On Coincidence of Feedback and Global Stackelberg Equilibria in a Class of Differential Games[☆]

Guiomar Martín-Herrán[a,1,*], Santiago J. Rubio[b,2]

[a]*IMUVA, Universidad de Valladolid. Spain.*
[b]*Department of Economic Analysis and ERI-CES, University of Valencia. Spain*

Abstract

This paper shows for a class of differential games that the global Stackelberg equilibrium (GSE) coincides with the feedback Stackelberg equilibrium (FSE), although the GSE assumes that the leader/regulator announces at the initial time the regulatory instrument rule she will follow for the rest of the game, while in the FSE, the regulator at any time chooses the optimal level of the regulatory instrument rate. This coincidence is based on the fact that the FSE is calculated using dynamic programming what implies that although the regulator chooses the regulatory instrument rate level that maximizes social welfare, the first-order condition for the maximization of the right-hand side of the Hamilton-Jacobi-Bellman equation implicitly defines a rule for the regulatory instrument. Then, as the regulatory instrument rule defined by the FSE implements the efficient outcome as the GSE does, the rules defined by both equilibria must be the same. In the second part of the paper, we check that this is the case for two examples. The first is an operations research model, while the second is an economic model. The first example fits in a linear-state differential game structure, while the second example presents a linear-quadratic specification. In both cases the regulatory instrument rules for both equilibria (GSE and FSE) are calculated and identical expressions are obtained.

Keywords: Game Theory, Dynamic Games, Global Stackelberg Equilibrium, Feedback Stackelberg Equilibrium, Supply Chain Management, Polluting Oligopoly

1. Introduction

The implementation of efficiency-inducing regulatory instruments can be modelled as a policy game where the regulator is the leader and the other players are the followers. In a one-shot game, the first mover is said to be the leader and the second mover is called the follower. However, it is not easy to translate the first-mover advantage to a dynamic setting. A first extension of this equilibrium concept for dynamic games was the open-loop Stackelberg equilibrium (OLSE) where at the initial time, the leader announces the control path she is going to follow for the rest of the game, and the followers taking this path as given, choose their control path as to maximize their integral of discounted payoffs. However, as is well known this type of equilibrium can be time inconsistent. Thus, the focus moved to look for feedback Stackelberg equilibria (FBSE). This kind of equilibria requires that the leader must use a feedback strategy: her control at any time depends on the observed pair (date, state), such that starting at any (date, state) pair, the continuation of the optimal strategy remains optimal for the leader. Strategies that specify the control at any time exclusively as a function of the observed pair (date, state) are called Markovian strategies.[3]

Unlike OLSE, there is no established general methodology for finding a FBSE. In the second edition of the excellent book on dynamic games by Başar and Olsder (1999), the authors mention two approaches that allow to calculate two FBSE in continuous time called by the authors the *feedback Stackelberg equilibrium* (FSE) and the *global Stackelberg equilibrium* (GSE).[4] The first one, that was originally defined by Başar and Haurie (1984), assumes that the leader has a stagewise first-mover advantage over the follower(s) at each stage of the game, which, in continuous time, means that the leader has an instantaneous advantage at each point in time. As shown by Başar and Haurie (1984), the FSE can be conceptualized as the limit of a sequence of feedback Stackelberg solutions of discretized (in time) versions of the original differential game, which become the collection of pointwise Stackelberg solutions of coupled Hamilton-Jacobi-Bellman equations.[5] The computation of the FSE involves two stages. First, the pointwise (in time) best response of the follower(s) to the leader's strategy is computed. Second, the leader, taking into account the followers' best response solves, again pointwise in time, her maximization problem. Therefore, to find the leader's feedback strategy, backward induction is applied, substituting the followers' instantaneous reaction function in the leader's Hamilton-Jacobi-Bellman (HJB) equation, and computing the leader's optimal strategy by maximizing the right-hand side (RHS) of this equation. Thus, the leader at any time selects the

---

[3]The concept of feedback Stackelberg equilibrium in multi-period games was first introduced by Simaan and Cruz (1973).

[4]A definition of the FSE for differential games may be also found in Haurie et al. (2012) and Başar and Zaccour (2018).

[5]Başar and Haurie (1984) establish a link between the feedback Stackelberg solution and an equilibrium solution for a multi-stage game having an asymmetric information structure. Then using the $G(\delta)$-game approach applied by Friedman (1971), they extend this equilibrium concept to the differential game setting. We remind the interested reader to this paper for the technical details of the analysis.

optimal *level* of the control, but the first-order condition defines a Markovian strategy. For this kind of equilibria, no commitment is required for the entire temporal horizon, but only a *stagewise* first-mover advantage. These equilibria are subgame perfect, and consequently, time consistent.

One of the first applications of the FSE is Başar et al. (1985) on the dynamic inefficiency of capitalism. Since then this equilibrium concept has been applied for the analysis of different issues in a dynamic context in continuous time. See, for instance, Cohen and Michel (1988) for the design of economic policy, Tahvonen (1996), Rubio and Escriche (2001) and Wirl (2012) for the analysis of the strategic trade policy with polluting nonrenewable resources, Jørgensen et al. (2003), Karray and Martín-Herrán (2009) and Chutani and Sethi (2012) for applications of the FSE in supply and marketing channels. More recently, Colombo and Labrecciosa (2019) have generalized it for an oligopoly exploiting a common-pool renewable resource with an arbitrary numbers of leaders and followers.

The second Stackelberg equilibrium concept, the GSE, assumes that the leader announces at the initial time the feedback strategy she is going to follow for the rest of the game, and the followers taking this strategy as given, choose their control path as to maximize their integral of discounted payoffs. In principle, this procedure yields the followers' reaction functions depending on the leader's feedback strategy. Then, the leader chooses the *rule* that maximizes her objective function among all possible feedback strategies. However, since the feedback strategy can be any function, it is not clear how such an optimal rule can be obtained in practice.

An example where the GSE can be found relatively easily is Benchekroun and Long (1998) where the case of a polluting oligopoly is studied. In this paper, for a linear-quadratic differential game with infinite time horizon, the authors show that there exists a stationary linear tax rule that supports the efficient outcome as a GSE. The procedure used by the authors follows the *classical* approach of market regulation. First, the output strategy that maximizes net social welfare is obtained assuming that the regulator can control the actions of the firms acting as a *social planner*. Second, the Markov-perfect Nash equilibrium (MPNE) of the game played by the firms is calculated for the emission tax rule announced by the regulator. Thus, the optimal tax rule selected by the regulator is the one for which the regulated market implements the efficient outcome. In other words, the parameters of the tax rule are selected to achieve that the Markov-perfect Nash equilibrium of the game played by the firms replicates the output strategy obtained by the social planner.

In this paper, we show that for a class of differential games where each player has only one control variable and there is only one state variable we show that the difference between the two equilibrium concepts is spurious: the FSE coincides with the GSE. In other words, there is no difference between choosing a *regulatory instrument rule* at the initial time that implements the efficient outcome or choosing the *level of the regulatory instrument rate* that maximizes social welfare at each period of time. The coincidence is explained by the fact that although for the FSE the regulator selects the level of the regulatory instrument rate, as the equilibrium is computed by dynamic programming, the maximization of the RHS of the HJB equation yields a regulatory instrument rule. Notice that

according to the first-order condition for the maximization of the RHS of the HJB equation, the optimal level of the regulatory instrument rate depends on the first derivative of the value function which is a function of the state variable. Thus, once a solution for the value function is found, the FSE gives a regulatory instrument rule. Then, if this regulatory instrument rule maximizes social welfare it must coincide with the regulatory instrument rule given by the GSE that also maximizes social welfare. Thus, we can conclude that the first-mover advantage of the regulator in the choice of the level of the policy instrument just for a period of time, that in continuous time translates into an instantaneous advantage at each time, yields the same equilibrium as the one obtained when the regulator commits at the initial time to follow a policy rule. We use two different models for illustrating this coincidence. The first model is borrowed from the operations research literature, in particular, from the marketing channel literature (Jørgensen et al. (2003)). The second model deals with environmental regulation and is borrowed from Benchekroun and Long (1998). In this case the coincidence between the GSE and the FSE also tells us that there are not differences between the *strategic* approach of market regulation that is behind the FSE and the *classical* approach followed by Benchekroun and Long (1998) to calculate the GSE, i.e., if the social planner that maximizes the net social welfare selecting the output and implements it announcing a tax rule at the beginning of the game becomes the leader of a policy game that maximizes net social welfare choosing the tax rate at each period of time, there is no difference in the optimal policy that both decision agents will apply.

Literature Review

The issue of coordination has often been in focus in studies of marketing channels using game theoretic methods. The literature has proved that in a marketing channel cooperation can have a positive impact, not only on channel members' profits but also on consumer welfare. Marketing science literature has focussed on the design of coordinating mechanisms that will reduce channel inefficiencies. The literature that deals with differential games and coordination in a marketing channel proposed mainly three approaches: cooperative advertising, channel leadership and incentive equilibrium (Jørgensen and Zaccour (2004), Chapter 5). Recent papers that use these approaches are, for example, De Giovanni et al. (2019), Martín-Herrán et al. (2011), Martín-Herrán & Taboubi (2015) and Taboubi (2019). As far as we know the approach we use in this paper has never been used before in the marketing channel literature. We do not know any work that uses regulatory instruments to implement the coordinated (efficient) solution as a regulated Nash equilibrium among the players.

On the contrary, the literature dealing with environmental regulation is extensive. Particularly, the literature addressing the regulation of firms with market power in the context of stock dynamics includes Bergstrom et al. (1981), Karp and Livernois (1992), Karp (1992), for the case of non-renewable resources and Xepapadeas (1992), Benchekroun and Long (1998, 2002), Stimming (1999), Feenstra et al. (2001), Yanese (2009) and more recently Wirl (2014) and Martín-Herrán and Rubio (2018a, 2018b) for the case of polluting firms. Bergstrom et al. (1981) showed that a regulator can induce a monopolist to extract the resource at the optimal rate by imposing a time path of subsidies. They

also show that there is a family of such optimal time paths. Karp and Livernois (1992) pointed out that Bergstrom et al. (1981) implicitly assumed that the regulator can commit to her chosen time path of subsidies. In other words, the time-dependent open-loop subsidy proposed by these authors is not generally subgame perfect and time consistency is not guaranteed. Karp and Livernois (1992) then showed that there exists a family of linear Markov subsidy rules that are subgame perfect and that induce the monopolist to extract efficiently. Karp (1992) extended the analysis to the case of a common property oligopoly. In both papers, a GSE is calculated following the procedure explained above. Benchekroun and Long (1998) applied this procedure to the case of a polluting oligopoly and Benchekroun and Long (2002) analyzed the case of a polluting monopoly.

