Ruan

Purdue University Northwest

Benteng Zou

University of Luxembourg

Research Article

Keywords: Coalition splitting, environmental agreements, constitutional vs technological heterogeneity, differential games, multistage optimal control

Posted Date: March 8th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2643650/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Dynamic Games and Applications on February 21st, 2024. See the published version at <https://doi.org/10.1007/s13235-024-00559-6>.

How do coalitions break down? An alternative view*

Raouf Boucekkine[†] Carmen Camacho[‡] Weihua Ruan[§] Benteng Zou[¶]

Abstract

We propose an alternative dynamic theory of coalition breakdown. Motivated by recent coalition splitting events through unilateral countries' withdrawals, we assume that: i) the payoff sharing rule within coalitions is not necessarily set according to any optimality and/or stability criterion, and, ii) players initially behave *as if* the coalition will last forever. If the sharing rule is non-negotiable or if renegotiation is very costly, compliance to these rules may become unbearable for a given member because the rule, being too rigid, would make exit preferable as time passes. We examine this endogenous exit problem in the case of time-invariant sharing rules. Assuming a Nash non-cooperative game after (potential) splitting where players play Markovian, we characterize the solutions of the endogenous exit problem in a linear-quadratic frame with endogenous splitting time. We find that splitting countries are precisely those which use to benefit the most from the coalition. Suitable sharing rules should be used to prevent coalition splitting. When initial pollution is high, all shares should be low enough and none of the players should detain a payoff share larger than 1/2. If initial pollution is small, we provide with an explicit interval for the sharing rule values preventing the collapse of the coalition.

Keywords: Coalition splitting; environmental agreements; constitutional vs technological heterogeneity; differential games; multistage optimal control.

JEL classification: C61, C73, D71.

*In memory of Ngo Van Long. Financial support from the French National Agency (grant ANR-17-EURE-0001) is gratefully acknowledged by Carmen Camacho.

[†]Corresponding author. Centre for Unframed Thinking (CUT), Rennes School of Business, France. E-mail: raouf.boucekkine@rennes-sb.com

[‡]Paris School of Economics and CNRS, France. E-mail: carmen.camacho@psemail.eu

[§]Purdue University Northwest, USA; Senior Fellow, Centre for Unframed Thinking (CUT), Rennes School of Business, France. E-mail: Wruan@pnw.edu

[¶]DEM, University of Luxembourg, Luxembourg. E-mail: benteng.zou@uni.lu.

1 Introduction

Recently, numerous withdrawals of countries from international organizations and agreements have been witnessed: (1) the most recent happens on July 7, 2020, when the Trump administration formally notified the United Nations that the U.S. was pulling out of the World Health Organization, which became effective as of July 6th, 2021;¹ (2) during the same Trump presidency, on June 1, 2017, President Trump announced that the U.S. would resume its participation in the 2015 Paris Agreement on climate change mitigation until some fair conditions to the U.S. would be negotiated; (3) the United Kingdom withdrew from the European Union on January 31, 2020; (4) Canada withdrew from Kyoto Protocol on December 13, 2011...etc.

Because of this strong trend, while the classical literature on coalitions had essentially addressed the question of coalition formation and stability (with a few exceptions though, see Bolton et al., 1996), numerous papers on coalition break-ups have been written in the last few years. Just as a way of illustration let us mention papers investigating the impact of Brexit (Sampson, 2017; La Torre et al., 2020; the special issue of the Oxford Review of Economic Policy, vol 33, 2017; etc.), or the economic consequences of the U.S. withdrawal from the Kyoto Protocol and the Paris agreement (Bucher et al., 2002; Dai et al., 2017; Nong and Siriwardana, 2018; ...).² This paper is a methodological contribution to this rising literature.

The traditional game-theoretical settings proposed to study the design of international agreements and the stability of coalitions are quite diverse: they range from cooperative to non-cooperative, from static to dynamic through repeated games or fully dynamic set-ups, and they often include interesting procedural ingredients, typically on enforceability of the agreements. Regarding the particular problem of coalition splitting, one finds two main conceptual settings. The first one is based on the theory of coalition stability. This particular part of the theory is anchored in the cooperative games literature with different concepts of coalition stability. This line of research can in turn be divided in two. The first will look at static or repeated games as explained by Tulkens (1998) and surveyed by Bréchet et al. (2011). Then, a second line of research deals instead with dynamic games (as ours) and, after Petrosjan (1977), searches for sharing mechanisms that will ensure the coalition stability. As summarized in Zaccour (2007), the main and most popular sharing mechanisms are the Shapley value, the kernel, the (Nash) bargaining set, the stable set and the core. A second type of setting uses the traditional Nash non-cooperative theory with individual strategies (with a Pareto-like criterion to evaluate efficiency).³

Our paper departs from the latter *optimal* coalition literature in two major ways:

¹Recall that Joe Biden reacted with the following Tweeter message: “On my first day as President, I will rejoin the WHO and restore our leadership on the world stage,” which he eventually did, although not on his first presidential day

²See also the empirical study of Mayer et al (2019) on the cost of being non-EU, and the general theoretical investigation of Gancia et al (2020) on the gain of being in some economic unions and partnerships.

³See for example, Van Long (2010).

- First of all, the sharing rule within coalitions is not necessarily set according to any optimality and/or stability criterion.
- Second, players initially behave *as if* the coalition will last for ever.

At first glance, the two departures seem rather realistic. The coalitions are typically based on a number of, say, constitutional rules specifying the individual duties and benefits accruing to each member of the coalition. Usually, the coalition may obviously entail large heterogeneities across players, technological, demographic or geographic notably. It's unlikely that the constitutional rules at the dawn of the coalition can cope with all these discrepancies, and meet any kind of optimality in the sense of the criteria given above (for example, Nash bargaining or Shapley value). It's also quite reasonable that when engaging in essential (and somehow existential) collective initiative like a strategic political alliance or a long-term environmental agreement, no player will start playing against it, and will rather act optimally, in accordance with the constitutional rules, *as if* the coalition will last till the end time agreed upon. If the constitutional rules are non-negotiable or if renegotiation is very costly, compliance to these rules may become unbearable as time passes, either because the rules, being too rigid, would make exit preferable (endogenous exit) or because an exogenous (symmetric or asymmetric) shock occurs undermining the rational behind the coalition creation. We shall consider the endogenous exit problem in this paper, specializing in the case of time-invariant constitutional rules. As we shall see, the problem remains largely nontrivial in this benchmark case.

In our theory, a non-negotiable coalition will split when an individual player quits. Two terms characterize our setting, and need further explanation: splitting and non-negotiable. To characterize the coalition splitting we will first of all provide with the exact time the coalition ends, the optimal joint trajectories for individual actions and joint resources, and very importantly we will define the follow-up of the coalition. That is, we will redefine the game played by two distinct players, including the new interaction rules and the players' strategic spaces. Regarding the second term, non-negotiable here means that the core principles establishing the initial coalition (say, the sharing rules) cannot be redefined to cope with the preferences of a single country. For instance, we have in mind the case of the U.S. splitting from the 2015 Paris agreement or the case of UK splitting from the EU. We do precisely formalize this problem by designing a dynamic game-theoretic frame, which allows to obtain analytical results, which usually are a daunting task in dynamic game settings. Thus the assumption of time-invariance of the sharing rules under coalition in our framework.

In this paper, we specialize in the theoretical literature of environmental agreements to fix the ideas. In environmental economics, fundamental work has already been accomplished on these issues in a variety of frameworks ranging from multistage games *à la* Carraro and Siniscalco (1993) to dynamic games (Hoel, 1993; Xepapadeas, 1995; Dutta and Radner, 2009). In this paper, we

take a different approach to better suit the question of coalition splitting as we have posed it just above. We consider a set of countries which are initially bound in a coalition, and whose aim is to maximize a given joint payoff subject to a public bad (or good). The coalition is based on constitutional rules, namely the sharing quotas of the benefits and the costs of the coalition. We assume that these rules are time-invariant, and of course, need not be optimal according to any of the criteria listed above. Given all these ingredients, under which conditions a country initially belonging to this coalition may eventually optimally decide to split at a finite date, and **when**? What are the determinants (Constitution, technology,...etc) of splitting and of the duration of the coalition? Is it possible to identify time-invariant constitutional rules which prevent the coalition breakdown?

In our framework, the splitting time is an explicit optimal control in the hands of any member of initial coalitions. The recent contribution of Colombo et al. (2022) is close to our setting in investigating partial cooperation in international environmental agreements. There, all players are identical and so is the share of each player in joint welfare (regardless of whether there is full or partial cooperation). In their setting, the coalition that is optimally set at the initial time will last forever. Our work is also related to the seminar contribution of Benchekroun et al (2006), who study the temporary natural resource cartels where the cartels' ending time is known for all players at the beginning of the game, while in our setting which is an endogenously optimal choice of one player. More precisely, we assume that initially players (countries) agree to manage cooperatively the public bad or good, the stock of pollution in our working example. As a shortcut to the constitutional aspects of the coalition, we assume that each country enters the coalition with a given fixed share of the (intertemporal) payoff of the coalition. We do not include splitting costs (formalizing possible penalties paid by the splitting country) for simplicity. As it will be clear in the main text and in the Appendix, the algebraic developments needed without this additional ingredient are already huge. Note that a splitting fixed sunk cost can hardly change the qualitative results in terms of the sustainability of time-invariant constitutional rules.

If a country splits at time T , a non-cooperative game sets in between the country and the group of countries remaining in the coalition. Within a fully linear-quadratic model, we characterize the optimal affine Markovian subgame perfect strategies for a given split time T . We later solve for the whole sequence starting with the initial cooperative game phase where all coalition members play *as if* the coalition will last for ever.⁴ We ultimately uncover the conditions under which splitting occurs at finite time. In particular, we study in depth the determinants of the duration of the coalition with particular attention paid to the role of technological vs constitutional heterogeneity across players. It's worth pointing out that the choice of strategies can be extended to non-Markovian ones. For example, open-loop Nash equilibria or heterogeneous Nash equilibria *à la* Zou (2016) may match better some situations. For example, after the U.S. withdrawal from the Paris agreement,

⁴Arguably, if this assumption fails, that is, if it is clear (or even uncertain) for the players that the coalition cannot last forever, then a different solution setting via backward induction should be applied.

the remaining coalition stayed committed to the initial decarbonization objectives. Nevertheless, the techniques developed in this paper are general enough to study different kinds of choices of strategic spaces.⁵

Technically, our analytical approach requires the tools of multistage optimal control together with the typical techniques needed to solve differential games. There exist an increasing number of papers using multistage optimal control to characterize optimal/ equilibrium regime transitions and the inherent timings (Boucekkine et al., 2013; Moser et al., 2014; Saglam, 2011; Zampolli et al. 2016; etc.).⁶ In contrast, much fewer papers merging multi-stage optimal control and dynamic games have come out.⁷ We shall show how the latter avenue can also be taken safely in our paper.

Three key aspects drive the paper’s results: the technological gap as an indicator of heterogeneity across players, the Constitution of the coalition (captured by a single parameter, the payoff share accruing to countries under coalitions) and the pollution damage. Thanks to this parsimonious specifications, we are able to provide with a full analytical solution to the two-stage differential game under scrutiny. We do cover all the set of parameterizations taken by the three indicators listed above, which results in a highly nontrivial mathematical analysis (despite parsimony). In particular, we characterize the intermediate parametric cases leading to optimal finite time splitting. A key characteristic of these finite-time-lived coalitions is the requirement of the payoff share accruing to the splitting country to be large enough. Incidentally, our two-stage differential game setting reaches the conclusion that splitting countries are precisely those which use to benefit the most from the coalition. Constraining the payoff share to be low by Constitution may lead to optimal everlasting coalitions only provided initial pollution is high enough, which may cover the emergency cases we are witnessing nowadays.

The paper is organized as follows. Section 2 briefly presents the general specification of our environmental game-theoretical setting. Section 3 fully analyzes a specialized linear-quadratic version of the game, providing in particular the optimal players’ strategies for given splitting time. Section 4 characterizes the existence of an optimal (interior) splitting time, discusses its constitutional and technological drivers and delivers some policy insights. Finally, Section 5 concludes.

2 The model of optimal splitting

Suppose that at time 0 there is a given coalition of players, such as a confederation of regions or a pro-environmental coalition like those which resulted from the Kyoto Protocol or the Paris climate Agreement. Suppose that one of the bloc’s players, named player i , who could eventually quit the

⁵In Appendix B, we illustrate this point by showing how our setting can be readily adapted to heterogeneous strategies after the splitting (if any), one player choosing to play Markovian while the other stays committed to the pre-splitting coalition decision making.

