A REVIEW OF DYNAMIC STACKELBERG GAME MODELS

TAO LI*

Santa Clara University
500 El Camino Real
Santa Clara, CA 95053, USA

SURESH P. SETHI

The University of Texas at Dallas
800 W Campbell Rd
Richardson, TX 75080, USA

ABSTRACT. Dynamic Stackelberg game models have been used to study sequential decision making in noncooperative games in various fields. In this paper we give relevant dynamic Stackelberg game models, and review their applications to operations management and marketing channels. A common feature of these applications is the specification of the game structure: a decentralized channel consists of a manufacturer and independent retailers, and a sequential decision process with a state dynamics. In operations management, Stackelberg games have been used to study inventory issues, such as wholesale and retail pricing strategies, outsourcing, and learning effects in dynamic environments. The underlying demand typically has a growing trend or seasonal variation. In marketing, dynamic Stackelberg games have been used to model cooperative advertising programs, store brand and national brand advertising strategies, shelf space allocation, and pricing and advertising decisions. The demand dynamics are usually extensions of the classic advertising capital models or sales-advertising response models. We begin each section by introducing the relevant dynamic Stackelberg game formulation along with the definition of the equilibrium used, and then review the models and results appearing in the literature.

1. **Introduction.** In 1934, H. von Stackelberg introduced a concept of a hierarchical solution for markets in which some firms have dominating power over the others ([57]). This solution concept is now known as the *Stackelberg equilibrium* or the *Stackelberg solution* (we use equilbrium and solution interchangably in this paper) which, in the context of two-person nonzero-sum static games, involves players with asymmetric roles, one *leading* (called *the leader*, (she)) and the other *following* (called *the follower*, (he)). The static game as formulated by von Stackelberg, proceeds with the leader announcing her action first, and the follower reacting to it by optimizing his performance index under the leader's announced policy. Of course, the leader anticipates this response (assuming that she knows the objective function of the follower) and picks the action that optimizes her performance index given the follower's rational response. Assuming that the follower's optimal (rational)

response is unique to each announced policy of the leader (that is, he has a unique rational response curve), then the leader's best policy is the one that optimizes her performance index on the rational response curve of the follower, which together with the corresponding unique action of the follower is known as the Stackelberg solution. If the follower's response is not unique, however, then the rational response curve is replaced with a rational response set, in which case taking a pessimistic approach on the part of the leader, her optimization problem is to find the best policy under the worst choices by the follower (*worst* from the point of view of the leader) from the rational response set; such a solution is known as the *generalized Stackelberg solution* ([39]; [4]).

The notion of the Stackelberg solution was later extended to multistage settings in the early 1970's by [51] and [52], who also introduced the notion of a *feedback Stackelberg solution* in which the leader dictates her policy choices on the follower only stagewise, but not globally. Such a solution concept requires (in a dynamic game setting) that the players know the current state of the game in every period, and its derivation involves a backward recursion (as in dynamic programming), where at every step of the recursion the Stackelberg solution of a static game is obtained. In this equilibrium concept, the leader has only stagewise advantage over the follower. On the other hand, if the leader has dynamic information, and is able to announce her policy for the entire duration of the dynamic game ahead of time (and not stagewise), then the Stackelberg solution, even though well defined as a concept, is generally very difficult to obtain, because the underlying optimization problems are on the policy spaces of the two players, with the response sets or functions generally being infinite dimensional. In such games, the leader has global advantage over the follower. Derivation of such *global Stackelberg solutions* for dynamic games with dynamic information structures also has connections to incentive design or mechanism design problems, and is still an active research area; see [9] and the text by [4] for an overview of various types of Stackelberg solutions.

It is possible to extend the feedback Stackelberg solution concept to continuous-time dynamic games, called *differential games*. [3], and [4] argue that the continuous-time problem can be viewed as the limit of a sequence of discrete-time games as the number of stages becomes unbounded in any finite interval of the differential game. Since any two consecutive decision points get arbitrarily close to one another, the stagewise advantage of the leader (on the follower) introduced in the discrete-time games turns into instantaneous advantage in a differential game. For futher details, see [3] who provide the associated coupled system of HJB equations to characterize the feedback Stackelberg equilibrium. These equations are parabolic partial differential equations. As a special case, they consider a finite-horizon linear-quadratic game, derive the corresponding Riccati equation, and show the existence of a solution for it in the case with a sufficiently small horizon.

[8] consider an infinite-horizon stochastic Stackelberg differential game involving Brownian motion, and obtain a sufficient condition for a feedback Stackelberg equilibrium. In doing this, they note that in the case of a Nash differential game, the associated HJB system of equations provides the equilibrium in terms of feedbacks, and these feedbacks are obtained from the Nash equilibrium at the level of Hamiltonian at each instant. In their setting of Stackelberg differential games, they write an analogous system by using the feedbacks obtained from the Stackelberg equilibrium at the level of Hamiltonian at each instant, and show that this approach provides a feedback Stackelberg equilibrium for the game under consideration. In contrast

to [3], the HJB equations in their case turn out to be elliptic partial differential equations due to the infinite horizon nature of their problem.

The feedback Stackelberg solution concept has been applied in applications in economics and management science ([19]; [26], [28]; [36]). For work on global Stackelberg solutions in continuous time, see [46] and [9].

Stackelberg game models have been used to study conflicts and coordination issues associated with inventory and production policies, outsourcing, capacity and shelf space allocation decisions, pricing strategies, learning effects, dynamic competitive advertising strategies, pricing for new products, etc. These studies are the ones we shall review in this paper. Particularly related to our review are applications of Stackelberg game models in operations management and marketing channels. We do not review applications in economics, for which we refer the interested readers to [2].

It should be noted that the literatures on operations management and marketing channels are closely related, since both deal with physical delivery of the product from suppliers to end users through intermediaries and in the most efficient manner. However, the operations management literature has been more concerned with production quantity and inventory decisions, and demand uncertainty, whereas marketing papers have looked at pricing and promotions decisions, customer heterogeneity, and brand positioning. Competition and channel coordination are important issues in both fields.

In operations management, Stackelberg game models have been applied to the study of topics such as inventory and production issues, wholesale and retail pricing strategies, and outsourcing in dynamic environments. The underlying demand may be growing over time or have seasonal variation. In the marketing area, we discuss applications such as cooperative advertising programs, shelf space allocation, and price and advertising decisions. The demand dynamics are usually advertising capital models (e.g., [45]) or sales-advertising response models (e.g., [56] and [49]). In the former, advertising is considered as an investment in the stock of goodwill, and in the latter, sales-advertising response is specified as a direct relation between the rate of change in sales and advertising.

The fact that the papers that have applied Stackelberg game models to operations management and marketing channels problems are relatively few in number and recent, suggests that the literature is still at a formative stage. Although, over the last two decades, supply chain management as a whole has garnered much research attention, most studies of strategic interactions between the channel members are based on the static newsvendor framework. It should be noted that while the static models may have two stages, one where the leader moves and the second where the follower moves, thereby modeling sequential decision making in the channel, they overlook strategic issues that arise when the firms interact with each other repeatedly over time and make decisions in a dynamic fashion. We therefore hope that this review, by focusing on dynamic interactions between the channel members, generates further interest in this emerging area of research.

This paper is organized as follows. In Section 2.1 we give the formulations of Stackelberg games in continuous time. In this section, we first introduce the basic concepts of the components involved in Stackelberg games, and then discuss in more details on the Stackelberg equilibria obtained under two different information structures: open-loop and feedback information structures. We introduce the definitions

of these Stackelberg equlibria and explain what procedures should be taken to obtain them. All the applications of open-loop Stackelberg equlibrium reviewed in this paper are in deterministic settings, so we formulate the open-loop Stackelberg equilibrium in a deterministic setting. Some of the applications of feedback Stackelberg equlibrium reviewed in this paper are in stochastic settings, so we formulate the feedback Stackelberg equilibrium in a stochastic setting. In Sections 2.2 and 2.3, we review the applications of Stackelberg game models in continuous time in the area of operations management and marketing channels. For each application, we briefly explain the key features of the model, how it fits into the Stackelberg game framework, and discuss its main results. To accommodate the frequent usage of Stackelberg games in discrete time in the operations management literature, in Section 2.4 we give the formulation of the Stackelberg game in discrete time. We also introduce the definitions of the feedback Stackelberg equlibrium in discrete time. In Section 2.5, we review the applications of Stackelberg game models in discrete time in the area of operations management. Once again, we briefly explain the key features of each model, how it fits into the Stackelberg game framework, and discuss its main results. Section 3 concludes the paper and discusses future research directions. Table 1 summarizes the notations used in the paper and Table 2 summarizes the descriptions of the models reviewed in this paper.

2. The formulations and applications of Stackelberg games.

2.1. The formulations of Stackelberg games in continuous time.
A Stackelberg game has the following structure: (1) The state of the dynamic system at any time t is characterized by a set of variables called the state variables. Typical state variables are market share, sales, inventory, production cost, goodwill, etc. (2) There are decision variables called controls that are chosen by the game players. For example, order/production quantity, pricing, advertising, and shelf-space decisions, etc. (3) The evolution of the state variables over time is described by a set of differential equations involving both state and control variables. (4) Each player has an objective function (e.g., the present value of the profit stream over time) to optimize by its choice of controls.

There are several different concepts of equilibria in Stackelberg games. Like the static Stackelberg game, due originally to [53], a dynamic Stackelberg game is hierarchical (or sequential) in nature. We consider a dynamic Stackelberg game involving two players labeled as L (the leader) and F (the follower) making decisions over a finite horizon T. We shall also remark on what needs to be done when $T = \infty$. Let $x(t) \in \mathbb{R}^n$ denote the vector of state variables, $u \in \mathcal{U} \subset \mathbb{R}^{m_1}$ denote the leader's control vector, and $v \in \mathcal{V} \subset \mathbb{R}^{m_2}$ denote the follwer's control vector. The sequence of the game is as follows. The leader announces her strategies/actions first, and then the follower responds to the leader's strategies/actions by choosing his optimal decision. A Stackelberg equilibrium is obtained when the leader solves her problem taking into account the follower's optimal response.

Further specification is required in terms of the information structure of the players in order to precisely define a Stackelberg equilibrium. In this section, we shall consider two different information structures. The first one, called the *open-loop* information structure, assumes that the player must formulate their strategies/actions at time t only with the knowledge of the initial condition of the state at time zero. The second one, called the *feedback* information structure, assumes that the players use their knowledge of the current state at time t in order to formulate their

decisions at time t. We give the procedure for obtaining an open-loop solution in Section 2.1.1. The feedback solution procedure will be described in Section 2.1.2.

2.1.1. *Open-loop solution.* The solution procedure requires the leader to first anticipate the follower's best response to her announced policy. The anticipation is derived from solving the follower's optimization problem given the leader's decisions. We then substitute the follower's response function into the leader's problem and solve for the leader's optimal decisions. The decisions of the leader, together with the follower's best response to those decisions, constitute a Stackelberg equilibrium; See [9] for further details. For the open-loop solution, given an announced control $u\left(\cdot\right) = \left(u\left(t\right), 0 \leq t \leq T\right)$ by the leader, the follower's optimal control problem is

$$\max_{v(\cdot)} J_F\left(t, x_0, v\left(\cdot\right); u\left(\cdot\right)\right) = \int_0^T e^{-\rho t} \pi_F\left(t, x\left(t\right), u\left(t\right), v\left(t\right)\right) dt + e^{-\rho T} S_F\left(x\left(T\right)\right)$$

subject to the state equation

$$\dot{x}\left(t\right) = f\left(t, x\left(t\right), u\left(t\right), v\left(t\right)\right), \quad x\left(0\right) = x_0, \tag{1}$$

where $\rho > 0$ is the follower's discount rate, π_F is his instantaneous profit rate function, $S_F\left(x\left(T\right)\right)$ is the salvage value, and x_0 is the initial state. Recall that the open-loop information structure for both players means that the controls $u\left(t\right)$ and $v\left(t\right)$ at time t depend only on the initial state x_0. We assume that $f\left(t, \cdot, u, v\right)$, $\pi_F\left(t, \cdot, u, v\right)$, and $S_F\left(\cdot\right)$ are continuously differentiable on $\mathbb{R}^n$, $\forall t \in [0, T]$.