In Xepapadeas (1992), Stimming (1999) and Feenstra et al. (2001) the environmental policy is *exogenously* determined and the research assessed the effects of a stricter environmental policy and the comparison of taxes versus emission standards. The first paper we know where is explicitly recognized that the regulator is a player in a policy game that selects the levels of the policy instruments in each period of time to maximize net social welfare is Yanese (2009). The author examined a differential policy game between national governments in a model of international pollution control in which duopolists compete myopically in quantities in a third country with product differentiation, and expense resources in abatement activities. The same approach is followed by Wirl (2014) to investigate a differential policy game between a monopoly that provides a clean technology for a polluting competitive industry and a regulator that uses an emission tax or standard to control a flow pollutant. For this model, the MPNE when the regulator applies a tax and when she uses a standard are compared. Wirl's analysis showed that the efficient outcome cannot be implemented using only one instrument but that the tax and the standard are equivalent.[6] The first paper where a FSE is used to characterize the first-best policy for a polluting firm is Martín-Herrán and Rubio (2018a). They analyzed the case of a polluting monopoly that spends resources in abatement activities. They also calculated the first-best policy, consisting of a tax on emissions and a subsidy on output, for a linear-quadratic specification of the model. As the FSE supports the efficient outcome, the calculation of the GSE would yield the same policy rules. Finally, Martín-Herrán and Rubio (2018b) focused on the second-best taxation of a polluting monopoly that can invest in abatement capital. The FSE allows for a complete characterization of the optimal policy in a second-best framework.

The remainder of the paper is organized as follows. Section 2 presents a class of differential games for which the two Stackelberg equilibrium concepts (FSE and GSE) coincide. The ingredients of the (differential) policy game and the efficient conditions are presented and the differential equations that characterize the GSE and the FSE are obtained and the coincidence between the two equilibria

---

[6]In Wirl's (2014) policy game, the instantaneous reaction functions are orthogonal and consequently the FSE coincides with the MPNE. The conditions that yield this coincidence have been studied by Rubio (2006). Other examples where this coincidence occurs are for instance Başar et al. (1985), Tahvonen (1996) and Rubio and Escriche (2001).

is showed. In Section 3, the GSE and FSE are calculated for a distribution channel modelled as a linear-state differential game. In Section 4, the GSE and FSE are calculated for a linear-quadratic specification of an environmental regulation differential game. Finally, Section 5 offers some concluding remarks.

2. A Class of n-player differential games where the FSE and the GSE coincide

We assume n players and the instantaneous payoff of player i is given by function $F_i : \mathbb{R}_+^{n+1} \mapsto \mathbb{R}$. Each player has a control variable and we assume that functions F_i are strictly concave and at least twice differentiable. We denote by u_i the control variable of the i-th player and the vector of control variables by $\mathbf{u} = (u_1(t), u_2(t), \ldots, u_n(t))$. The state variable is represented by $x(t)$. The dynamics of the state variable is defined by the function $f(\mathbf{u}(t), x(t))$ which is twice differentiable and satisfies $\frac{\partial f}{\partial u_i} \neq 0$ for all $i = 1, 2, \ldots, n$.

The noncooperative problem of player i reads:

$$\max_{u_i(t)} \quad \int_0^\infty F_i(\mathbf{u}(t), x(t))e^{-rt} \, dt \tag{1}$$

$$\text{s.t.:} \quad \dot{x}(t) = f(\mathbf{u}(t), x(t)), \quad x(0) = x_0. \tag{2}$$

Any feedback Nash equilibrium of the noncooperative game described above must satisfy the following system of HJB equations:

$$rV_i(x(t)) = \max_{u_i(t)} \left\{ F_i(\mathbf{u}(t), x(t)) + V_i'(x(t))f(\mathbf{u}(t), x(t)) \right\}, \quad i = 1, 2, \ldots, n,$$

where $V_i(x(t))$ represents the i-th player value function for the current value, $x(t)$, of the state variable.

The following first-order conditions (FOC) for an interior Nash equilibrium are obtained from the maximization of the RHS of the HJB equations:

$$\frac{\partial F_i}{\partial u_i}(\mathbf{u}(t), x(t)) + V_i'(x(t))\frac{\partial f}{\partial u_i}(\mathbf{u}(t), x(t)) = 0, \quad i = 1, 2, \ldots, n. \tag{3}$$

On the other hand, the "efficient outcome" is defined as the solution of the following maximization problem:

$$\max_{u_1(t), u_2(t), \ldots, u_n(t)} \quad \int_0^\infty J(\mathbf{u}(t), x(t))e^{-rt} \, dt$$

$$\text{s.t.:} \quad \dot{x}(t) = f(\mathbf{u}(t), x(t)), \quad x(0) = x_0,$$

where function $J : \mathbb{R}_+^{n+1} \mapsto \mathbb{R}$ represents social welfare. This could be just the addition of the individual payoffs (as in the example analyzed in Section 3) or it could add other elements as the consumer's surplus (as in the example analyzed in Section 4).

Solving by dynamic programming, the solution to this dynamic optimization problem must comply the following HJB equation:

$$rW(x(t)) = \max_{u_1(t),u_2(t),\ldots,u_n(t)} \left\{ J(\mathbf{u}(t),x(t)) + W'(x(t))f(\mathbf{u}(t),x(t)) \right\},$$

where $W(x(t))$ represents the value function for the efficient problem for the current value, $x(t)$, of the state variable.

The maximization of the RHS of the HJB equation yields the following first-order conditions (FOC) for an interior solution

$$\frac{\partial J}{\partial u_i}(\mathbf{u}(t),x(t)) + W'(x(t))\frac{\partial f}{\partial u_i}(\mathbf{u}(t),x(t)) = 0, \quad i = 1,2,\ldots,n. \tag{4}$$

These conditions implicitly define the optimal feedback strategies $\mathbf{u}^* = \boldsymbol{\chi}^*(x)$.

The comparison of conditions (3) and (4) clearly shows that, in general, the Nash equilibrium does not satisfy the efficiency conditions characterizing the efficient solution. As J is built on F_i, we can claim that the following condition must be satisfied by definition

$$\frac{\partial J}{\partial u_i}(\mathbf{u}(t),x(t)) = \frac{\partial F_i}{\partial u_i}(\mathbf{u}(t),x(t)) + d_i(\mathbf{u}(t),x(t)), \tag{5}$$

where the structure of $d_i(\mathbf{u}(t),x(t))$ will depend on the particular model that is being analyzed. This term is originated by the inefficiency of the Nash equilibrium. Notice that if this term is zero, there is no difference between the efficient solution and the Nash equilibrium[7].

The regulator, acting as the leader of a Stackelberg game, can use a regulatory instrument to implement the efficient solution as a regulated Nash equilibrium among the n players, acting as the followers of the Stackelberg game. We will assume that the regulatory function is linear in the policy instrument for each control variable: $\tau_i(t)u_i(t)$. These functions can take positive or negative values and the instruments can be understood, for example, as a tax or a subsidy depending on the problem at hand.

We want to show how the efficient outcome can be implemented both as a *feedback Stackelberg equilibrium* (FSE) and as a *global Stackelberg equilibrium* (GSE). In the next subsection, we show how the efficient outcome can be implemented as FSE.

2.1. The Feedback Stackelberg Equilibrium

In this section we are interested in characterizing a subgame-perfect feedback equilibrium of the differential game under regulation. Specifically, the equilibrium concept we apply is the stagewise

---

[7] Rincón-Zapatero et al. (2000) and Martín-Herrán and Rincón-Zapatero (2005) present a method for the characterization of Markov-perfect Nash equilibria being Pareto efficient in nonlinear differential games.

feedback Stackelberg equilibrium. Such an equilibrium concept requires that the n players and the regulator know the current state of the game at every period. The equilibrium derivation involves a backward recursion, where at every step the Stackelberg equilibrium of the "static/one-shot game" is calculated, i.e., it is assumed that the leader has a stagewise first-mover advantage over the followers at each stage of the game, that, in continuous time, means that the leader has an instantaneous advantage at each time t. For this kind of equilibria, no commitment is required for the entire temporal horizon.[8]

We assume that both the leader's (the regulator) and the followers' (the n agents) strategies are stationary feedback strategies. Let $\Upsilon = \mathbb{R}$ and $\Phi = \mathbb{R}_+$ denote the action spaces of the leader and followers, respectively. The strategies spaces are sets of functions mapping to Υ and Φ, respectively. A stationary feedback strategy of the leader is a function $\boldsymbol{\phi} : S \times S \times \cdots \times S \mapsto \Upsilon$, where $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_n)$, $S = \mathbb{R}$ is the state space, and a stationary feedback strategy of the i-th follower is a function $\psi_i : S \times \Upsilon \mapsto \Phi$. Thus, it follows that $\tau_i(t) = \phi_i(x(t))$, $i \in \{1, 2, \ldots, n\}$, for the regulator and $u_i(t) = \psi_i(x(t), \boldsymbol{\tau}(t))$ for the followers, $i \in \{1, 2, \ldots, n\}$. We assume that strategies be well-defined over the entire state space and continuous. At any t, the game is solved by backward induction.

Let $Z_i(x)$ stand for the value function for the current value, x, of the state variable for the i-th follower and $W(x)$ for the value function for the current value, x, of the state variable for the leader. The FSE can be obtained using the HJB equations[9]

$$
\begin{aligned}
rZ_i(x) &= \max_{\{u_i \in \Phi\}} \{F_i(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x) + \phi_i(x)u_i \\
&+ Z_i'(x)f(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x)\}, \quad i = 1, \ldots, n, \quad (6) \\
rW(x) &= \max_{\{\tau_1, \tau_2, \ldots, \tau_n \in \Upsilon\}} \{J(\psi_1(x, \boldsymbol{\tau}), \psi_2(x, \boldsymbol{\tau}), \ldots, \psi_i(x, \boldsymbol{\tau}), \ldots, \psi_n(x, \boldsymbol{\tau}), x) \\
&+ W'(x)f(\psi_1(x, \boldsymbol{\tau}), \psi_2(x, \boldsymbol{\tau}), \ldots, \psi_i(x, \boldsymbol{\tau}), \ldots, \psi_n(x, \boldsymbol{\tau}), x)\}, \quad (7)
\end{aligned}
$$

under the additional condition that, $\forall \boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_n), \tau_i \in \Upsilon, \; \forall u_i \in \Phi,$

$$
\begin{aligned}
&F_i(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, \psi_i(x, \boldsymbol{\tau}), \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x) + \phi_i(x)\psi_i(x, \boldsymbol{\tau}) \\
&+ Z_i'(x)f(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, \psi_i(x, \boldsymbol{\tau}), \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x) \\
&\geq F_i(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x) + \phi_i(x)u_i \\
&+ Z_i'(x)f(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x), \quad i = 1, 2, \ldots, n, \quad (8)
\end{aligned}
$$

and $\lim_{t \to \infty} e^{-rt}Z_i(x^*(t)) = 0$, $i \in \{1, 2, \ldots, n\}$, $\lim_{t \to \infty} e^{-rt}W(x^*(t)) = 0$ where $x^*(t)$ denotes the equilibrium path of x. Notice that $\psi_i(x, \boldsymbol{\tau})$ represents the instantaneous reaction function of the i-th follower to the actions of the leader.