⁶Commonly, all these studies rely on the seminal exploration of Tomiyama (1984).

⁷An exception is for example Boucekkine et al. (2011).

bloc at some future date T , where $0 \leq T \leq \infty$. The rest of the bloc is named player J . The two players, i and J , share a common variable, $y \in [0, Y]$.

When players act as one bloc, players i and J choose jointly the level of variables $x_i, x_J \in [0, X] \subset [0, +\infty)$, which provide them with a joint utility or payoff. Their choices for x_i and x_J may increase or lower the level of the common variable y , which may induce a loss or gain in utility. In this sense, y is a public good or a public bad. To fix ideas, let us assume that the players' actions increase the level of y , resulting in a drop in utility. This mimics the typical public bad (pollution) model in environmental economics, which will be our workhorse along this paper. In our model, we assume that at time 0, players play cooperatively until time T , when player i decides to quit the bloc. Note that at time T , player J may also switch her strategy in response. We shall concentrate on the Markovian perfect equilibria of the game.

Coming back to our workhorse pollution model, suppose there exists a unique final good, which requires only a polluting resource as input. With a quantity of pollution x_j , player j produces $f_j(x_j)$ units of the final good for its own consumption. The consumption of this final good provides both players with utility, but at the same time the production process increases the level of CO₂ emissions by exactly $x_i + x_J$. Obviously, the level of CO₂, y , affects both players. More on functions f_j will be added later on. In the end, player j can obtain utility directly from the consumption of x_j , but she also suffers from pollution, since y brings disutility.

Initially, the objective of the players in the coalition is to maximize joint overall welfare or payoff, which is defined as

$$\max_{x_i, x_J} W(\infty) = \int_0^{+\infty} e^{-rt} [u_i(x_i) + u_J(x_J) - c_i(y) - c_J(y)] dt, \quad (1)$$

where r is the time discount rate, u_i and u_J are utility functions of players i and J , respectively, which are strictly increasing and concave; and $c_i(y), c_J(y)$ are their respective individual disutility due to pollution, which are strictly increasing and convex functions. Note that the objective function is simply the aggregate payoff of the two players. We shall discuss just below how the optimal payoff is shared across players when we come to the constitutional bases of the initial coalition.

Finally, decisions are subject to the dynamic constraint:

$$\dot{y}(t) = f(x_i, x_J, y) = x_i + x_J - \delta y(t), \quad y(0) = y_0 \text{ given}, \quad (2)$$

and $\delta \in [0, 1]$ is the depreciation rate. In our example, y stands for the stock of CO₂, so that δ would stand for the natural reabsorption rate of CO₂ in the atmosphere.

We get now to the constitutional aspects of the coalition. Essential aspects are sharing rules (of the aggregate payoff) and penalties in case of splitting. As argued in the Introduction, we shall focus

on the first aspect. Concretely, we suppose that player i 's share in the total payoff is $\alpha \in (0, 1)$ and the remaining share, $1 - \alpha$, belongs to the rest of the bloc, i.e., welfare of players i and J are

$$W_i = \alpha W(\infty) \quad \text{and} \quad W_J = (1 - \alpha)W(\infty).$$

Two points are worth doing at this stage, both deriving from the fact that we consider a fixed share α . First of all, while α is defined as a fraction of the intertemporal and discounted payoff, it indeed applies at any period of time since it's constant. We can therefore interpret it also as an instantaneous share. This said, the payoff to be shared in our game-theoretic framework is the intertemporal one: when splitting is an option, player i will consider the share of the intertemporal payoff from the start of the coalition to its (potential) end at date T , that is $\alpha W(T)$ where $W(T) = \int_0^T e^{-rt} [u_i(x_i) + u_J(x_J) - c_i(y) - c_J(y)] dt$. Second, beside this technical point, one would inquire which particular meaning such a sharing would mean concretely in a public bad problem like this one. One would be naturally tempted to bring this aspect closer to the standard literature of environmental agreements where the enforcement of a Pareto efficiency criterion would require transfers from certain countries to others (see the early contribution of Tahvonen, 1994). However, this is precisely what we don't do in this model, the weight α is by no way a Pareto weight nor the Shapley value: it is *stricto sensu* a constitutional parameter, it's fixed initially with the birth of the coalition according to the initial political, demographic or economic relative powers of the members.

Of course, α would be generally dependent on the characteristics of each country member of the coalition, that is, on the shapes of the national preferences and technologies. But in reality each country's weight also depends on more complex characteristics like global and regional history, geography and the resulting regional and global geopolitics, which can hardly be recovered unequivocally from technological differentials or cultural differences. In this paper, we define the constitutional rule as being independent from the $u_j(\cdot)$ and $c_j(\cdot)$ to clearly discriminate between the constitutional foundations of the coalition and the more purely technological diversity. We consider this case as the natural benchmark to deeply explore.

Moreover, consistently with the concrete splitting cases put forward in the Introduction, we assume that renegotiation (of α) is impossible or too costly, which is more than realistic.⁸ Note that this could be one of the reasons why player i could decide to quit the bloc at some future date T and play non-cooperatively. We shall see that this outcome is far from automatic and it is only obtained under highly nontrivial conditions.

We now move to some preliminary technical considerations. If player i quits the bloc at time T , then as already said, she obtains a share α of overall welfare until time T . From time T onwards,

⁸Obviously, one way to avoid that player i quits the bloc is to redefine α as the Shapley value, allow renegotiation and letting α be a function of the common resource, i.e. $\alpha = \alpha(y)$.

player i 's objective becomes

$$W_{i,II} = \max_{x_i} \int_T^{+\infty} e^{-rt} [u_i(x_i) - c_i(y)] dt, \quad (3)$$

and player J faces

$$W_{J,II} = \max_{x_J} \int_T^{+\infty} e^{-rt} [u_J(x_J) - c_J(y)] dt, \quad (4)$$

subject to the same state equation:

$$\dot{y}(t) = f(x_i, x_J, y) = x_i + x_J - \delta y(t), \quad t \geq T, \quad (5)$$

where initial condition $y(T)$ comes from the outcome of the first period.

The optimal switching time for player i is defined as

$$\max_T \left(\alpha W(T) + \int_T^{+\infty} e^{-rt} [u_i(x_i^*) - c_i(y^*)] dt \right) = \max_T (\alpha W(T) + W_{i,II}(T)), \quad (6)$$

where $W(T)$ is the same integral in (1) but over time interval $[0, T]$:

$$\int_0^T e^{-rt} [u_i(x_i^*) + u_J(x_J^*) - c_i(y^*) - c_J(y^*)] dt,$$

and x_i^* and y^* are the equilibrium choices and the resulting state, respectively.

Intuitively, a coalition between i and J established at time 0 can last over the period of time $[0, T]$ if the first term in (6) is non-decreasing in T . That is, the longer player i stays in, the higher is joint social welfare. Otherwise, if the first term in (6) is decreasing in T , then player i would exit immediately the coalition with J and T will be 0. Similarly, player i may consider to quit the coalition with J if the second term in (6) is non-increasing with T . Otherwise, if the second term was also increasing in T , then it would always be optimal to set $T = +\infty$ and somehow very surprisingly, this a priori non-optimal (almost ad-hoc) coalition would be stable and last forever.

Obviously, the precise optimal choice of T relies on the game that is played after the splitting, or more precisely, on the strategy space after the splitting. As one can deduce from all the above, the optimal choice of T can be 0, ∞ or take any other finite value between 0 and ∞ , depending on the parameter set. If it exists, the interior optimal switching time T is obtained by taking the first order condition of (6), that is, T is the solution to

$$\alpha \frac{dW(T)}{dT} + \frac{dW_{i,II}}{dT} = 0, \quad (7)$$

provided the second order optimality condition

$$\alpha \frac{d^2 W(T)}{dT^2} + \frac{d^2 W_{i,II}}{dT^2} < 0$$

holds.

In the next section and using a linear-quadratic model, we first study the situation in which the coalition lasts forever, that is, when $T = +\infty$. Then, using those results we will analyse the conditions that ensure the existence of a unique interior solution for T , $0 < T < +\infty$. If splitting occurs in finite time, then applying the implicit function theorem to (7) shows that

$$\frac{\partial T}{\partial \alpha} = - \frac{\frac{dW(T)}{dT}}{\alpha \frac{d^2 W(T)}{dT^2} + \frac{d^2 W_{i,II}}{dT^2}} > 0.$$

That is, the larger share of the aggregate welfare player i detains, the later date she will quit the coalition. Similarly, the smaller the stake of player i in the coalition, the sooner she quits the bloc to gain some freedom of choice and free from commitments.

It is worth mentioning that time T is by no means a truncated terminal time, rather it is the time at which the splitting happens. We would also like to underline that differently from most of the optimal switching literature, the players before and after time T are different in our setting. Indeed, before time T , there is a single player: the coalition. After T , there are two competing players. Thus special care should be taken when employing the usual necessary optimal switching conditions at T . These difficulties come mainly from the choice of different strategic spaces after the coalition splits. In the following sections, we present some choices for the strategic spaces after the split.

3 The linear-quadratic case two-stage differential game

To hopefully solve in a fully analytical way the above problem, we resort as it is traditional in differential games to linear-quadratic functional forms (see Dockner and Van Long, 1993; Dockner et al, 2000; Bertinelli et al, 2014; etc). In this section, we focus on the strategies at each stage of the game for a given splitting time. Then Section 4 will address the optimal splitting time issues.

In our linear-quadratic setting, the utility functions are given by

$$u_i(x_i) = a_i x_i - \frac{x_i^2}{2}, \quad u_J(x_J) = a_J x_J - \frac{x_J^2}{2}.$$

If x_j is the pollution emission that player j employs to produce the final consumption good, then a_j is here the efficiency parameter which converts pollution into the consumption good. Note

that a higher a_j indicates a more advanced economy, meaning that it can generate more of the consumption good from the same unit of pollution.

The pollution damage functions are

$$c_j(y) = \frac{by^2}{2}, \quad j = i, J.$$

That is, the pollution damage is the same for both players independent of any individual characteristic and in particular, independent of the agent's development level. In the following, we assume that coefficients a_j , $j = i, J$, are sufficiently large, such that the utility functions are always positive and increasing in x_j and that the long-run steady state of pollution is positive.

3.1 The cooperative stage

Let us start by solving the problem of the ever lasting coalition. As motivated in the Introduction, each coalition member acts *as if* the coalition lasts forever, and will do so till splitting (if any) occurs. The joint payoff function is

$$\max_{x_i, x_J} W(\infty) = \int_0^{+\infty} e^{-rt} \left[a_i x_i + a_J x_J - \frac{x_i^2 + x_J^2}{2} - by^2 \right] dt, \quad (8)$$

subject to the following state equation:

$$\dot{y}(t) = x_i + x_J - \delta y(t), \quad y(0) = y_0 \text{ given}, \quad (9)$$

and as mentioned above $\delta \in [0, 1]$ is the depreciation rate. In our example, y stands for the stock of CO₂, so that δ stands for the natural reabsorption rate of CO₂ in the atmosphere.

We can readily summarize the main results of the optimization problem faced by any coalition member acting *as if* coalition is everlasting, in the following proposition.

Proposition 1 *For any positive constants b, r, δ , then for any state trajectory $y(t)$, the choices for player i and J are*

$$x_j^*(y) = a_j + B + Cy, \quad j = i, J,$$

where

$$C = \frac{r + 2\delta - \sqrt{(r + 2\delta)^2 + 16b}}{4} (< 0), \quad B = \frac{(a_i + a_J)C}{r + \delta - 2C} (< 0).$$

The trajectory of state is: $\forall t \geq 0$,

$$y(t) = (y_0 - y^*)e^{(2C - \delta)t} + y^*$$

where y^* is the asymptotically stable long-run steady state given by

$$y^* = \frac{a_i + a_J + 2B}{\delta - 2C} (> 0).$$

The long-run steady state y^* depends on all the parameters, especially the sum of the technology levels, a_i and a_J . A higher technology level, which translates into higher consumption, leads to a higher level of long-run pollution. Consistently with the standard linear-quadratic model considered, the convergence speed, $(2C - \delta)$, is independent of the technology levels, it rather depends on time preference r , the unit damage of pollution, b , and nature's regeneration rate δ .