Let $v^0\left(\cdot\right)$ denote the follower's optimal response to the announced control $u\left(\cdot\right)$ by the leader. The follower's optimal response must be made at time zero and will, in general, depend on the entire announced control path $u\left(\cdot\right)$ by the leader and the initial state x_0. Clearly, given $u\left(\cdot\right)$, which can be expressed as a function of x_0 and t, the follower's problem reduces to a standard optimal control problem. This problem is solved by using the maximum principle (see e.g., [50], and [22]), which requires introducing an adjoint variable or a shadow price that decouples the optimization problem over the interval $[0, T]$ into a sequence of static optimization problems, one for each instant $t \in [0, T]$. Note also that the shadow price at time t denotes the per unit value of a change in the state $x(t)$ at time t, and the change $\dot{x}(t)$ in the state at time t requires only the time t control $u\left(t\right)$ from the leader's announced control trajectory $u\left(\cdot\right)$. Nevertheless, we shall see later that the follower's optimal response $v^0(t)$ at any given time t will depend on the entire $u\left(\cdot\right)$ via the solution of the two-boundary problem resulting from the application of the Maximum Principle. Let us now proceed to solve the follower's problem given $u\left(\cdot\right)$.

The follower's current-value Hamiltonian is

$$H_F\left(t, x, \lambda, u, v\right) \quad = \quad \pi_F\left(t, x, u, v\right) + \lambda' f\left(t, x, u, v\right), \tag{2}$$

where $\lambda \in \mathbb{R}^n$ is the (column) vector of the shadow prices associated with the state variable x, the (row) vector λ' denotes its transpose, and it satisfies the adjoint equation

$$\dot{\lambda}' \quad = \quad \rho\lambda' - \frac{\partial}{\partial x} H_F\left(t, x, \lambda, u, v\right), \quad \lambda'\left(T\right) = \frac{\partial}{\partial x} S_F\left(x(T)\right). \tag{3}$$

Here we have suppressed the argument t as is standard in the control theory literature, and we will do this whenever convenient and when there arises no confusion in doing so. Note that the gradient $\partial H_F / \partial x$ is a row vector by convention.

The necessary condition for the follower's optimal response denoted by v^0 is that it maximizes H_F with respect to $v \in \mathcal{V}$, i.e.,

$$v^0(t) = \arg \max_{v \in \mathcal{V}} H_F(t, \lambda(t), x(t), u(t), v). \tag{4}$$

Clearly, it is a function that depends on t, x, λ, and u. A usual procedure is to obtain an optimal response function $r(t, x, \lambda, u)$ so that $v^0(t) = r(t, x(t), \lambda(t), u(t))$. In the case when $v^0(t)$ is an interior solution, r satisfies the first-order condition

$$\frac{\partial}{\partial v} H_F(t, x, \lambda, u, r) = 0. \tag{5}$$

If the Hamiltonian H_F defined in (2) is jointly concave in the variables x and v for any given u, then the condition (5) is also sufficient for the optimality of the response r for the given u.

It is now possible to complete our discussion of the dependence of the follower's optimal response r on $u(\cdot)$ and x_0. Note that the time t response of the follower is given by

$$v^0(t) = r(t, x(t), \lambda(t), u(t)). \tag{6}$$

But the resulting two-point boundary value problem given by (25), (2) with $r(t, x, \lambda, u)$ in place of v makes it clear that the shadow price $\lambda(t)$ at time t depends on the entire $u(\cdot)$ and x_0. Thus, we have the dependence of $v^0(t)$, and therefore of $v^0(\cdot)$, on $u(\cdot)$ and x_0.

If we can explicitly solve for the optimal response $r(t, x, \lambda, u)$, then we can specify the leader's problem to be

$$\max_{u(\cdot)} \quad J_L(t, x, u(\cdot)) = \int_0^T e^{-\rho t} \pi_L(t, x, u, r(t, x, \lambda, u))\, dt + e^{-\rho T} S_L(x(T)), \tag{7}$$

$$\dot{x} = f(t, x, u, r(t, x, \lambda, u)), \quad x(0) = x_0, \tag{8}$$

$$\dot{\lambda} = \rho \lambda' - \frac{\partial}{\partial x} H_F(t, x, \lambda, u, r(t, x, \lambda, u)), \quad \lambda'(T) = \frac{\partial}{\partial x} S_F(x(T)), \tag{9}$$

where $\rho > 0$ is the leader's discount rate and the differential equations in (8) and (9) are obtained by substituting the follower's best response $r(t, x, \lambda, u)$ for v in the state equation (25) and the adjoint equation (3), respectively. We formulate the leader's Hamiltonian

$$
\begin{aligned}
\bar{H}_L(t, x, \lambda, \phi, \psi, u) \quad &= \pi_L(t, x, u, r(t, x, \lambda, u)) + \phi' f(t, x, u, r(t, x, \lambda, u)) \\
&\quad + \psi' \left[\rho \lambda - \left(\frac{\partial}{\partial x} H_F(t, x, \lambda, u, r(t, x, \lambda, u)) \right)' \right], \tag{10}
\end{aligned}
$$

where $\phi \in \mathbb{R}^n$ and $\psi \in \mathbb{R}^n$ are the vectors of the shadow prices associated with x and λ, respectively, and they satisfy the adjoint equations

$$
\begin{aligned}
\dot{\phi}' &= \rho \phi' - \frac{\partial}{\partial x} \bar{H}_L(t, x, \lambda, \phi, \psi, u, r(t, x, \lambda, u)) \\
&= \rho \phi' - \frac{\partial}{\partial x} \pi_L(t, x, u, r(t, x, \lambda, u)) - \phi' \frac{\partial}{\partial x} f(t, x, u, r(t, x, \lambda, u)) \\
&\quad + \psi' \frac{\partial^2}{\partial x^2} H_F(t, x, \lambda, u, r(t, x, \lambda, u)), \\
\phi'(T) &= \frac{\partial}{\partial x} S_L(x(T)) - \frac{\partial^2}{\partial x^2} S_F(x(T))\, \psi(T), \tag{11}
\end{aligned}
$$

$$
\begin{aligned}
\dot{\psi}' \;=\;& \rho\psi' - \frac{\partial}{\partial\lambda}\bar{H}_L\left(t,x,\lambda,\phi,\psi,u,r\left(t,x,\lambda,u\right)\right)\\
=\;& \rho\psi' - \frac{\partial}{\partial\lambda}\pi_L\left(t,x,u,r\left(t,x,\lambda,u\right)\right) - \phi'\frac{\partial}{\partial\lambda}f\left(t,x,u,r\left(t,x,u,\lambda\right)\right)\\
&+\psi'\frac{\partial^2}{\partial x^2}H_F\left(t,x,\lambda,u,r\left(t,x,\lambda,u\right)\right),\;\;\psi\left(0\right)=0.
\end{aligned}
\tag{12}
$$

In (10), we are using the notation $\bar{H}_L$ for the Hamiltonian to recognize that it uses the optimal response function of the follower for v in its definition. In (11) and (12), we have used the envelope theorem (see, e.g., [16]) in taking the derivative $\partial\bar{H}_L/\partial x$.

Finally, it is important to remark that unlike in Nash differential games, the adjoint equation and the transversality condition $\phi'\left(T\right)$ in (11) have a second-order term each on account of the facts that the adjoint equation (9) is a state equation in the leader's problem and that it has a state-dependent terminal condition, also in (9), arising from the sequential game structure of Stackelberg differential games. The necessary optimality condition for the leader's optimal control $u^*\left(t\right)$ is that it maximizes $\bar{H}_L$ with respect to $u\in\mathcal{U}$, i.e.,

$$
u^*\left(t\right) = \arg\max_{u\in\mathcal{U}}\bar{H}_L\left(t,x\left(t\right),\lambda\left(t\right),\phi\left(t\right),\psi\left(t\right),u\right).
\tag{13}
$$

In the case when $u^*\left(t\right)$ is an interior solution, this is equivalent to the condition

$$
\frac{\partial}{\partial u}\bar{H}_L\left(t,x\left(t\right),\lambda\left(t\right),\phi\left(t\right),\psi\left(t\right),u^*\left(t\right)\right)=0.
\tag{14}
$$

Also if $\bar{H}_L$ is jointly concave in the variables x, λ, and u, then (14) is also sufficient for optimality of $u^*\left(t\right)$. Once we have obtained $u^*\left(t\right)$ from (14), we can substitute it into r to obtain the optimal control $v^*\left(t\right)$ of the follower.

Under the open-loop information structure, the leader's information set is $\{x_0\}$ and the follower's information set is $\{x_0,u(\cdot)\}$, where $u(\cdot)$ is the leader's strategy, since the follower makes his decision after the leader announces her whole strategy over $[0,T]$. The leader's policy $u(t)$ depends on x_0. The value $v(t)$ of the follower's response strategy at a future time t depends on the leader's whole strategy $u(\cdot)$. Then, the leader's strategy space and the follower's response strategy space are

$$
\mathcal{U} = \{u|\;u:\{x_0\}\times[0,T]\to U\}
$$

and

$$
\mathcal{V} = \{v|\;v:\{x_0\}\times[0,T]\times\mathcal{U}\to V\},
$$

respectively.

Definition 2.1. A pair of strategies $(u^*,v^*)\in\mathcal{U}\times\mathcal{V}$ is called an open-loop Stackelberg equilibrium if the following holds:

$$
J_L(u^*,v^*(u^*)) \ge J_L(u,v^*(u)),\;\forall\;u\in\mathcal{U},
$$
$$
J_F(u,v^*(u)) \ge J_F(u,v(u)),\;\;\forall\;(u,v)\in\mathcal{U}\times\mathcal{V}.
$$

In many cases, it may not be possible to have an explicit expression for the response function r. Then, the alternate method would be to impose (5) as an equality constraint on the leader's optimization problem (7)-(9). It is easy to redo the above procedure in these cases and derive the following result. Note that this will require a Lagrange multiplier function $\mu\left(\cdot\right)\in\mathbb{R}^n$ associated with constraint (5).

Theorem 2.2. *For the Stackelberg differential games under consideration in this subsection, assume that*

i) $f(t, \cdot, u, v)$ is twice continuously differentiable on $\mathbb{R}^n$, $\forall t \in [0, T]$.

ii) $\pi_F(t, \cdot, u, v)$ and $S_F(\cdot)$ are twice continuously differentiable on $\mathbb{R}^n$, $\forall t \in [0, T]$.

iii) $\pi_L(t, \cdot, u, v)$ and $S_L(\cdot)$ are twice continuously differentiable on $\mathbb{R}^n$, $\forall t \in [0, T]$.

iv) $H_F(t, x, u, \cdot)$ is continuously differentiable and strictly convex on $\mathcal{V}$.

v) $f(t, x, \cdot, v)$, $\pi_i(t, x, \cdot, v)$, $i = 1, 2$, are continuously differentiable on $\mathcal{U}$.

If $u^*(t)$ and $v^*(t), 0 \le t \le T$, provide an open-loop Stackelberg solution for the leader and the follower, $u^*(t)$ is interior to the sets $\mathcal{U}$ and $\mathcal{V}$, and $\{x^*(t), 0 \le t \le T\}$ denotes the corresponding state trajectory, then there exist continuously differentiable functions $\lambda(\cdot)$, $\phi(\cdot)$, $\psi(\cdot)$: $[0, T] \to \mathbb{R}^n$ and a continuous function $\mu(\cdot)$: $[0, T] \to \mathbb{R}^n$, such that the following relations are satisfied:

$$\dot{x}^* = f(t, u^*, v^*), \ x^*(0) = x_0,$$

$$\dot{\lambda}' = \rho\lambda' - \frac{\partial}{\partial x}H_F(t, x^*, \lambda, u^*, v^*), \ \lambda'(T) = \frac{\partial}{\partial x}S_F(x^*(T)),$$

$$\dot{\phi}' = \rho\phi' - \frac{\partial}{\partial x}H_L(t, x^*, \lambda, \phi, \psi, \mu, u^*, v^*),$$

$$\phi'(T) = \frac{\partial}{\partial x}S_L(x^*(T)) - \frac{\partial^2}{\partial x^2}S_F(x^*(T))\psi(T),$$

$$\dot{\psi}' = \rho\psi' - \frac{\partial}{\partial \lambda_F}H_L(t, x^*, \lambda, \phi, \psi, \mu, u^*, v^*), \ \psi(0) = 0,$$

$$\frac{\partial}{\partial u}H_L(t, x^*, \lambda, \phi, \psi, \mu, u^*, v^*) = 0, \ \frac{\partial}{\partial v}H_L(t, x^*, \lambda, \phi, \psi, \mu, u^*, v^*) = 0,$$

$$\frac{\partial}{\partial v}H_F(t, x^*, \lambda, u^*, v^*) = 0,$$

where H_F is defined by (2) and H_L is defined by

$$H_L(t, x, \lambda, \phi, \psi, \mu, u, v) = \pi_L(t, x, u, v(t, x, \lambda, u)) + \phi' f(t, x, u, v(t, x, \lambda, u))$$

$$+\psi'\left[\rho\lambda - \left(\frac{\partial}{\partial x}H_F(t, x, \lambda, u, v)\right)'\right] + \frac{\partial}{\partial v}H_F(t, x, \lambda, u, v)\mu.$$

We note that in infinite horizon problems, i.e., when $T = \infty$, it is usually assumed that the salvage values $S_L = S_F = 0$. Then the practice is to replace the terminal conditions on λ, ϕ and ψ by $\lim_{T \to \infty} e^{-\rho T}\lambda(T) = 0$, $\lim_{T \to \infty} e^{-\rho T}\phi(T) = 0$, and $\lim_{T \to \infty} e^{-\rho T}\psi(T) = 0$, respectively. We should note that these conditions are not necessary, but are sufficient when coupled with appropriate concavity conditions.

