

Non-linear incentive equilibrium strategies for a transboundary pollution differential game

Javier de Frutos and Guiomar Martín-Herrán

Abstract In this paper we apply non-linear incentive strategies to sustain over time an agreement. We illustrate the use of these strategies in a linear-quadratic transboundary pollution differential game. The incentive strategies are constructed in such a way that in the long run the pollution stock (the state variable) is close to the steady state of the pollution stock under the cooperative mode of play. The non-linear incentive functions depend on the emission rates (control variables) of both players and on the current value of the pollution stock. The credibility of the incentive equilibrium strategies is analyzed and the performance of open-loop and feedback incentive strategies is compared in their role of helping to sustain an agreement over time. We present numerical experiments to illustrate the results.

1 Introduction

This paper revisits one of the mechanisms already proposed in the literature to ensure the sustainability over time of an agreement reached at the starting date of a game. An agreement will last for its whole intended duration if, at any intermediate instant of time, each player stands to receive a greater payoff being part of the agreement rather than leaving it. A first approach proposed in the literature to sustain over time an agreement is the design of an agreement which is time-consistent or agreeable. In a time-consistent agreement the coordinated payoffs-to-go are greater than the non-cooperative ones along the cooperative state trajectory, and hence, no player finds it optimal to switch to his non-cooperative control at any intermedi-

Javier de Frutos

IMUVA and Dpto. Matemática Aplicada, Facultad de Ciencias, Universidad de Valladolid, Campus Miguel Delibes, Paseo Belén 7, 47011 Valladolid, Spain. e-mail: frutos@mac.uva.es

Guiomar Martín-Herrán

IMUVA and Dpto. Economía Aplicada, Facultad de Ciencias Económicas y Empresariales, Universidad de Valladolid, Avda. Valle Esgueva 6, 47011 Valladolid, Spain. e-mail: guiomar@eco.uva.es

ate instant of time. An agreement is agreeable if the comparison condition holds along any state trajectory. Time-consistent agreements are analyzed, among others, in Petrosjan (1997), Petrosjan & Zenkevich (1996), Petrosjan & Zaccour (2003) and Jørgensen & Zaccour (2001a, 2002); and agreeable agreements are studied, for example, in Kaitala & Pohjola (1990, 1995) and Jørgensen et al. (2003, 2005).

The second approach that can be generally called equilibrium approach is to embody the cooperative solution with an equilibrium property such that, by definition each player will find it individually rational to stick to his part of the coordinated solution. One option to build a cooperative equilibrium is to use so-called trigger strategies (Tolwinski et al. (1986), Haurie & Pohjola (1987)). These are strategies based on the past actions in the game and they include a threat to punish, credibly and effectively, any player who cheats on the agreement. These strategies are non-Markovian because they are based on all past information of the game evolution to the current time.

This paper examines the use of non-linear incentive strategies as another option to implement cooperative solutions by means of non-cooperative play. Ehtamo & Hämmäläinen (1986, 1989, 1993, 1995), Jørgensen & Zaccour (2001b), Martín-Herrán & Zaccour (2005, 2009) and De Frutos & Martín-Herrán (2015)), among others, propose incentive strategies to support the cooperative solution in two-player differential games. The incentive is designed in such a way that a coordinated outcome becomes a Nash equilibrium. Incentive strategies are functions of the possible deviation of the other player and recommend to each player to implement his part of the coordinated or agreed solution whenever the other player is doing so. One important characteristic of the incentive strategies is their credibility. The credibility property of incentive strategies requires that each player sticks to the agreed-upon incentive strategy and does not revert to the cooperative solution, even when the other player chooses to break the agreement. If the incentive strategies are credible no player will be tempted for unilaterally deviate from the agreed decision.

Incentive strategies have been extensively used in the differential games literature in different areas, especially environmental economics (Jørgensen & Zaccour (2001b), Martín-Herrán & Zaccour (2005, 2009), Breton et al. (2008), and De Frutos & Martín-Herrán (2015)) and marketing (Jørgensen & Zaccour (2003), Martín-Herrán & Taboubi (2005), Jørgensen et al. (2006), Buratto & Zaccour (2009), De Giovanni et al. (2016), Taboubi (2017) and De Giovanni (2018)). Many of these works have not addressed the analysis of the credibility of the incentive strategies. In general, this property cannot be studied analytically, even for games that belong to the class of solvable games such as linear-state and linear-quadratic differential games. All the papers previously cited, except De Frutos & Martín-Herrán (2015), study games belonging to these classes and the incentive strategies are constructed in such a way that the incentive equilibrium is the cooperative solution. Furthermore, the strategies are assumed to be linear and decision-dependent, i.e., each player makes his current decision contingent on the current decision of his opponent. When the credibility property is analyzed the common conclusion is that credibility is assured only for sufficiently small values of the deviation from the agreed solution.

This result led us to consider whether the definition of less restrictive strategies could help to guarantee the sustainability of an agreement over time. In De Frutos & Martín-Herrán (2015) we consider state-dependent and decision-dependent equilibrium strategies defined as non-linear functions of the control variables of both players and the current value of the state variable. More importantly, we look for an incentive strategy equilibrium such that the steady state of the optimal state trajectory is close enough but not necessarily identical to the steady state of the state variable under the cooperative mode of play. We show that the incentive equilibrium is credible in a larger region than the one associated with the usual linear incentive strategies.

The focus of this follow-up paper is to analyze nonlinear incentive strategies if the players use open-loop strategies instead of stationary Markovian strategies as previously assumed. We compare the performance of open-loop and feedback incentive strategies when maintaining an agreement over time and we study the credibility of the incentive strategies when one of the players deviates from the incentive equilibrium. The two information structures are compared for the well-known linear-quadratic transboundary pollution differential game proposed in Ploeg & Zeeuw (1992) and Dockner & Long (1993). We present numerical experiments to illustrate the results.

The rest of the paper is organized as follows. In Section 2 we briefly recall the formulation of the linear-quadratic transboundary pollution differential game, its cooperative solution, the open-loop non-cooperative Nash strategies, as well as the steady-state pollution stocks under cooperative and non-cooperative modes of play. In Section 3 we define the incentive strategies and equilibrium, and in Section 4 we analyze their credibility. Section 5 concludes.