We notice that the two players' aggregate consumption at the long-run steady state is obviously a function of y^* and it is always positive:

$$x_i^*(y^*) + x_J^*(y^*) = a_i + a_J + 2B + 2Cy^* = \frac{\delta}{\delta - 2C}(a_i + a_J + 2B) > 0,$$

where the last inequality comes from the fact that $a_i + a_J + 2B > 0$. Furthermore, aggregate consumption is always positive along the optimal trajectory path, that is,

$$\begin{aligned} x_i^*(y) + x_J^*(y) &= a_i + a_J + 2B + 2Cy = a_i + a_J + 2B + 2C[(y_0 - y^*)e^{(2C-\delta)t} + y^*] \\ &= x_i^*(y^*) + x_J^*(y^*) + 2C(y_0 - y^*)e^{(2C-\delta)t} > 0 \end{aligned}$$

$\forall y$ and $\forall r, \delta, a_i, a_J, b > 0$, provided $y_0 < y^*$, which is a natural assumption.

If there existed a unique finite solution to (7), and a result, player i quit the bloc at time T , then, the above optimal choice will be enforced from time 0 until $t = T$ and pollution accumulation would reach the value $y(T)$:

$$y(T) = (y_0 - y^*)e^{(2C-\delta)T} + y^* \quad (10)$$

where T is yet to be determined using (7), and it will crucially depend on the choice of strategic space after the splitting.

The total payoff of player i just before the splitting is thus

$$\alpha W(T) = \alpha \left[\frac{(a_i^2 + a_J^1 - 2B^2)(1 - e^{-rT})}{2r} - 2BC \int_0^T e^{-rt} y(t) dt - (C^2 + b) \int_0^T e^{-rt} y^2(t) dt \right]. \quad (11)$$

It is straightforward that

$$\frac{dW(T)}{dT} = e^{-rT} \left[\frac{a_i^2 + a_J^2 - 2B^2}{2} - 2BCy(T) - (C^2 + b)y^2(T) \right] > 0 \quad (12)$$

if and only if

$$y(T) = (y_0 - y^*)e^{(2C-\delta)T} + y^* \in (0, \underline{y})$$

where

$$\underline{y} = \frac{-2BC + \sqrt{4B^2C^2 + 2(C^2 + b)(a_i^2 + a_J^2 - 2B^2)}}{2(C^2 + b)} (> 0). \quad (13)$$

Obviously, the analysis above provides an upper-bound condition for remaining in the bloc in terms of pollution:

Corollary 1 *Under the assumptions of Proposition 1, and provided that $\alpha > 0$:*

- *if the initial condition checks $y_0 > \underline{y}$, then $T = 0$;*
- *if the coalition potential long-run steady state checks $(y_0 <)y^* < \underline{y}$, then $T = +\infty$.*

The above corollary can be written in a more compact manner:

$$T \begin{cases} = 0, & \text{if } \underline{y} < y_0, \\ \in (0, +\infty), & \text{if } y_0 < \underline{y} < y^*, \\ = +\infty, & \text{if } \underline{y} > y^*. \end{cases}$$

From (13), it is easy to see that $\lim_{a_i, a_J \rightarrow 0} \underline{y} = 0$, for all $b > 0$. By continuity when a_i, a_J are sufficiently small, and for any $y_0 > 0$, it follows that $\underline{y} < y_0$, thus according to Corollary 1 it must necessarily be that $T = 0$. In other words, when both players in the initial coalition have a low enough development level, the coalition hardly exists. It is intuitive to see why. Both players here are heavy polluters, in the sense that they do not extract much consumption from the pollutant. None of them can make a remarkable effort to reduce pollution and alleviate the damage from the common bad. Since player i does not perceive the gains of staying in the coalition, she will exit immediately. Note that here the motive for quitting the coalition is not free-riding in a strict sense since player i also takes into account J 's welfare when they remain in the coalition. Here the coalition does not last because player i does not perceive any advantage for neither of them.

Additionally, it can be shown that for any a_i, a_J not both zero, $\lim_{b \rightarrow 0} \underline{y} \geq +\infty > \lim_{b \rightarrow 0} y^* = \frac{a_i + a_J}{\delta}$.⁹ When pollution damage is low and a_i and a_J are not both close to zero, the coalition will remain

⁹It is easy to check that

$$\begin{aligned} \underline{y} &= \frac{\sqrt{2(C^2 + b)(a_i + a_J) - 4bB^2} - 2BC}{2(C^2 + b)} \\ &\geq \left[\frac{\sqrt{C^2 + b - (4bC^2/(r + \delta - 2C)^2) - 2C^2}}{2(C^2 + b)} \right] (a_i + a_J) \\ &\equiv S(b)(a_i + a_J) \end{aligned}$$

together forever. Intuitively, when pollution damage is negligible, then one would have expected T to be 0 since there is no incentive to stay in the coalition. But, what we prove here is just the opposite. The reason for this seemingly contradicting result comes from the fact that when b is close to zero, both cooperation in the coalition and competition bring nearly identical welfare to player i . Indeed, note that the limit case shows $\lim_{b \rightarrow 0} x_j^* = a_j = \lim_{b \rightarrow 0} x_j^m$ for both $j = i, J$. Therefore, when damage from pollution is low and the coalition already established, then it will last forever. However, if there was no coalition at time 0, then there is no incentive to form one either.

Let us add some further comments on our results. A given level of pollution stock provides on the one hand with utility via consumption, but on the other hand, it also generates disutility. The positive effect on welfare is linear on emissions while the disutility generated by the stock of pollution is quadratic. Hence, participating in a coalition makes sense especially when b is small enough relative to the technological parameters a_i and a_J . As a result, if there is no technological progress (at least in one player), then coalitions make no sense if $b > 0$ whatever $\alpha \neq 0$. Conversely, if b tends to zero, and technological progress is nonzero, then coalitions would last forever whatever $\alpha \neq 0$. These mechanisms will play an important role in all the cases we study in the following sections.

To study the interior situation where splitting happens in finite time, we impose the following condition on the parameters:

Assumption 1 *The model parameters ensure that the following inequality holds:*

$$y_0 < \underline{y} < y^*.$$

Unfortunately, it's not possible to explore analytically how this condition relates to the deep parameters of the model given the expressions of \underline{y} , y^* , B and C . We shall fortunately obtain interpretable expressions for the interior splitting conditions (and optimality) in the next section under the above assumption.¹⁰

3.2 Optimal strategies in the non-cooperative stage (after T)

Suppose player i quits the coalition at time T , and that after that, only Markovian subgame perfect strategies are considered. Here the strategy is defined as that choice variable x_j for player $j = i, J$, which depends on time and the current state: $x_i(t) = x_i(t, y(t))$, for all y . Furthermore, the game

and

$$y^* = \frac{r + \delta}{\delta(r + \delta) - 4b}(a_i + a_J).$$

Thus, one sufficient condition for $\underline{y} > y^*$ is $S(b) > y^*$, for any a_i and a_J not both zero. By l'Hopital's rule, $\lim_{b \rightarrow 0} S(b) = +\infty$.

¹⁰It's not difficult to see that Assumption 1 is not void through numerical exercises.

is autonomous at all t and defined over infinite time horizon, $[T, +\infty)$. Under these assumptions we can directly study the stationary Markovian perfect equilibrium (MPE) via stationary Bellman value functions.

Define the Bellman Value function of player $j = i, J$ as $U_j(y)$, which must check the following HJB equation for $t \geq T$

$$rU_j(y) = \max_{x_j} \left[a_j x_j - \frac{x_j^2}{2} - \frac{b y^2}{2} + U_j'(y) (x_i + x_j - \delta y) \right], \quad j = i, J.$$

From these HJB equations, Appendix A.1 demonstrates the following existence results of MPE.

Proposition 2 *Suppose that player i quits the coalition at a finite time T , and that after the splitting both players i and J adopt Markovian strategies. Then there exists a stable affine Markovian subgame perfect Nash equilibrium (MPE)*

$$(x_i^m, x_J^m) = (a_i + B^m + C^m y, a_J + B^m + C^m y), \quad \forall y,$$

with coefficients

$$C^m = \frac{(r + 2\delta) - \sqrt{(r + 2\delta)^2 + 12b}}{6} (< 0), \quad B^m = \frac{(a_i + a_J)C^m}{r + \delta - 3C^m} (< 0).$$

For a given initial condition at T , the corresponding optimal state trajectory is

$$y^m(t) = (y(T) - \widehat{y}^m) e^{(2C^m - \delta)(t-T)} + \widehat{y}^m, \quad \forall t \geq T,$$

where $\widehat{y}^m = \frac{a_i + a_J + 2B^m}{\delta - 2C^m} (> 0)$ is the asymptotically stable long-run steady state.

The following corollary obtains from the above two propositions. Its proof is provided in detail in Step 3 of Appendix A.1.

Corollary 2 *Under the assumptions of Propositions 1 and 2, it follows*

$$y^* < \widehat{y}^m, \quad \forall r, \delta, b > 0.$$

In other words, even if the coalition is dissolved at T because staying in the coalition does not provide player i with higher welfare, the coalition does better in terms of the stock of pollution. This is a standard outcome in the environmental agreements literature (see for example, Calvo and Rubio, 2013). Obviously, the decision to split cannot be entirely determined by the steady state of pollution as we will see in the following section.

We have so far characterized the optimal strategies for each stage of the game, and identified some useful properties along the way. We now move to the determination of the optimal splitting time (if it exists).

4 Optimal splitting time and its constitutional vs technological drivers

Under the above Markovian perfect Nash equilibrium, it is easy to obtain player i 's welfare in second period

$$\begin{aligned} W_{i,II}^m &= \int_T^{+\infty} e^{-rt} \left(a_i x_i - \frac{x_i^2}{2} - \frac{by^2}{2} \right) dt \\ &= \frac{a_i^2 - (B^m)^2}{2} \int_T^{+\infty} e^{-rt} dt - B^m C^m \int_T^{+\infty} e^{-rt} y^m(t) dt - \frac{((C^m)^2 + b)}{2} \int_T^{+\infty} e^{-rt} (y^m)^2 dt, \end{aligned} \quad (14)$$

where $y^m(t)$ also depends on the splitting time T . In order to assess how the splitting time affects player i 's welfare, we compute $\frac{dW_{i,II}^m(y(T))}{dT}$ using (14):¹¹

$$\begin{aligned} \frac{dW_{i,II}^m(y(T))}{dT} &= \{[-rA_i^m + B^m(a_i + a_J + 2B)] + [B^m(2C - \delta - r) + (a_i + a_J + 2B)C^m]y(T) \\ &\quad + [(2C - \delta) - r/2]C^m y^2(T)\} e^{-rT} = [\hat{a} + \hat{b}y(T) + \hat{c}y^2(T)]e^{-rT} \\ &\quad \begin{cases} < 0 & \text{for } 0 \leq y(T) < \bar{y}, \\ > 0 & \text{for } y(T) > \bar{y}, \end{cases} \end{aligned} \quad (15)$$

with \bar{y} the positive root of the second degree polynomial $\frac{dW_{i,II}^m(y(T))}{dT} = 0$ ¹² and $A_i^m = \frac{a_i^2}{2r} + \frac{(a_i + a_J)B^m}{r} + \frac{3(B^m)^2}{2r}$.

If $y(T) > \bar{y}$, then $\frac{dW_{i,II}^m(y(T))}{dT} > 0$. Hence the later the splitting happens, if there is splitting at all, the higher player i 's welfare in the second period. If this is the case, then player i would postpone the splitting as late as possible. In other words, splitting never happens when the stock of pollution is sufficiently high, namely if $y(t) > \bar{y}$.

Recall that Assumption 1 explicitly states the precise condition under which the stock of pollution

¹¹See Appendix A.2 for the details.

¹²The positive root is given by

$$\bar{y} = \frac{-\hat{b} - \sqrt{\hat{b}^2 - 4\hat{a}\hat{c}}}{2\hat{a}}$$

where $\hat{a} = [-rA_i^m + B^m(a_i + a_J + 2B)] < 0$, $\hat{c} = [(2C - \delta) - r/2]C^m > 0$ and $\hat{b} = B^m(2C - \delta - r) + (a_i + a_J + 2B)C^m$.

is increasing over time. Thus if we assume that players i and J are initially in a coalition and that $y_0 > \bar{y}$, then $y(t) > \bar{y}$ for all $t \geq 0$, and there will be no splitting. We conclude in the following corollary

Corollary 3 *Suppose Assumption 1 holds and that $y_0 > \bar{y}$, then splitting will never happen, that is, $T = +\infty$.*

Combining the above two corollaries, Corollaries 1 and 3, it turns out that the relationship between \bar{y} and \underline{y} is not essential in determining the optimal splitting time T , as long as they both are located between y_0 and y^* . In that case, then there may be splitting in finite time. We study this issue in detail in the next sub-section.