It is known that in general open-loop Stackelberg equilibria are not time consistent. This means that given an opportunity to revise her strategy at any future time after the initial time, the leader would benefit by choosing another strategy than the one she chose at the initial time. Thus, an open-loop Stackelberg equilibrium only makes sense if the leader can credibly pre-commit at time zero her strategy for the entire duration of the game. In many management settings such commitment is not observed in practice. Nevertheless, there is a considerable literature dealing with open-loop Stackelberg equilibrium, on account of its mathematical tractability.

Most of the literature on Stackelberg games, with a few exceptions, has assumed that the roles of the players are fixed at the outset, and the leader remains a leader for the entire duration of the game; likewise follower remains as follower.

Perhaps the first paper that brought up the issue of whether being a leader is always advantageous to a player is Başar (1973). It turns out that leadership is not always a preferred option for the players: for some classes of games, leadership of one specific player is preferred by both (in the sense that both players collect highest utilities when compared with what they would collect under other combinations), which is a stable situation, whereas there are other classes of games where either both players prefer their own leadership or neither do, which leads to a stalemate situation. Sometimes, the players do not have the option of leadership open to them, but leadership is governed by an exogenous process (say a Markov chain), which determines who should be the leader and who should be the follower at each stage of the game, perhaps based on the history of the game or the current state of the game. Başar and Haurie (1984) have shown that feedback Stackelberg solution is a viable concept for such (stochastic) dynamic games also, and have obtained recursions for the solution.

More recently, Başar et al. (2010) introduced the notion of mixed leadership in nonzero-sum differential games with open-loop information patterns, where the same player can act as leader in some decisions and as follower in others, depending on the instrument variables he is controlling. This kind of game proceeds as follows. The two players first announce their decisions both as leaders, simultaneously, and then both of them respond simultaneously, in the role of followers, to maximize their corresponding objective functions in the sense of Nash equilibrium. Given the rational response, the two players again act as leaders to maximize their respective performance index also in the sense of Nash equilibrium. Therefore, the mixed differential game consists of two Nash games (the parallel play) and one Stackelberg game (hierarchical play). Başar et al. (2010) used the Maximum Principle to obtain a set of algebraic equations and differential equations with mixed-boundary conditions that characterize an optimal open-loop Stackelberg solution. In particular, a linear-quadratic differential game was discussed as a specific case and related coupled Riccati equations with mixed boundary conditions were derived by which the optimal Stackelberg solution can be represented in terms of the system state. While this is not the same as feedback Stackelberg solution, the representation in terms of the system state may perform better in presence of not explicitly modeled small noise.

[7] study open-loop Stackelberg equilibria of two-player linear-quadratic differential games with mixed leadership. They show that, under some appropriate assumptions on the coefficients, there exists a unique Stackelberg solution to such a differential game. Moreover, by means of the close interrelationship between the Riccati equations and the set of equations satisfied by the optimal open-loop control, they provide sufficient conditions to guarantee the existence and uniqueness of solutions to the associated Riccati equations with mixed boundary conditions. As a result, the players' open-loop strategies can be represented in terms of the system state.

2.1.2. *Feedback solution.* In the preceding section, we have formulated the open-loop Stackelberg solution concept. We shall now develop the procedure to obtain a feedback Stackelberg solution. While open-loop Stackelberg solutions can be said to be static in the sense that decisions can be derived at the initial time, without regard to the state variable evolution beyond that time, feedback Stackelberg equilibrium strategies at any time t are functions of the values of the state variables at that time. They are perfect state-space equilibria because the necessary optimality conditions

are required to hold for all values of the state variables, and not just the values that lie on the optimal state-space trajectories. Therefore, the solutions obtained continue to remain optimal at each instant of the time after the game has begun. Thus, they are known as *subgame perfect* because they do not depend on the initial conditions.

Next, we formulate the continuous-time feedback Stackelberg game in a stochastic setting. We consider a stochastic differential game with its state evolving as

$$\begin{cases} dx(t) = f(x(t), u(t), v(t))dt + \sigma(x(t))dW(t), \\ x(0) = x, \end{cases}$$

where W is a d-dimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$,

$$f : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R}^n \text{ and } \sigma : \mathbb{R}^n \to \mathbb{R}^{n \times d}$$

are Lipschitz continuous functions, and u and v are the decision variables of the leader and the follower, respectively. Since the applications we review in this paper are infinite-horizon setups, we formulate the feedback Stackelberg game in continuous time in infinite horizon in this section. The feedback Stackelberg game in continuous time in finite horizon can be formulated in a corresponding manner. The objective for the leader and the follower to maximize are as follows:

$$J_L = E \int_0^\infty e^{-\rho t} \pi_L(x(t), u(t), v(t))dt,$$

$$J_F = E \int_0^\infty e^{-\rho t} \pi_F(x(t), u(t), v(t))dt.$$

with

$$\pi_L : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R} \text{ and } \pi_F : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \to \mathbb{R}.$$

In a continuous-time feedback Stackelberg game, the leader determines her instantaneous strategy of the form $u(x)$ in accordance with the feedback information structure, and the follower also makes his instantaneous decision $v(x, u(x))$ based on the observed state x and the leader's instantaneous action as the game evolves. Therefore, the admissible strategy spaces for the leader and the follower are

$$\mathcal{U} = \{u | u : \mathbb{R}^n \to U \text{ and } u(x) \text{ is Lipschitz continuous in } x\},$$
$$\mathcal{V} = \{v | v : \mathbb{R}^n \times U \to V \text{ and } v(x, u) \text{ is Lipschitz continuous in } (x, u)\},$$

where U and V are given subsets in $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$, respectively.

For a pair of strategies $(u, v) \in \mathcal{U} \times \mathcal{V}$, we denote by $x^{t,x}(\cdot; u, v)$, the solution of the parameterized state equation

$$\begin{cases} dx(s) = f(x(s), u(x(s)), v(x(s), u(x(s))))ds + \sigma(x(s))dW(s), \\ x(t) = x, \end{cases}$$

and let $J_L^{t,x}(u(\cdot), v(\cdot, u(\cdot)))$ and $J_F^{t,x}(u(\cdot), v(\cdot, u(\cdot)))$ represent the corresponding objectives of the two players, respectively, i.e.,

$$J_L^{t,x}(u(\cdot), v(\cdot, u(\cdot))) =$$
$$E \int_t^\infty e^{-\rho(s-t)} \pi_L(x^{t,x}(s; u, v), u(x^{t,x}(s; u, v)), v(x^{t,x}(s; u, v), u(x^{t,x}(s; u, v))))ds,$$

$$J_F^{t,x}(u(\cdot), v(\cdot, u(\cdot))) =$$

$$E \int_t^\infty e^{-\rho(s-t)} \pi_F(x^{t,x}(s; u, v), u(x^{t,x}(s; u, v)), v(x^{t,x}(s; u, v), u(x^{t,x}(s; u, v)))) ds,$$

where we should stress that $u(\cdot)$, $v(\cdot, u(\cdot))$ evaluated at any state y are $u(y)$, $v(y, u(y))$, respectively.

Definition 2.3. A pair of strategies $(u^*, v^*) \in \mathcal{U} \times \mathcal{V}$ is called a feedback Stackelberg equilibrium if the following holds:

$$\begin{aligned} J_L^{0,x}(u^*(\cdot), v^*(\cdot, u^*(\cdot))) &\geq J_L^{0,x}(u(\cdot), v^*(\cdot, u(\cdot))), \ \forall u \in \mathcal{U}, \ \forall x \in \mathbb{R}^n, \\ J_F^{0,x}(u^*(\cdot), v^*(\cdot, u^*(\cdot))) &\geq J_F^{0,x}(u^*(\cdot), v(\cdot, u^*(\cdot))), \ \forall v \in \mathcal{V}, \ \forall x \in \mathbb{R}^n. \end{aligned} \tag{15}$$

Remark 1. Since we consider an infinite-horizon Stackelberg game in which the coefficients are independent of t, it is easy to see that for any $(u, v) \in \mathcal{U} \times \mathcal{V}$ and $(t, x) \in [0, \infty) \times \mathbb{R}^n$,

$$\begin{aligned} J_L^{t,x}(u(\cdot), v(\cdot, u(\cdot))) =& J_L^{0,x}(u(\cdot), v(\cdot, u(\cdot))), \ J_F^{t,x}(u(\cdot), v(\cdot, u(\cdot))) \\ =& J_F^{0,x}(u(\cdot), v(\cdot, u(\cdot))). \end{aligned}$$

Therefore, (26) is equivalent to

$$\begin{aligned} J_L^{t,x}(u^*(\cdot), v^*(\cdot, u^*(\cdot))) &\geq J_L^{t;x}(u(\cdot), v^*(\cdot, u(\cdot))), \ \forall u \in \mathcal{U}, \ \forall(t, x) \in [0, \infty) \times \mathbb{R}^n, \\ J_F^{t,x}(u^*(\cdot), v^*(\cdot, u^*(\cdot))) &\geq J_F^{t,x}(u^*(\cdot), v(\cdot, u^*(\cdot))), \ \forall v \in \mathcal{V}, \ \forall(t, x) \in [0, \infty) \times \mathbb{R}^n. \end{aligned}$$

Remark 2. A remarkable difference between Definitions 1 and 2 is in the second inequality which is for the follower's best response. For an open-loop Stackelberg equilibrium, the follower's best response is for the leader's (announced) strategy $u \in \mathcal{U}$ for the entire game; but for a feedback Stackelberg equilibrium, the follower's best response is for the optimal (announced) decision by the leader, i.e., $u^*(\cdot)$.

Remark 3. If we change the definition of $\mathcal{V}$ in Definition 2 from $\mathcal{V} = \{v | v : \mathbb{R}^n \times U \to V$ and $v(x, u)$ is Lipschitz continuous in $(x, u)\}$ to $\mathcal{V} = \{v | v : \mathbb{R}^n \to V$ and $v(x)$ is Lipschitz continuous in $x\}$, that is, if $v(x, u)$ is independent of u, then Definition 2 is reduced to the definition for feedback Nash equilibrium.

2.2. **Applications in operations management.** We review applications in which the leader is the supplier/manufacturer (M) who decides on variables such as the wholesale price and/or production rate, and the follower is the retailer (R) whose decision variables can be, for example, retail price and shelf-space allocation. In the reminder of the paper, subscripts M and R will be associated with the manufacturer and the retailer, respectively.

2.2.1. *Eliashberg and Steinberg (1987): Pricing and production with constant wholesale price.* [20] study a decentralized assembly system composed of a manufacturer and a retailer. The retailer processes the product and his demand has seasonal fluctuations. As an aside, [47] uses a general time-varying demand function $D(t) = a(t) - b(t) p(t)$, where $a(t)$ is the market potential, $b(t)$ is the coefficient of price sensitivity, and $p(t)$ is the retailer's price. In [20], it is assumed that $b(t)$ is constant and $a(t)$ has the form $a(t) = -\alpha_1 t^2 + \alpha_2 t + \alpha_3, 0 \leq t \leq T$, where α_1, α_2 and α_3 are positive parameters and $T = \alpha_2 / \alpha_1$.

The manufacturer and the retailer play a Stackelberg differential game in which the manufacturer is the leader and the retailer is the follower. The retailer decides his processing, pricing, and inventory policies. The manufacturer chooses her

production rate and a constant wholesale price. The retailer's problem is

$$\max_{I_R(\cdot),p(\cdot)} \left\{ J_R = \int_0^T \left[p(t) D(t) - w Q_R(t) - C_R(Q_R(t)) - h_R I_R(t) \right] dt \right\},$$

$$\dot{I}_R(t) = Q_R(t) - D(t), \ I_R(0) = I_R(T) = 0, \ I_R(t) \geq 0,$$

where $Q_R(t)$ is his processing rate, $C_R(Q_R)$ is his processing cost function, h_R is his inventory holding cost per unit, and $I_R(t)$ is his inventory level. Similar to [47], a linear holding cost function is assumed. It is also assumed that the processing cost function is increasing and strictly convex.