2 A linear-quadratic transboundary pollution differential game

For simplicity in the exposition and with the objective of comparing our results with those obtained in De Frutos & Martín-Herrán (2015), we focus on a particular linear-quadratic model that has been extensively studied in the environmental economics literature. The formulation is borrowed from Ploeg & Zeeuw (1992) and Dockner & Long (1993). Let player i 's optimization problem be given by¹:

$$\max_{u_i} \left\{ W_i(u_1, u_2, x_0) := \int_0^{\infty} \left[u_i \left(A_i - \frac{1}{2} u_i \right) - \frac{1}{2} \varphi_i x^2 \right] e^{-\rho t} dt \right\} \quad (1)$$

$$\text{s.t.: } \dot{x} = \beta(u_1 + u_2) - \alpha x, \quad x(0) = x_0, \quad (2)$$

where β, A_i and $\varphi_i, i \in \{1, 2\}$ are positive parameters and $0 < \alpha < 1$. The control variable u_i is the emissions of player (country) i and the state variable x represents the accumulated stock of pollution and its dynamics is defined by the linear ordinary

¹ To simplify the notation we will drop the explicit dependence on the time variable when no confusion can arise.

differential equation (2), where parameter α denotes the natural absorption rate. The state dynamics says that the variation in the pollution stock level is the sum of emissions, scaled by parameter β , minus what is absorbed by nature. Assuming that emissions are a proportional by-product of industrial activities, the objective of player i is given by the difference between revenues from industrial activities and pollution damage costs. Function $u_i(A_i - \frac{1}{2}u_i)$ represents the concave revenue function of player i . Pollution induces damage costs, given by $\frac{1}{2}\varphi_i x^2$, assumed to depend on accumulated pollution. Parameter ρ is a positive constant discount rate.

If countries (players) use open-loop strategies they choose a time profile of actions at the beginning of the game and commit themselves to retain these preannounced profiles from the rest of the game. If countries choose state-dependent decision rules as their strategies, they choose emission strategies that are functions of the pollution stock. State-dependent Markovian strategies imply that, whenever country i makes a decision that results in a change in the pollution stock, country j immediately reacts. This action and reaction pattern implies more competitive behavior and the outcome of the game is further from the cooperative level

In this paper we assume that the players restrict their-selves to open-loop strategies (Haurie et al. (2012)), meaning that the players base their decisions only on time and an initial condition. An open-loop strategy selects the control action according to a decision rule μ_i , which is a function of the initial state x_0 : $u_i(t) = \mu_i(x_0, t)$. Because the initial state is fixed, there is no need to distinguish between $u_i(t)$ and $\mu_i(x_0, t)$. Using an open-loop strategy means that the player commits, at the initial time, to a fixed time path for his control, i.e., his choice of control at each instant of time is predetermined. More precisely, the set of admissible controls for Player i , $i = 1, 2$, \mathcal{U}_i is defined as the set of non negative absolutely continuous functions $u_i = u_i(t)$ defined in $\mathbb{R}_+ = [0, +\infty)$ with values in \mathbb{R}_+ such that if $(u_1, u_2) \in \mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$, the initial value problem (2) possesses a unique solution defined in \mathbb{R}_+ . The pair $(u_1^N, u_2^N) \in \mathcal{U}$ is an open-loop Nash equilibrium for the differential game (1)-(2) if

$$W_1(u_1^N, u_2^N) \geq W_1(u_1, u_2^N), \quad W_2(u_1^N, u_2^N) \geq W_2(u_1^N, u_2)$$

for all u_1 and u_2 such that $(u_1, u_2^N) \in \mathcal{U}$ and $(u_1^N, u_2) \in \mathcal{U}$.

If players agree to cooperate they solve an optimal control problem in which they jointly maximize the aggregate payoff

$$W_1 + W_2 = \int_0^\infty \sum_{i=1}^2 \left[u_i \left(A_i - \frac{1}{2}u_i \right) - \frac{1}{2}\varphi_i x^2 \right] e^{-\rho t} dt,$$

subject to dynamics (2). Martín-Herrán & Zaccour (2009) proved that the cooperative optimal controls read $u_i^c = \gamma_i + \beta(a^c x + b^c)$ where superscript c stands for cooperation and coefficients a^c and b^c are the quadratic and linear coefficients of the quadratic value function. These coefficients can be found in Martín-Herrán & Zaccour (2009) (page 272) and allow us to compute the steady-state pollution stock under cooperation:

$$x_{ss}^c = \frac{\beta(\rho + \alpha)(A_1 + A_2)}{(\rho + \alpha)\alpha + 2\beta^2(\varphi_1 + \varphi_2)}. \quad (3)$$

Next proposition characterizes the Nash equilibrium if the players do not cooperate and use open-loop strategies.

Proposition 1. *Assuming interior solutions, the pair $(u_1^N(t), u_2^N(t))$ is an open-loop Nash equilibrium of the differential game (1)-(2), where*

$$u_i^N(t) = A_i - \frac{\varphi_i x_{ss}^N}{\rho + \alpha} - \beta(x_0 - x_{ss}^N) \frac{\varphi_i}{\rho + \alpha - \xi} e^{\xi t},$$

$$\xi = \frac{\rho - \sqrt{(\rho + 2\alpha)^2 + (\varphi_1 + \varphi_2)\beta^2}}{2},$$

and superscript N stands for Nash equilibrium.

The optimal state trajectory is

$$x^N(t) = (x_0 - x_{ss}^N) e^{\xi t} + x_{ss}^N,$$

where x_{ss}^N denotes the steady state given by:

$$x_{ss}^N = \frac{(\rho + \alpha)\beta(A_1 + A_2)}{(\rho + \alpha)\alpha + \beta^2(\varphi_1 + \varphi_2)}. \quad (4)$$

Proof. We define the current-value Hamiltonian of player i

$$H^i(x, u_1, u_2, \lambda_i) = u_i \left(A_i - \frac{1}{2} u_i \right) - \frac{1}{2} \varphi_i x^2 + \lambda_i (\beta(u_1 + u_2) - \alpha x),$$

where λ_i is the i -th player costate variable associated with the state variable, x .

Assuming interior solution, the sufficient conditions for optimality derived from the Pontryagin maximum principle include

$$\begin{aligned} \frac{\partial H^i}{\partial u_i} &= A_i - u_i + \beta \lambda_i = 0, \\ \dot{x} &= \beta(u_1 + u_2) - \alpha x, \quad x(0) = x_0, \\ \dot{\lambda}_i &= \rho \lambda_i - \frac{\partial H^i}{\partial x} = (\rho + \alpha)\lambda_i + \varphi_i x, \quad \lim_{t \rightarrow \infty} e^{-\rho t} \lambda_i(t) = 0. \end{aligned}$$

Solving the system of linear ordinary differential equations and taking into account the initial and transversality conditions, the expressions of $u_i^N(t)$ and $x^N(t)$ in the statement can be easily derived. \square

As usual in this kind of models the non-cooperative solution leads to emission levels greater than those prescribed by the cooperative solution. The comparison of (3) and (4) clearly shows that the steady state of the pollution stock is lower if players cooperate than if they do not. Then, one can assume that if one player deviates from the cooperative solution, he is choosing an emission level greater than

that corresponding to the cooperative solution. The general main objective of this paper is the design of an incentive strategy implying that the players will not depart importantly from their part of the coordinated solution. Specifically, the long-run pollution stock will be near the long-run pollution stock attained under cooperation.