---

[8]The presentation of this equilibrium concept we do in this paper is based on Colombo and Labrecciosa (2019).

[9]From now on, the time argument is removed when no confusion may arise.

The functionals for the leader and for the i-th follower when the leader plays $\boldsymbol{\phi}^* = (\phi_1^*, \phi_2^*, \ldots, \phi_n^*)$ and the i-th follower plays ψ_i^* are given by

$$J^L(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*) = \int_t^\infty e^{-r(s-t)} \left(J(\boldsymbol{\psi}^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))), x^*(s)) \right) ds$$

and

$$J_i^F(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*) = \int_t^\infty e^{-r(s-t)} \left(F_i(\boldsymbol{\psi}^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))), x^*(s)) + \phi_i^*(x^*(s))(\psi_i^*(x^*(s), \boldsymbol{\phi}^*(x^*(s)))) \right) ds$$

respectively, where $\boldsymbol{\psi}^* = (\psi_1^*, \psi_2^*, \ldots, \psi_n^*)$, $\boldsymbol{\phi}^* = (\phi_1^*, \phi_2^*, \ldots, \phi_n^*)$ and

$$\boldsymbol{\psi}^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))) = (\psi_1^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))), \psi_2^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))), \ldots, \psi_n^*(x^*(s), \boldsymbol{\phi}^*(x^*(s)))),$$

$$\boldsymbol{\phi}^*(x^*(s)) = (\phi_1^*(x^*(s)), \phi_2^*(x^*(s)), \ldots, \phi_n^*(x^*(s))).$$

In order for $\boldsymbol{\phi}^*$ and ψ_i^* to be equilibrium strategies, $J^L(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*)$ must be greater than the functional resulting from any deviation from the equilibrium path and the same must occur for the functional $J_i^F(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*)$ of each player $i = 1, 2, \ldots, n$. Denote by $\hat{\boldsymbol{\phi}}$ a unilateral deviation by the leader, and by $\hat{\psi}_i$ a unilateral deviation by the i-th follower. Also, let $\hat{\psi}_i^*$ denote the strategy of the i-th follower when the leader plays $\hat{\boldsymbol{\phi}}$. The functional for the leader unilaterally deviating from $\boldsymbol{\phi}^*$ is given by

$$J^L(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\psi}}^*) = \int_t^\infty e^{-r(s-t)} \left(J(\hat{\boldsymbol{\psi}}^*(\hat{x}^*(s), \boldsymbol{\phi}^*(\hat{x}^*(s))), \hat{x}^*(s)) \right) ds$$

where $\hat{\boldsymbol{\psi}}^* = (\hat{\psi}_1^*, \hat{\psi}_2^*, \ldots, \hat{\psi}_n^*)$,

$$\hat{\boldsymbol{\psi}}^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))) = (\hat{\psi}_1^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))), \hat{\psi}_2^*(x^*(s), \boldsymbol{\phi}^*(x^*(s))), \ldots, \hat{\psi}_n^*(x^*(s), \boldsymbol{\phi}^*(x^*(s)))),$$

and $\hat{x}^*(s)$ stands for the path of x induced by $(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\psi}}^*)$. The functional for the i-th follower unilaterally deviating from ψ_i^* reads

$$
\begin{aligned}
J_i^F(\boldsymbol{\phi}^*, \hat{\psi}_i, \boldsymbol{\psi}_{-i}^*) &= \int_t^\infty e^{-r(s-t)} \Big(F_i(\psi_1^*(\hat{x}^{**}(s), \boldsymbol{\phi}^*(\hat{x}^{**}(s))), \ldots, \hat{\psi}_i(\hat{x}^{**}(s), \boldsymbol{\phi}^*(\hat{x}^{**}(s))), \ldots, \psi_n^*(\hat{x}^{**}(s), \boldsymbol{\phi}^*(\hat{x}^{**}(s))), \hat{x}^{**}(s)) \\
&\quad + \phi_i^*(\hat{x}^{**}(s))(\hat{\psi}_i^*(\hat{x}^{**}(s), \boldsymbol{\phi}^*(\hat{x}^{**}(s)))) \Big) ds,
\end{aligned}
$$

where $\boldsymbol{\psi}_{-i}^* = (\hat{\psi}_1^*, \ldots, \hat{\psi}_{i-1}^*, \hat{\psi}_{i+1}^*, \ldots, \hat{\psi}_n^*)$ and $\hat{x}^{**}(s)$ represents the path of x induced by $(\boldsymbol{\phi}^*, \hat{\psi}_i, \boldsymbol{\psi}_{-i}^*)$.

Definition 1. The $2n$-tuple of strategies $(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*)$ defines a feedback Stackelberg equilibrium if

$$J^L(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*) \geq J^L(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\psi}}^*),$$

$$J_i^F(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*) \geq J_i^F(\boldsymbol{\phi}^*, \hat{\psi}_i, \boldsymbol{\psi}_{-i}^*), \quad i = 1, 2, \ldots, n.$$

We may point out that the feedback equilibrium strategies are perfect state-space equilibria because the necessary optimality conditions are required to be satisfied for all values of the state variable, and not only for the values that lie on the optimal state-space paths. Therefore, the solutions obtained continue to remain optimal at each point in time after the game has begun and consequently they are subgame perfect, and hence, time consistent.

2.1.1. The Optimal Policy

Let $Z_i(x)$ and $W(x)$ denote the value functions of the $i-th$ $(i \in \{1,2,\ldots,n\})$ player and the regulator, respectively. Then the $2n$-tuple of strategies $(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*)$ constitutes a feedback Stackelberg equilibrium if there exist $n+1$ continuously differentiable value functions satisfying the HJB Eqs. (6) and (7) together with inequalities in (8) and the transversality conditions $\lim_{t \to \infty} e^{-rt} Z_i(x^*(t)) = \lim_{t \to \infty} e^{-rt} W(x^*(t)) = 0$. Let

$$R_i(x, \boldsymbol{\phi}(x), \psi_{-i}(x, \boldsymbol{\phi}(x))) = \arg \max_{\{u_i \in \Phi\}} \{F_i(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x)$$

$$+ \phi_i(x)u_i + Z_i'(x)f(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x)\}, \quad i = 1, 2, \ldots, n.$$

Thus, the argument that maximizes the RHS of the HJB equation for the followers' problem (6) must satisfy the following conditions

$$\frac{\partial F}{\partial u_i}(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x) + \phi_i(x)$$

$$+ \quad Z_i'(x)\frac{\partial f}{\partial u_i}(\psi_1(x, \boldsymbol{\phi}(x)), \psi_2(x, \boldsymbol{\phi}(x)), \ldots, u_i, \ldots, \psi_n(x, \boldsymbol{\phi}(x)), x) = 0, \, i = 1, 2, \ldots, n, \quad (9)$$

where $Z_i'(x)$ is the *private* shadow price of the state variable when a regulatory instrument is used. From (9), we obtain the instantaneous reaction function of the players to the action profile of the leader: $u_i = \psi_i(x, \boldsymbol{\phi}(x)), i = 1, 2, \ldots, n$.

On the other hand, the FOC for the maximization of the RHS of the leader's HJB (7) requires that the efficient condition (4) holds

$$\left(\frac{\partial J}{\partial u_i}(\boldsymbol{\psi}(x, \boldsymbol{\tau}), x) + W'(x)\frac{\partial f}{\partial u_i}(\boldsymbol{\psi}(x, \boldsymbol{\tau}), x)\right)\frac{\partial \psi_i}{\partial \tau_i} = 0, \quad i = 1, 2, \ldots, n. \quad (10)$$

Then, taking into account that τ_i is equal to $\phi_i(x)$, the efficient condition implicitly defines the optimal regulatory instrument rules, $\boldsymbol{\phi}^*(x)$. Next, after substitution of these rules in the agents' reaction functions, we obtain the equilibrium feedback strategy for the agents $\boldsymbol{u}^* = \boldsymbol{\psi}^*(x, \boldsymbol{\phi}^*(x))$ which implies that the FSE supports the efficient outcome: $\mathbf{u}^* = \boldsymbol{\chi}^*(x) = \boldsymbol{\psi}^*(x, \boldsymbol{\phi}^*(x))$.

2.1.2. The Characterization of the Optimal Policy Instrument

If condition (5) is taking into account, the FOC (9) can be rewritten as follows for the optimal solution

$$\frac{\partial J}{\partial u_i}(\psi^*(x, \phi^*(x)), x) - d_i(\psi^*(x, \phi^*(x)), x) + \phi_i^*(x) + Z_i'(x)\frac{\partial f}{\partial u_i}(\psi^*(x, \phi^*(x)), x) = 0.$$

Then, using (10) we obtain that the policy instrument is given by the following expression

$$\tau_i^f = \phi_i^*(x) = d_i(\psi^*(x, \phi^*(x)), x) + (W'(x) - Z_i'(x))\frac{\partial f}{\partial u_i}(\psi^*(x, \phi^*(x)), x). \tag{11}$$

Thus, the optimal policy presents two components.[10] The first one is the term that establishes the difference between the efficient outcome and the Nash equilibrium defined in (5). The second one is the difference between the social shadow price of state variable and its private shadow price multiplied by the effect that one variation in the control variable has on the state variable.

2.2. The Global Stackelberg Equilibrium

In this case, instead of choosing the level of the regulatory instrument at each moment, the leader announces at time zero the policy rule that she will use throughout the game. Taking into account the reaction of the followers to this rule, the leader chooses among all possible rules the one that maximizes her payoffs.

However, as we pointed out in the introduction since the rule can be any function, it is not clear how such an optimal rule can be derived. In this section, we present the standard procedure proposed in the literature via the use of dynamic programming. Given the regulatory instrument rule, $\tau_i(x)$, announced by the regulator at time $t = 0$, the n players compete à la Nash as in the second stage of the previous section. Suppose that player i knows that all the other players use Markovian strategies: $u_j = \chi_j(x)$, $j \neq i$. Then, the Markov-perfect Nash equilibrium (MPNE) of the game played by the agents must satisfy the HJB equations

$$rZ_i(x) = \max_{\{u_i \in \Phi\}} \{F_i(\chi_1(x), \chi_2(x), \ldots, u_i, \ldots, \chi_n(x), x) + \tau_i(x)u_i$$
$$+ Z_i'(x)f(\chi_1(x), \chi_2(x), \ldots, u_i, \ldots, \chi_n(x), x)\}, \ i = 1, 2, \ldots, n,$$

that yield the following FOCs

$$\frac{\partial F_i}{\partial u_i}(\chi_1(x), \chi_2(x), \ldots, u_i, \ldots, \chi_n(x), x) + \tau_i(x) + Z_i'(x)\frac{\partial f}{\partial u_i}(\chi_1(x), \chi_2(x), \ldots, u_i, \ldots, \chi_n(x), x) = 0, \tag{12}$$

$i = 1, 2, \ldots, n.$

---

[10]Superscript f stands for the regulatory instrument supported by the feedback Stackelberg equilibrium.