This paper shows that the retailer follows a two-part processing strategy. During the first part of the processing schedule, he processes at a constant increasing rate. This policy builds up inventory initially and then draws down inventory until it reaches zero at a time t_R^*. During the second part, which begins at the stockless point t_R^*, he processes at precisely the market demand rate. Pricing also follows a two-part strategy. The price charged by the retailer is first increasing at a decreasing rate and then decreasing at an increasing rate. The inventory builds up for a while and then reaches zero. From then on, the retailer processes just enough to meet demand.

An intuitive interpretation is as follows. The retailer, facing a seasonal demand that increases and then decreases, can smooth out processing operations by carrying inventory initially due to the assumption of convex processing cost. In other words, if he did not carry any inventory throughout the entire horizon, he could incur higher costs due to the convexity of the processing cost function.

Turning to the manufacturer's problem, let Q_M, I_M, h_M, and $C_M(Q_M)$ denote her production rate, inventory level, inventory holding cost per unit, and the production cost function, respectively. The manufacturer's problem is formulated as

$$\max_{Q_M(\cdot),w} \left\{ J_M = \int_0^T \left[(w - c_M) Q_R(w,t) - C_M(Q_M(t)) - h_M I_M(t) \right] dt \right\},$$

$$\dot{I}_M(t) = Q_M(t) - Q_R(w,t), \ I_M(0) = I_M(T) = 0,$$

$$Q_M(t) \geq Q_R(t) \geq 0, \ I_M(t) \geq 0,$$

where c_M is her cost per unit transferred to the retailer and $Q_R(w,t)$ is the best response of the retailer at time t given w. It is assumed that $w > c_M$ and that the production cost function is quadratic.

The manufacturer's policies are characterized as following a two-part production policy. During the first part, she produces at a constantly increasing rate. During the second part, which begins at the manufacturer's stockless point t_M^*, she produces at exactly the retailer's processing rate. In general, if the manufacturer's inventory holding cost per unit is sufficiently low and the retailer's processing efficiency and inventory holding cost per unit are high, then the manufacturer can also smooth out her operations.

2.2.2. Desai (1992): Pricing and production with retailer carrying no inventory.
[17] analyzes the production and pricing decisions in a supply chain in which a manufacturer produces the goods and sells them through a retailer who serves the final market. This paper differs from [20] in three ways. First, it allows the manufacturer to change the wholesale price over time. Second, the retailer is not allowed to carry inventory. Third, a quadratic holding cost function is assumed. The retailer faces a price-dependent seasonal demand. As in [20], [17] uses a time

varying demand function $D(t) = a(t) - bp(t)$, where $a(t) = \alpha_1 + \alpha_2 \sin(\pi t/T)$ and T is the duration of the season. The retailer (the follower) decides on the pricing policy, and his problem is given by

$$\max_{p(\cdot)} J_R = \int_0^T [p(t) - w(t)] D(t) \, dt,$$

where $p(t)$ is the retail price and $w(t)$ is the wholesale price announced by the manufacturer. Note that [20] assume the wholesale price to be constant. Here, the manufacturer (the leader) decides on the wholesale price $w(t)$ and her production rate $Q_M(t)$. Her problem is formulated as

$$\max_{w(\cdot), Q_M(\cdot)} J_M = \int_0^T \left[w(t) D(t) - c_M Q_M^2(t) - h_M I_M^2(t) \right] dt + S_M I_M(T),$$

where h_M is the inventory holding cost, c_M is the per unit production cost, $I_M(t)$ is the inventory level at time t, and S_M is the per unit salvage value. The inventory dynamics is

$$\dot{I}_M(t) = \frac{dI_M}{dt} = Q_M(t) - D(t), \ I_M(0) = I_{M0},$$

where $Q_M(t) \geq 0$ and $D(t) \geq 0$.

The paper demonstrates that once the production rate becomes positive, it does not become zero again, which implies production smoothing. However, none of the gains of production smoothing are passed on to the retailer. The optimal production rate and the inventory policy are a linear combination of the nominal demand rate, the peak demand factor, the salvage value, and the initial inventory. In the scenario where the retailing operation does not require an effort, the pricing policies of the manufacturer and the retailer and the production policy of the manufacturer have the synergistic effect that an increase in the manufacturer's price or production rate or the retailer's price leads to an increase in the rate of change of inventory. However, in the scenario where the retailing operation does benefit from the effort, the retailer's pricing policy may not necessarily be synergistic with the other policies.

2.2.3. *Desai (1996): Pricing and production with further processing by the retailer.* This work extends [17] by requiring the retailer to process the goods received from the manufacturer before they can be sold in the final market. The manufacturer makes the production and pricing decisions while the retailer decides on the processing rate and pricing policies. The paper investigates optimal policies under three types of contracts: contracts under which the manufacturer charges a constant price throughout a season, contracts under which the retailer processes at a constant rate throughout the season, and contracts under which the manufacturer and retailer cooperate to make decisions jointly. It is shown that the type of contract does not significantly impact the retailer's price. However, the type of contract has an impact on the manufacturer's price and the production rate as well as the retailer's processing rate. If the demand is not highly seasonal, a constant processing rate contract will lead to higher production and processing rates, and a lower manufacturer's price compared to a constant manufacturer's price contract.

2.2.4. *Kogan and Tapiero (2007a): Inventory game with endogenous demand.* This paper considers a supply chain that consists of a manufacturer (leader) and a retailer (follower) facing a time-dependent endogenous demand depending on the retail

price. Furthermore, the retailer has a finite processing capacity, which requires consideration of the inventory effect. Therefore, the retailer must decide on the retail price $p(t)$ and the order quantity $Q_R(t)$. The manufacturer, on the other hand, has ample capacity and decides the wholesale price $w(t)$ only. It is assumed that the game is played over a season of length T which includes a short promotional period $[t_s, t_f]$ such as the Christmas holiday season, during which both the demand potential $a(t)$ and the customer price sensitivity $b(t)$ are high. Specifically, the demand $D(t) = a(t) - b(t)p(t)$, where

$$a(t) = \begin{cases} a_1, \ t < t_S \ and \ t \geq t_f \\ a_2, \ t_S \leq t < t_f \end{cases}$$

$$b(t) = \begin{cases} b_1, \ t < t_S \ and \ t \geq t_f \\ b_2, \ t_S \leq t < t_f \end{cases}$$

with $a_2 > a_1$ and $b_2 > b_1$. The manufacturer is assumed to be restricted to setting a constant wholesale price w_1 in the regular periods and $w_2 \leq w_1$ in the promotion period. The manufacturer has ample capacity and produces exactly based on the retailer's order to maximize her profits:

$$\max_{w(\cdot)} \int_0^T [w(t)Q_R(t) - c_M Q_R(t)]\, dt$$

subject to $w(t) \geq c_M$, where c_M is the per unit production cost.

The retailer's decides his order quantity $Q_R(t)$ and the retail price $p(t)$, $0 \leq t < T$, by solving the following problem:

$$\max_{Q_R(\cdot),p(\cdot)} \int_0^T \{p(t)[a(t) - b(t)p(t)] - c_R Q_R(t) - w(t)Q_R(t) - h(X(t))\}\, dt,$$
$$\dot{I}_R(t) = Q_R(t) - [a(t) - b(t)p(t)],$$
$$0 \leq Q_R(t) \leq \bar{Q}; \ a(t) - b(t)p(t) \geq 0; \ p(t) \geq 0,$$

where $\bar{Q}$ is the retailer's maximum processing rate.

The optimal solution to the centralized channel as well as the Stackelberg equilibrium is obtained. The analysis requires the system to begin in a steady state at time 0, go to a transient state in response to promotional decisions, and then revert back to the steady state by the end of the season at time T. Thus, the solution is meant to be implemented in a rolling horizon fashion.

Under reasonable conditions on the parameters, formulas for the equilibrium values of the regular and promotional wholesale prices for the manufacturer are derived. It is shown that it is beneficial for the retailer to change pricing and processing policies in response to the reduced promotional wholesale price and the increased customer price sensitivity during the promotion. The change is characterized by instantaneous jump upward in quantities ordered and downward in retailer prices at the instant when the promotion period starts and vice versa just when the promotion ends. In fact, the retailer starts lowering his prices sometime before the promotion starts. This causes a greater demand when the promotion period begins, thereby taking advantage of the reduced wholesale price during the promotion. This is accomplished gradually to strike a trade-off between the surplus/ backlog cost and the wholesale price over time. Specifically, any reduction in the wholesale price results first in backlog and then in surplus. An opposite scenario takes place on the side when the promotion period ends.

In the symmetric case when unit backlog and surplus costs are equal, the typical equilibrium as shown in Figure 1 is obtained. As can be seen, due to symmetric costs, the transient solution is symmetric with respect to the midpoint of the promotion phase.

Finally, due to inventory dynamics, the traditional two-part tariff does not coordinate the supply chain as it does in the static case. This is because the manufacturer when setting the promotional wholesale price ignores not only the retailer's profit margin from sales, but also the profit margin from handling inventories. However, in the special case where the manufacturer fixes a wholesale price throughout the season, the retailer's problem becomes identical to the centralized problem and the supply chain is coordinated.
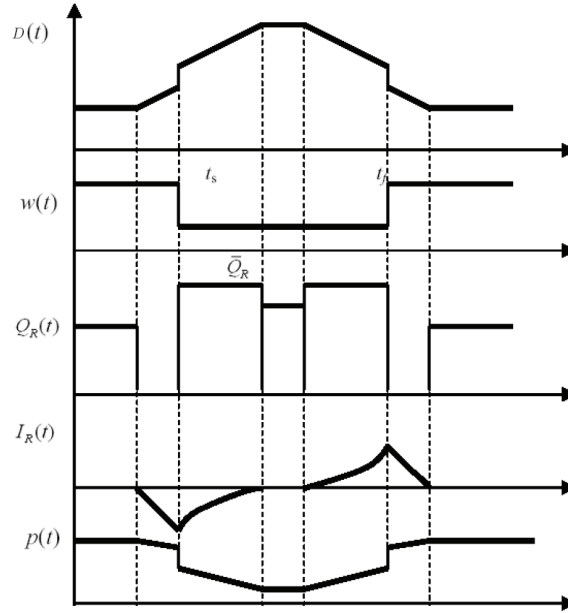

FIGURE 1. Optimal Policies with Promotion

2.2.5. *Kogan and Tapiero (2007b): Production balancing with subcontracting.* This paper studies a supply chain with one manufacturer (follower) and one supplier (leader). The supplier has ample capacity and so the inventory dynamics is not an issue. The manufacturer, on the other hand, has a limited capacity, and his decisions are dependent on the available inventory. The product demand rate at time t is $a(t)\tilde{D}$, where $\tilde{D}$ is a random variable and $a(t)$ is known as the demand shape parameter. The realization of $\tilde{D}$ is observed only at the end of a short selling season, such as in the case of fashion goods, and as a result, the manufacturer can only place an advance order to obtain an initial end-product inventory, which is then used to balance production over time with the limited in-house capacity.

Therefore, this problem is a dynamic version of the newsvendor problem that incorporates production control. The supplier's problem is to decide a constant

wholesale price to maximize her profit from the advance order from the manufacturer. The manufacturer decides on the advance order and the production rate over time in order to minimize his total expected cost of production, inventory/backlog, and advance order. The authors show that it is possible to transform this problem to a deterministic optimal control problem that can be solved to obtain the manufacturer's best response to the supplier's announced wholesale price.

This paper assumes that the unit in-house production cost is greater than the supplier's cost, and shows that if the supplier makes profit (i.e., has a positive margin), then the manufacturer produces more in-house and buys less than in the centralized solution. This is due to the double marginalization observed in the static newsvendor problem. Furthermore, if the manufacturer is myopic, he also orders less than the centralized solution even though he does not produce in-house, since he does not take into account the inventory dynamics.

While the optimal production rate over time would depend on the nature of the demand, it is clear that the optimal production policy will have intervals of zero production, maximum production, and a singular level of production. The authors also solve a numerical example and obtain the equilibrium wholesale price, the manufacturer's advance order quantity, and the production rate over time.

2.2.6. *Kogan and Tapiero (2007c): Outsourcing game.* This paper considers a supply chain that has one producer and multiple suppliers, all with limited production capacities. The suppliers are the leaders and choose their wholesale prices over time to maximize profits. Following this, the producer determines his in-house production rate and supplements this by ordering from a selection of suppliers over time in order to meet a random demand at the end of a planning period T. The producer incurs a penalty for any unmet demand. Unlike in [36], no assumption regarding the cost of in-house production and the supplier's production cost is made here. The producer's objective is to minimize his expected cost.

As in [36], the producer's problem can be transformed into a deterministic optimal control problem. Since the producer's problem is linear in his decisions, the production rate has one of three regimes as in the previous model, and his ordering rate from any chosen supplier will also have similar three regimes.