3 Incentive Equilibria

For the sake of completeness, let us recall the definition of an incentive equilibrium. The admissible incentive strategies for Player i are functions ψ_i defined in $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ such that for every $(u_1, u_2) \in \mathcal{U}$, $(\Psi_1(t), \Psi_2(t)) \in \mathcal{U}$ where $\Psi_i(t) = \psi_i(u_1(t), u_2(t), x(t))$, $i = 1, 2$, with $x(t)$ the solution of (2). We denote by Γ_i the set of admissible strategies for Player i .

Definition 1. A pair $\psi_1(v_1, v_2, x)$, $\psi_2(v_1, v_2, x)$ with $\psi_i \in \Gamma_i$, $i = 1, 2$, is an incentive equilibrium at $(u_1^*, u_2^*) \in \mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2$ iff for all $u_1 \in \mathcal{U}_1$ and $u_2 \in \mathcal{U}_2$,

$$W_1(u_1^*, u_2^*) \geq W_1(u_1, \psi_2(u_1, u_2^*, \hat{x})), \quad W_2(u_1^*, u_2^*) \geq W_2(\psi_1(u_1^*, u_2, \check{x}), u_2),$$

where \hat{x} and \check{x} satisfy $\dot{\hat{x}} = \beta(u_1 + \psi_2(u_1, u_2^*, \hat{x})) - \alpha\hat{x}$, and $\dot{\check{x}} = \beta(\psi_1(u_1^*, u_2, \check{x}) + u_2) - \alpha\check{x}$, respectively, with $\hat{x}(0) = \check{x}(0) = x_0$. Furthermore, $u_1^* = \psi_1(u_1^*, u_2^*, x^*)$, $u_2^* = \psi_2(u_1^*, u_2^*, x^*)$, where $\dot{x}^* = \beta(u_1^* + u_2^*) - \alpha x^*$, with $x^*(0) = x_0$.

An incentive equilibrium is thus characterized by the following pair of optimal control problems

$$\max_{u_i \in \mathcal{U}_i} W_i(u_i, u_j^*) = \int_0^\infty \left(u_i \left(A_i - \frac{1}{2} u_i \right) - \frac{1}{2} \phi_i x^2 \right) e^{-\rho t} dt, \quad (5)$$

$$\text{s.t.: } \dot{x} = \beta(u_i + \psi_j(u_i, u_j^*, x)) - \alpha x, \quad x(0) = x_0, \quad (6)$$

with $u_i^* = \arg \max_{u_i} W_i(u_i, u_j^*)$, $i, j = 1, 2$, $i \neq j$. The equilibrium condition $u_i^* = \psi_i(u_i^*, u_j^*, x^*)$, $i, j = 1, 2$, $i \neq j$, has to be satisfied.

The linear incentive strategies previously proposed in the literature are a particular case of Definition 1. This literature (except De Frutos & Martín-Herrán (2015)) assumes that the incentive equilibrium is the cooperative solution, (u_1^c, u_2^c) , and the following affine function has been usually proposed as incentive strategy (see, for example, Ehtamo & Hämäläinen (1986) and Martín-Herrán & Zaccour (2005, 2009)):

$$\psi_j(u_i, u_j, x) = \psi_j(u_i) = u_j^c + D_j(u_i - u_i^c), \quad i, j = 1, 2, i \neq j, \quad (7)$$

with D_j , $j = 1, 2$, denoting an appropriate non-zero constant.

However, in this paper, as in De Frutos & Martín-Herrán (2015), we look for an incentive strategy equilibrium (u_1^*, u_2^*) such that the steady state of the pollution stock of the system when the incentive strategies are used, x_{ss}^* , is greater but close to

this value under cooperation, x_{ss}^c , and lower than the steady-state value under non-cooperation, x_{ss}^N . In De Frutos & Martín-Herrán (2015) the players use feedback strategies,² while in this paper they use open-loop strategies and the results under both information structures are compared.

For the sake of completeness, let us recall the form of the incentive functions ψ_j , $j = 1, 2$, in Definition 1 we choose to attain the purpose:

$$\psi_j(u_i, u_j, x) = (u_j^c + D_j(u_i - u_i^c))\phi(x - x_{ss}^c, \varepsilon) + u_j(1 - \phi(x - x_{ss}^c, \varepsilon)), \quad (8)$$

where $\varepsilon > 0$ is a small positive parameter and $\phi(x, \varepsilon)$ is a smooth function satisfying

$$\phi(x, \varepsilon) = 0, \text{ if } x \leq \varepsilon; \quad \phi(x, \varepsilon) = 1, \text{ if } x \geq 2\varepsilon. \quad (9)$$

The definition of the non-linear incentive in (8) and the cutoff function in (9) show that the incentive strategy is exclusively implemented if one player deviates from the cooperative outcome (and emits at a greater level) and, therefore, at some time t the trajectory $x(t)$ is above x_{ss}^c . If $x(t)$ is far from x_{ss}^c , then the linear incentive in (7) applies and pushes the players' emissions in such a way that the pollution path returns close to the steady-state value under cooperation, x_{ss}^c . Conversely, the non-linear incentive strategy (8) when $x(t)$ is close enough to x_{ss}^c (the distance measured by parameter ε) allows the players to choose any time path.

The next proposition characterizes the incentive equilibrium if the non-linear incentive in (8) is used and the players restrict themselves to open-loop strategies.

Proposition 2. *If the incentive strategy is defined by (8), then the open-loop interior incentive equilibrium $(u_1^*(t), u_2^*(t))$ satisfies the equilibrium conditions $u_i^* = \psi_i(u_i^*, u_j^*, x^*)$, $i, j = 1, 2, i \neq j$ together with the following set of optimality conditions:*

$$\begin{aligned} u_i &= A_i + \beta(1 + D_j\phi(x - x_{ss}^c, \varepsilon))\xi_i, \\ \dot{x} &= \beta(u_i + \psi_j(u_i, u_j, x)) - \alpha x, \quad x(0) = x_0, \\ \dot{\xi}_i &= \left(\rho + \alpha - \beta \frac{\partial \psi_j}{\partial x}(u_i, u_j, x) \right) \xi_i + \varphi_i x, \quad \lim_{t \rightarrow \infty} e^{-\rho t} \xi_i(t) = 0, \quad i = 1, 2, i \neq j, \end{aligned} \quad (10)$$

where ξ_i denotes the costate variable of player i .