This approach gives the same FOCs as the ones obtained for the FSE except that the strategies that support the MPNE do not include directly the dependence of the control variable with respect to the regulatory instrument.

To calculate the regulatory instrument rule that implements the efficient outcome, we begin differentiating with respect to x the FOCs

$$\sum_{j=1}^{n} \frac{\partial^2 F_i}{\partial u_i \partial u_j} \chi_j'(x) + \frac{\partial^2 F_i}{\partial u_i \partial x} + \tau_i'(x) + Z_i''(x) \frac{\partial f}{\partial u_i} + Z_i'(x) \left(\sum_{j=1}^{n} \frac{\partial^2 f}{\partial u_i \partial u_j} \chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x} \right) = 0, \quad (13)$$

$i = 1, 2, \ldots, n$, where the partial derivatives $\partial^2 G/(\partial u_i \partial u_j)$ and $\partial^2 G/(\partial u_i \partial x)$ for $G = F_i, f$ are evaluated at $(\boldsymbol{\chi}(x), x)$.

On the other hand, differentiating the HJB equation, and making use of the envelope theorem, we obtain

$$r Z_i'(x) = \sum_{j=1}^{n} \frac{\partial F_i}{\partial u_j} \chi_j'(x) + \frac{\partial F_i}{\partial x} + \tau_i'(x) \chi_i(x) + \tau_i(x) \chi_i'(x) + Z_i''(x) f(\boldsymbol{\chi}(x), x) + Z_i'(x) \left(\sum_{j=1}^{n} \frac{\partial f}{\partial u_j} \chi_j'(x) + \frac{\partial f}{\partial x} \right),$$

where

$$\frac{\partial F_i}{\partial u_i} \chi_i'(x) + \tau_i(x) \chi_i'(x) + Z_i' \frac{\partial f}{\partial u_i} \chi_i'(x) = \left(\frac{\partial F_i}{\partial u_i} + \tau_i(x) + Z_i' \frac{\partial f}{\partial u_i} \right) \chi_i'(x) = 0$$

because of the FOC (12). Yielding

$$r Z_i'(x) = \sum_{j=1,j\neq i}^{n} \frac{\partial F_i}{\partial u_j} \chi_j'(x) + \frac{\partial F_i}{\partial x} + \tau_i'(x) \chi_i(x) + \tau_i(x) \chi_i'(x) + Z_i''(x) f(\boldsymbol{\chi}(x), x)$$

$$+ Z_i'(x) \left(\sum_{j=1,j\neq i}^{n} \frac{\partial f}{\partial u_j} \chi_j'(x) + \frac{\partial f}{\partial x} \right),$$

that using (13) allows us to eliminate Z_i'' if $\frac{\partial f}{\partial u_i} \neq 0$:

$$Z_i''(x) = -\frac{1}{\frac{\partial f}{\partial u_i}} \left(\sum_{j=1}^{n} \frac{\partial^2 F_i}{\partial u_i \partial u_j} \chi_j'(x) + \frac{\partial^2 F_i}{\partial u_i \partial x} + \tau_i'(x) + Z_i'(x) \left(\sum_{j=1}^{n} \frac{\partial^2 f}{\partial u_i \partial u_j} \chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x} \right) \right).$$

The result is

$$\left[r + \frac{1}{\frac{\partial f}{\partial u_i}} \left(\sum_{j=1}^{n} \frac{\partial^2 f}{\partial u_i \partial u_j} \chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x} \right) f(\boldsymbol{\chi}(x), x) - \left(\sum_{j=1,j\neq i}^{n} \frac{\partial f}{\partial u_j} \chi_j'(x) + \frac{\partial f}{\partial x} \right) \right] Z_i'(x)$$

$$= \sum_{j=1,j\neq i}^{n} \frac{\partial F_i}{\partial u_j}\chi_j'(x) + \frac{\partial F_i}{\partial x} + \tau_i'(x)\chi_i(x) - \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 F_i}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 F_i}{\partial u_i \partial x} + \tau_i'(x)\right)f(\boldsymbol{\chi}(x),x).$$

Next we eliminate Z_i' using the FOCs (12) to obtain

$$\left[r + \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 f}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x) - \left(\sum_{j=1,j\neq i}^{n}\frac{\partial f}{\partial u_j}\chi_j'(x) + \frac{\partial f}{\partial x}\right)\right]\left(-\frac{\frac{\partial F_i}{\partial u_i}+\tau_i(x)}{\frac{\partial f}{\partial u_i}}\right)$$

$$= \sum_{j=1,j\neq i}^{n} \frac{\partial F_i}{\partial u_j}\chi_j'(x) + \frac{\partial F_i}{\partial x} + \tau_i'(x)\chi_i(x) - \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 F_i}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 F_i}{\partial u_i \partial x} + \tau_i'(x)\right)f(\boldsymbol{\chi}(x),x),$$

that reordering terms yields

$$-\left[r + \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 f}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x) - \left(\sum_{j=1,j\neq i}^{n}\frac{\partial f}{\partial u_j}\chi_j'(x) + \frac{\partial f}{\partial x}\right)\right]\left(\frac{\frac{\partial F_i}{\partial u_i}}{\frac{\partial f}{\partial u_i}}\right)$$

$$-\left[r + \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 f}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x) - \left(\sum_{j=1,j\neq i}^{n}\frac{\partial f}{\partial u_j}\chi_j'(x) + \frac{\partial f}{\partial x}\right)\right]\frac{\tau_i(x)}{\frac{\partial f}{\partial u_i}}$$

$$= \sum_{j=1,j\neq i}^{n} \frac{\partial F_i}{\partial u_j}\chi_j'(x) + \frac{\partial F_i}{\partial x} + \tau_i'(x)\left(\chi_i(x) - \frac{f(\boldsymbol{\chi}(x),x)}{\frac{\partial f}{\partial u_i}}\right) - \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 F_i}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 F_i}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x).$$

Last expression is a system of differential equations for τ_i, $i = 1, 2, \ldots, n$, that can be summarized as follows

$$G_i(x)\tau_i(x) + H_i(x)\tau_i'(x) = I_i(x), \ i = 1, 2, \ldots, n, \tag{14}$$

where

$$G_i(x) = -\left[r + \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 f}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x) - \left(\sum_{j=1,j\neq i}^{n}\frac{\partial f}{\partial u_j}\chi_j'(x) + \frac{\partial f}{\partial x}\right)\right]\frac{1}{\frac{\partial f}{\partial u_i}},$$

$$H_i(x) = -\left(\chi_i(x) - \frac{f(\boldsymbol{\chi}(x),x)}{\frac{\partial f}{\partial u_i}}\right),$$

$$I_i(x) = \sum_{j=1,j\neq i}^{n}\frac{\partial F_i}{\partial u_j}\chi_j'(x) + \frac{\partial F_i}{\partial x} - \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 F_i}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 F_i}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x)$$

$$+ \left[r + \frac{1}{\frac{\partial f}{\partial u_i}}\left(\sum_{j=1}^{n}\frac{\partial^2 f}{\partial u_i \partial u_j}\chi_j'(x) + \frac{\partial^2 f}{\partial u_i \partial x}\right)f(\boldsymbol{\chi}(x),x) - \left(\sum_{j=1,j\neq i}^{n}\frac{\partial f}{\partial u_j}\chi_j'(x) + \frac{\partial f}{\partial x}\right)\right]\left(\frac{\frac{\partial F_i}{\partial u_i}}{\frac{\partial f}{\partial u_i}}\right).$$

All the partial derivatives in the expressions above are evaluated at $(\boldsymbol{\chi}(x),x)$.

Then, a policy instrument rule that guides the agents to achieve the efficient path as a MPNE is the solution of the system of differential equations (14) for $\boldsymbol{\chi}(x) = \boldsymbol{\chi}^*(x) = \mathbf{u}^*$. Under this condition the system of differential equations (14) will give the optimal rules for the policy instruments that support the efficient outcome as a "regulated" MPNE.

2.2.1. The Characterization of the Optimal Policy Instrument

As occurs with the FSE, if condition (5) is considered, the FOC (12) for the optimal solution can be rewritten as follows

$$\frac{\partial J}{\partial u_i}(\boldsymbol{\chi}^*(x), x) - d_i(\boldsymbol{\chi}^*(x), x) + \tau_i^*(x) + Z_i'(x)\frac{\partial f}{\partial u_i}(\boldsymbol{\chi}^*(x), x) = 0,$$

But, if the policy instrument implements the efficient condition then, according to condition (4), the following condition must be satisfied

$$\frac{\partial J}{\partial u_i}(\boldsymbol{\chi}^*, x) + W'(x)\frac{\partial f}{\partial u_i}(\boldsymbol{\chi}^*, x) = 0.$$

This allows us to write the following expression for the regulatory instrument

$$\tau_i^g = \tau_i^*(x) = d_i(\boldsymbol{\chi}^*(x), x) + (W'(x) - Z_i'(x))\frac{\partial f}{\partial u_i}(\boldsymbol{\chi}^*(x), x). \tag{15}$$

Thus, if we take into account that $\boldsymbol{\chi}^*(x) = \mathbf{u}^*(x) = \boldsymbol{\psi}^*(x, \boldsymbol{\phi}^*(x))$, we obtain that conditions (11) and (15) are the same so that we can conclude that[11] $\tau_i^g = \tau_i^*(x) = \phi_i^*(x) = \tau_i^f$ and we have that

Proposition 1. *There is no difference between choosing a regulatory instrument rule at $t = 0$ that implements the efficient outcome or choosing the regulatory instrument rate that maximizes social welfare at each period of time, i.e. the GSE of the policy game and the FSE coincide.*

In the next two sections we illustrate our main result by analyzing two examples. The first is an operations research model, while the second is an economic example.

3. An operations research example

We consider the model proposed by Jørgensen et al. (2003). There is a distribution channel composed by one manufacturer M and one retailer R. The manufacturer controls the rate of advertising effort, $A(t)$, in national media for the brand, while the retailer decides the local promotional activities, $P(t)$, for the brand. The manufacturer's advertising effort affects positively the brand image. The

---

[11]Superscript g stands for the regulatory instrument defined by the global Stackelberg equilibrium.

retailer's promotional activities have a positive impact on current sales, but a negative impact on the brand image. The following differential equation describes the dynamics of the brand image, $G(t)$:

$$\dot{G}(t) = \alpha A(t) - \beta P(t) - \delta G(t), \quad G(0) = G_0 \geq 0,$$

with α, β and δ positive parameters.

The sales revenue rate of the product is given by $Q(P(t), G(t)) = \gamma P(t) + \theta G(t)$, with γ and θ positive parameters. Advertising and promotions cost functions are quadratic: $C(A) = \frac{\mu_A}{2} A^2, \quad C(P) = \frac{\mu_P}{2} P^2$. The manufacturer gets a constant and predetermined share $\pi \in (0,1)$ of the total channel revenue, and the retailer gets $1 - \pi$.