The authors show that the greater the wholesale price of a supplier, the longer the producer waits before ordering from that supplier. This is because the producer has an advantage over that supplier, up to and until a breakeven point in time for outsourcing to this supplier is reached. As in [36], here also when a supplier sets her wholesale price strictly above her cost over the entire horizon, then the outsourcing order quantity is less than that in the centralized solution.

Again, the supply chain can be coordinated if the suppliers set their wholesale prices equal to their unit costs and get lump-sum transfers from the producer.

2.2.7. *Kogan and Tapiero (2007d): Pricing game.* This paper concerns a Stackelberg differential game between a supplier (the leader) who sets a wholesale price for a product and a retailer (the follower) who responds with a retail product price. The authors incorporate learning on the part of the supplier whose production cost declines as more units are produced. The demand for the product is assumed to be decreasing in the retail price. The authors show that myopic pricing is optimal for the retailer. Under a certain profitability condition, the retail price is greater than in the centralized solution because of double marginalization. However, the gap is larger than that in the static pricing game. This is because the higher retail price

implies less cost reduction learning in the dynamic setting in comparison to that in the centralized solution, whereas in the static framework no learning is involved in the centralized solution and in the Stackelberg solution.

2.2.8. *Kogan and Tapiero (2007e): Co-investment in supply chains.* This paper examines a supply chain with N firms. Let $K(t)$ denote the current level of the supply chain infrastructure capital, $L = (L_1, ..., L_N)$ denote a vector of the labor force, $I = (I_1, ..., I_N)$ denote a vector of investment policy, and $Q_j = f(K, L_j)$ denote an aggregate production function of the jth firm, where $\frac{\partial f}{\partial K} \geq 0$, $\frac{\partial f}{\partial L_j} > 0$ for $L \neq 0$, $\frac{\partial f(K,0)}{\partial L_j} = 0$, and $\frac{\partial^2 f}{\partial L_j^2} < 0$. The process of capital accumulation is modeled as

$$\dot{K}(t) = -\delta K(t) + \sum_{j=1}^{N} I_j(t), \ K(0) = K_0, \ f(K(t), \ L_j(t)) \geq I_j(t) \geq 0, j = 1, ...N.$$

The jth firm's objective is to maximize its discounted total profit, i.e.,

$$\max_{L_j(\cdot), I_j(\cdot)} \left\{ \int_0^\infty e^{-\rho_j t} \left[p_j(t) f\left(K(t), L_j(t)\right) - c_j(t) L_j(t) - C_I\left((1-\theta) I_j(t)\right) \right] dt \right\},$$

$j = 1, ...N$, where p_j is the price, c_j is the unit labor cost, $C_I(\cdot)$ is the investment cost function, and θ is the portion of the cost that is subsidized.

The authors derive the Nash strategy as well as the Stackelberg strategy for the supply chain where firms are centralized and controlled by a "supply chain manager". Their results show that the Stackelberg strategy applied to consecutive subsets of firms will result in an equilibrium identical to that obtained in case of a Nash supply chain. The implication is that it does not matter who the leader is and who the followers are.

2.2.9. *Gutierrez and He (2011): Life-cycle channel coordination.* This paper investigates the intertemporal channel coordination issues in an innovative durable product (IDP) supply chain composed of a manufacturer who produces the IDP and a retailer who serves the final market.

The demand in this paper evolves according to a [5] type diffusion process. Specifically, the demand is affected by both the external and internal market influences (i.e., word-of-mouth) as well as the retail price. The word-of-mouth creates an interdependence between the current and future demand.

The manufacturer leads by announcing her wholesale price strategy. The retailer follows by deciding on the retail price. While the manufacturer is far-sighted, i.e., she maximizes her life-cycle profits, two scenarios for the retailer are studied: (1) a far-sighted strategy of maximizing the life-cycle profit, and (2) a myopic strategy of maximizing the instantaneous profit rate at any time t. They address the following research questions: Does the manufacturer prefer the retailer to be far-sighted or myopic? Does the retailer prefer the manufacturer to assume a far-sighted or myopic retailer? This paper derives open-loop Stackelberg pricing equilibrium for both players.

Far-sighted Retailer. When the retailer is far-sighted, for a given wholesale price strategy $w(\cdot)$, he decides in response a retail price path $p(\cdot)$ by solving the problem

$$\max_{p(\cdot)} \int_0^T \left\{ [p(t) - w(t) - c_R] \dot{X}(t) \right\} dt \tag{16}$$

$$\dot{X}(t) = (M - X(t))(\alpha + \beta X(t))(1 - bp(t)), X(0) = X_0, \qquad (17)$$

where c_R is the per unit selling cost including any opportunity cost, α and β are internal and external influence parameters, b is the price sensitivity parameter, and X_0 is the initial number of adopters. The manufacturer's problem is

$$\max_{w(\cdot)} \int_0^T [w(t) - c_M] \dot{X}(t) dt, \qquad (18)$$

$$\dot{X}(t) = \frac{1}{2} F(X(t)) \{1 - b[w(t) + c_R - \lambda_R(t)]\}, X(0) = X_0, \qquad (19)$$

$$\dot{\lambda}_R(t) = -\frac{1}{4b}[-\alpha + \beta M - 2\beta X(t)]\{1 - b[w(t) + c_R - \lambda_R(t)]\}^2, \lambda_R(T) = 0, (20)$$

where c_M is the per unit production cost and λ_R is the retailer's shadow price. Note that X and λ_R are the manufacturer's two state variables, and their evolution incorporates the retailer's best response.

Myopic Retailer. For each $t \in [0, T]$, the retailer determines his response $p(t)$ in order to maximize his instantaneous profit rate at time t subject to the state equation (17). Accordingly, the retailer's shadow price λ_R is removed from the retailer's best response and then obtain the manufacturer's optimization problem as (18) and (19) with $\lambda_R(t)$ removed from (19).

This paper shows that the manufacturer does not always find it more profitable for the retailer to be far-sighted, and may sometimes benefit from a myopic retailer. On the other hand, both the manufacturer and the retailer are better off if the retailer is far-sighted when the final market is insufficiently penetrated. However, if the market is close to saturation such as at the end of the planning horizon, the manufacturer will shift her preference and become better off with a myopic retailer, while the retailer prefers the manufacturer to set the wholesale prices assuming that the retailer is far-sighted. However, monitoring the retailer sales volume or retail price becomes an implementation necessity when the manufacturer offers a wholesale price contract assuming the retailer is myopic.

2.2.10. *He and Sethi (2008): Pricing and slotting decisions.* This paper extends the work of [25] by considering the impact of shelf space allocation (or another promotional device for that matter) on retail demand. In this paper, it is assumed that retail demand is a concave increasing function of the shelf-space of merchandise displayed on the shelf. This is operationalized by introducing a multiplicative term $\sqrt{S(t)}$ to the right-hand side of (17), where $S(t)$ is the shelf space allocated to the product at time t, and by including a linear cost of shelf space in the retailer's objective function (16). The solution is for the equilibrium wholesale and retail pricing and slotting decisions. However, in connection with the strategic profitability of the retailer, myopic or far-sighted, the obtained results are similar to those in [25].

2.3. **Applications in marketing channel.** In the marketing literature, differential game models have been used to study dynamic advertising strategies in competitive markets. These papers mainly focus on the horizontal interactions such as the advertising competition between brands, and accordingly seek Nash equilibria ([14], [15], [55], [21], [13], [23]).

In the recent years, some Stackelberg differential game models have been formulated to study the vertical interactions in marketing channels. Depending on the dynamics, these models can be categorized into two groups: advertising capital models and sales-advertising response models. Advertising capital models consider

advertising as an investment in the stock of goodwill $G(t)$ as in the model of [45] (NA, thereafter). The advertising capital is affected by the current and past advertising expenditures by a firm on the demand for its products. It changes over time according to

$$\dot{G}(t) = A(t) - \delta G(t), \; G(0) = G_0 \geq 0, \tag{21}$$

where $A(t)$ is the current advertising investment (in dollars) and δ is a constant positive decay rate.

Sales-advertising response models characterize a direct relation between the rate of change in sales and advertising. The Vidale-Wolfe (1957) (VW, hereafter) advertising model is given by

$$\dot{x}(t) = \frac{dx(t)}{dt} = rA(t)(1 - x(t)) - \delta x(t), \; x(0) = x_0,$$

where x is the market share, A is the advertising rate, and r is the effectiveness of advertising. The Sethi (1983) model is a variation of the VW model, and it is given by

$$\dot{x}(t) = rA(t)\sqrt{1 - x(t)} - \delta x(t), \; x(0) = x_0. \tag{22}$$

[49] also gave a stochastic extension of his model. [42] discusses some of the desirable features of the advertising models that have appeared in the literature.

2.3.1. *Jørgensen, Sigue, and Zaccour (2000): Dynamic cooperative advertising.* This paper considers a channel where both the manufacturer and the retailer can make advertising expenditures that have both long-term and short-term impacts on the retail demand. Specifically, the long-term advertising influences the carry-over effect of advertising, while the short term advertising impacts the current retail sales only.

The manufacturer decides her rate of long-term advertising effort $A(t)$ and short-term advertising effort $\hat{A}(t)$. The retailer sets his long-term advertising rate $B(t)$ and short-term advertising rate $\hat{B}(t)$. An extended NA model describes the dynamics of the goodwill as

$$\dot{G}(t) = a_l A(t) + b_l B(t) - \delta G(t), \; G(0) = G_0, \tag{23}$$

where a_l and b_l are positive parameters that capture the effectiveness of the long-term advertising of the manufacturer and the retailer, respectively. At any instant of time, the demand D is given by

$$D\left(\hat{A}, \hat{B}, G\right) = \left(a_s \hat{A} + b_s \hat{B}\right)\sqrt{G},$$

where a_s and b_s are parameters that capture the effectiveness of the manufacturer's and retailer's short-term advertising, respectively.

Suppose the manufacturer and the retailer can enter into a cooperative advertising program in which the manufacturer pays a certain share of the retailer's advertising expenditure. The manufacturer is the Stackelberg game leader in designing the program. She announces her advertising strategies and support rates for the retailer's long-term and short-term advertising efforts.

The manufacturer's problem is

$$\max_{A(\cdot),\hat{A}(\cdot),\theta(\cdot),\hat{\theta}(\cdot)} J_M =$$
$$\int_0^\infty \frac{1}{2} e^{-\rho t} \left\{ 2m_M D - c_M A^2(t) - \hat{c}_M \hat{A}^2(t) - \theta(t) c_R B^2(t) - \hat{\theta}(t)\hat{c}_R(t)\hat{B}^2(t) \right\} dt,$$

and the retailer's problem is

$$\max_{B(\cdot),\hat{B}(\cdot)} J_R =$$
$$\int_0^\infty \frac{1}{2} e^{-\rho t} \left\{ 2m_R D - (1 - \theta(t)) c_R B^2(t) - \left(1 - \hat{\theta}(t)\right) c_R(t) \hat{B}^2(t) \right\} dt,$$

where m_M and m_R are the manufacturer's and the retailer's margins, respectively, $\theta(t)$ and $\hat{\theta}(t)$ are the percentages that the manufacturer pays of the retailer's long-term and short-term advertising costs, respectively.

This paper shows that both the manufacturer and the retailer prefer full support to either long-term or short-term support alone, which in turn is preferred to no support at all.

2.3.2. *Jørgensen, Sigue, and Zaccour (2001) : Impact of leadership on channel efficiency.* This paper examines the effects of strategic interactions in both pricing and advertising in a channel that consists of a manufacturer and a retailer. It studies three settings: the manufacturer and the retailer decide their margins and advertising rates simultaneously; sequentially with the retailer as the leader; and sequentially with the manufacturer as the leader. The manufacturer determines her margin $m_M(t)$ and the rate of advertising $A(t)$. The retailer chooses his margin $m_R(t)$ and the advertising rate $B(t)$. The demand rate $D(t)$ follows a NA type of dynamics, and is given by

$$D(t) = B(t)(a - bp(t))\sqrt{G(t)},$$

where $p(t)$ is the retail price, a and b are positive parameters, and $G(t)$ is the stock of brand goodwill given by (23). It is assumed that the retailer is myopic, meaning that he is only concerned with the short-term effects of his pricing and advertising decisions. The manufacturer is concerned with her brand image. The manufacturer's objective functional to maximize is

$$J_M = \int_0^\infty e^{-\rho t} \left[m_M(t) B(a - bp(t))\sqrt{G(t)} - \frac{1}{2}c_M A^2(t) \right] dt$$

and the retailer's is

$$J_R = \int_0^\infty e^{-\rho t} \left[m_R(t) B(t)(a - bp(t))\sqrt{G(t)} - \frac{1}{2}c_R B^2(t) \right] dt.$$

The paper shows that the manufacturer's leadership and the retailer's leadership in a channel are not symmetric as in pure pricing games. The manufacturer's leadership improves channel efficiency and is desirable in terms of consumer welfare, but the retailer's leadership is not desirable for channel efficiency and for consumer welfare.