Proof. Let define the Hamiltonian for player i as

$$H_i(u_i, u_j, \xi_i, x) = u_i \left(A_i - \frac{1}{2} u_i \right) - \frac{1}{2} \varphi_i x^2 + \xi_i (\beta(u_i + \psi_j(u_i, u_j, x)) - \alpha x).$$

Assuming interior strategies, the maximum principle optimality conditions read:

² The pollution stock and emission time-paths for the Nash and incentive equilibria as well as the payoffs when the players use feedback strategies presented later in this paper have been taken from De Frutos & Martín-Herrán (2015).

$$\begin{aligned} \frac{\partial H_i}{\partial u_i}(u_i, u_j, \xi_i, x) &= 0, \\ \dot{\xi}_i &= \rho \xi_i - \frac{\partial H_i}{\partial x}, \quad \lim_{t \rightarrow \infty} e^{-\rho t} \xi_i(t) = 0, \quad i = 1, 2, i \neq j, \\ \dot{x} &= \beta(u_i + \psi_j(u_i, u_j, x)) - \alpha x, \quad x(0) = x_0. \end{aligned}$$

From these conditions, those in the statement of the proposition immediately follow. \square

The non-linear incentive equilibrium cannot be analytically characterized and numerical methods are required for the analysis of these incentive strategies. To numerically solve the system of optimality conditions (10) we first introduce a large $T > 0$ and substitute the transversality condition $\lim_{t \rightarrow \infty} e^{-\rho t} \xi_i(t) = 0$ by the approximate boundary condition $\xi_i(T) = 0$, $i = 1, 2$. The resulting boundary value problem is solved by means of a collocation method implemented in the MATLAB subroutine `bvp4c.m`, see Kierzenka & Shampine (2001). The procedure is repeated with a larger T until no differences between approximate solutions are found.

For illustration purposes we consider a symmetric example and fix the following values of the parameters: $A_1 = A_2 = 0.5$, $\varphi_1 = \varphi_2 = 1$, $\alpha = 0.2$, $\beta = 1$, $\rho = 0.1$. The threshold ε in the cut off function in (9) is set to $\varepsilon = 0.025$ and the initial pollution stock was set to $x_0 = 0$. The parameter D_j in (8) was set to $D_j = 1$, $j = 1, 2$ as in Martín-Herrán & Zaccour (2009).

Using the above parameters values, we have represented in Figure 1 the phase diagram of the optimality system (10) in the symmetric case. The variables are the pollution stock in the abscissas axis and emissions in the ordinates axis. The system possesses a unique steady state (represented by a circle in the figure) which is a saddle point. In the figure the stable variety is highlighted in red colour. This curve represents, for a given initial condition, the unique nonlinear symmetric open-loop incentive equilibrium. Note that, for simplicity, in the figure the positivity conditions on emissions have not been imposed.

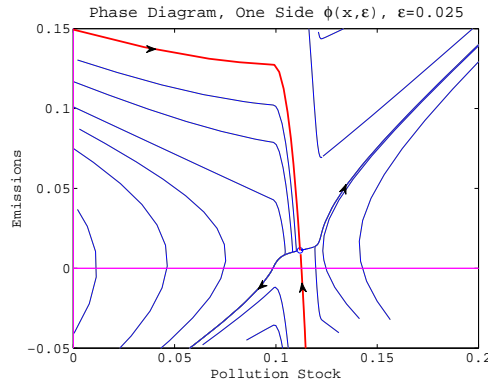


Fig. 1 Pollution stock-emissions phase diagram in the symmetric case. In red the stable variety of the unique equilibrium of (10). Parameters values $A_1 = A_2 = 0.5$, $\varphi_1 = \varphi_2 = 1$, $\alpha = 0.2$, $\beta = 1$, $\rho = 0.1$, $\varepsilon = 0.025$, $D_j = 1$, $j = 1, 2$.

Figure 2 shows the optimal pollution and emission time-paths for five different modes of play. The solid red lines represent the cooperative state (left) and control (right) optimal time-paths. The optimal time-paths corresponding to the non-cooperative Nash equilibrium are represented using solid (black) and dashed (black) lines for the feedback and open-loop information structures, respectively. Finally, solid (green) and dashed (green) lines are the optimal time-paths for the open-loop and feedback incentive equilibrium strategies, respectively. A first message from Figure 2 (left) is that the incentive strategies attain their objective of approaching the long-run cooperative level of the pollution stock. Furthermore, for the same threshold ε in the cutoff function the open-loop incentive equilibrium is closer to the cooperative optimal time-paths than the feedback incentive equilibrium. The transition to the steady state in the open-loop case is smoother than in the feedback case. This effect is clearer in the case of emissions (right chart). The main difference is in the short run. The emissions in the feedback incentive equilibrium are initially very high, while they are much lower in the open-loop case. However, after a very short period of time, the first ones decrease sharply, while the second ones begin their fall later. The times at which the emissions start to decline towards their stationary levels depend on the value of parameter ε in the cutoff function. The smaller ε , the earlier the emissions should begin their fall in order to attain a steady state closer to the cooperative steady-state. In the long run the emissions in both open-loop and feedback incentive equilibria are very similar to the cooperative emissions.

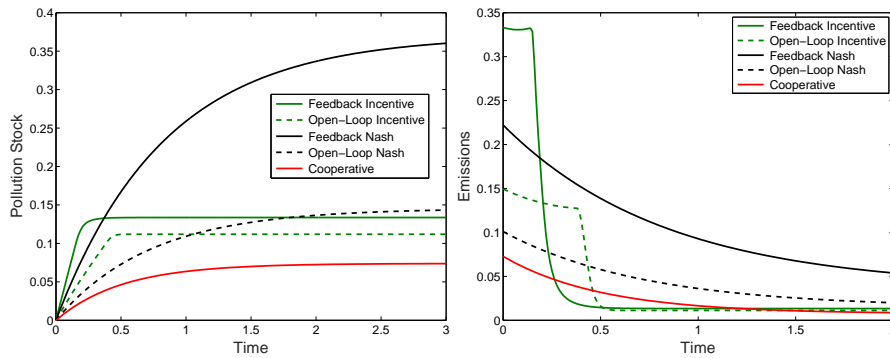


Fig. 2 Nash, Incentive and Cooperative pollution stock and emission time-paths

4 Credibility

In this section we focus on the study of the credibility of the incentive strategies. The incentive strategies are credible if player j deviates unilaterally from his incentive equilibrium action, $u_j = u_j^*(x)$, then, it will be more beneficial for player i to follow

the incentive strategy, rather than to stick to $u_i = u_i^*(x)$. Hence, if the credibility property is satisfied there will not be any temptation for unilateral deviation from the pair $u_j = u_j^*(x)$, $j = 1, 2$.