When there is no regulation, the manufacturer's objective functional is

$$J_M(A(\cdot)) = \int_0^\infty e^{-\rho t} \left[\pi(\gamma P(t) + \theta G(t)) - \frac{\mu_A}{2} A(t)^2 \right] dt \tag{16}$$

and the retailer's

$$J_R(P(\cdot)) = \int_0^\infty e^{-\rho t} \left[(1 - \pi)(\gamma P(t) + \theta G(t)) - \frac{\mu_P}{2} P(t)^2 \right] dt. \tag{17}$$

Jørgensen et al. (2003) characterize the Nash equilibrium advertising and promotion strategies (Proposition 1, page 398). These strategies are constant over time since the differential game presents a linear-state structure, and hence, value functions are linear in the state variable, G.

The efficient outcome in this model corresponds to the solution of the cooperative game defined as the joint maximization as follows:

$$J_C(A(\cdot), P(\cdot)) = \int_0^\infty e^{-\rho t} \left[\gamma P(t) + \theta G(t) - \frac{\mu_A}{2} A(t)^2 - \frac{\mu_P}{2} P(t)^2 \right] dt.$$

Solving by dynamic programming, the solution to this dynamic optimization problem must comply the following HJB equation:

$$\rho W(G(t)) = \max_{A(t), P(t)} \left\{ \gamma P(t) + \theta G(t) - \frac{\mu_A}{2} A(t)^2 - \frac{\mu_P}{2} P(t)^2 + W'(G(t)) \left(\alpha A(t) - \beta P(t) - \delta G(t) \right) \right\},$$

where $W(G(t))$ represents the cooperative value function for the current value, $G(t)$, of the state variable.

The maximization of the RHS of the HJB equation yields the following first-order conditions (FOC) for an interior solution

$$-\mu_A A(t) + \alpha W'(G(t)) = 0, \quad \gamma - \mu_P P(t) - \beta W'(G(t)) = 0. \tag{18}$$

Therefore,[12]

$$A = \frac{\alpha W'(G)}{\mu_A}, \quad P = \frac{\gamma - \beta W'(G)}{\mu_P} \tag{19}$$

Inserting (19) on the RHS of the HJB equation provides

$$\rho W(G) = (\theta - \delta W'(G))G + \frac{\gamma^2 \mu_A + W'(G)^2 \left(\alpha^2 \mu_P + \beta^2 \mu_A\right) - 2\beta\gamma\mu_A W'(G)}{2\mu_A\mu_P}.$$

Since the problem is linear state, we define a linear value function as follows $W(G) = m_1 G + m_2$. Substituting $W(G)$ and its derivative, into the last expression and identifying coefficients, we obtain m_1 and m_2:

$$m_1 = \frac{\theta}{\delta + \rho}, \quad m_2 = \frac{\theta^2 \alpha^2 \mu_P + \mu_A(\gamma(\delta + \rho) - \theta\beta)^2}{2\mu_A\mu_P\rho(\delta + \rho)^2}. \tag{20}$$

Then, the final expressions of the cooperative advertising and promotion strategies read[13]:

$$A^C = \frac{\alpha\theta}{\mu_A(\rho + \delta)}, \quad P^C = \begin{cases} \frac{\gamma(\rho+\delta)-\beta\theta}{\mu_P(\rho+\delta)} & \text{if} \quad \beta < \frac{\gamma(\rho+\delta)}{\theta}, \\ 0 & \text{otherwise.} \end{cases}$$

The comparison of the Nash and cooperative strategies shows that, as expected, these strategies are different, the Nash equilibrium is not efficient, and since $\pi \in (0,1)$, both the advertising and promotion strategies are lower under the Nash equilibrium than under the cooperative solution. Note that for this example the difference functions defined in condition (5) are constant and given by $d_M \equiv 0$ and $d_R \equiv \pi\gamma$ for the manufacturer and the retailer, respectively.

In what follows we show how a regulator, acting as the leader of a Stackelberg game, can use a subsidy as a regulatory instrument to implement the efficient solution (the cooperative solution in the model at hand) as a regulated Nash equilibrium among the manufacturer and the retailer, acting as the followers of the Stackelberg game. The main objective of the subsidy is to push the manufacturer and the retailer to increase the advertising and promotion rates, respectively. Under regulation the noncooperative game described by (16) and (17) translates into:

$$J_M^R(A(\cdot)) = \int_0^\infty e^{-\rho t} \left[\pi(\gamma P(t) + \theta G(t)) - \frac{\mu_A}{2}A(t)^2 + \tau_M(t)A(t)\right] dt, \tag{21}$$

$$J_R^R(P(\cdot)) = \int_0^\infty e^{-\rho t} \left[(1 - \pi)(\gamma P(t) + \theta G(t)) - \frac{\mu_P}{2}P(t)^2 + \tau_R(t)P(t)\right] dt, \tag{22}$$

where $\tau_M(t)$ and $\tau_R(t)$ stand for the subsidies that the manufacturer and the retailer receive from the regulator, respectively.

---

[12]From now on, the time argument is removed when no confusion may arise.
[13]Superscript C stands for cooperative solution.

We want to show how the cooperative solution can be implemented both as a *feedback Stackelberg equilibrium* (FSE) and as a *global Stackelberg equilibrium* (GSE). In the next subsection, we show how the efficient outcome can be implemented as FSE.

3.1. The Feedback Stackelberg Equilibrium

In this subsection we are interested in characterizing the stagewise feedback Stackelberg equilibrium of the differential game under regulation where the objectives of the manufacturer and the retailer are described in (21) and (22). We use the same notation as in Subsection 2.1 and denote by $\tau_i(t) = \phi_i(G(t))$ with $i \in \{M, R\}$ the stationary feedback strategy of the regulator, and by $A(t) = \psi_M(G(t), \boldsymbol{\tau}(t))$ and $P(t) = \psi_R(G(t), \boldsymbol{\tau}(t))$ the stationary feedback strategy of the followers, the manufacturer (M) and the retailer (R), respectively, with $\boldsymbol{\tau}(t) = (\tau_M(t), \tau_R(t))$ and $\boldsymbol{\phi}(G(t)) = (\phi_M(G(t)), \phi_R(G(t)))$.

Let $Z_i(G)$ and $W(G)$ stand for the value function for the current value, G, of the state variable for the i-th follower, $i \in \{M, R\}$ and the leader, respectively. The FSE can be obtained using the HJB equations

$$
\rho Z_M(G) = \max_{\{A \in \Phi\}} \left\{ \pi(\gamma \psi_R(G, \boldsymbol{\phi}(G)) + \theta G) - \frac{\mu_A}{2} A^2 + \phi_M(G) A \right.
$$
$$
+ \; Z_M'(G) \left(\alpha A - \beta \psi_R(G, \boldsymbol{\phi}(G)) + \theta G \right) \Big\}, \tag{23}
$$

$$
\rho Z_R(G) = \max_{\{P \in \Phi\}} \left\{ (1 - \pi)(\gamma P + \theta G) - \frac{\mu_P}{2} P^2 + \phi_R(G) P \right.
$$
$$
+ \; Z_R'(G) \left(\alpha \psi_M(G, \boldsymbol{\phi}(G)) - \beta P + \theta G \right) \Big\}, \tag{24}
$$

$$
\rho W(G) = \max_{\{\tau_M, \tau_R \in \Upsilon\}} \left\{ \gamma \psi_R(G, \boldsymbol{\tau}) + \theta G - \frac{\mu_A}{2} (\psi_M(G, \boldsymbol{\tau}))^2 - \frac{\mu_P}{2} (\psi_R(G, \boldsymbol{\tau}))^2 \right.
$$
$$
+ \; W'(G) \left(\alpha \psi_M(G, \boldsymbol{\tau}) - \beta \psi_R(G, \boldsymbol{\tau}) + \theta G \right) \Big\}, \tag{25}
$$

The 4-tuple of strategies $(\boldsymbol{\phi}^*, \boldsymbol{\psi}^*)$ constitutes a feedback Stackelberg equilibrium if there exist 3 continuously differentiable value functions satisfying the HJB Eqs. (23), (24) and (25) and the transversality conditions $\lim_{t \to \infty} e^{-\rho t} Z_i(G^*(t)) = \lim_{t \to \infty} e^{-\rho t} W(G^*(t)) = 0$, with $i \in \{M, R\}$. Let

$$
R_M(G, \phi_M(G), \psi_R(G, \boldsymbol{\phi}(G))) = \arg \max_{\{A \in \Phi\}} \left\{ \pi(\gamma \psi_R(G, \boldsymbol{\phi}(G)) + \theta G) - \frac{\mu_A}{2} A^2 + \phi_M(G) A \right.
$$
$$
+ \; Z_M'(G) \left(\alpha A - \beta \psi_R(G, \boldsymbol{\phi}(G)) + \theta G \right) \Big\},
$$

$$
R_R(G, \phi_R(G), \psi_M(G, \boldsymbol{\phi}(G))) = \arg \max_{\{P \in \Phi\}} \left\{ (1 - \pi)(\gamma P + \theta G) - \frac{\mu_P}{2} P^2 + \phi_R(G) P \right.
$$
$$
+ \; Z_R'(G) \left(\alpha \psi_M(G, \boldsymbol{\phi}(G)) - \beta P + \theta G \right) \Big\}.
$$

Thus, the argument that maximizes the RHS of the HJB equations for the followers' problem (23),

(24) must satisfy the following conditions

$$-\mu_A A + \phi_M(G) + \alpha Z'_M(G) = 0, \tag{26}$$

$$(1-\pi)\gamma - \mu_P P + \phi_R(G) - \beta Z'_R(G) = 0, \tag{27}$$

where $Z'_i(G)$ is the *private* shadow price of the state variable when the problem is regulated through a subsidy. From (26) and (27), we obtain the instantaneous reaction functions of the manufacturer and the retailer to the action profile of the leader:

$$A = \psi_M(G, \phi(G)) = \frac{\alpha Z'_M(G) + \phi_M(G)}{\mu_A},$$

$$P = \psi_R(G, \phi(G)) = \frac{(1-\pi)\gamma + \phi_R(G) - \beta Z'_R(G)}{\mu_P}.$$

On the other hand, the FOCs for the maximization of the RHS of the leader's HJB (25) requires that the efficient conditions (18) hold, and the feedback Stackelberg equilibrium sustains the efficient solution.