2.3.3. *Jørgensen, Taboubi, and Zaccour (2003): Retail promotions with negative brand image effects.* In this paper, the manufacturer advertises in the national media to build up the image for her brand. The retailer promotes the brand locally (by such means as local store displays and advertising in local newspapers) to increase sales revenue, but these local promotional efforts are assumed to be harmful to the brand image. This paper studies two firms in a cooperative program in which the manufacturer supports the retailer's promotional efforts by paying part of the cost incurred by the retailer when promoting the brand. The two firms play a Stackelberg differential game in which the manufacturer is the leader and the retailer is the follower. This paper addresses the question of whether the cooperative promotion program is profitable and whether the retailer's choice on being myopic or far-sighted will impact the implementation of a cooperative program.

Let $A(t)$, $B(t)$, and $G(t)$ denote the manufacturer's advertising rate, the retailer's promotional rate, and the brand image, respectively. The dynamics of $G(t)$ is modeled by the differential equation

$$\dot{G}(t) = aA(t) - bB(t) - \delta G(t), G(0) = G_0 > 0,$$

where a and b are positive parameters measuring the impact of the manufacturer's advertising and retailer's promotion, respectively, on the brand image. The margin on the product is $m(B(t), G(t)) = dB(t) + eG(t)$, where d and e are positive parameters that represent the effects of promotion and brand image on the current sales revenue. With this formulation of the demand and the revenue, the retailer faces a trade-off between the sales volume and the negative impact of the local advertising on the brand image.

The manufacturer and the retailer incur quadratic advertising and promotional costs $C(A(t)) = c_A A^2/2$ and $C(B(t)) = c_B B^2/2$, respectively. The manufacturer's objective is to maximize

$$J_M = \int_0^\infty \left\{ e^{-\rho t} q \left[dB(t) + eG(t) \right] - \frac{c_A A(t)^2}{2} - \frac{c_B \theta(t) B(t)^2}{2} \right\} dt$$

and the retailer's objective is to maximize

$$J_R = \int_0^\infty \left\{ e^{-\rho t} (1-q) \left[dB(t) + eG(t) \right] - \frac{c_B (1 - \theta(t)) B(t)^2}{2} \right\} dt,$$

where q is the manufacturer's fraction of the margin and $\theta(t)$ is the fraction the manufacturer contributes to the retailer's promotion cost.

This paper demonstrates that a cooperative program is implementable if the initial value of the brand image G_0 is sufficiently small, and if the initial brand image is "intermediate" but promotion is not "too damaging" (i.e., b is small) to the brand image.

2.3.4. *Karray and Zaccour (2005): Advertising for national and store brands.* This paper studies a marketing channel that consists of a national manufacturer and a retailer who sells the manufacturer's national product (labeled as 1) and may also introduce a private label (labeled as 2) at a lower price than the manufacturer's brand. This paper finds that while the retailer benefits from introducing the private label, the manufacturer is worse off. Furthermore, the paper examines whether a cooperative advertising program can help the manufacturer to alleviate the negative impact of the private label.

The manufacturer sets on the national advertising $A(t)$. The retailer determines the promotion efforts $B_1(t)$ for the national brand and efforts $B_2(t)$ for the store brand. The retailer's effort has an immediate impact on sales, but it does not affect the dynamics of the goodwill of the national brand. The goodwill $G(t)$ of the national brand evolves according to the NA dynamics given by equation (21). The demand functions D_1 for the national brand and D_2 for the store brand are as follows:

$$D_1(B_1, B_2, G) = aB_1 - b_1 B_2 + e_1 G - dG^2,$$
$$D_2(B_1, B_2, G) = aB_2 - b_2 B_1 - e_2 G,$$

where a, b_1, b_2, e_1, e_2, and d are positive parameters.

[35] study three games:

1) Game N: The retailer carries only the national brand and there is no cooperative advertising program. They show that the retailer promotes the national brand at a positive constant rate and the advertising strategy is decreasing in the goodwill.

2) Game S: The retailer carries both the national and the store brands and there is no cooperative advertising program. They show that the retailer will promote the national brand if the marginal revenue from doing so exceeds the marginal loss on the store brand.

3) Game C: The retailer carries both the national and the store brands and the manufacturer proposes to the retailer a cooperative advertising program. In this game, the retailer will always promote the store brand. The retailer will also promote the national brand, but only under certain conditions specified in the paper.

They conclude that the introduction of store brand always hurts the manufacturer. The cooperative advertising program is profit Pareto-improving for both firms.

2.3.5. *Martin-Herran and Taboubi (2005): Shelf-space allocation.* This paper investigates a supply chain that consists of one retailer and two manufacturers. The retailer has limited shelf-space and must decide on the allocation of the shelf-space to two products. Let $S_1(t)$ denote the fraction of the shelf-space allocated to product 1, then $S_2(t) = 1 - S_1(t)$. At time t, each manufacturer $i \in \{1, 2\}$ decides on its advertising strategy $A_i(t)$ and a shelf-space dependent display allowance $W_i(t) = \omega_i(t) S_i(t)$, where $\omega_i(t)$ is the coefficient of the incentive strategy. The retail demand is a function of the shelf-space and goodwill given by

$$D_i(t) = S_i(t) \left[a_i G_i(t) - \frac{1}{2} b_i S_i(t) \right],$$

where G_i is the goodwill for brand i evolving according to the NA dynamics. This model implies that the shelf-space has a diminishing marginal effect on the sales.

It is assumed that the retailer and two manufacturers play an Stackelberg differential game with the retailer as the follower, while the two manufacturers lead by playing a Nash game, i.e., they announce simultaneously their advertising and incentive strategies ω_1 and ω_2.

This paper shows that manufacturers can affect the retailer's shelf-space allocation decisions through the use of incentive strategies (push) and/or advertising investment (pull). Depending on the system parameters, the manufacturer should choose between incentive strategies and/or advertising investment.

2.3.6. *Jørgensen, Taboubi, and Zaccour (2006): Incentives for retail promotions.* This paper considers a channel composed of a manufacturer selling a particular brand through a retailer. The manufacturer invests in national advertising, which improves the image of her brand, while the retailer makes local promotions for the brand. The manufacturer and the retailer play a Stackelberg differential game with the manufacturer as the leader. The dynamics of the goodwill stock follows the classical NA dynamics given in (21). [34] assume that the manufacturer and retailer apportion a fixed share of the total revenue. They consider two settings: joint maximization and individual maximization. This paper shows that the manufacturer advertises more in the joint maximization setting than in the individual maximization setting. This result does not depend on whether or not the manufacturer supports the retailer's promotion in the individual maximization setting.

2.3.7. *Breton, Jarrar, and Zaccour (2006): Feedback Stackelberg equilibrium with a Lanchester-type model and empirical application.* This paper studies dynamic equilibrium advertising strategies in a duopoly. A Lanchester-type model provides the market share dynamics for the two competitors in the duopoly. Let $A_i(t)$ denote the advertising expenditure of player $i \in \{1, 2\}$ at any instant of time $t \in [0, \infty)$ and let $x(t)$ denote the market share of player 1 (the leader) at time t. The market share of player 2 (the follower) is thus $(1 - x(t))$. The market share evolves as follows

$$\dot{x}(t) = r_1 \sqrt{A_1(t)} (1 - x(t)) - r_2 \sqrt{A_2(t)} x(t), \ x(0) = x_0, \tag{24}$$

where the positive constant r_i, $i \in \{1, 2\}$, denotes the advertising effectiveness of player i and x_0 is player 1's initial market share.

[12] use a discrete-time version of (24). They assume the following sequence of the game. At stage k, the leader observes the state of the system and chooses an optimal advertising level $A_1(k)$; the follower does not play at this stage. At the next stage, the follower observes the state of the system and chooses an optimal advertising level $A_2(k+1)$; the leader does not play at this state. This procedure leads to a feedback Stackelberg equilibrium (FSE).

The authors empirically test the discrete-time model specification by using a dataset of Coke and Pepsi advertising expenditures. They compare the fit of the FSE against the method in [30] for the feedback Nash equilibrium (FNE). Note that in the continuous-time game, the FNE coincides with the FSE, and [48], in a more general setting, also observes this coincidence. However, in the discrete-time version, the solutions need not coincide. The authors find that the FSE fits the actual cola industry advertising expenditures better than the FNE, which suggests that the advertising decision making in this industry follows a sequential rather than simultaneous pattern. However, they find that there is no significant difference in being a leader or a follower.

2.3.8. *He, Prasad, and Sethi (2009): Cooperative advertising.* In a decentralized channel, while the retailer often incurs the full cost of advertising, he only captures a portion of the benefits. This creates an incentive for the retailer to under-advertise. Several papers, such as [11], [10] and [29], have studied the use of cooperative advertising programs to improve channel performance. All of these papers derive the results in a static setup. In contrast, [28] study cooperative advertising in a dynamic supply chain. The manufacturer announces a participation rate, i.e., she will provide a proportion of the retailer's advertising expenditure. In addition, she

also announces the wholesale price. In response, the retailer chooses his optimal advertising and retail pricing policies. This paper formulate this problem as a Stackelberg differential game and provides a feedback solution to the optimal advertising and pricing policies for the manufacturer and the retailer. The market potential dynamics is given by the Sethi model (22). The retailer's objective function is

$$J_R = \int_0^\infty e^{-\rho t} \left[(p(t) - w(t)) D(p(t)) x(t) - (1 - \theta(t)) A^2(t) \right] dt,$$

where $w(t)$ and $\theta(t)$ are the manufacturer's announced wholesale price and participation rate given at time t, and $p(t)$ and $A(t)$ denote the retailer's retail price and advertising effort at time t. The manufacturer's objective function is

$$J_M = \int_0^\infty e^{-\rho t} \left[(w(t) - c) D(p(t)) x(t) - \theta(t) A^2(t) \right] dt,$$

where $p(t)$ and $A(t)$ are the retailer's optimal responses to the manufacturer's announced $w(t)$ and $\theta(t)$.

The authors also solve the model for a vertically integrated channel. Comparing its results to those for the decentralized channel, they find that the decentralized channel has higher optimal prices and lower optimal advertising. Whereas wholesale price by itself cannot correct for these problems, they show that a combination of wholesale price and co-op advertising allows the channel to achieve coordination. Therefore, it is beneficial for the manufacturer to jointly make co-op advertising and pricing decisions. The authors also solve the stochastic extension of the problem in which the dynamics is an Ito equation version of (22) developed in [49].

2.3.9. *Bensoussan, Chen, Chutani, and Sethi (2016): Mixed leadership game in cooperative advertising.* For mixed leadership games as introduced in Section 2.1.1, [6] develop a procedure to obtain a feedback Stackelberg-Nash equilibria where decision makers act as leaders on some decisions and as followers on the other decisions.

The class of games under consideration is given by the state equation

$$dx(t) = f(x(t), u_1(t), v_1(t), u_2(t), v_2(t))dt + \sigma(x(t))dW(t), \qquad x(0) = x, \quad (25)$$

and the profit functions $J_i(u_1, v_1, u_2, v_2) = E \int_0^\infty e^{-\rho t} l_i(x(t), u_1(t), v_1(t), u_2(t), v_2(t))dt$, where $i = 1, 2$, x is the initial state known by both players, W is a one-dimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$, and $(u_i(t), v_i(t)) \in \mathbb{U}_i \times \mathbb{V}_i \subseteq R^{m_i} \times R^{n_i}$, $t \in [0, +\infty)$, are the decision variables of the two players. The underlying information structure is feedback for both players, and u_i and v_i are the lead and follow decision variables, respectively.

It is assumed that $f : \mathbb{R}^n \times \mathbb{R}^{m_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{m_2} \times \mathbb{R}^{n_2} \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n \times 1}$ satisfy

i) f is Lipschitz continuous in (x, u_1, u_2, v_1, v_2) and σ is Lipschitz continuous in x.

ii) f is continuously differentiable and σ is twice continuously differentiable.