Definition 2. A pair of incentive equilibrium strategies $\psi_1(v_1, v_2, x)$, $\psi_2(v_1, v_2, x)$, with $\psi_i \in \Gamma_i$, $i = 1, 2$, is credible in a set $U_1 \times U_2 \subset \mathcal{U}_1 \times \mathcal{U}_2$ iff given $u_1 \in U_1$ and $u_2 \in U_2$ there exist $\hat{u}_1 \in \mathcal{U}_1$ and $\check{u}_2 \in \mathcal{U}_2$ such that

$$W_1(\psi_1(\hat{u}_1, u_2, \hat{x}), u_2) \geq W_1(u_1^*, u_2), \quad W_2(u_1, \psi_2(u_1, \check{u}_2, \check{x})) \geq W_2(u_1, u_2^*), \quad (11)$$

where \hat{x} and \check{x} satisfy $\dot{\hat{x}} = \beta(\psi_1(\hat{u}_1, u_2, \hat{x}) + u_2) - \alpha\hat{x}$ and $\dot{\check{x}} = \beta(u_1 + \psi_2(u_1, \check{u}_2, \check{x})) - \alpha\check{x}$, respectively, with $\hat{x}(0) = \check{x}(0) = x_0$.

A sufficient, although obviously not necessary, condition for credibility is that for all $u_1 \in U_1$ and $u_2 \in U_2$

$$W_1(\psi_1(u_1^*, u_2, \hat{x}), u_2) \geq W_1(u_1^*, u_2); \quad W_2(u_1, \psi_2(u_1, u_2^*, \check{x})) \geq W_2(u_1, u_2^*),$$

where \hat{x} and \check{x} are defined as in Definition 2, with $\hat{u}_1 = u_1^*$ and $\check{u}_2 = u_2^*$.

Note that, in the case of linear incentive strategies, ψ_j is given by (7). Then, Definition 2 reduces to $W_i(\psi_i(u_j), u_j) \geq W_i(u_i^*, u_j)$, $\forall u_j \in U_j$ with $u_i^* = u_i^c$, for $i = 1, 2$, which is the credibility definition usually proposed in the literature (see, for example, Jørgensen and Zaccour (2003), Martín-Herrán and Zaccour (2005, 2009)).

Definition 2 requires conditions (11) to be checked in some subset of admissible controls $U_1 \times U_2 \subset \mathcal{U}_1 \times \mathcal{U}_2$. In order to be able to analyze the credibility properties of the non-linear incentive strategies we assume that the set of possible deviations is restricted to convex combinations of the cooperative control u_i^c and the non-cooperative Nash equilibrium $u_i^{(n)}$: $U_i = \{u_i = \theta u_i^c + (1 - \theta)u_i^N\}$. Next tables allow us to illustrate the credibility of the non-linear incentive strategies and to compare this property for the open-loop and feedback information structures.

Tables 1-4 show the players' payoffs when they play either open-loop (first two columns) or feedback (last two columns) strategies (superscripts *ol* and *f*, respectively). Each row presents the different strategies used by the players in each case.

	W_1^{ol}	W_2^{ol}	W_1^f	W_2^f
$U_1 = u_1^N, U_2 = u_2^N$	3.14×10^{-3}	3.14×10^{-3}	-3.68×10^{-1}	-3.68×10^{-1}
$U_1 = u_1^*, U_2 = u_2^*$	1.68×10^{-2}	1.68×10^{-2}	3.91×10^{-4}	3.91×10^{-4}
$U_1 = u_1^*, U_2 = u_2^N$	8.30×10^{-4}	2.16×10^{-2}	-5.81×10^{-1}	3.74×10^{-1}
$U_1 = \psi_1(u_1^*, u_2^N, x), U_2 = u_2^N$	2.16×10^{-3}	2.39×10^{-3}	-3.72×10^{-1}	-3.79×10^{-1}

Table 1 Players' payoffs under open-loop and feedback strategies. Player 2 deviates to u_2^N

The first row in Table 1 shows the players' payoffs when they do not cooperate and play the open-loop or the feedback Nash equilibrium. The second row presents these payoffs when they follow their open-loop or feedback incentive strategies.

Both players and under both information structures improve their payoffs with respect to the non-cooperative Nash levels. In the third row the payoffs are no longer identical for both players, when player 1 continues to play his part of the incentive equilibrium, while player 2 deviates to his part of the non-cooperative Nash equilibrium. Let us note that it could be considered irrational that player 2 implements a strategy that would provide him a smaller payoff than that associated with the incentive strategy. Table 1 shows that the deviation from the incentive equilibrium to u_2^N provides a greater payoff to the deviating player, i.e. $W_2(u_1^*, u_2^N) > W_2(u_1^*, u_2^*)$. The credibility of the incentive equilibrium requires the existence of a feasible u_1 such that $W_1(\psi_1(u_1, u_2^N, x), u_2^N) > W_1(u_1^*, u_2^N)$. The fourth row in Table 1 shows that last inequality is satisfied in particular for $u_1 = u_1^*$. Furthermore, the deviating player (player 2) is penalized because $W_2(\psi_1(u_1^*, u_2^N, x), u_2^N) < W_2(u_1^*, u_2^*)$. In fact, deviating from the incentive strategy player 2 even worsens his payoff compared to his payoff in the non-cooperative case: $W_2(\psi_1(u_1^*, u_2^N, x), u_2^N) < W_2(u_1^N, u_2^N)$. All these results apply when the players use either open-loop or feedback strategies.

Table 1 also allows us to compare the relative improvement of the players' payoffs when moving from their non-cooperative mode of play to the incentive equilibrium as well as the relative loss with respect to the cooperative solution both under open-loop and feedback strategies. Because we are analyzing a completely symmetric game, each player under cooperation receives half of the total cooperative payoff. The individual cooperative payoff is given by $W_1^c = W_2^c = 2.64 \times 10^{-2}$. Let us denote by W_i^{Nol} and W_i^{Nf} the players' payoffs when the open-loop and feedback Nash equilibrium, respectively, is played (first row in Table 1). Equivalently, let us denote by W_i^{*ol} and W_i^{*f} the players' payoffs when the open-loop and feedback incentive equilibrium, respectively, is implemented (second row in Table 1).

The comparison of the Nash equilibrium payoffs W_i^{Nol} and W_i^{Nf} with the individual cooperative payoff W_i^c shows that if the players implement the feedback Nash equilibrium the payoffs are much smaller than if the open-loop Nash equilibrium is played, which is, already, a 88.1% lower than the cooperative ones. This loss of welfare is a well-known consequence of the non-cooperative mode of play that can be mitigated if the players agree to follow incentive strategies (both open-loop or feedback), as can be seen in the second row of Table 1. The payoffs when the incentive equilibrium is played compared to the payoffs under cooperation account for a fall of a 36.7% and a 98.5%, under open-loop and feedback strategies, respectively. If the incentive equilibrium payoffs are compared to the non-cooperative Nash payoffs they show an increase of more than one order of magnitude in both cases, being four times greater in the open-loop than in the feedback case. These results clearly show that the players can find neatly advantageous the use of non-linear incentive strategies.