FOCs (26), (27) and (18) lead to the following conditions that must be satisfied at the feedback Stackelberg equilibrium:

$$-\alpha W'(G) + \alpha Z'_M(G) + \phi_M(G) = 0, \quad -\pi\gamma + \beta W'(G) - \beta Z'_R(G) + \phi_R(G) = 0. \tag{28}$$

Then, taking into account that τ_i is equal to $\phi_i(G)$, $i \in \{M, R\}$ the efficient conditions define the optimal subsidy rules, $\phi_i^*(G)$, $i \in \{M, R\}$ as follows[14]:

$$\tau_M^f(G) = \phi_M^*(G) = \alpha(W'(G) - Z'_M(G)), \tag{29}$$

$$\tau_R^f(G) = \phi_R^*(G) = \pi\gamma - \beta(W'(G) - Z'_R(G)). \tag{30}$$

Next, after substitution of the optimal subsidy rule in the manufacturer's and retailer's reaction functions, we obtain the equilibrium feedback strategy for the agents $\psi_i^*(G, \phi^*(G))$, $i \in \{M, R\}$:

$$A^* = \psi_M^*(G, \phi^*(G)) = \frac{\alpha W'(G)}{\mu_A}, \quad P^* = \psi_R^*(G, \phi^*(G)) = \frac{\gamma - \beta W'(G)}{\mu_P}.$$

Let us note that the equilibrium strategies exclusively depend on the first derivative of the regulator's value function, that coincides with the cooperative value function, and is given by m_1 in (20) because the regulated distribution channel replicates the efficient outcome. The final step to derive

---

[14]Superscript f again stands for the subsidy supported by the feedback Stackelberg equilibrium.

the final expressions of the optimal subsidy strategies is to solve the following system of HJB equations

$$
\begin{aligned}
\rho Z_M(G) &= \pi(\gamma \psi_R^*(G, \phi^*(G)) + \theta G) - \frac{\mu_A}{2}(\psi_M^*(G, \phi^*(G)))^2 + \phi_M^*(G)\psi_M^*(G, \phi^*(G)) \\
&+ Z_M'(G)\left(\alpha \psi_M^*(G, \phi^*(G)) - \beta \psi_R^*(G, \phi^*(G)) + \theta G\right), \tag{31} \\
\rho Z_R(G) &= (1-\pi)(\gamma \psi_R^*(G, \phi^*(G)) + \theta G) - \frac{\mu_P}{2}(\psi_R^*(G, \phi^*(G)))^2 + \phi_R^*(G)\psi_R^*(G, \phi^*(G)) \\
&+ Z_R'(G)\left(\alpha \psi_M^*(G, \phi^*(G)) - \beta \psi_R^*(G, \phi^*(G)) + \theta G\right). \tag{32}
\end{aligned}
$$

Since the differential game is linear state, we postulate linear value functions for the manufacturer and the retailer as we previously did for the cooperative value function as follows:

$$
Z_M(G) = k_1 G + k_2, \quad Z_R(G) = l_1 G + l_2.
$$

Substituting these value functions and their derivative in the HJB equations in (31) and (32), and identifying coefficients we obtain:

$$
k_1 = \frac{\pi\theta}{\delta + \rho}, \quad k_2 = \frac{\alpha^2\theta^2\mu_P + 2\pi\mu_A(\beta\theta - \gamma(\delta+\rho))^2}{2\mu_A\mu_P\rho(\delta+\rho)^2},
$$

$$
l_1 = \frac{(1-\pi)\theta}{\delta + \rho}, \quad l_2 = \frac{\mu_A(\beta\theta - \gamma(\delta+\rho))^2 + 2\theta^2(1-\pi)\alpha^2\mu_P}{2\mu_A\mu_P\rho(\delta+\rho)^2}.
$$

Once these coefficients are known, next proposition presents the final expressions of the optimal subsidy strategies in the FSE.

Proposition 2. *The optimal policy defined by the FSE is given by the following rules*

$$
\tau_M^f = \frac{(1-\pi)\alpha\theta}{\delta + \rho}, \quad \tau_R^f = \begin{cases} \frac{\pi(\gamma(\delta+\rho)-\beta\theta)}{\delta+\rho} & \text{if } \beta < \frac{\gamma(\rho+\delta)}{\theta} \\ 0 & \text{otherwise} \end{cases}.
$$

These constant subsidy rates τ_M^f and τ_R^f allow the regulator to implement the efficient solution as a FSE, such that the optimal advertising and promotions strategies coincide with the cooperative solution.

3.2. The Global Stackelberg Equilibrium

In this subsection we follow the lines in Subsection 2.2 to characterize the global Stackelberg equilibrium (GSE) via the use of dynamic programming. Given the subsidy rules, $\tau_M(G)$ and $\tau_R(G)$ announced by the regulator at time $t = 0$, the manufacturer and the retailer compete à la Nash as in the second stage of the previous section. Suppose that each player knows that the other player uses Markovian strategies, that is, the manufacturer knows that $P = \chi_R(G)$, and the retailer knows that $A = \chi_M(G)$. Then, the Markov-perfect Nash equilibrium (MPNE) of the game played by the

manufacturer and the retailer must satisfy the HJB equations

$$\rho Z_M(G) \;=\; \max_{\{A \in \Phi\}} \left\{ \pi(\gamma\chi_R(G) + \theta G) - \frac{\mu_A}{2}A^2 + \tau_M(G)A + Z'_M(G)\left(\alpha A - \beta\chi_R(G) + \theta G\right)\right\},$$

$$\rho Z_R(G) \;=\; \max_{\{P \in \Phi\}} \left\{ (1-\pi)(\gamma P + \theta G) - \frac{\mu_P}{2}P^2 + \tau_R(G)P + Z'_R(G)\left(\alpha\chi_M(G) - \beta P + \theta G\right)\right\},$$

that yield the following FOCs

$$-\mu_A A + \tau_M(G) + \alpha Z'_M(G) \;=\; 0, \tag{33}$$

$$(1-\pi)\gamma - \mu_P P + \tau_R(G) - \beta Z'_R(G) \;=\; 0, \tag{34}$$

This approach gives the same FOCs as the ones obtained for the FSE except that the strategies that support the MPNE do not include directly the dependence of the control variables, A and P, with respect to the subsidies, τ_M and τ_R.

To calculate the subsidy rules that implement the cooperative outcome, we first differentiate with respect to G the FOCs

$$-\mu_A\chi'_M(G) + \tau'_M(G) + \alpha Z''_M(G) \;=\; 0, \tag{35}$$

$$-\mu_P\chi'_R(G) + \tau'_R(G) - \beta Z''_R(G) \;=\; 0. \tag{36}$$

Differentiating the HJB equations, and making use of the envelope theorem, we obtain

$$(\rho + \delta + \beta\chi'_R(G))Z'_M(G) = \pi(\gamma\chi'_R(G) + \theta) + \tau'_M(G)\chi_M(G) + Z''_M(G)(\alpha\chi_M(G) - \beta\chi_R(G) - \delta G),$$

$$(\rho + \delta - \alpha\chi'_M(G))Z'_R(G) = (1-\pi)\theta + \tau'_R(G)\chi_R(G) + Z''_R(G)(\alpha\chi_M(G) - \beta\chi_R(G) - \delta G).$$

Next, we eliminate Z'_i and Z''_i for $i \in \{M, R\}$ using (33),(34), (35) and (36) and obtain the following equations:

$$\mu_A\chi'_M(G)(\alpha\chi_M(G) - \beta\chi_R(G) - \delta G) + (\tau_M(G) - \mu_A\chi_M(G))(\rho + \delta + \beta\chi'_R(G))$$
$$+\; \tau'_M(G)(\delta G + \beta\chi_R(G)) + \pi\alpha(\theta + \gamma\chi'_R(G)) = 0, \tag{37}$$

$$\mu_P\chi'_R(G)(\alpha\chi_M(G) - \beta\chi_R(G) - \delta G) + (\tau_R(G) - \mu_P\chi_R(G))(\rho + \delta - \alpha\chi'_M(G))$$
$$+\; \tau'_R(G)(\delta G - \alpha\chi_M(G)) + (1-\pi)(\gamma(\delta + \rho) - \beta\theta - \alpha\gamma\chi'_M(G)) = 0. \tag{38}$$

Then, subsidy rules that guide the manufacturer and the retailer to achieve the cooperative path as a MPNE are the solution of the system of differential equations (37) and (38) for $\chi_M(G) = A^C$ and

$\chi_R(G) = P^C$. Therefore, system (37) and (38) simplifies as follows:

$$\mu_P(\delta + \rho)[(\delta + \rho)\tau_M(G) - (1 - \pi)\alpha\theta] + (-\beta^2\theta + (\delta + \rho)(\mu_P\delta G + \beta\gamma))\tau_M'(G) = 0,$$

$$\mu_A(\delta + \rho)[(\delta + \rho)\tau_R(G) + \beta\theta - \gamma(\delta + \rho)] + (-\alpha^2\theta + (\delta + \rho)\mu_A\delta G)\tau_R'(G) = 0.$$

Notice that the two equations are independent. Given the structure of the model, we look for linear subsidy rules $\tau_M(G) = \sigma_1 G + \sigma_2$, $\tau_R(G) = \kappa_1 G + \kappa_2$. Substituting these expressions and identifying coefficients, we obtain:

$$\sigma_1 = 0, \quad \sigma_2 = \frac{(1 - \pi)\alpha\theta}{\delta + \rho}, \quad \kappa_1 = 0, \quad \kappa_2 = \frac{\pi((\delta + \rho)\gamma - \beta\theta)}{\delta + \rho}.$$

Therefore,[15]

$$\tau_M^g = \tau_M^*(G) = \frac{(1 - \pi)\alpha\theta}{\delta + \rho}, \quad \tau_R^g = \tau_R^*(G) = \begin{cases} \frac{\pi((\delta + \rho)\gamma - \beta\theta)}{\delta + \rho} & \text{if} \quad \beta < \frac{\gamma(\rho + \delta)}{\theta} \\ 0 & \text{otherwise} \end{cases}, \quad (39)$$

and we obtain that $\tau_i^f = \phi_i^*(G) = \tau_i^*(G) = \tau_i^g$, $i \in \{M, R\}$. Thus, this allows us to corroborate the result in Proposition 1, and there is no difference between choosing subsidy rules at $t = 0$ that implement the cooperative (efficient) outcome or choosing the subsidies that maximize the cooperative payoff at each period of time, i.e. the GSE of the policy game and the FSE coincide.

4. An economic example

We consider a linear-quadratic specification of the model proposed by Benchekroun and Long (1998). There is an oligopoly consisting of n identical firms that produce a homogenous good and compete in quantities. Each firm has a constant unit cost $c \geq 0$. Let $q_i(t)$ denote firm i 's production at time t. Industry output is $Q(t) = \sum_{i=1}^n q_i(t)$ and the linear inverse demand function is $P(Q(t)) = P(0) - bQ(t)$, $P(0) - c = a > 0$. The amount of pollutant emitted by firm i is $e_i(t) = q_i(t)$. We assume that the dynamics of the pollution stock, $S(t)$, obeys the law $\dot{S}(t) = Q(t) - \delta S(t)$, where $\delta > 0$ stands for the rate of natural purification. $D(S(t)) = (\gamma/2)S^2$ where $\gamma > 0$ is the damage function representing the harm causes by pollution.