The admissible strategy spaces are defined as follows:

$$\mathcal{U}_i = \{u_i | u_i : \mathbb{R}^n \to \mathbb{U}_i \text{ and } u_i(x) \text{ is Lipschitz continuous in } x\},$$
$$\mathcal{V}_i = \{v_i | v_i : \mathbb{R}^n \times \mathbb{U}_1 \times \mathbb{U}_2 \to \mathbb{V}_i \text{ and } v_i(x, \mu_1, \mu_2)$$
$$\text{is Lipschitz continuous in } (x, \mu_1, \mu_2)\},$$

where $\mathbb{U}_i$ and $\mathbb{V}_i$ are given subsets in $\mathbb{R}^{m_i}$ and $\mathbb{R}^{n_i}$, $i = 1, 2$, respectively. Conditions $i)$ and $ii)$ guarantee that the state equation (25) yields a unique state trajectory for admissible strategies (u_1, u_2, v_1, v_2). For two pairs of strategies $(u_i, v_i) \in \mathcal{U}_i \times \mathcal{V}_i$, $x(\cdot; u_1, v_1, u_2, v_2)$ denotes the solution of the state equation

$$\begin{cases} dx(s) = \\ f\Big(x(s), u_1(x(s)), v_1\big(x(s), u_1(x(s)), u_2(x(s))\big), u_2(x(s)), v_2\big(x(s), u_1(x(s)), u_2(x(s))\big)\Big) ds \\ \quad + \sigma(x(s)) dW(s), \\ x(0) = x. \end{cases}$$

Let $J_1^{0,x}\Big(u_1(\cdot), v_1\big(\cdot, u_1(\cdot), u_2(\cdot)\big), u_2(\cdot), v_2\big(\cdot, u_1(\cdot), u_2(\cdot)\big)\Big)$ and $J_2^{0,x}\Big(u_1(\cdot), v_1\big(\cdot, u_1(\cdot), u_2(\cdot)\big), u_2(\cdot), v_2\big(\cdot, u_1(\cdot), u_2(\cdot)\big)\Big)$ represent the corresponding profits of the two players, respectively, i.e.,

$$J_1^{0,x}\Big(u_1(\cdot), v_1\big(\cdot, u_1(\cdot), u_2(\cdot)\big), u_2(\cdot), v_2\big(\cdot, u_1(\cdot), u_2(\cdot)\big)\Big)$$
$$= E \int_0^\infty e^{-\rho s} l_1\big(x(s; u_1, v_1, u_2, v_2), u_1, v_1, u_2, v_2\big) ds,$$
$$J_2^{0,x}\Big(u_1(\cdot), v_1\big(\cdot, u_1(\cdot), u_2(\cdot)\big), u_2(\cdot), v_2\big(\cdot, u_1(\cdot), u_2(\cdot)\big)\Big)$$
$$= E \int_0^\infty e^{-\rho s} l_2\big(x(s; u_1, v_1, u_2, v_2), u_1, v_1, u_2, v_2\big) ds,$$

where u_i, v_i are evaluated at $x(s; u_1, v_1, u_2, v_2)$ with the forms $u_i\big(x(s; u_1, v_1, u_2, v_2)\big)$, $v_i\Big(x(s; u_1, v_1, u_2, v_2), u_1\big(x(s; u_1, v_1, u_2, v_2)\big), u_2\big(x(s; u_1, v_1, u_2, v_2)\big)\Big)$, $i = 1, 2$, respectively.

Definition 2.4. Two pairs of strategies $(u_i^*, v_i^*) \in \mathcal{U}_i \times \mathcal{V}_i$, $i = 1, 2$, are called a feedback Stackelberg-Nash equilibrium if the following holds:

$$J_1^{0,x}\Big(u_1^*(\cdot), v_1^*\big(\cdot, u_1^*(\cdot), u_2^*(\cdot)\big), u_2^*(\cdot), v_2^*\big(\cdot, u_1^*(\cdot), u_2^*(\cdot)\big)\Big)$$
$$\geq J_1^{0,x}\Big(u_1(\cdot), v_1\big(\cdot, u_1(\cdot), u_2^*(\cdot)\big), u_2^*(\cdot), v_2^*\big(\cdot, u_1(\cdot), u_2^*(\cdot)\big)\Big), \ \forall (u_1, v_1) \in \mathcal{U}_1 \times \mathcal{V}_1, \ \forall x \in \mathbb{R}^n,$$
$$J_2^{0,x}\Big(u_1^*(\cdot), v_1^*\big(\cdot, u_1^*(\cdot), u_2^*(\cdot)\big), u_2^*(\cdot), v_2^*\big(\cdot, u_1^*(\cdot), u_2^*(\cdot)\big)\Big)$$
$$\geq J_2^{0,x}\Big(u_1^*(\cdot), v_1^*\big(\cdot, u_1^*(\cdot), u_2(\cdot)\big), u_2(\cdot), v_2\big(\cdot, u_1^*(\cdot), u_2(\cdot)\big)\Big), \ \forall (u_2, v_2) \in \mathcal{U}_2 \times \mathcal{V}_2, \ \forall x \in \mathbb{R}^n.$$
$$(26)$$

In this mixed leadership feedback game, both players are leaders in u decisions, and followers in v decisions. In other words, at each instant of time, first player 1 (P1) decides u_1 and player 2 (P2) decides u_2, simultaneously. Then P1 and P2 follow with their decisions v_1 and v_2 at each instant of time, respectively and simultaneously. At the level of u decisions, P1 and P2 play a feedback Nash game. At the level of v decisions, P1 and P2 also play a feedback Nash game, with the additional information of the instantaneous actions from the level of u decisions. From a vertical view of the two Nash games, they are played with hierarchy and therefore constitute a feedback Stackelberg game. In addtion, since each player is both a leader in u decision and a follower in v decision, the roles of players in the game are mixed.

This paper considers a manufacturer-retailer supply chain, in a mature product category where sales, expressed as a fraction of the potential market, are positively

influenced through advertising spending by the supply chain partners. The authors use the mixed leadership game theoretic framework to study advertising cooperation between a manufacturer and a retailer. The cost of advertising is expressed as the square of the advertising effort. They model the sales-advertising dynamics as the following extension of the Sethi model to incorporate multiple advertising decisions:

$$
\frac{dx(t)}{dt} = \left(av_r(t) + bv_m(t) + k\sqrt{v_r(t)v_m(t)}\right)\sqrt{1 - x(t)} - \delta x(t) + \sigma(x(t))dz(t),
$$
$$
x(0) = x_0 \in [0, 1].
$$

The supply chain carries out its decisions in a Stackelberg framework. It is assumed that M and R announce their respective participation rates u_m and u_r first. Thus, the participation rates are the lead decisions. Given these, M and R decide on the national and local advertising rates, v_m and v_r, respectively. That is, the advertising rates are the follow-up decisions. Thus, a mixed leadership differential game is formulated. Accordingly, M and R play Nash to solve for v_m and v_r given u_m and u_r. This produces optimal responses, which are then used by M and R in solving for u_m and u_r, respectively, once again in the Nash differential game framework. The Stackelberg step is the use of the response function in the formulation of the Nash game for the lead decisions. The objective functions for M and R are, respectively, $J_M = E\left[\int_0^\infty e^{-\rho t}(Mx - u_m v_r^2 - (1 - u_r)v_m^2)dt\right]$ and $J_R = E\left[\int_0^\infty e^{-\rho t}(Rx - (1 - u_m)v_r^2 - u_r v_m^2)dt\right]$.

First, given the lead decisions u_m and u_r, the static Nash game is solved for the follow-up advertising rates decisions v_m and v_r. Then, by incorporating the responses $\bar{v}_m$ and $\bar{v}_r$, the static Nash game is solved for the optimal lead decisions u_m and u_r for the participation rates. A set of equations are obtained to characterize the optimal decisions. Using numerical studies, the impact of margins, advertising effectiveness, decay factor, as well as discount rate are also examined.

2.4. **The formulations of Stackelberg games in discrete time.** In this subsection, we consider a game with its state $x(t) \in \mathbb{R}^n$ evolving as

$$
\begin{cases}
x(t+1) - x(t) = f(x(t), u(t), v(t), \Lambda(t), t), \ t = 1, 2, ..., N, \\
x(1) = y,
\end{cases}
$$

where $\Lambda(t)$ is a random variable, and $u(t)$ and $v(t)$ are the decisions of the leader and the follower at time t, respectively. The profit functionals for the leader (L) and the follower (F) are as follows:

$$
J_L = E\left[\sum_{t=1}^N \pi_L(x(t), u(t), v(t), \Lambda(t), t) + S_L(x(N+1))\right],
$$

$$
J_F = E\left[\sum_{t=1}^N \pi_F(x(t), u(t), v(t), \Lambda(t), t) + S_F(x(N+1))\right],
$$

where S_L and S_F are known as salvage value functions.

In an Feedback Stackelberg Equilibrium (FSE), the leader determines her strategy in the feedback form as $u(x, t)$, and the follower's strategy is based on the state as well as the leader's decisions, and is therefore of the form $v(x, t, u(x, t))$, $t = 1, 2, ..., N$. Notationally, we specify the leader's strategy as $u = (u(x, 1), u(x, 2), ..., u(x, N))$ and the follower's strategy as $v = (v(x, 1, u(x, 1)), v(x, 2, u(x, 2)), ..., v(x, N,$

$u(x, N)))$. Let $\mathbb{U}$ and $\mathbb{V}$ denote the spaces of such strategies of the leader and the follower, respectively. Then given $u \in \mathbb{U}$ and $v \in \mathbb{V}$, we denote by $x^{t,y}(i; u, v)$, $i = t, t+1, ..., N+1$, the solution of the state equation

$$\begin{cases} x(i+1) = x(i) + f(x(i), u(i), v(i), \Lambda(i), i), \ i = t, t+1, ..., N, \\ x(t) = y, \end{cases}$$

and let $J_L^{t,y}(u, v) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ and $J_F^{t,y}(u, v) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ represent the corresponding profit functionals of the two players, i.e.,

$$J_L^{t,y}(u, v) = J_L^{t,y}(u(\cdot), v(\cdot, u(\cdot)))$$

$$= E[\sum_{i=t}^{N} \pi_L(x^{t,y}(i; u, v), u(x^{t,y}(i; u, v), i), v(x^{t,y}(i; u, v), i, u(x^{t,y}(i; u, v), i)), \Lambda(i))$$

$$+ S_L(x^{t,y}(N+1; u, v))] ,$$

$$J_F^{t,y}(u, v) = J_F^{t,y}(u(\cdot), v(\cdot, u(\cdot)))$$

$$= E[\sum_{i=t}^{N} \pi_F(x^{t,y}(i; u, v), u(x^{t,y}(i; u, v), i), v(x^{t,y}(i; u, v), i, u(x^{t,y}(i; u, v), i)), \Lambda(i))$$

$$+ S_F(x^{t,y}(N+1; u, v))],$$

$t = 1, 2, ..., N$, where we should note that the notation in the middle expression is introduced to facilitate the definition of the FSE to be introduced next.

Definition 2.5. A pair of strategies $u^* \in \mathbb{U}$ and $v^* \in \mathbb{V}$ is called a feedback Stackelberg equilibrium if the following holds:

$$J_L^{t,y}(u^*(\cdot), v^*(\cdot, u^*(\cdot))) \geq J_L^{t,y}(u(\cdot), v^*(\cdot, u(\cdot))), \forall \, u \in \mathbb{U}, \forall \, y \in \mathbb{R}^n, \forall \, t \in [1, N]$$

$$J_F^{t,y}(u^*(\cdot), v^*(\cdot, u^*(\cdot))) \geq J_F^{t,y}(u^*(\cdot), v(\cdot, u^*(\cdot))), \forall \, v \in \mathbb{V}, \forall \, y \in \mathbb{R}^n, \forall \, t \in [1, N].$$

Note that if $v(x, t, u(x, t))$ is independent of $u(x, t)$ for each t, then this definition reduces to the feedback Nash equilibrium. For further details on FSEs, see [4] and [7].

2.5. **Applications in operations management.**

2.5.1. *Anand, Anupindi, and Bassok (2008): Strategic inventory.* This paper considers a dynamic model of an upstream firm (manufacturer) and a downstream firm (retailer) who can carry inventories, and shows that the retailer's optimal strategy in equilibrium is to carry inventories for strategic considerations. These inventories have a significant impact on the equilibrium solution as well as the manufacturer, retailer, and channel profits. This paper also shows that two-part tariff contracts fail to achieve the first-best solution.

The authors formulate a dynamic (two-period) model of vertical contracting under full information and no uncertainty. The retailer may carry inventories from the first period to the second period at a linear holding cost h per unit. The baseline model assumes identical linear demand curves $p(q) = a - bq$ for both periods and linear wholesale prices. The unit production cost for the manufacturer is normalized to zero.