4.1 Optimal finite splitting time

In order to focus on the situation where splitting can happen in finite time, we complete Assumption 1 with the following:

Assumption 2 *Suppose the initial condition and the parameter set check*

$$y_0 < \bar{y}.$$

Obviously, if $y^* < \bar{y}$, then splitting may happen in finite time. If instead $y_0 < \bar{y} < y^*$, then splitting can only take place before the pollution stock reaches the upper limit \bar{y} . Otherwise, player i can not afford the damage cost from the accumulated pollution and would rather stay with player J .

Substituting the first and second periods' welfare derivatives with respect to T , i.e. (12) and (15), into the first order condition $\alpha \frac{dW(T)}{dT} + \frac{dW_{i,II}^m}{dT} = 0$, it follows that the above first order condition is equivalent to:

$$\Lambda (y_T^m)^2 + \Sigma y_T^m + \Gamma = 0, \quad (16)$$

where $y_T^m = y^m(T)$ stands for the stock of pollution at the switching time and the coefficients in (16) are

$$\begin{cases} \Lambda = -\alpha (C^2 + b) + C^m (2C - \delta - \frac{r}{2}), \\ \Sigma = -2\alpha BC + B^m (2C - \delta - r) + C^m (a_i + a_J + 2B), \\ \Gamma = \frac{1}{2}\alpha (a_i^2 + a_J^2 - 2B^2) - rA_i^m + B^m (a_i + a_J + 2B). \end{cases} \quad (17)$$

The roots of (16), if they exist, are given by

$$y_T^m = \frac{-\Sigma \pm \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda}. \quad (18)$$

Obviously, the existence of real roots is granted if and only if $\Sigma^2 - 4\Lambda\Gamma \geq 0$. The last inequality condition is ensured by $\Lambda \leq 0$ and $\Gamma \geq 0$, which are equivalent to

$$\alpha \geq \frac{C^m (2C - \delta - r/2)}{C^2 + b} \equiv G(b) \quad (19)$$

and

$$\alpha \geq \frac{1 + \left[\frac{3(C^m)^2}{(r+\delta-3C^m)^2} - \frac{4CC^m}{(r+\delta-3C^m)(r+\delta-2C)} \right] \left(\frac{a_J}{a_i} + 1 \right)^2}{\left(\frac{a_J}{a_i} \right)^2 + 1 - \left[\frac{2C^2}{(r+\delta-2C)^2} \right] \left(\frac{a_J}{a_i} + 1 \right)^2} \equiv F \left(\frac{a_J}{a_i}, b \right). \quad (20)$$

In other words, there exists a finite splitting time $T^m \in (0, \infty)$ if the payoff share is large enough, for given b . It can be readily shown that $G(b)$ is increasing in b and that $G(0) = \frac{1}{2}$ (see the Appendix for all related computations). Consequently, the larger b , the larger the share α needed to make finite time splitting possible. Moreover, as function $G(\cdot)$ is increasing, α should be always bigger than one-half whatever b . As explained below Corollary 1, the larger b , the more reluctant player i is to stay in the coalition. Only when the payoff share α is large enough, would player i remain in the coalition (even if she will eventually leave it at some later time).

Again, as for condition (19), finite time splitting is granted if the payoff share is large enough, although the involved lower bound in this case is different: in contrast to condition (19), the lower bound also depends on the technological gap, $\frac{a_J}{a_i}$. We shall examine the implications below.

It should be noted that the above conditions result from the first-order condition, and we now move to the analysis of the second-order condition. Given that parameters Λ, Σ and Γ are independent of the switching time T , the second-order sufficient condition, $\alpha \frac{d^2 W(T)}{dT^2} + \frac{d^2 W_{i,II}^m}{dT^2} < 0$, holds if and only if

$$[2\Lambda y^m(T) + \Sigma] \frac{\partial y^m(T)}{\partial T} < 0.$$

Given that pollution accumulation is increasing over time, we also have that $\frac{\partial y^m(T)}{\partial T} > 0$ for any T . Then the second-order sufficient condition holds if and only if

$$2\Lambda y_T^m + \Sigma < 0. \quad (21)$$

The above second order condition is equivalent to

$$\pm \sqrt{\Sigma^2 - 4\Lambda\Gamma} < 0.$$

If $\Lambda < 0$, then the only optimizer of $\alpha W(T) + W_{i,II}^m(T)$ is at

$$y_T^m = \frac{-\Sigma - \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda}.$$

So the optimal finite splitting time, T , is unique. At the minute notice that combining Conditions (19) and (20) ensures the existence of a unique optimal finite splitting time T . We summarize this important result in the following proposition, and its detailed proof is reported in Appendix A.3.

Proposition 3 *Let Assumptions 1 and 2 hold. Suppose player i quits the coalition at time T , and after that players i and J adopt MPE given by Proposition 2. Suppose the sharing parameter α checks*

$$\max \left\{ F \left(\frac{a_J}{a_i}, b \right), G(b) \right\} < \alpha < 1 \quad (22)$$

where functions $G(b)$ and $F \left(\frac{a_J}{a_i}, b \right)$ are defined in (19) and (20). Furthermore, suppose that the pollution quantity

$$y_T^m = \frac{-\Sigma - \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda} \quad (23)$$

satisfies

$$y_0 < y_T^m < y^*,$$

where Λ , Σ , Γ are given by (17). Then player i optimally quits the coalition at a finite time T :

$$T = \frac{1}{2C - \delta} \ln \left(\frac{y_T^m - y^*}{y_0 - y^*} \right). \quad (24)$$

If Condition (22) fails, then it can happen that either $T = 0$ and splitting is immediate, or $T = +\infty$ and there is no splitting at all. We shall pay more attention to the “corner” solution, $T = +\infty$ in the policy implications part of this section.

Note that Proposition 3 delivers an explicit solution for the optimal splitting time as none of the terms involved in (24) depends on T . It should also be noted that both the splitting time and the stock of pollution depend on the sharing parameter. Indeed, the optimal splitting time, given by (24), depends on α through y_T^m ; and so do the level of the pollution stock at T and the conditions that ensure its existence (in short, the second-order optimality conditions). Interestingly enough, Condition (22) shows that the three fundamental ingredients of our model do matter in the duration of coalitions: the sharing parameter, the pollution damage parameter, b , and the technological gap, $\frac{a_J}{a_i}$.

The economic interpretation of Condition (22) involves the technological gap: for given technological gap and pollution damage parameter, the payoff share under coalition is required to be large enough for player i to engage in a coalition and to stay in for a finite time. Again, the constitutional parameter α is key in the optimal institutional dynamics: it’s key for the existence of an optimal finite time splitting, and it’s also key for the duration of the coalition (through the level of pollution at the splitting time, y_T^m as explained above). We shall devote the next subsection to the latter point. In the meanwhile, we shall clarify the implications of Proposition 3 by specializing in two

cases depending on the ratio $\frac{a_J}{a_i}$. A more general result is stated and proved in Lemma 1 in the Appendix.

We shall indeed isolate two cases of a finite lived coalition here, first the case of a technologically lagged country i , $\frac{a_J}{a_i} > 1$, and then a case where this country is advanced.

Corollary 4 *Under the assumptions of Proposition 3, and provided $\frac{a_J}{a_i} > 1$ and $\alpha > \frac{1}{2}$, player i optimally quits the coalition at time T if and only if $\alpha > G(b)$.*

A technologically lagged country may remain in the coalition for any value of b provided the reward, as captured by α , is large enough, and in any case, larger than $\frac{1}{2}$. Notice that this is true whatever the value of the technological gap, provided it's bigger than one. Suppose now that country i is more advanced than J , what would the outcome be? We provide below a simple illustrative case.

Corollary 5 *Let the assumptions of Proposition 3 be satisfied and let $\alpha > \frac{1}{2}$. Assume that*

$$\frac{a_J}{a_i} < \frac{\sqrt{3} - 2 + \sqrt{12 - 2\sqrt{3}}}{2 + \sqrt{3}} \approx 0.7, \quad (25)$$

then

$$F\left(\frac{a_J}{a_i}, b\right) \geq G(b)$$

holds for all $b \geq 0$. Therefore, (22) holds if and only if

$$F\left(\frac{a_J}{a_i}, b\right) < \alpha.$$

Condition (22) is the single most important requirement in Proposition 3 since it provides a range for α for the player to remain in the coalition, even if temporarily. The condition clearly states that player i needs to retain a sufficiently large share of the total payoff to stay. In the case of a technologically advanced country i , the technological gap shows up in the existence and optimality Condition (22) contrary to the case of the lagged player i studied in Corollary 4. This is hardly surprising: as we said, a lagged country will always benefit from a coalition if the pollution damage is small enough, at least for a while, if her payoff share under coalition is good enough. However, as Corollary 5 shows, the tradeoffs are more involved if the country is more advanced than the other coalition members (on average). Suppose that the pollution damage parameter, b , is small, then the benefits for player i to remain in the coalition are rather thin. In this case one may expect that the higher player i 's technology, that's the smaller the technological gap $\frac{a_J}{a_i}$, the larger the payoff share requested by player i to remain in the coalition. The contrary also holds true, if the pollution damage is large enough, then the more advanced the country, the lower the payoff share requested

to remain in the coalition.¹³ Proposition 3 and its corollaries show that our model is indeed able to generate finely all the possible institutional configurations depending on three key parameters, $(\alpha, b, \frac{a_i}{a_i})$.

4.2 Payoff sharing and the duration of coalitions

This section clarifies how the sharing parameter α affects the duration of a coalition. We already know that the duration of the coalition is increasing in y_T^m , the pollution stock at the splitting time, according to equation (24). A critical point, already also outlined in the section above is that the pollution stock at the splitting time, y_T^m , is increasing in α if $y_T^m \leq \underline{y}$ and decreasing in α if $y_T^m > \underline{y}$, where \underline{y} is defined in (13) and y_T^m in (23). We can go a step further, Appendix A.5 shows the following results.

Theorem 1 *Suppose there exists $\alpha_0 \in [0, 1]$, such that, $y_T^m(\alpha_0)$ satisfies $0 \leq y_T^m(\alpha_0) < \underline{y}$, then $y_T^m(\alpha)$ is increasing in α in a neighborhood of α_0 in $[0, 1]$. Similarly, if there exists a value α_0 such that $y_T^m(\alpha_0) > \underline{y}$, then $y_T^m(\alpha)$ is decreasing in α in a neighborhood of α_0 in $[0, 1]$.*

To better grasp the importance of Theorem 1, let us link Assumption 1 and the last statement in Proposition 3. Assumption 1 delivers a condition under which a coalition may split in finite time. The first assumption of Theorem 1 states that if there exists a sharing parameter, α_0 , such that the coalition can split in finite time, then increasing player i 's payoff share raises the stock of pollution upon splitting y_T^m , leading in turn to a more durable coalition (by Proposition 3). Therefore, when the stock of pollution generated by the coalition is relatively small, one obtains that the larger the payoff share of player i , the later splitting (and the larger the subsequent pollution stock at the splitting time). Recall that postponing T increases the joint payoff in the first stage of the game, thus lengthening the coalition duration benefits both players i and J .

The opposite also holds true. If the payoff share of player i is such that the stock of pollution y_T^m is relatively high, then the coalition is actually no longer beneficial (for player i)¹⁴ and an increase in α fastens the splitting process.

The above results are based on the existence of one particular sharing strategy α_0 , which may be difficult to find. The following corollary extends the conditions in Theorem 1 from one particular point into an interval, which is easier to find and apply. A detailed proof is given in Appendix A.6.

Corollary 6 *Suppose $y_T^m(\alpha)$ is real and nonnegative in a subinterval $(\alpha_1, \alpha_2) \subset [0, 1]$. Then, either $y_T^m(\alpha) \leq \underline{y}$ in the entire subinterval (α_1, α_2) , or $y_T^m(\alpha) \geq \underline{y}$ in the entire subinterval (α_1, α_2) . In*

¹³We refer the reader to Appendix A.4 for a complete description and more general results.

¹⁴See the discussion after (12)

particular, if

$$\max \left\{ \frac{C^m (2C - \delta - r/2)}{C^2 + b}, \frac{2 [r A_i^m - B^m (a_i + a_J + 2B)]}{a_i^2 + a_J^2 - 2B^2} \right\} < 1,$$

then either $y_T^m(\alpha) \leq \underline{y}$ or $y_T^m(\alpha) \geq \underline{y}$ for all α that satisfies (22).