Tables 2-4 present the different payoffs when player 2 deviates from the incentive equilibrium to different convex combinations of his cooperative control, u_2^c , and his part of the non-cooperative Nash equilibrium, u_2^N . In Tables 2, 3 and 4 the weight assigned to the non-cooperative part decreases from 0.75 to 0.5 and to 0.25, respectively. Concerning the rationality property of player 2, these tables show that it is rational that player 2 deviates from the incentive equilibrium when he changes to

a strategy in which no-cooperation is weighted at least as cooperation. However, Table 4 shows that it is irrational that player 2 deviates to $u_2^{(3)} = 0.75u_2^c + 0.25u_2^N$ or to $u_2^{(3)} = u_2^c$ not showed in the table. These results are applicable both for the open-loop and feedback incentive strategies.

The last two rows in Tables 2-4 allow to analyze the credibility of the incentive strategies when player 2 deviates to $u_2^{(i)}$, $i = 1, 2, 3$. Player 1 implements $U_1 = \psi_1(u_1, u_2^{(i)}, x)$ with either $u_1 = u_1^*$ (third row) or $u_1 = u_1^N$ (fourth row). Comparing the entries in the third (fourth) row with those in the second, it can be deduced that the incentive equilibrium is credible in the same scenarios when it is rational for player 2 to deviate (Tables 2 and 3) and is not credible when the deviation is irrational for player 2 (Table 4). The results are qualitatively similar for open-loop and feedback strategies.

Finally, from the comparison of $W_2(\psi_1(u_1, u_2^{(i)}, x), u_2^{(i)})$ and $W_2(u_1^*, u_2^*)$ for $u_1 \in \{u_1^*, u_1^N\}$ in Tables 2-3 we can deduce that deviating from the incentive strategy player 2 always worsens his payoff compared to his payoff in the incentive equilibrium. More precisely, if $u_2 = u_2^{(1)} = 0.25u_2^c + 0.75u_2^N$ both choices for player 1, $u_1 = u_1^*$ or $u_1 = u_1^N$ lead to a payoff for player 2 lower than $W_2(u_1^*, u_2^*)$ regardless of whether players use open-loop or feedback strategies. If $u_2 = u_2^{(2)} = 0.5u_2^c + 0.5u_2^N$, the same comment applies for the feedback case. In the open-loop case the choice $u_1 = u_1^*$ for player 1, although being credible, it does not penalize enough player 2 with respect to his payoff in the incentive equilibrium. However, if player optimizes his choice by moving to $u_1 = u_1^N$, then $W_1(\psi_1(u_1^N, u_2^{(2)}, x), u_2^{(2)}) > W_1(\psi_1(u_1^*, u_2^{(2)}, x), u_2^{(2)}) > W_1(u_1^*, u_2^*)$ and $W_2(\psi_1(u_1^N, u_2^{(2)}, x), u_2^{(2)}) < W_2(u_1^*, u_2^*)$.

	W_1^{ol}	W_2^{ol}	W_1^f	W_2^f
$U_1 = u_1^*, U_2 = u_2^*$	1.68×10^{-2}	1.68×10^{-2}	3.91×10^{-4}	3.91×10^{-4}
$U_1 = u_1^*, U_2 = u_2^{(1)}$	1.08×10^{-2}	2.01×10^{-2}	-3.58×10^{-1}	2.70×10^{-1}
$U_1 = \psi_1(u_1^*, u_2^{(1)}, x), U_2 = u_2^{(1)}$	1.23×10^{-2}	1.22×10^{-2}	-1.14×10^{-1}	-1.54×10^{-1}
$U_1 = \psi_1(u_1^N, u_2^{(1)}, x), U_2 = u_2^{(1)}$	1.36×10^{-2}	1.01×10^{-2}	-9.22×10^{-2}	-3.92×10^{-1}

Table 2 Players' payoffs under open-loop and feedback strategies. Player 2 deviates to $u_2^{(1)} = 0.25u_2^c + 0.75u_2^N$

	W_1^{ol}	W_2^{ol}	W_1^f	W_2^f
$U_1 = u_1^*, U_2 = u_2^*$	1.68×10^{-2}	1.68×10^{-2}	3.91×10^{-4}	3.91×10^{-4}
$U_1 = u_1^*, U_2 = u_2^{(2)}$	2.01×10^{-2}	1.76×10^{-2}	-1.54×10^{-1}	1.34×10^{-1}
$U_1 = \psi_1(u_1^*, u_2^{(2)}, x), U_2 = u_2^{(2)}$	2.02×10^{-2}	1.75×10^{-2}	-3.02×10^{-2}	-4.75×10^{-2}
$U_1 = \psi_1(u_1^N, u_2^{(2)}, x), U_2 = u_2^{(2)}$	2.26×10^{-2}	5.01×10^{-3}	1.23×10^{-1}	-4.46×10^{-1}

Table 3 Players' payoffs under open-loop and feedback strategies. Player 2 deviates to $u_2^{(2)} = 0.5u_2^c + 0.5u_2^N$

	W_1^{ol}	W_2^{ol}	W_1^f	W_2^f
$U_1 = u_1^*, U_2 = u_2^*$	1.68×10^{-2}	1.68×10^{-2}	3.91×10^{-4}	3.91×10^{-4}
$U_1 = u_1^*, U_2 = u_2^{(3)}$	2.88×10^{-2}	1.44×10^{-2}	5.29×10^{-2}	-4.87×10^{-2}
$U_1 = \psi_1(u_1^*, u_2^{(3)}, x), U_2 = u_2^{(3)}$	2.87×10^{-2}	1.45×10^{-2}	2.08×10^{-2}	-6.00×10^{-3}
$U_1 = \psi_1(u_1^N, u_2^{(3)}, x), U_2 = u_2^{(3)}$	3.12×10^{-2}	-1.41×10^{-4}	2.91×10^{-1}	-5.19×10^{-1}

Table 4 Players' payoffs under open-loop and feedback strategies. Player 2 deviates to $u_2^{(3)} = 0.75u_2^* + 0.25u_2^N$

Figure 3 for the open-loop strategies and Figure 4 for the feedback strategies show the time paths of the emission rates and the pollution stock for the non-cooperative, cooperative and incentive strategies used by the players, as well as the strategies when player 2 deviates to his part of the non-cooperative equilibrium, while player 1 sticks to his part of the incentive equilibrium. These time paths allow us to analyze the credibility of the incentive strategies. The optimal paths when the non-cooperative Nash equilibrium or the cooperative solution is played are represented using solid black and red lines, respectively. Those associated with the incentive equilibrium are depicted using solid blue line. Finally, the solid green line shows the time paths when player 2 deviates while player 1 plays the incentive strategy.