If the market is not regulated, firms do not take into account the environmental damages caused by the pollution stock and choose the output that maximizes their current profits given by $\pi_i(t) = \{P(q_i(t) + Q_{-i}(t)) - c\} q_i(t)$, where $Q_{-i}(t) = \sum_{j=1, j \neq i}^n q_j(t)$. The maximization of profits requires that the marginal revenue be equal to the marginal cost. If we focus on a symmetric equilib-

---

[15]Superscript g again stands for the subsidies defined by the global Stackelberg equilibrium.

rium this condition reads[16]

$$Q^m = \frac{a}{b}\frac{n}{n+1}, \quad q^m = \frac{a}{b}\frac{1}{n+1}. \tag{40}$$

In this model the market equilibrium is inefficient for two reasons. On the one hand, the firms have market power that will cause a contraction of output and pollution with respect to the efficient levels. On the other hand, the output generates an externality that will cause an expansion of output and pollution with respect to the efficient levels. Thus, to characterize the efficient outcome we have to maximize the net social welfare defined as the consumers' surplus plus the producers' surplus minus environmental damages.[17]

$$\max_{\{Q(t)\geq 0\}} \int_0^\infty \left\{ aQ(t) - \frac{b}{2}(Q^{(}t))^2 - \frac{\gamma}{2}(S(t))^2 \right\} e^{-rt}dt,$$

$$s.t. \ \dot{S}(t) = Q(t) - \delta S(t), \quad S(0) = S_0.$$

Solving by dynamic programming, the solution to this dynamic optimization problem must comply the following HJB equation:

$$rW(S(t)) = \max_{\{Q(t)\geq 0\}} \left\{ aQ(t) - \frac{b}{2}(Q(t))^2 - \frac{\gamma}{2}(S(t))^2 + W'(S(t))\left(Q(t) - \delta S(t)\right) \right\}, \tag{41}$$

where $W(S(t))$ represents the maximum discounted present value of the flow of net social welfare for the current value, $S(t)$, of the pollution stock.

The maximization of the RHS of the HJB equation yields the following FOC for an interior solution

$$a - bQ(t) = -W'(S(t)) \quad \text{or equivalently} \quad P(0) - bQ(t) = c - W'(S(t)). \tag{42}$$

The FOC establishes that the price must be equal to the full marginal cost that includes the marginal cost of production and the *social* pollution shadow price. The latter is given by the reduction in the discounted present value of the flow of net social welfare because of an increase in the pollution stock caused by an increase in output.

To proceed with the computation of the efficient solution, we substitute (42) in the HJB equation (41)

$$rW(S) = a\frac{a + W'(S)}{b} - \frac{b}{2}\left(\frac{a + W'(S)}{b}\right)^2 - \frac{\gamma}{2}S^2 + W'(S)\left(\frac{a + W'(S)}{b} - \delta S\right),$$

---

[16]The superscript m represents the market equilibrium values.

[17]Note that for this example the difference functions defined in condition (5) is given by bq for the symmetric case and is explained by the difference between the price and the marginal revenues that appears when there is no perfect competition.

that yields the following non-linear differential equation

$$rW(S) = \frac{(a + W'(S))^2}{2b} - \frac{\gamma}{2}S^2 - \delta SW'(S). \tag{43}$$

In order to find a solution for this equation, in view of the linear-quadratic structure of the problem we guess a quadratic representation for the value function W

$$W(S) = \frac{A_r}{2}S^2 + B_r S + C_r,$$

which implies that $W'(S) = A_r S + B_r$ and where A_r, B_r, and C_r are unknowns to be determined.

The substitution of $W(S)$ and $W'(S)$ into Eq. (43) gives a system of Riccati equations that must be satisfied for every S. Selecting the stable solution of the system, we obtain the following values for the coefficients of the regulator's value function

$$A_r = b\rho < 0, \quad B_r = \frac{a\rho}{r + \delta - \rho} < 0, \tag{44}$$

where ρ is the negative root of equation

$$\rho^2 - (r + 2\delta)\rho - \frac{\gamma}{b} = 0. \tag{45}$$

Then, the optimal strategy for production reads

$$Q^*(S) = \frac{a(r + \delta)}{b(r + \delta - \rho)} + \rho S, \tag{46}$$

that establishes that the output decreases with respect to the pollution stock and that the optimal strategy for the firms is

$$q = \frac{a(r + \delta)}{bn(r + \delta - \rho)} + \frac{\rho}{n}S. \tag{47}$$

In what follows we show how a regulator, acting as the leader of a Stackelberg game, can use tax emissions to implement the efficient solution as a regulated Nash equilibrium among the n firms, acting as the followers of the Stackelberg game. We aim showing how the efficient solution can be implemented both as a FSE and as a GSE.

4.1. The Feedback Stackelberg Equilibrium

Next, we calculate the FSE of the linear-quadratic differential game we have just introduced. For this game the HJB equations (6) and (7) are

$$rZ_i(S) = \max_{\{q_i \in \Phi\}} \left\{ \left(a - b \left(\sum_{j \neq i} \psi_j(S, \phi(S)) + q_i \right) - \phi(S) \right) q_i + Z_i'(S) \left(q_i + \sum_{j \neq i} \psi_j(S, \phi(S)) - \delta S \right) \right\},$$
(48)

$$rW(S) = \max_{\{\tau \in \Upsilon\}} \left\{ a \sum_{i=1}^{n} \psi_i(S, \tau) - \frac{b}{2} \left(\sum_{i=1}^{n} \psi_i(S, \tau) \right)^2 - \frac{\gamma}{2} S^2 + W'(S) \left(\sum_{i=1}^{n} \psi_i(S, \tau) - \delta S \right) \right\}.$$
(49)

Focusing on interior solutions, the FOCs for the maximization of the RHS of (48) are

$$a - b \left(\sum_{j \neq i} \psi_j(S, \phi(S)) + 2q_i \right) - \phi(S) + Z_i'(S) = 0,$$

that for a symmetric equilibrium where for all j, $q_j = \psi(S, \phi(S))$ yield

$$q = \psi(S, \phi(S)) = \frac{a - \phi(S) + Z'(S)}{b(n+1)}.$$

Being the total output of the industry

$$Q = n\psi(S, \phi(S)) = \frac{n(a - \phi(S) + Z'(S))}{b(n+1)}.$$
(50)

On the other hand, the FOC for the maximization of the RHS of (49) obviously yields the efficient condition in (42), where $Q = \sum_{i=1}^{n} \psi_i(S, \tau)$. Then, taking into account that τ is equal to $\phi(S)$ and using (50) we obtain the equilibrium strategy for the tax

$$a - \frac{n(a - \phi(S) + Z'(S))}{b(n+1)} + W'(S) = 0,$$

$$\tau^* = \phi^*(S) = -\frac{1}{n} \left(a - nZ'(S) + (n+1)W'(S) \right).$$
(51)

Next, substituting $\phi^*(S)$ in (50) the equilibrium strategy for total output is obtained

$$Q^* = n\psi^*(S) = \frac{a + W'(S)}{b},$$
(52)

that gives the equilibrium strategy for the firms' output

$$q^* = \psi^*(S) = \frac{a + W'(S)}{bn}.$$
(53)

24

Observe that this strategy does not depend on the private shadow price of pollution, but only depends on the social pollution shadow price that satisfies the efficient condition in (42) and that this shadow price is completely characterized by (44).

Therefore, the optimal tax can be computed using expression (51) and eliminating $W'(S)$ one has

$$\tau = \phi^*(S) = -\frac{1}{n}\left(\frac{a(r+\delta+n\rho)}{r+\delta-\rho} + (n+1)b\rho S - nZ'(S)\right). \tag{54}$$

Thus, in order to obtain a complete characterization of the tax, we need to solve the representative firm's HJB equation given by the following expression

$$rZ(S) = (a - bQ^*(S) - \phi(S))\frac{Q^*(S)}{n} + Z'(S)(Q^*(S) - \delta S).$$

Substituting the output and the tax rate by (46) and (54) respectively, the following differential equation is obtained

$$rZ(S) = \frac{b}{n^2}\left(\frac{a(r+\delta)}{b(r+\delta-\rho)} + \rho S\right)^2 + Z'(S)\frac{n-1}{n}\frac{a(r+\delta)}{b(r+\delta-\rho)} + \left(\frac{n-1}{n}\rho - \delta\right)SZ'(S). \tag{55}$$

In order to solve this equation, we also guess a quadratic representation of the representative firm's value function

$$Z(S) = \frac{A_f}{2}S^2 + B_f S + C_f,$$

that yields $Z'(S) = A_f S + B_f$. The substitution of $Z(S)$ and $Z'(S)$ into Eq. (55) gives a system of Riccati equations whose solution for the first two parameters is

$$A_f = \frac{2b\rho^2}{n(n(r+2\delta) - 2(n-1)\rho)} > 0,$$

$$B_f = \frac{2a\rho(r+\delta)(n(r+2\delta) - (n-1)\rho)}{n(r+\delta-\rho)(n(r+2\delta) - 2(n-1)\rho)(n(r+\delta) - (n-1)\rho)} < 0.$$

The private pollution shadow price is given by the following expression

$$Z'(S) = \frac{2\rho}{n(n(r+2\delta) - 2(n-1)\rho)}\left(\frac{a(r+\delta)(n(r+2\delta) - (n-1)\rho)}{(r+\delta-\rho)(n(r+\delta) - (n-1)\rho)} + b\rho S\right).$$

Then, eliminating $Z'(S)$ of expression (54), the optimal tax is obtained using (45) to simplify the expression of the independent term.