4.3 Policy insights

We shall now point at some potentially interesting policy insights one can extract from our theoretical analysis. Let's first mention that while our model may seem too stylized to tackle "real" splitting problems, it's not more stylized than the typical two-stage or repeated games devoted to this topic, surveyed in the Introduction: while it has some ingredients of differential games, it's essentially a two-stage game, the first stage being the coalition stage. There are two main differences with respect to the vast majority of dynamic or two-stage games in this topic. First, the initial coalition is defined by a profit sharing scheme that is not necessarily an equilibrium of any sort, so that the coalition is not necessarily stable. And second, the duration of the first stage is itself the direct result of the individual decision of a particular player. It's very hard to argue that this latter trait is unrealistic, this is exactly how the environmental agreements/protocols and other political coalitions have been unravelled.

Our analysis brings several interesting results along Sections 3, 4.1 and 4.2. Let's stress one of them: as clearly stated in Corollaries 4 and 5, only "big" enough countries may under certain conditions quit the coalition. More precisely, optimal finite time splitting requires $\alpha > \frac{1}{2}$. Recall that the payoff share is determined by the "constitution" of the coalition, reflecting in particular the relative historical, geographic, demographic and economic weight of the countries. If the constitution is also meant to guarantee no-splitting, two avenues can be taken within our framework. One is to counterbalance the impact of too large payoff share (in the sense of Corollaries 4 and 5) by adding penalties to the constitution, making sure that penalties are increasing enough in the payoff share to discourage splitting. The second (non-exclusive) solution is to limit by constitution the payoff share of all individual players, which guarantees that everlasting coalitions are the unique optimal institutional arrangement. In our theory, such an arrangement can be possible under the following conditions summarized in the following proposition.

Proposition 4 *Suppose α is chosen according to*

$$\alpha < \min \left\{ G(b), F\left(\frac{a_J}{a_i}, b\right) \right\},$$

thus violating Condition (22). Then, the coalition optimally lasts forever provided the initial pollution, y_0 is large enough.

Technically speaking, Proposition 4 delivers conditions under which the “corner” solution, $T = \infty$, is optimal. As already mentioned, the full analysis of the corner solutions, including the one singled out in Proposition 4, is given in the Appendix. With respect to the optimal finite time splitting case, not only Condition (22) is violated: the stability of the everlasting coalition also requires the initial pollution to be large enough, which indeed also violates the second condition stated in Proposition 3 (that is, $y_0 < y_T^m < y^*$).¹⁵ Therefore, a constitution which limits the payoff share (to the upper bound identified in Proposition 4 depending on the pollution damage and the relative technological position of each country would, in our model, allow to block splitting if the level of pollution is threatening enough. Of course, if the pollution level is small, then coalitions are much less attractive, and they would not even exist. That is, the corner solution $T = 0$ would arise in such a case.

Finally, one can show that there is an interval of α on which the coalition lasts forever, without assuming any lower bound on initial pollution. Instead, we require a lower bound for the share α . Let

$$\underline{\alpha} = \frac{C^m (2C - \delta - r/2)}{C^2 + b}$$

and let $\hat{\alpha}$ satisfies the linear equation

$$\Lambda(\hat{\alpha})(y^*)^2 + \Sigma(\hat{\alpha})y^* + \Gamma(\hat{\alpha}) = 0$$

where y^* is defined by 23). Furthermore, let $\tilde{\alpha}$ be the first root greater than $\underline{\alpha}$ such that

$$(\Sigma(\tilde{\alpha}))^2 - 4\Lambda(\tilde{\alpha})\Gamma(\tilde{\alpha}) = 0.$$

If there is no such $\tilde{\alpha}$, we let $\tilde{\alpha} = +\infty$. Then, it can be shown that

$$0 < \underline{\alpha} < \min\{1, \hat{\alpha}, \tilde{\alpha}\}$$

and

Proposition 5 *Suppose $y_0 < y^*$. Then $T = \infty$ if α satisfies*

$$\underline{\alpha} < \alpha \leq \min\{1, \hat{\alpha}, \tilde{\alpha}\}. \tag{26}$$

The proof is given in Appendix A.8.

¹⁵In the Appendix, we show that y_0 large enough corresponds indeed to $y_0 > y_T^m$

5 Conclusion

In this paper, we have presented an alternative view of the coalition breakdown problem. Motivated by the most recent related events (in particular, unilateral withdrawal of several countries from institutional or environmental agreements), we build up an alternative framework in which the sharing rules under the initial coalitions may not be necessarily optimal, and in which, because the coalitions are initially views are essential or existential, members act *as if* they will stay for ever. We have specialized in the case where the sharing rules are time-invariant and we have studied in detail the corresponding endogenous exit problem (for a given member of the initial coalition) in a linear-quadratic two-stage differential game frame where players (the splitting member and the remaining, presumably homogenous, coalition) play Markovian.

We have, in particular, solved the specific two-stage differential game induced by our alternative theory, and derived several highly nontrivial results on the technological/constitutional characteristics of the splitting coalitions and on the way the sharing rules could be set in order to prevent coalition breakdown. There are obvious interesting extensions of this work. For example, one may consider the case where, as mentioned in the Introduction, a particular coalition member may not act *as if* the coalition will last forever, and will plan her actions accordingly from the start of the coalition. Uncertainty (for example on the unit cost of pollution) is also an important ingredient to insert into the picture. We are making progress in both research avenues.

A Appendix: Proof

A.1 Proof of Proposition 2 and Corollary 2

The proof is completed in three steps: step 1 demonstrates the existence of affine-linear Markovian Nash equilibrium; step 2 shows the stability and step 3 provides steady states comparison.

Step 1. Existence of Markovian Nash equilibrium

Define the Bellman Value function of player $j = i, J$ as $U_j(y)$, which must check the following HJB equation: for $t \geq T$,

$$rU_j(y) = \max_{x_j} \left[a_j x_j - \frac{x_j^2}{2} - \frac{b y^2}{2} + U_j'(y) (x_i + x_J - \delta y) \right], \quad j = i, J.$$

Then the first order condition yields

$$x_j^m(t) = a_j + U_j'(y(t)). \tag{27}$$

Guess

$$U_j(y) = A_j + B_j y + \frac{C_j}{2} y^2, \quad \text{and } j = i, J,$$

then

$$U'_j(y) = B_j + C_j y.$$

Substituting $x_i = a_i + B_i + C_i y$ and $x_J^m(t) = a_J + B_J + C_J y(t)$ into the HJB equations, comparing coefficients on both hand sides, it yields

$$\begin{cases} rA_i = \frac{(a_i+B_i)^2}{2} + (a_J + B_J)B_i, \\ (r + \delta - C_i - C_J)B_i = C_i(a_i + a_J + B_J), \\ (r + 2\delta)C_i = C_i^2 + 2C_i C_J - b, \end{cases} \quad (28)$$

and

$$\begin{cases} rA_J = \frac{(a_J+B_J)^2}{2} + (a_i + B_i)B_J, \\ (r + \delta - C_i - C_J)B_J = C_J(a_i + a_J + B_i), \\ (r + 2\delta)C_J = C_J^2 + 2C_i C_J - b, \end{cases} \quad (29)$$

Remark. More generally, if $b_i \neq b_J$, then the b in the last two equations should be b_i and b_J respectively.

Solving the above two group equations system simultaneously, it follows that the only coefficients which yields valid Bellman value functions are

$$C_i = C_J = \frac{(r + 2\delta) - \sqrt{(r + 2\delta)^2 + 12b}}{6} \equiv C^m, \quad \text{and } B_i = B_J = \frac{(a_i + a_J)C^m}{r + \delta - 3C^m} \equiv B^m,$$

and

$$A_j^m = \frac{a_j^2}{2r} + \frac{(a_i + a_J)B^m}{r} + \frac{3(B^m)^2}{2r}, \quad j = i, J.$$

Thus the Markovian Nash equilibrium is given by

$$(x_i^m, x_J^m) = (a_i + U'_i(y), a_J U'_J(y)) = (a_i + B^m + C^m y, a_J + B^m + C^m y), \quad \forall y.$$

Step 2 Stability

The stability is straightforward by substituting the above Markovian Nash equilibrium into the state equation, it yields

$$\dot{y} = (a_i + a_J + 2B^m) + (2C^m - \delta)y \quad \forall t \geq T$$

with $y(T)$ coming from the first period cooperation and T unknown. The explicit solution is thus straightforward as given in the Proposition. Furthermore, it is easy to obtain that long-run steady state

$$y^m(t) = (y(T) - \widehat{y}^m)e^{(2C^m - \delta)t} + \widehat{y}^m (> 0).$$

Given $2C^m - \delta < 0$, for any $y(T)$, the trajectory asymptotically converges to this steady state.

Step 3. Proof of Corollary 2

We can easily rewrite two steady states of pollution as

$$y^* = \frac{a_i + a_J + 2B}{\delta - 2C} = (a_i + a_J) \frac{r + \delta}{(\delta - 2C)(r + \delta - 2C)}$$

and

$$\widehat{y}^m = \frac{a_i + a_J + 2B^m}{\delta - 2C^m} = (a_i + a_J) \frac{r + \delta - C^m}{(\delta - 2C^m)(r + \delta - 3C^m)}.$$

Therefore, in order to compare which steady state yields higher pollution, we only need to compare

$$h_1 \equiv \frac{r + \delta}{(\delta - 2C)(r + \delta - 2C)} = \frac{r + \delta}{\delta(r + \delta) - 2(r + 2\delta)C + 4C^2}$$

and

$$h_2 \equiv \frac{r + \delta - C^m}{(\delta - 2C^m)(r + \delta - 3C^m)} = \frac{r + \delta - C^m}{\delta(r + \delta) - 2(r + 2\delta)C^m + 6(C^m)^2 - \delta C^m}.$$

It is easy to see that

$$C < C^m < 0, \quad \forall r, \delta, b > 0.$$

Thus,

$$-C > -C^m > 0, \quad C^2 > (C^m)^2 > 0,$$

and

$$-2C(r + 2\delta) + 4C^2 > -2C^m(r + 2\delta) + 4(C^m)^2.$$

So it is straightforward that

$$h_1 = \frac{r + \delta}{\delta(r + \delta) - 2C(r + 2\delta) + 4C^2} < \frac{r + \delta}{\delta(r + \delta) - 2C^m(r + 2\delta) + 4(C^m)^2}.$$

Furthermore, simple algebra yields that

$$\frac{r + \delta}{\delta(r + \delta) - 2C^m(r + 2\delta) + 4(C^m)^2} < \frac{r + \delta - C^m}{\delta(r + \delta) - 2(r + 2\delta)C^m + 6(C^m)^2 - \delta C^m} = h_2.$$

Hence,

$$h_1 < h_2 \text{ and } y^* < \widehat{y^m}, \quad \forall r, \delta, b > 0.$$

That completes the proof.

A.2 The first order condition

In this section, we obtained the derivative of second period's welfare with respect to time, (15), and the first order condition at the same time.

Suppose $rA_i^m > B^m(a_i + a_J + 2B)$ and Assumption 1 holds. The switching time T must be given by the FOC

$$\alpha \frac{dW(T)}{dT} + \frac{dW_{i,II}^m}{dT} = 0,$$

provided the second order sufficient condition holds:

$$\alpha \frac{d^2W(T)}{dT^2} + \frac{d^2W_{i,II}^m(T)}{dT^2} < 0.$$

By the definition,

$$\begin{aligned} W_{i,II}^m &= \int_T^{+\infty} e^{-rt} \left(a_i x_i - \frac{x_i^2}{2} - \frac{by^2}{2} \right) dt \\ &= \int_T^{+\infty} \frac{e^{-rt}}{2} \left(a_i^2 - (B^m)^2 - 2B^m C^m y(t) - \left((C^m)^2 + b \right) y(t)^2 \right) dt \end{aligned}$$

where $y(t)$ is a function of $y^m(t)$, which depends on T . Direct calculation yields:

$$\begin{aligned} \frac{dW_{i,II}^m}{dt} &= \frac{e^{-rT}}{2} \left[-a_i^2 + (B^m)^2 + 2B^m C^m y(T) - \left((C^m)^2 + b \right) y(T)^2 \right] \\ &\quad + \int_T^{+\infty} \frac{e^{-rt}}{2} \left[-2B^m C^m \frac{\partial y^m}{\partial T} - 2 \left((C^m)^2 + b \right) y^m \frac{\partial y^m}{\partial T} \right] dt. \end{aligned}$$

In order to obtain explicit result, we try to get rid of the term $\frac{\partial y^m}{\partial T}$ in the above first order derivative.