Figure 3 shows that when player 2 deviates from the incentive equilibrium and follows his part of the non-cooperative Nash equilibrium the emission time-paths are described by the black solid line starting around 0.1. When player 1 responds using the incentive strategy $U_1 = \psi(u_1^*, u_2^N, x)$ his emission time-path initially starts around 0.15 and follows the incentive equilibrium strategy, decreasing up to a minimum level. When player 1 realizes that the pollution stock is far from the pollution stock under cooperation departs from his part of the incentive equilibrium and increases the emissions up to a level similar to the non-cooperative level. From this point on the emission time-paths run very closely. This behavior translates into a pollution stock time-path associated with the strategies $U_1 = \psi(u_1^*, u_2^N, x)$ and $U_2 = u_2^N$ that

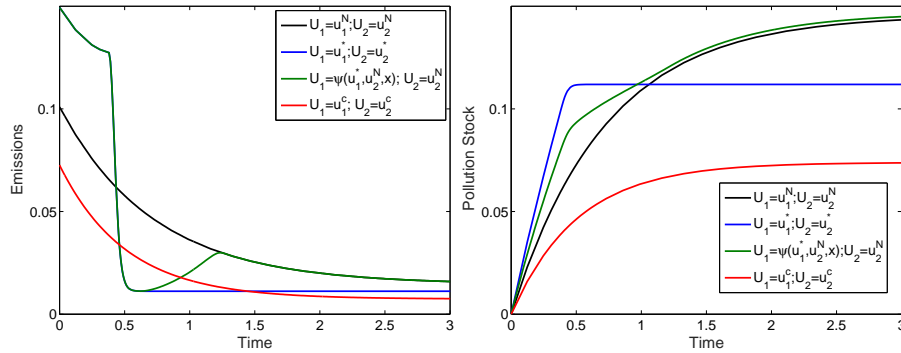


Fig. 3 Credibility open-loop incentive strategies

initially evolves between the time-paths for the non-cooperative Nash equilibrium and the incentive equilibrium. The time path initially increases at a speed greater than that corresponding to the Nash equilibrium. From a time on the incentive strategy used by player 1, $U_1 = \psi(u_1^*, u_2^N, x)$, allows to approximate the pollution stock time-path under non-cooperation.

Figure 4 collects the different feedback strategies. The deviating player (player 2) initially follows the discontinuous green line starting around 0.22, corresponding to the emission time-path for $U_1 = \psi(u_1^*, u_2^N, x)$ and $U_2 = u_2^N$. This time path evolves quite close to the non-cooperative emission time-path. Let us note that although in Figure 4 we display the time paths associated with the different strategies, the players are using feedback strategies, and as such they are taking their optimal decisions depending on the value of the state variable (the pollution stock). Hence, the discontinuous green line does not coincide with the continuous black line because player 1 is playing $U_1 = \psi(u_1^*, u_2^N, x)$ instead of u_1^N , implying a different value of the pollution stock. Player 1's emissions (continuous green line) start around 0.33 and sharply decrease imitating the incentive equilibrium time-path up to a point in time where the trajectory reaches the discontinuous green line. From this time on the emission time-paths for both players coincide and evolve close to the non-cooperative emission time-path. The main difference with respect to the previous case is that in the case of open-loop strategies player 1 reduces his emissions too sharply and then he has to raise them during an intermediate period of time to follow a trajectory similar to the non-cooperative case. However, in the case of feedback, the decrease in emissions towards values close to those of the non-cooperative scenario is monotonous and smoother. As the right chart in Figure 4 shows the pollution stock when $U_1 = \psi(u_1^*, u_2^N, x)$ and $U_2 = u_2^N$ is not far away from the non-cooperative pollution stock even in the short term.

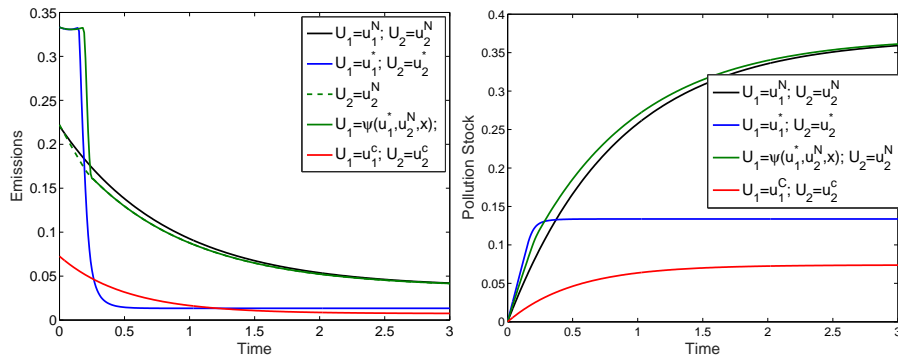


Fig. 4 Credibility feedback incentive strategies

5 Concluding remarks

This paper examines the use of non-linear incentive strategies as another option to sustain over time an agreement by means of non-cooperative play. Incentive strategies have been extensively used in the differential games literature in different areas and have been proposed to support the cooperative solution in two-player differential games. The incentive is designed in such a way that a coordinated outcome becomes a Nash equilibrium. If the incentive strategies are credible no player will be tempted for unilaterally deviate from the agreed decision. As far as we know all the previous literature on incentive strategies, except De Frutos & Martín-Herrán (2015), studies games belonging to the linear-state or linear-quadratic classes and the incentive strategies are constructed in such a way that the incentive equilibrium is the cooperative solution. Furthermore, the strategies are assumed to be linear and decision-dependent, i.e., each player makes his current decision contingent on the current decision of his opponent. Most of the previous works have not addressed the analysis of the credibility of the incentive strategies.

This paper is a follow-up of De Frutos & Martín-Herrán (2015) and assumes state-dependent and decision-dependent equilibrium strategies defined as non-linear functions of the control variables of both players and the current value of the state variable. We relax the definition of incentive equilibrium in the sense that we look for an incentive strategy equilibrium such that the steady state of the optimal state trajectory is close enough, but not necessarily identical, to the steady state of the state variable under cooperation. We show that the definition of less restrictive strategies helps to guarantee the sustainability of an agreement over time. We analyze a well-known linear-quadratic transboundary pollution differential game and present numerical experiments to illustrate the results. We compare the incentive equilibrium strategies, its credibility and the players' payoff when players use open-loop strategies and when they focus on stationary Markovian strategies.