Proposition 3. *The optimal policy defined by the FSE is given by the following rule* $\tau^* = \phi^*(S) = \eta^s + \alpha^s S$, *where*

$$\alpha^s = \frac{(1+n)\gamma + (n-1)b\rho^2}{n(r+2\delta) - 2(n-1)\rho} > 0, \tag{56}$$

and

$$\eta^s = \frac{aF}{(r + \delta - \rho)(n(r + 2\delta) - 2(n-1)\rho)(n(r+\delta) - (n-1)\rho)}, \tag{57}$$

where

$$\begin{aligned}
F &= (n-1)(2\delta + nr)\frac{\gamma}{b} - n(r+\delta)^2(r+2\delta) \\
&\quad - \left(\left(\frac{2\gamma}{b} + (2\delta + r)\delta\right)n^2 - \left(\frac{4\gamma}{b} + 2r^2 + 7r\delta + 8\delta^2\right)n + r^2 + 3r\delta + 4\delta^2 + \frac{2}{b}\gamma\right)\rho. \tag{58}
\end{aligned}$$

The slope of the tax rule is positive indicating that the tax increases with the pollution stock. However, η could be negative. Nevertheless, it is easy to check that if damages are high enough this term will be positive and the optimal policy consists of taxing emissions for any level of the pollution stock.[18]

4.2. The Global Stackelberg Equilibrium

In order to calculate the tax rule corresponding to the GSE, we impose that $\chi(S) = \psi^*(S)$ since we look for the tax rule that implements the efficient outcome. Then, taking into account that according to (47), the slope of the efficient strategy for output is $(\psi^*)' = \rho/n$ so that differential equation (14) gives the following expression for the linear-quadratic model at hand

$$(r + \delta - (n-1)\frac{\rho}{n})\tau(S) + ((\delta - (n-1)\frac{\rho}{n})S - \frac{(n-1)a(r+\delta)}{bn(r+\delta - \rho)})\tau'(S)$$

$$= \frac{a}{n(r + \delta - \rho)}\left((r+\delta)\rho + (n-1)\rho^2 - (r+\delta)^2\right) + (2n\rho - 2(n+1)\delta - r(1+n))\frac{b}{n}\rho S. \tag{59}$$

Given the structure of the model, we look for a linear tax rule $\tau(S) = \eta + \alpha S$ that satisfies this equation. By substitution, we obtain the following expression

$$(r + \delta - (n-1)\frac{\rho}{n})(\eta + \alpha S) + ((\delta - (n-1)\frac{\rho}{n})S - \frac{(n-1)a(r+\delta)}{bn(r+\delta - \rho)})\alpha$$

$$= \frac{a}{n(r + \delta - \rho)}\left((r+\delta)\rho + (n-1)\rho^2 - (r+\delta)^2\right) + (2n\rho - 2(n+1)\delta - r(1+n))\frac{b}{n}\rho S.$$

This expression yields the same values for α and η as the ones obtained for the FSE.

The procedure used by Benchekroun and Long (1998) to calculate the tax rule, although yields the same results, is not exactly the same we have just presented. They assume not only that the tax rule is linear, but also that firms use a linear Markovian strategy $\chi(S) = X + YS$, and calculate the Markov-Perfect Nash Equilibrium of the game played by the firms. Solving in this way, the coefficients

---

[18]Observe that F is increasing in γ.

of the output strategies depend both on the parameters of the model and the coefficients of the tax rule. Then doing $\chi(S) = \psi^*(S)$, the coefficients of the tax rule can be obtained. Obviously, the tax rule they obtain is the same we have derived in this paper.[19]

5. Conclusions

This paper claims that the GSE and the FSE coincide if we focus on the design of the first-best policy that implements the efficient outcome as a regulated market equilibrium. For a class of differential games we show that the regulatory instrument rule given by the FSE is the same rule that implements the efficient outcome as a GSE. This coincidence is based on the fact that although the FSE assumes that the regulator selects the level of the regulatory instrument rate at each period of time, as the equilibrium is calculated using dynamic programming, the first-order condition for the maximization of the RHS of the HJB equation implicitly defines a policy rule. Notice that according to this condition the optimal regulatory instrument rate level is dependent on the first derivative of the value function which in turn depends on the state variable. Thus, once we have a solution for the value function, we obtain a policy rule that induces the players to act efficiently. Then, as the GSE also implements the efficient outcome, the regulatory instrument rules defined by both equilibria must be identical. In the second part of the paper, we check that this is the case for two examples. The first example is an operations research model, while the second is an economic example. The first example deals with coordination in a marketing channel and the model is borrowed from Jørgensen et al. (2003). The second example deals with environmental regulation in a polluting oligopoly and we use Benchekroun and Long's (1998) model. First, we calculate the subsidy rule for both equilibria in the marketing channel example and we obtain exactly the same expressions. Second, we compute the tax rule for both equilibria for a linear-quadratic specification of the polluting oligopoly and show that both expressions coincide. Moreover, although we illustrate our argument using these two examples, as we show in Section 2 the argument will work for other examples of market regulation provided that we are interested in the design of the first-best policy for the class of differential games defined in this paper. Notice that as the first-best policy implements the efficient outcome, the policy rule supported by the FSE that is calculated to maximize social welfare must be the same as the policy rule that induces the players to act efficiently in the GSE.

This result has two clear implications. The first one is that it does not matter whether the regulator commits at the beginning of the game with a policy rule or she has a commitment just for a period of time to choose the level of the regulatory instrument rate. In both cases the optimal policy rule is the same. The second implication is that the *classical* approach of market regulation where the regulator

---

[19]Indeed we have confirmed with the authors that the term η in Prop. 6 of page 338 contains a typo and should read $\eta = [\gamma/(r+\delta) - [\delta b(r+\delta) + (nr + (n+1)\delta - (n-1)\rho)\alpha]/(n(r+\delta) - (n-1)\rho]\hat{S}_\infty$. Using this expression, it can be checked that the two tax rules coincide.

acts as a social planner coincides with the *strategic* approach where the regulator acts as the leader of a policy game.

Finally, we would want to highlight that not only the FSE coincides with the GSE for the characterization of the first-best policy, but that it offers a clear methodology to calculate second-best policies, as Martín-Herrán and Rubio (2018b) showed, while this is not the case for the GSE.

References

[1] Başar, T., & Haurie, A. (1984). Feedback equilibria in differential games with structural and modal Uncertainties. In J. B. Cruz, Jr. (Ed.), *Advances in large scale systems 1* (pp.163-201). Greenwich, CT: JAI Press.

[2] Başar, T., Haurie, A., & Ricci, G. (1985). On the dominance of capitalists leadership in a 'Feedback-Stackelberg' solution of a differential game model of capitalism. *Journal of Economic Dynamics and Control*, 9, 101-125.

[3] Başar, T., & Olsder, G. J. (1999). *Dynamic noncooperative game theory*. (2nd ed.). Philadelphia: SIAM.

[4] Başar, T., & Zaccour, G. (2018). *Handbook of dynamic game theory*. Springer International Publishing.

[5] Benchekroun, H., & Long, N. V. (1998). Efficiency inducing taxation for polluting oligopolists. *Journal of Public Economics*, 70, 325-342.

[6] Benchekroun, H., & Long, N. V. (2002). On the multiplicity of efficiency-inducing tax rules. *Economic Letters*, 76, 331-336.

[7] Bergstrom, T. C., Gross, J. G., & Porter, R. C. (1981). Efficiency-inducing taxation for a monopolistically supplied depletable resource. *Journal of Public Economics*, 15, 23-32.

[8] Chutani, A., & Sethi, S. P. (2012). Optimal advertising and pricing in a dynamic durable goods supply chain. *Journal of Optimization Theory and Applications*, 154, 615-643.

[9] Cohen, D., & Michel, P. (1988). How should control theory be used to calculate a time-consistent government policy?. *The Review of Economic Studies*, 55(2), 263-274.

[10] Colombo, L., & Labrecciosa, P. (2019). Stackelberg versus Cournot: A differential game approach. *Journal of Economic Dynamicc and Control*, 101, 239-261.

[11] De Giovanni, P., Karray, S., & Martín-Herrán, G. (2019). Vendor management inventory with consigment contracts and the benefits of cooperative advertising. *European Journal of Operational Research*, 272, 465-480.

[12] Feenstra, T., Kort, P. M., & de Zeeuw, A. (2001). Environmental policy instrument in an international duopoly with feedback investment strategies. *Journal of Economic Dynamics and Control*, 25, 1665-1687.

[13] Friedman, A. (1971). *Differential games.* New York: Wiley-Interscience.

[14] Haurie, A., Krawczyk, J. B., & Zaccour, G. (2012). *Games and dynamic games.* Singapore: World Scientific.

[15] Jørgensen, S., Taboubi, S., & Zaccour, G. (2003). Retail promotions with negative brand image effects: Is cooperation possible? *European Journal of Operational Research*, 150, 395-405.

[16] Jørgensen, S., & Zaccour, G. (2004). *Differential games in marketing.* Massachussets: Kluwer Academic Publishers.

[17] Karp, L. (1992). Efficiency inducing tax for a common property oligopoly. *Economic Journal*, 102, 321-332.

[18] Karp, L., & Livernois, J. (1992). On efficiency-inducing taxation for a non-renewable resource monopolist. *Journal of Public Economics*, 49, 219-239.

[19] Karray, S., & Martín-Herrán, G. (2009). A dynamic model for advertising and pricing competition between national and store brands. *European Journal of Operational Research*, 193, 451-467.

[20] Martín-Herrán, G., & Rincón-Zapatero, J. P. (2005). Efficient Markov perfect Nash equilibria: theory and application to dynamic fishery games. *Journal of Economic Dynamics & Control*, 29, 1073-1096.

[21] Martín-Herrán, G., & Rubio, S. J. (2018a). Optimal environmental policy for a polluting monopoly with abatement costs: Taxes versus standards. *Environmental Modelling & Assessment*, 23, 671-689.

[22] Martín-Herrán, G., & Rubio, S. J. (2018b). Second-best taxation for a polluting monopoly with abatement investment. *Energy Economics*, 73, 178-193.

[23] Martín-Herrán, G., Sigué, S. P. & Zaccour, G. (2011). Strategic interactions in traditional franchise systems: Are franchisors always better off?. *European Journal of Operational Research*, 213, 526-537.

[24] Martín-Herrán, G., & Taboubi, S. (2015). Price coordination in distribution channels: A dynamic perspective. *European Journal of Operational Research*, 240, 401-414.

[25] Rincón-Zapatero, J. P., Martín-Herrán, G., & Martínez, J. (2000). Identification of efficient subgame-perfect Nash equilibria in a class of differential games. *Journal of Optimization Theory and Applications*, 104, 235-242.

[26] Rubio, S. J. (2006). On coincidence of feedback Nash equilibria and Stackelberg equilibria in economic applications of differential games. *Journal of Optimization Theory and Applications*, 128, 203-221.

[27] Rubio, S. J., & Escriche, L. (2001). Strategic Pigouvian taxation, stock externalities and polluting non-renewable resources. *Journal of Public Economics*, 79, 297-313.

[28] Simaan, M., & Cruz, J. B. Jr. (1973). Additional aspects of the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11, 613-626.

[29] Stimming, M. (1999). Capital-accumulation games under environmental regulation and duopolistic competition. *Journal of Economics*, 69, 267-287.

[30] Taboubi, S. (2019). Incentive mechanisms for price and advertising coordination in dynamic marketing channels. *International Transactions in Operational Research*, 26(6), 2281-2304.

[31] Tahvonen, O. (1996). Trade with polluting nonrenewable resources. *Journal of Environmental Economics and Management*, 30, 1-17.

[32] Wirl, F. (2012). Global warming: Prices versus quantities from a strategic point of view. *Journal of Environmental Economics and Management*, 64, 217-229.

[33] Wirl, F. (2014). Taxes versus permits as incentive for the temporal supply of a clean technology by a monopoly. *Resource and Energy Economics*, 36, 248-269.

[34] Xepapadeas, A. P. (1992). Environmental policy, adjustment costs, and behavior of the firm. *Journal of Environmental Economics and Management*, 23, 258-275.

[35] Yanase, A. (2009). Global environment and dynamic games of environmental policy in an international duopoly. *Journal of Economics*, 97, 121-140.