Let $V_j^\sigma(y)$ be the value function of Player j in Mode σ when the value of the state variable is y , for $j = i, J$ and $\sigma = 1, 2$, where $\sigma = 1$ represents the mode before splitting, and $\sigma = 2$ after splitting. In Mode 1 the players have the unchangeable Markovian strategies

$$x_j = a_j + B + Cy \quad \text{for } j = i, J.$$

In addition, Player i has the impulse control in Mode 1 to exit the bloc. In Mode 2, the players

have the value functions $V_j^2(y) = U_j(y)$ for $j = i, J$. The optimal strategy of Player i 's impulse control results in maximization of the value $V_i^1(y)$ for $y < y^m$, where y^m is the value of y when Player i breaks up with the bloc. Hence, we have

$$V_i^1(y^m) = V_i^2(y^m) \equiv U_i(y^m) \quad (30)$$

and $\partial_{y^m} V_i^1(y) = 0$ for any $y < y^m$. Since in Mode 1, Player i has the instantaneous utility

$$\alpha \left[a_i x_i + a_J x_J - \frac{x_i^2 + x_J^2}{2} - b y^2 \right] = \alpha \left[a_i^2 + a_J^2 - B^2 - 2BCy - (C^2 + b) y^2 \right],$$

and the equation of dynamics is

$$\dot{y} = x_i + x_J - \delta y \equiv a_i + a_J + 2B + (2C - \delta) y.$$

The HJB equation for V_i^1 is

$$rV_i^1 = \alpha \left[a_i^2 + a_J^2 - B^2 - 2BCy - (C^2 + b) y^2 \right] + \frac{dV_i^1}{dy} \left[a_i + a_J + 2B + (2C - \delta) y \right] \quad \text{for } y < y^m$$

subject to the terminal condition (30). The solution can be written in the integral form

$$\begin{aligned} V_i^1(y) &= U_i(y^m) + \int_{y^m}^y \frac{dV_i^1(z)}{dy} dz \\ &= U_i(y^m) + \int_{y^m}^y \frac{rV_i^1(z) - \alpha \left[a_i^2 + a_J^2 - B^2 - 2BCz - (C^2 + b) z^2 \right]}{a_i + a_J + 2B + (2C - \delta) z} dz \end{aligned}$$

for $y < y^m$. Differentiating the both sides with respect to y^m and using the condition $\partial_{y^m} V_i^1(y) = 0$ we find

$$0 = U_i'(y^m) - \frac{rV_i^1(y^m) - \alpha \left[a_i^2 + a_J^2 - B^2 - 2BCy^m - (C^2 + b) (y^m)^2 \right]}{a_i + a_J + 2B + (2C - \delta) y^m}.$$

Since

$$\begin{aligned} V_i^1(y^m) &= U_i(y^m) = A_i^m + B^m y^m + \frac{C^m}{2} (y^m)^2 \quad \text{and} \\ U_i'(y^m) &= B^m + C^m y^m, \end{aligned}$$

we obtain that the first order condition is equivalent to

$$\begin{aligned} (B^m + C^m y^m) [a_i + a_J + 2B + (2C - \delta) y^m] &= r \left(A_i^m + B^m y^m + \frac{C^m}{2} (y^m)^2 \right) \\ &\quad - \alpha \left[a_i^2 + a_J^2 - B^2 - 2BCy^m - (C^2 + b) (y^m)^2 \right], \end{aligned}$$

and

$$\begin{aligned} \frac{dW_{i,II}^m}{dT} &= \left\{ -r \left(A_i^m + B^m y(T) + \frac{C^m}{2} y(T)^2 \right) \right. \\ &\quad \left. + (B^m + C^m y(T)) (a_i + a_J + 2B + (2C - \delta) y(T)) \right\} e^{-rT}. \end{aligned}$$

Combining with (12), we obtain

$$\alpha \frac{dW(T)}{dT} + \frac{dW_{i,II}^m}{dT} = e^{-rT} \left\{ \Lambda y^m(T)^2 + \Sigma y^m(T) + \Gamma \right\}$$

where the coefficients are given in (17).

That completes the proof.

A.3 Proof of Proposition 3

In the last subsection, the above first order condition can be rewritten as the following second degree polynomial in term of $y(T)$:

$$\Lambda y^m(T)^2 + \Sigma y^m(T) + \Gamma = 0.$$

The roots, if exist, they are given by

$$y^m(T) = \frac{-\Sigma \pm \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda}. \quad (31)$$

Given parameters Λ, Σ and Γ are independent of switching time T , the second order sufficient condition holds if and only if

$$(2\Lambda y(T) + \Sigma) y'(T) < 0.$$

Given the assumption that pollution accumulation is increasing over time, that is, $y'(T) > 0$ is true always, then the second order sufficient condition holds if and only if

$$2\Lambda y(T) + \Sigma < 0. \quad (32)$$

Combing the second order condition (32) and the explicit solution (18), it follows that

$$2\Lambda y(T) + \Sigma = \pm \sqrt{\Sigma^2 - 4\Lambda\Gamma} < 0$$

if and only if the negative sign is taken in the explicit solution (18). Taking into account that only

positive pollution level is possible, then

$$y^m(T) = -\frac{\Sigma + \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda} > 0,$$

which is true if $\Lambda < 0$, $\Gamma > 0$ (these two inequality implicitly guarantee the existence of real positive solution from FOC) and regardless the sign of Σ . Condition $\Lambda < 0$ is equivalent to

$$\alpha > \frac{C^m (2C - \delta - r/2)}{C^2 + b} \equiv G(b)$$

and $\Gamma > 0$ if and only if

$$\alpha > \frac{2[rA_i^m - B^m(a_i + a_J + 2B)]}{a_i^2 + a_J^2 - 2B^2} \equiv F(a_J, a_i, b).$$

To finish the proof, from the explicit solution,

$$y(T) = (y_0 - y^*)e^{(2C-\delta)T} + y^* = y^m.$$

rearranging terms, it yields that

$$T = \frac{1}{2C - \delta} \ln \left(\frac{y^m - y^*}{y_0 - y^*} \right).$$

Recall Assumption 1 guarantees that $y_0 < y(T) < y^*$, thus, $0 < \frac{y^m - y^*}{y_0 - y^*} < 1$ and

$$T \in (0, +\infty).$$

That completes the proof.

A.4 Profs of Corollaries 4 and 5.

We first prove

Lemma 1 *Let the assumptions of Proposition 3 be satisfied. Then the following properties hold.*

1. For a_J and a_i that satisfies (25), relation

$$F\left(\frac{a_J}{a_i}, b\right) \geq G(b)$$

holds for all $b \geq 0$.

2. For a_J and a_i that satisfies

$$\frac{\sqrt{3} - 2 + \sqrt{12 - 2\sqrt{3}}}{2 + \sqrt{3}} < \frac{a_J}{a_i} < 1, \quad (33)$$

there is $b^*(a_J/a_i)$ such that

$$F\left(\frac{a_J}{a_i}, b\right) > G(b) \quad \text{if and only if } b < b^*\left(\frac{a_J}{a_i}\right).$$

Hence, (22) holds if

$$\begin{aligned} F\left(\frac{a_J}{a_i}, b\right) &< \alpha && \text{for } b < b^*\left(\frac{a_J}{a_i}\right) \text{ and} \\ G(b) &< \alpha && \text{for } b > b^*\left(\frac{a_J}{a_i}\right). \end{aligned}$$

3. For a_J and a_i that satisfies

$$\frac{a_J}{a_i} \geq 1 \quad (34)$$

relation

$$F\left(\frac{a_J}{a_i}, b\right) \leq G(b)$$

holds for all $b > 0$.

Proof. Part 1. We rewrite $F(a_J, a_i, b)$ as

$$F\left(\frac{a_J}{a_i}, b\right) = \frac{1 + \left[\frac{3(C^m)^2}{(r+\delta-3C^m)^2} - \frac{4CC^m}{(r+\delta-3C^m)(r+\delta-2C)} \right] \left(\frac{a_J}{a_i} + 1\right)^2}{\left(\frac{a_J}{a_i}\right)^2 + 1 - \left[\frac{2C^2}{(r+\delta-2C)^2} \right] \left(\frac{a_J}{a_i} + 1\right)^2}.$$

For shorter notation, denote $x = \frac{a_J}{a_i}$, $H = \frac{3(C^m)^2}{(r+\delta-3C^m)^2}$, $K = \frac{4CC^m}{(r+\delta-3C^m)(r+\delta-2C)}$ and $L = \frac{2C^2}{(r+\delta-2C)^2}$, then

$$F(x, b) = F\left(\frac{a_J}{a_i}, b\right) = \frac{1 + (1+x)^2(H-K)}{x^2 + 1 - (1+x)^2L}, \quad (35)$$

with $H - K < 0$ and $L < 1/2$. Thus, the condition on α can be shorten as:

$$\max\{G(b), F(x, b)\} < \alpha < 1.$$

Straightforward algebra yields that $\forall r, \delta > 0$,

$$\lim_{b \rightarrow 0} G(b) = \frac{1}{2}, \quad \lim_{b \rightarrow +\infty} G(b) = \frac{1}{\sqrt{3}}, \quad \text{and} \quad \frac{dG(b)}{db} > 0, \quad \lim_{b \rightarrow 0} \frac{dG(b)}{db} = 0,$$

thus

$$G(b) \in \left(\frac{1}{2}, \frac{1}{\sqrt{3}} \right), \quad \forall b > 0.$$

Again straightforward, though cumbersome, algebra yield that

$$\lim_{b \rightarrow 0} F(x, b) = \frac{1}{x^2 + 1}, \quad \lim_{b \rightarrow +\infty} F(x, b) = \frac{1 - (1+x)^2/3}{x^2 + 1 - (x+1)^2/2}. \quad (36)$$

Furthermore,

$$\frac{\partial F(x, b)}{\partial b} > 0 \quad \text{if } x < \frac{\sqrt{2}}{2}, \quad \frac{\partial F(x, b)}{\partial b} < 0 \quad \text{if } x > \sqrt{\frac{3}{5}}$$

for all $b > 0$ and there is $\hat{b} > 0$ such that

$$\frac{\partial F(x, b)}{\partial b} = \begin{cases} > 0, & \text{if } 0 < b < \hat{b}, \\ < 0, & \text{if } b > \hat{b}, \end{cases} \quad \text{if } \frac{\sqrt{2}}{2} \leq x \leq \sqrt{\frac{3}{5}}. \quad (37)$$

Note that

$$\frac{1 - (1+x)^2/3}{x^2 + 1 - (x+1)^2/2} > \frac{1}{\sqrt{3}}$$

if and only if (25) holds with $x = a_J/a_i$. Therefore

$$F(x, b) > \frac{1}{\sqrt{3}} \geq G(b) \quad \text{for any } b > 0$$

if (25) holds. This proves Part 1.

Part 2. Since $x = a_J/a_i$ satisfies the reversed inequality in (25),

$$\lim_{b \rightarrow \infty} F(x, b) = \frac{1 - (1+x)^2/3}{x^2 + 1 - (x+1)^2/2} < \frac{1}{\sqrt{3}}.$$

In addition

$$\lim_{b \rightarrow 0} F(x, b) = \frac{1}{x^2 + 1} > \frac{1}{2} = \lim_{b \rightarrow 0} G(b),$$

by the intermediate value theorem, there is a $b^*(x)$ such that

$$G(b) = F(x, b^*(x)).$$

We show that $b^*(x)$ is the only solution to the above equation and

$$\begin{cases} G(b) < F(x, b), & \text{if } 0 < b < b^*(x), \\ G(b) > F(x, b), & \text{if } b > b^*(x). \end{cases} \quad (38)$$

Since

$$\frac{1}{\sqrt{2}} < \frac{\sqrt{3} - 2 + \sqrt{12 - 2\sqrt{3}}}{2 + \sqrt{3}} < \sqrt{\frac{3}{5}},$$

$F(x, b)$ is either decreasing in b for all b or there is a $\hat{b} > 0$ such that (37) holds if

$$x > \frac{\sqrt{3} - 2 + \sqrt{12 - 2\sqrt{3}}}{2 + \sqrt{3}}.$$

In the former case, since $G(b)$ is increasing, it is obvious that $b^*(x)$ is the only solution and (38) holds. In the latter case, $F(x, b)$ is bell-shaped. Note that

$$\lim_{b \rightarrow 0} F(x, b) = \frac{1}{x^2 + 1} > \frac{5}{8} > \frac{1}{\sqrt{3}} = \lim_{b \rightarrow \infty} G(b) \quad \text{if } x < \sqrt{\frac{3}{5}},$$

it follows that

$$F(x, b) \geq \lim_{b \rightarrow 0} F(x, b) > \lim_{b \rightarrow \infty} G(b) \quad \text{if } b < \hat{b}.$$

Therefore, $b^*(x) > \hat{b}$. Since $F(x, b)$ is decreasing in b for $b > \hat{b}$, we again find $b^*(x)$ is the only solution and (38) holds.