Acknowledgements This research is partially supported by MINECO under projects MTM2016-78995-P (AEI) (Javier de Frutos) and ECO2014-52343-P and ECO2017-82227-P (AEI) (Guiomar Martín-Herrán) and by Junta de Castilla y León VA024P17 and VA105G18 co-financed by FEDER funds (EU).

References

1. Breton, M., Sokri, A., Zaccour, G.: Incentive equilibrium in an overlapping-generations environmental game. *European Journal of Operational Research* **185**, 687–699 (2008)
2. Buratto, A., Zaccour, G.: Coordination of advertising strategies in a fashion licensing contract. *Journal of Optimization Theory and Applications* **142**, 31–53 (2009)
3. De Frutos, J., Martín-Herrán, G.: Does flexibility facilitate sustainability of cooperation over time? A case study from environmental economics. *Journal of Optimization Theory and Applications* **165**, 657–677 (2015)

4. De Giovanni, P.: A joint maximization incentive in closed-loop supply chains with competing retailers: The case of spent-battery recycling. *European Journal of Operational Research* **268**, 128–147 (2018)
5. De Giovanni, P., Reddy, P.V., Zaccour, G.: Incentive strategies for an optimal recovery program in a closed-loop supply chain. *European Journal of Operational Research* **249**, 605–617 (2016)
6. Dockner, E.J., Long, N.V.: International pollution control: cooperative versus noncooperative strategies. *Journal of Environmental Economics and Management* **25**, 13–29 (1993)
7. Ehtamo, H., Hämäläinen, R.P.: On affine incentives for dynamic decision problems. In: Basar, T. (ed.) *Dynamic Games and Applications in Economics*, pp. 47–63. Springer-Verlag, Berlin (1986)
8. Ehtamo, H., Hämäläinen, R.P.: Incentive strategies and equilibria for dynamic games with delayed information. *Journal of Optimization Theory and Applications* **63**, 355–369 (1989)
9. Ehtamo, H., Hämäläinen, R.P.: Cooperative incentive equilibrium for a resource management problem. *Journal of Economic Dynamics and Control* **17**, 659–678 (1993)
10. Ehtamo, H., Hämäläinen, R.P.: Credibility of linear equilibrium strategies in a discrete time fishery management game. *Group Decision and Negotiation* **4**, 27–37 (1995)
11. Haurie, A., Pohjola, M.: Efficient equilibria in a differential game of capitalism. *Journal of Economic Dynamics and Control* **11**, 65–78 (1987)
12. Haurie, A., Krawczyk, J.B., Zaccour, G.: *Games and Dynamic Games*. World Scientific, Singapore (2012)
13. Jørgensen, S., Martín-Herrán, G., Zaccour, G.: Agreeability and time-consistency in linear-state differential games. *Journal of Optimization Theory and Applications* **119**, 49–63 (2003)
14. Jørgensen, S., Martín-Herrán, G., Zaccour, G.: Sustainability of cooperation overtime in linear-quadratic differential games. *International Game Theory Review* **7**, 395–406 (2005)
15. Jørgensen, S., Taboubi, S., Zaccour, G.: Incentives for retailer promotion in a marketing channel. In: Haurie, A., Muto S., Petrosjan, L.A. and Raghavan, T.E.S. (eds.) *Annals of the International Society of Dynamic Games*, Vol. 8, pp. 365–378. Birkhäuser, Boston (2006)
16. Jørgensen, S., Zaccour, G.: Time consistent side payments in a dynamic game of downstream pollution. *Journal of Economic Dynamics and Control* **25**, 1973–1987 (2001)
17. Jørgensen, S., Zaccour, G.: Incentive equilibrium strategies and welfare allocation in a dynamic game of pollution control. *Automatica* **37**, 29–36 (2001)
18. Jørgensen, S., Zaccour, G.: Time consistency in cooperative differential games. In: Zaccour, G. (ed.) *Decision and Control in Management Science (in Honor of Alain Haurie)*, pp. 349–366. Kluwer Academic Publishers, Dordrecht (2002)
19. Jørgensen, S., Zaccour, G.: Channel coordination over time: incentive equilibria and credibility. *Journal of Economic Dynamics and Control* **27**, 801–822 (2003)
20. Kaitala, V., Pohjola, M.: Economic development and agreeable redistribution in capitalism: Efficient game equilibria in a two-class Neoclassical growth model. *International Economic Review* **31**, 421–437 (1990)
21. Kaitala V., Pohjola, M.: Sustainable international agreements on greenhouse warming: A game theory study. In: Carraro, C., Filar, J.A. (eds.) *Annals of the International Society of Dynamic Games*, Vol. 2, pp. 67–87. Birkhäuser, Boston (1995)
22. Kierzenka, J., Shampine L.F.: A BVP Solver based on Residual Control and the MATLAB PSE. *ACM TOMS* **27**, 299–316 (2001)
23. Martín-Herrán, G., Taboubi, S.: Incentive strategies for shelf-space allocation in duopolies. In: Haurie, A., and Zaccour, G. (eds.) *Dynamic Games: Theory and Applications*, GERAD 25th Anniversary Series, pp. 231–253. Springer, New York (2005)
24. Martín-Herrán, G., Zaccour, G.: Credibility of incentive equilibrium strategies in linear-state differential games. *Journal of Optimization Theory and Applications* **126**, 1–23 (2005)
25. Martín-Herrán, G., Zaccour, G.: Credible linear incentive equilibrium strategies in linear-quadratic differential games. In: Bernhard, P., Gaitsgory V. and Pourtallier, O. (eds.) *Annals of the International Society of Dynamic Games*, vol. 10, pp. 261–291. Birkhäuser, Boston (2009)

26. Petrosjan, L.A.: Agreeable solutions in differential games. *International Journal of Mathematics, Game Theory and Algebra* **7**, 165–177 (1997)
27. Petrosjan, L.A., Zaccour, G.: Time-consistent Shapley value allocation of pollution cost control. *Journal of Economic Dynamics and Control* **27**, 381–398 (2003)
28. Petrosjan, L.A., Zenkevich, N.A.: *Game Theory*. World Scientific, Singapore (1996).
29. Taboubi, S.: Incentive mechanisms for price and advertising coordination in dynamic marketing channels. *International Transactions in Operational Research* (2017) doi: 10.1111/itor.12467
30. Tolwinski, B., Haurie, A., Leitmann, G.: Cooperative equilibria in differential games. *Journal of Mathematical Analysis and Applications* **119**, 192–202 (1986)
31. Van der Ploeg, F., De Zeeuw, A. J.: International aspects of pollution control. *Environmental and Resource Economics* **2**, 117–139 (1992)