Part 3. Since $x \geq 1$, it follows that

$$\lim_{b \rightarrow 0} F(x, b) = \frac{1}{x^2 + 1} \leq \frac{1}{2} = \lim_{b \rightarrow 0} G(b).$$

Furthermore, since $x > \sqrt{3/5}$, $F(x, b)$ is decreasing in b . Hence, since $G(b)$ is increasing in b , we find

$$F(x, b) \leq \lim_{b \rightarrow 0} F(x, b) \leq \lim_{b \rightarrow 0} G(b) \leq G(b).$$

This completes the proof.

Now, Corollary 4 follows directly from Part 3 of Lemma 1, and Corollary 5 follows directly from Part 1 of Lemma 1.

A.5 Proof of Theorem 1

We first show that y is a positive real number. Since $C < 0$, it follows that

$$|B| = \frac{(a_i + a_j)|C|}{r + \delta + 2|C|} < \frac{a_i + a_j}{2}.$$

Hence,

$$a_i^2 + a_j^2 - 2B^2 > a_i^2 + a_j^2 - \frac{(a_i + a_j)^2}{2} = \frac{1}{2}(a_i - a_j)^2 \geq 0.$$

Therefore the numerator in (13) is positive. This proves the assertion.

Note that Λ , Σ , Γ and y^m all depend on α . We use $\Lambda(\alpha)$, $\Sigma(\alpha)$, $\Gamma(\alpha)$, and $y^m(\alpha)$ to indicate the dependence.

By Proposition 3, $y^m(\alpha)$ is a solution to the quadratic equation

$$\Lambda(\alpha)(y^m)^2 + \Sigma(\alpha)y^m + \Gamma(\alpha) = 0.$$

Let

$$F(y, \alpha) = \Lambda(\alpha)y^2 + \Sigma(\alpha)y + \Gamma(\alpha).$$

Then, $F(y^m(\alpha), \alpha) = 0$ and

$$\frac{dy^m(\alpha)}{d\alpha} = -\frac{F_\alpha(y^m(\alpha), \alpha)}{F_y(y^m(\alpha), \alpha)}. \quad (39)$$

By differentiation,

$$\begin{aligned} F_\alpha(y^m(\alpha), \alpha) &= \Lambda_\alpha(\alpha)(y^m(\alpha))^2 + \Sigma_\alpha(\alpha)y^m(\alpha) + \Gamma_\alpha(\alpha) \\ &= -(C^2 + b)y^m(\alpha)^2 - 2BCy^m(\alpha) + \frac{1}{2}(a_i^2 + a_j^2 - 2B^2), \\ F_y(y^m(\alpha), \alpha) &= 2\Lambda(\alpha)y^m(\alpha) + \Sigma(\alpha). \end{aligned}$$

It is shown in the Proof of Proposition 3 that $2\Lambda(\alpha)y^m(\alpha) + \Sigma(\alpha) < 0$ (see (21)). Note that \underline{y} is the only positive solution to the quadratic equation

$$-(C^2 + b)y^2 - 2BCy + \frac{1}{2}(a_i^2 + a_j^2 - 2B^2) = 0.$$

Hence, since $y^m(\alpha_0) < \underline{y}$,

$$F_\alpha(y^m(\alpha_0), \alpha_0) = -(C^2 + b)(y^m(\alpha_0))^2 - 2BCy^m(\alpha_0) + \frac{1}{2}(a_i^2 + a_j^2 - 2B^2) > 0.$$

It follows that $dy^m/d\alpha > 0$ at α_0 . Hence $y^m(\alpha)$ is nondecreasing in α in a neighborhood of α_0 .

This proof for the other case is similar. This completes the proof.

A.6 Proof of Corollary6

Suppose there are points $\hat{\alpha}, \tilde{\alpha} \in (\alpha_1, \alpha_2)$ such that $y^m(\hat{\alpha}) < \underline{y} < y^m(\tilde{\alpha})$. Then $y^m(\alpha)$ is increasing in α at $\hat{\alpha}$ and it is decreasing in α at $\tilde{\alpha}$. It is not possible that $\tilde{\alpha} < \hat{\alpha}$ because otherwise there is an $\bar{\alpha} < \bar{\alpha} < \hat{\alpha}$ such that $y^m(\bar{\alpha})$ is the minimum of $y^m(\alpha)$ between $\tilde{\alpha}$ and $\hat{\alpha}$. Thus $dy^m(\bar{\alpha})/d\alpha = 0$.

However, by (39), $F_\alpha(y^m(\bar{\alpha}), \bar{\alpha}) = 0$, which implies that $y^m(\bar{\alpha}) = \underline{y}$. Therefore, $y^m(\bar{\alpha}) = \underline{y} > y^m(\hat{\alpha})$, which is a contradiction. So it is necessary that $\hat{\alpha} < \tilde{\alpha}$. In this case there is $\bar{\alpha}$ such that

$\hat{\alpha} < \bar{\alpha} < \tilde{\alpha}$ and $y^m(\bar{\alpha})$ is the maximum of $y^m(\alpha)$ between $\hat{\alpha}$ and $\tilde{\alpha}$. Hence, again $dy^m(\bar{\alpha})/d\alpha = 0$ and we have $y^m(\bar{\alpha}) = \underline{y} < y^m(\tilde{\alpha})$. This is again a contradiction. Therefore no such points $\hat{\alpha}$ and $\tilde{\alpha}$ exist.

This completes the proof.

A.7 Optimal splitting time: interior vs corner solutions

The optimal splitting time, T , satisfies the equation

$$\alpha \frac{dW(T)}{dT} + \frac{dW_{i,II}^m(T)}{dT} = 0.$$

By (11) and (14), the left-hand side can be written as

$$e^{-rT} \left[\Lambda y^m(T)^2 + \Sigma y^m(T) + \Gamma \right]$$

where Λ , Σ , and Γ are given by (16). We define

$$\eta(y) = \Lambda y^2 + \Sigma y + \Gamma.$$

In the case where $\Lambda \neq 0$, $\eta(y)$ has two roots

$$y_1^m = \frac{-\Sigma - \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda}, \quad y_2^m = \frac{-\Sigma + \sqrt{\Sigma^2 - 4\Lambda\Gamma}}{2\Lambda}$$

which are real if

$$\Sigma^2 \geq 4\Lambda\Gamma,$$

and in the case where $\Lambda = 0$, $\eta(y)$ has one root

$$y_0^m = -\Gamma/\Sigma$$

provided that $\Sigma \neq 0$.

Using $G(b)$ and $F(a_J/a_i, b)$ defined in (19) and (20),

$$\Lambda < 0 \iff \alpha > G(b), \quad \Gamma > 0 \iff \alpha > F\left(\frac{a_J}{a_i}, b\right).$$

There are four possible cases.

Case 1:

$$\alpha > \max\{G(b), F(a_J/a_i, b)\}.$$

In this case, $\Lambda < 0$ and $\Gamma > 0$. So η has one positive root, y_1^m . Furthermore, $\eta(y) > 0$ if $y < y_1^m$ and $\eta(y) < 0$ if $y > y_1^m$. So if $y(0) < y_1^m$, coalition can be formed and lasts until $y(T) = y_1^m$, and if $y(0) \geq y_1^m$, coalition cannot be formed.

Case 2:

$$F(a_J/a_i, b) \leq \alpha \leq G(b).$$

In this case, $\Gamma \geq 0$ and $\Lambda \geq 0$. There are four subcases, (1) $\Lambda = 0$ and $\Sigma < 0$, (2) $\Lambda = 0$ and $\Sigma \geq 0$, (3) $\Lambda > 0$ and $\Sigma^2 \leq 4\Lambda\Gamma$, and (4) $\Lambda > 0$ and $\Sigma^2 > 4\Lambda\Gamma$.

In subcase (1), $\eta(y)$ is linear and has one positive root, y_0^m . Also, $\eta(y) > 0$ if $y < y_0^m$ and $\eta(y) < 0$ if $y > y_0^m$. So if $y(0) < y_0^m$ coalition lasts until $y(T) = y_0^m$, and if $y(0) \geq y_0^m$ coalition cannot be formed.

In subcases (2) and (3), $\eta(y) \geq 0$ for all $y > 0$. Therefore, coalition lasts forever.

In subcase (4), $\eta(y)$ has two positive roots y_1^m and y_2^m if $\Sigma < 0$. It is clear that $y_1^m < y_2^m$, and $\eta(y) > 0$ for $y < y_1^m$ or $y > y_2^m$, and $\eta(y) < 0$ for $y_1^m < y < y_2^m$. So if $y(0) < y_1^m$, coalition continues until pollution reaches y_1^m . If $y_1^m \leq y(0) \leq y_2^m$, coalition cannot be formed, and if $y(0) > y_2^m$, coalition lasts forever.

Case 3:

$$G(b) \leq \alpha \leq F(a_J/a_i, b).$$

In this case $\Lambda \leq 0$ and $\Gamma \leq 0$. There are four subcases, (1) $\Lambda = 0$ and $\Sigma \leq 0$, (2) $\Lambda = 0$ and $\Sigma > 0$, (3) $\Lambda < 0$ and $\Sigma^2 \leq 4\Lambda\Gamma$, and (4) $\Lambda < 0$ and $\Sigma^2 > 4\Lambda\Gamma$.

In subcases (1) and (3), $\eta(y)$ is nonpositive for all $y > 0$. So coalition cannot be formed.

In subcase (2) $\eta(y) < 0$ if $y < y_0^m$ and $\eta(y) > 0$ if $y > y_0^m$. So if $y(0) < y_0^m$ coalition cannot be formed, and if $y(0) > y_0^m$ coalition lasts forever.

In subcase (4), both y_1^m and y_2^m are nonnegative, and $\eta(y) < 0$ for $y < y_1^m$ or $y > y_2^m$ and $\eta(y) > 0$ for $y_1^m < y < y_2^m$. So if $y(0) < y_1^m$ or $y(0) > y_2^m$, coalition cannot be formed, and If $y_1^m \leq y(0) \leq y_2^m$, coalition continues until $y(T) = y_2^m$.

Case 4:

$$\alpha < \min \{G(b), F(a_J/a_i, b)\}.$$

In this case $\Gamma < 0$ and $\Lambda > 0$. So η has one positive root, y_2^m . Also, $\eta(y)$ changes from negative to positive as y passes through y_2^m . So if $y(0) < y_2^m$, coalition cannot be formed, and if $y(0) \geq y_2^m$, coalition lasts forever.

Note that Proposition 4 follows from this conclusion.

A.8 Proof of Proposition 5

By computation,

$$\underline{\alpha} = \frac{r + 2\delta + \sqrt{(r + 2\delta)^2 + 16b}}{2 \left[r + 2\delta + \sqrt{(r + 2\delta)^2 + 12b} \right]}. \quad (40)$$

It can be shown that the right-hand side is increasing in b . Hence

$$\frac{1}{2} < \underline{\alpha} < \frac{1}{\sqrt{3}} < 1.$$

We show that $\Sigma(\underline{\alpha}) > 0$. This is equivalent to showing that

$$B^m(2C - \delta - r) + C^m(a_i + a_J + 2B) > 2\hat{\alpha}BC. \quad (41)$$

By computation

$$\begin{aligned} B^m(2C - \delta - r) &= \frac{2b(a_i + a_J) \left[r + \sqrt{(r + 2\delta)^2 + 16b} \right]}{\left[r + \sqrt{(r + 2\delta)^2 + 12b} \right] \left[r + 2\delta + \sqrt{(r + 2\delta)^2 + 12b} \right]}, \\ C^m(a_i + a_J + 2B) &= \frac{-4b(a_i + a_J)(r + \delta)}{\left[r + \sqrt{(r + 2\delta)^2 + 16b} \right] \left[r + 2\delta + \sqrt{(r + 2\delta)^2 + 12b} \right]}, \\ BC &= \frac{32b^2(a_i + a_J)}{\left[r + \sqrt{(r + 2\delta)^2 + 16b} \right] \left[r + 2\delta + \sqrt{(r + 2\delta)^2 + 16b} \right]^2}. \end{aligned}$$

Substituting the above and (40) into (41) and cancel common factors, the inequality is equivalent to

$$\frac{\left(r + \sqrt{(r + 2\delta)^2 + 16b} \right)^2}{r + \sqrt{(r + 2\delta)^2 + 12b}} - \frac{16b}{r + 2\delta + \sqrt{(r + 2\delta)^2 + 16b}} > 2(r + \delta). \quad (42)$$

Note that

$$-\frac{16b}{r + 2\delta + \sqrt{(r + 2\delta)^2 + 16b}} = r + 2\delta - \sqrt{(r + 2\delta)^2 + 16b}.$$

The left-hand side of (42) can be written as

$$\frac{\left(r + \sqrt{(r+2\delta)^2 + 16b}\right)^2 + \left[r + 2\delta - \sqrt{(r+2\delta)^2 + 16b}\right] \left[r + \sqrt{(r+2\delta)^2 + 12b}\right]}{r + \sqrt{(r+2\delta)^2 + 12b}}. \quad (43)$$

Since

$$r + 2\delta - \sqrt{(r+2\delta)^2 + 16b} < 0, \quad \sqrt{(r+2\delta)^2 + 12b} < \sqrt{(r+2\delta)^2 + 16b},$$

it follows that

$$\begin{aligned} & \left[r + 2\delta - \sqrt{(r+2\delta)^2 + 16b}\right] \left[r + \sqrt{(r+2\delta)^2 + 12b}\right] \\ & > \left[r + 2\delta - \sqrt{(r+2\delta)^2 + 16b}\right] \left[r + \sqrt{(r+2\delta)^2 + 16b}\right] \\ & = \left[r^2 - (r+2\delta)^2 - 16b\right] + 2\delta \left[r + \sqrt{(r+2\delta)^2 + 16b}\right]. \end{aligned}$$

Thus, the numerator of the quotient in (43) is greater than

$$\begin{aligned} & r^2 + (r+2\delta)^2 + 16b + 2r\sqrt{(r+2\delta)^2 + 16b} + \left[r^2 - (r+2\delta)^2 - 16b\right] + 2\delta\sqrt{(r+2\delta)^2 + 16b} \\ & = 2r^2 + 2(r+\delta)\sqrt{(r+2\delta)^2 + 16b} + 2\delta r = 2(r+\delta) \left[r + \sqrt{(r+2\delta)^2 + 16b}\right] \\ & > 2(r+\delta) \left[r + \sqrt{(r+2\delta)^2 + 12b}\right]. \end{aligned}$$

As a result, the quotient in (43) is greater than $2(r+\delta)$. This proves (42), which is equivalent to $\Sigma(\underline{\alpha}) > 0$.

Since $\Sigma(\underline{\alpha}) > 0$, $\Lambda(\underline{\alpha}) = 0$ and $\Lambda(\alpha) < 0$ for $\alpha > \underline{\alpha}$, it follows from (23) that $y_T^m(\alpha)$ exists for $\alpha > \underline{\alpha}$ and is near $\underline{\alpha}$. In addition,

$$\lim_{\alpha \rightarrow \underline{\alpha}^+} y_T^m(\alpha) = \infty.$$

Hence, for $\alpha > \underline{\alpha}$ and is near $\underline{\alpha}$, $y_T^m(\alpha) > \underline{y}$. By Theorem 1 and corollary 6, $y_T^m(\alpha)$ is decreasing in α for all $\alpha > \underline{\alpha}$ at which $y_T^m(\alpha)$ exists. As α increases, $y_T^m(\alpha)$ continues to exist until either α reaches $\tilde{\alpha}$ if $\tilde{\alpha}$ is finite, or α reaches $\hat{\alpha}$, or α reaches 1, whichever is earlier. If $\tilde{\alpha} \leq \min\{\hat{\alpha}, 1\}$, then for any $\alpha \in (\underline{\alpha}, \tilde{\alpha})$

$$y_T^m(\alpha) > y^*. \quad (44)$$

By Proposition 3, $T = \infty$. If $\hat{\alpha} \leq \min\{\tilde{\alpha}, 1\}$, then (44) holds for $\alpha \in (\underline{\alpha}, \hat{\alpha})$. Hence, $T = \infty$. Finally, if $1 \leq \min\{\hat{\alpha}, \tilde{\alpha}\}$, (44) holds for $\alpha \in (\underline{\alpha}, 1)$. So, $T = \infty$.

This completes the proof.

B Heterogeneous strategies

In this section, we illustrate one “out of equilibrium” outcome in which one player re-optimizes its behaviour after the fall of a coalition but the remaining members don’t. The motivation draws from the facts that Canada withdrew from Kyoto Protocol on December 13, 2011, the U.S. ceased its participation in the 2015 Paris Agreement on climate change mitigation in 2017 and the United Kingdom withdrew from the European Union on January 31, 2020. In all three examples, the remaining coalition did not change its strategy: neither did so the CO₂ emission targets in the Kyoto or Paris agreements, nor the trading rules within the EU.

More precisely, suppose player J stays with her original commitment to the coalition after T , $x_J^*(t)$. Thus, the differential game is reduced after T^i to a standard optimal control problem for player i :

$$\max_{x_i} W_{II}^i \equiv \int_{T^i}^{+\infty} e^{-rt} \left(a_i x_i - \frac{x_i^2}{2} - \frac{b y^2}{2} \right) dt,$$

subject to

$$\dot{y} = x_i + x_J^* - \delta y, \quad \forall t \geq T^i,$$

with $y(T^i) = (y_0 - y^*)e^{(2C-\delta)T^i} + y^*$ and $x_J^*(y) = a_J + B + Cy$ given. The system is still autonomous and it is defined over an infinite time horizon. The same calculation as above yields the following:

Proposition 6 *Suppose that player i quits the coalition at time T^i , and that player J keeps her initial commitment. Then for any $t \geq T^i$, the optimal Markovian strategy of player i is*

$$x_i^i(y) = a_i + B^i + C^i y, \quad \forall y. \quad (45)$$

Furthermore, given the initial condition $y(T^i)$, the corresponding state variable $y^i(t)$ is given by

$$y^i(t) = [y(T) - \hat{y}^i] e^{(C^i + C - \delta)(t - T^i)} + \hat{y}^i \quad \forall t \geq T^i,$$

where \hat{y}^i is the asymptotically stable long-run steady state and it is given by

$$\hat{y}^i = \frac{a_i + a_J + B + B^i}{\delta - C - C^i},$$

and parameters

$$C^i = \frac{-2(C - \delta) - \sqrt{4(C - \delta)^2 + 4b(1 - r)}}{2(1 - r)} (< 0),$$

$$B^i = \frac{(a_i + a_J + B)C^i}{r + \delta - C - C^i} (< 0).$$

Obviously, since player J does not update her choice after the collapse of the coalition, the pair

(x_J^*, x_i^i) may not be a Nash equilibrium given that x_J^* may not be the optimal response from player J 's point of view, after the collapse of the coalition. Nonetheless, as discussed in Section 1, this is one possible choice among others.

Straightforwardly, the separation time T^i can be obtained via the same technique as the one of Proposition 3 with similar, though different, parameter conditions.

The detailed calculations of the separation time T^i , as well as the comparison between T^i and the Markovian splitting time, T , together with their respective impacts on welfare can be found in Boucekkine et al (2022).

References

- [1] Benchenkroun, H., G. Gaudet, N. Van Long (2006). Temporary natural resource cartels. *Journal of Environmental Economics and Management*, 52, 663–674.
- [2] Bertinelli L., C. Camacho and B. Zou (2014). Carbon capture and storage and transboundary pollution: A differential game approach. *European Journal of Operational Research*, 237, 721-728.
- [3] Bolton P., G. Roland and E. Spolaore (1996). Economic theories of the break-up and integration of nations. *European Economic Review*, 3(40), 697-705.
- [4] Boucekkine R., J. Krawczyk and T. Vallée (2011). Environmental quality versus economic performance: a dynamic game approach. *Optimal Control Applications and Methods* 32, 29-46.
- [5] Boucekkine R., Pommeret A. and F. Prieur (2013). Optimal regime switching and threshold effects. *Journal of Economic Dynamics and Control*, 37, 2979-2997.
- [6] Boucekkine, R., C. Camacho, W. Ruan and B. Zou (2022). Optimal coalition splitting with heterogenous strategies. DEM, DP 2022-13.
- [7] Bréchet, T., F. Gerard and H. Tulkens (2011). Efficiency vs. stability in climate coalitions: A conceptual and computational appraisal. *The Energy Journal*, 32, 49-75.
- [8] Bucher B. C. Carraro and I. Cersosimo (2002). Economic consequences of the US withdrawal from the Kyoto/Bonn Protocol. *Climate Policy*, 2 (4), 273-292.
- [9] Carraro, C., and D. Siniscalco (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, 52, 309-328.
- [10] Colombo, L., P. Labrecciosa, N. Van Long (2022). A dynamic analysis of international environmental agreements under partial cooperation. *European Economic Review* 143, 104036.

- [11] Dockner E. and N. Van Long (1993). International pollution control: cooperative versus noncooperative strategies. *Journal of Environmental Economics and Management* 24, 13-29.
- [12] Dockner E., S. Jorgensen, N. Van Long , and G. Sorger (2000). *Differential Games in Economics and Management*. Cambridge University Press.
- [13] Dutta, P., and R. Radner (2009). A strategic analysis of global warming: Theory and some numbers. *Journal of Economic Behavior and Organization*, 71, 187-209.
- [14] Hattori, T. (2001). Renegotiations in the post-agreement period: the future of the Kyoto protocol. *Interdisciplinary Environmental Review*, 3, 64-80.
- [15] Hoel, M. (1993). Intertemporal properties of an international carbon tax. *Resource and Energy Economics*, 15, 51-70.
- [16] Latorre M., Z. Olekseyuk and H. Yonezawa (2020). Trade and foreign direct investment-related impacts of Brexit. *The World Economy*. 43, 2–32.
- [17] Moser E., A. Seidel and G. Feichtinger (2014). History-dependence in production-pollution-trade-off models: a multi-stage approach. *Annals of Operations Research* 222, 455-481
- [18] Nong D. and M. Siriwardana (2018). Effects on the U.S. economy of its proposed withdrawal from the Paris Agreement: A quantitative assessment. *Energy*, 159, 621-629.
- [19] Petrosjan, L. (1977). “Stable Solutions of Differential Games with Many Participants,” *Viestnik of Leningrad University* 19, 46–52.
- [20] Saglam C. (2011). Optimal pattern of technology adoptions under embodiment: A multi-stage optimal control approach. *Optimal Control Applications and Methods* 32, 574-586
- [21] Sampson T. (2017). Brexit: The Economics of International Disintegration. *Journal of Economic Perspectives* 31 (4), 163–184.
- [22] Tahvonen, O. (1994). Carbon dioxide abatement as a differential game. *European Journal of Political Economy* 10, 685-705.
- [23] Tomiyama K. (1985). Two-stage optimal control problems and optimality conditions. *Journal of Economic Dynamics and Control* 9, 317-337
- [24] Tulkens, H. (1998). Cooperation vs. free riding in international environmental affairs: two approaches, in: N. Hanley and H. Folmer (eds), “Game Theory and the Environment”, Chapter 2, 30- 44, Edward Elgar, Cheltenham.
- [25] Van Long, N. (2010). A Survey on dynamic games in environmental economics. *Surveys on Theories in Economics and Business Administration*, 35-70. World Scientific.

- [26] Xepapadeas A. (1995). Induced technical change and international agreements under greenhouse warming. *Resource and Energy Economics*, 17(1), 1-23.
- [27] Zaccour, G. (2008). Time consistency in cooperative differential games: A tutorial. *INFOR: Information Systems and Operational Research*, 46(1), 81-92.
- [28] Zampolli F. (2006). Optimal monetary policy in a regime switching economy: the response to abrupt shifts in exchange rate dynamics. *Journal of Economic Dynamics and Control* 30, 1527–1567.
- [29] Zou B. (2016). Differential games with (a)symmetric payers and heterogeneous strategies. *Journal of Reviews on Global Economics* 5, 171-179.