The decentralized channel in this paper is formulated as a two-period Stackelberg game between the manufacturer as the leader and the retailer as the follower. In this game, the manufacturer determines her first-period and second-period wholesale prices $w_1 \geq 0$ and $w_2 \geq 0$, and the retailer determines his first and second-period

order quantities $Q_1 \geq 0$ and $Q_2 \geq 0$ and retail prices $p_1 \geq 0$ and $p_2 \geq 0$, respectively. The objective functions for the manufacturer and the retailer, respectively, are as follows:

$$
\begin{aligned}
J_M &= w_1 Q_1 + w_2 Q_2, \\
J_R &= p_1 \min\{Q_1, q_1\} - w_1 Q_1 + p_2 \min\{Q_2, q_2\} - w_2 Q_2 - hI,
\end{aligned}
$$

subject to the stochastic state dynamics $I = Q_1 - q_1$.

This model is therefore a special case of the general model presented in Section 2.4. In this model, $n = 1$, $N = 2$, $S_M = S_R = 0$, $u(t) = w_t$, $v(t) = (Q_t, p_t)$, $t = 1, 2$, and there is no uncertainty so Λ does not appear in the expression for J_M and J_R. Note also that we have switched the notation so that the time period now appears as a subscript and not in a parenthesis as earlier in the section. The authors present two types of contracts depending on when the manufacturer announces the wholesale prices: dynamic contract (feedback Stackelberg equilibrium) and commitment contract (open-loop Stackelberg equilibrium).

By comparing the two contracts, the authors show that when the retailer carries strategic inventories, he forces the manufacturer to price only for the retailer's residual demand, leading to a lower second-period wholesale price w_2. Thus, the retailer curtails the manufacturer's monopoly power in the second period, by inducing (Cournot-like) supply-side competition between the manufacturer and the retailer's inventories. They also demonstrate that strategic inventories reduce the double marginalization effect.

2.5.2. *Gray, Tomlin, and Roth (2009): The effect of learning-by-doing.* This paper investigates the impact of learning-by-doing on an original equipment manufacturer's (OEM) outsourcing decisions in the presence of a powerful contract manufacturer (CM). The authors consider a two-period, game theoretic model in which both parties can reduce their production costs through learning-by-doing. They find that partial outsourcing, wherein the OEM simultaneously outsources and produces in-house, can be an optimal strategy. Also, they find that the OEM's outsourcing strategy may be dynamic – i.e., change from period to period. Furthermore, they show both that the OEM may engage in production for leverage (i.e., produce internally when at a cost disadvantage) and that the CM may engage in low balling.

The authors formulate a dynamic (two-period) model of vertical contracting under full information and no uncertainty. The inventory cannot be carried from one period to the next. The paper assumes identical linear demand curves $p(q) = (a - q)/b$ for both periods and linear wholesale prices. Discount factor is δ for the second period.

In the basic model, it is assumed that the OEM's (CM's) production cost decreases linearly in its production quantity, i.e., $c_{i2} = \max\{c_{i1} - \gamma_i q_{i1}, \underline{c}_i\}$, where $i = F$ (for OEM) or L (for CM), and $\underline{c}_i$ is the lower bound of production cost.

The decentralized channel in this paper is formulated as a two-period Stackelberg game between the CM as the leader and the OEM as the follower. In this game, the CM determines the first-period and second-period wholesale prices $w_1 \geq 0$ and $w_2 \geq 0$, and the OEM determines his first and second-period in-house production quantities $q_{F1} \geq 0$ and $q_{F2} \geq 0$ and outsourcing quantities $q_{L1} \geq 0$ and $q_{L2} \geq 0$, respectively. The objective functions for the CM and the OEM, respectively, are as follows:

$$
J_L = (w_1 - c_{L1})q_{L1} + \delta[(w_2 - c_{L2})q_{L2}],
$$

$$
\begin{aligned}
J_F \quad = \quad & (q_{L1} + q_{F1})(a - q_{L1} - q_{F1})/b - c_{F1}q_{F1} - w_1 q_{L1} \\
& + \delta[(q_{L2} + q_{F2})(a - q_{L2} - q_{F2})/b - c_{F2}q_{F2} - w_2 q_{L2}],
\end{aligned}
$$

subject to the stochastic state dynamics $c_{i2} = \max\{c_{i1} - \gamma_i q_{i1}, \underline{c}_i\}$.

This model is therefore a special case of the general model presented in Section 2.4. In this model, $n = 1$, $N = 2$, $S_L = S_F = 0$, $u(t) = w_t$, $v(t) = (q_{Lt}, q_{Ft})$, $t = 1, 2$, and there is no uncertainty so Λ does not appear in the expression for J_L and J_F. Note also that we have switched the notation so that the time period now appears as a subscript and not in a parenthesis as earlier in the section. The authors analyze the dynamic contract (feedback Stackelberg equilibrium) in this paper. Different dynamic strategies and partial strategies are analyzed and discussed. The results are also generalized to incorporate non-linear learning, in-period learning benefit, or private information about costs and learning rates. Other games structures (dominant OEM feedback Stackelberg game, Nash bargaining) are analyzed in this paper as well.

2.5.3. *Li, Sethi, and He (2015): Pricing and coordination with stochatic learning.* This paper considers a decentralized two-period supply chain in which a manufacturer produces a product with benefits of cost learning, and sells it through a retailer facing a price-dependent demand. The manufacturer's second-period production cost declines linearly in the first-period production, but with a random learning rate. The manufacturer may or may not have the inventory carryover option. If the manufacturer has the inventory carryover option, then a linear holding cost h per unit is assumed. The baseline model assumes asymmetric linear demand curves $D_i(p_i) = a_i - bp_i$ for period $i = 1, 2$ and linear wholesale prices. The manufacturer's second-period unit production cost is given by the random variable $C_2 = c_1 - \Lambda Q_1$, where c_1 is her first-period unit production cost. Λ is realized and observed at the beginning of the second period, and therefore C_2 is observed. Let c_2 and λ denote the realizations of C_2 and Λ, respectively.

Without the manufacturer's inventory carryover option, the decentralized channel in this paper is formulated as a two-period Stackelberg game between the manufacturer as the leader and the retailer as the follower. In this game, the manufacturer determines her first-period and second-period wholesale prices $w_1 \geq 0$ and $w_2 \geq 0$ and production quantities $Q_1 \geq 0$ and $Q_2 \geq 0$, and the retailer determines his first and second-period order quantities $0 \leq q_1 \leq Q_1$ and $0 \leq q_2 \leq Q_2$ and retail prices $p_1 \geq 0$ and $p_2 \geq 0$, respectively. The objective functions for the manufacturer and the retailer, respectively, are as follows:

$$
\begin{aligned}
J_M \quad &= \quad E\left[w_1 q_1 - c_1 Q_1 + w_2 q_2 - C_2 Q_2\right], \\
J_R \quad &= \quad E\left[p_1 \min\{D_1(p_1), q_1\} - w_1 q_1 + p_2 \min\{D_2(p_2), q_2\} - w_2 q_2\right],
\end{aligned}
$$

subject to the stochastic state dynamics $C_2(c_1) = c_1 - \Lambda Q_1(c_1)$.

Their model is therefore a special case of the general model presented above. In this case, $n = 1$, $N = 2$, $S_M = S_R = 0$, $u(t) = (w_t, Q_t)$, $v(t) = (q_t, p_t)$, $t = 1, 2$, $\Lambda(1) = \Lambda$, and Λ does not appear directly in the expression for J_M and J_R. Note also that we have switched the notation so that the time period now appears as a subscript and not in a parenthesis as earlier in the section.

In this paper, the authors also consider the problems in which the manufacturer has the inventory carryover option, but not the backlog option. With the inventory option, the manufacturer has stronger incentive to produce more than the order quantity in the first period to take advantage of the learning effect. Then, in

the first period, when the manufacturer's production quantity is greater than the
retailer's order quantity, i.e., $Q_1 \geq q_1$, the manufacturer carries over the leftover
units $Q_1 - q_1$ to the second period with the holding cost of $h(\geq 0)$ per unit. Let
$I_2 = \max\{0, I_1 + Q_1 - q_1\}$ denote the inventory that is carried over to the second
period and I_1 is the initial inventory at the beginning of the first period. To avoid
trivial cases, they assume that $I_1 \leq (a_1 - bc_1)/2 + (a_2 - bc_1)/2$, where $(a_i - bc_1)/2$ is
the centralized firm's period i's optimal production quantity without the presence
of learning.

With the manufacturer's inventory carryover option, the decentralized channel is
formulated as a feedback Stackelberg game between the manufacturer as the leader
and the retailer as the follower. In this game, the manufacturer determines her
first-period and second-period wholesale prices $w_1 \geq 0$ and $w_2 \geq 0$ and production
quantities $Q_1 \geq 0$ and $Q_2 \geq 0$, and the retailer determines his first and second-
period order quantities $0 \leq q_1 \leq Q_1 + I_1$ and $0 \leq q_2 \leq Q_2 + I_2$ and retail prices
$p_1 \geq 0$ and $p_2 \geq 0$, respectively. The objective functions for the manufacturer and
the retailer, respectively, are as follows:

$$J_M = E\left[w_1 q_1 - c_1 Q_1 + w_2 q_2 - C_2 Q_2 - h I_2\right],$$
$$J_R = E\left[p_1 \min\{D_1(p_1), q_1\} - w_1 q_1 + p_2 \min\{D_2(p_2), q_2\} - w_2 q_2\right],$$

subject to the stochastic state dynamics $C_2 = c_1 - \Lambda Q_1$ and $I_2 = \max\{0, I_1 + Q_1 - D_1(p_1)\}$..

This model is therefore a special case of the general model presented above. In
this case, $n = 2$, $N = 2$, $S_M = S_R = 0$, $u(t) = (w_t, Q_t)$, $v(t) = (q_t, p_t)$, $t = 1, 2$,
$\Lambda(1) = \Lambda$, and Λ does not appear directly in the expressions for J_M and J_R. Note
also that we have switched the notation so that the time period now appears as a
subscript and not in a parenthesis as earlier in the section.

To obtaining a feedback Stackelberg equilibrium, the authors assume that the
leader determines her wholesale price strategy in the feedback form $w_1(c_1, I_1)$,
$Q_1(c_1, I_1)$, $w_2(c_2, I_2)$, and $Q_2(c_2, I_2)$. The follower's strategy is based on the state
as well as the leader's decisions, and is therefore of the form $q_1(c_1, I_1, w_1, Q_1)$,
$p_1(c_1, I_1, w_1, Q_1)$, $q_2(c_2, I_2, w_2, Q_2)$, and $p_2(c_2, I_2, w_2, Q_2)$. Let M and R denote the
spaces of such admissible strategies of the manufacturer and the retailer, respec-
tively. Then given $(w_1, Q_1, w_2, Q_2) \in M$ and $(p_1, q_1, p_2, q_2) \in R$, the manufacturer's
and the retailer's profit functions can be defined similar to the ones in Section 2.4.

This paper shows that as the mean learning rate or the learning rate variabil-
ity increases, the traditional double marginalization problem becomes more severe,
leading to greater efficiency loss in the channel. The authors obtain revenue shar-
ing contracts that can coordinate the dynamic supply chain. In particular, when
the manufacturer may hold inventory, the authors identify two major drivers for
inventory carryover: market growth and learning rate variability. Finally, the au-
thors demonstrate the robustness of the results by examining a model in which cost
learning takes place continuously.

Before concluding this review, we note that [38], [54], [43] and [41] formulate a
(decentralized) supply chain without providing a solution for the underlying two-
period Stackelberg game, solve the corresponding centralized problem, and develop
a coordinating contract for the supply chain. These papers assume that the man-
ufacturer announces her first- and second-period wholesale prices at the beginning
of the game in the open-loop fashion, and the retailer can determine his feedback
ordering quantities by solving a dynamic programming problem given the wholesale

prices. Then, with the retailer's best response in hand, the manufacturer can find her optimal wholesale prices by solving a simple static optimization problem. We should note that the decentralized games they formulate (without solving) are of a hybrid Stackelberg variety requiring that the manufacturer commit to the announced wholesale prices.

3. **Conclusions.** Dynamic Stackelberg game formulations discussed in this paper include deterministic open-loop Stackelberg games in continuous time, stochastic feedback Stackelberg games in continuous and discrete times, and the mixed leadership differential games in continuous time. Definitions are provided for equilibria under different information structures. For each game type, we review its applications to operations management and marketing areas. We briefly explain the key features of each application, how it fits into the Stackelberg game framework, and discuss its main results.

Open-loop Stackelberg equilibria are obtained by using the Maximum Principle and numerical analysis is often used to gain insights into the impact of the key parameters on the issues under investigation. A major drawback of the open-loop Stackelberg equilibria is that in general they are not time consistent. Furthermore, in stochastic settings, they also do not respond to the realizations of the underlying random processes over time. On the other hand, the feedback Stackelberg equilibria are obtained using backward recursion, and are time consistent.

Future research directions may include making the deterministic application more realistic by introducing inherent uncertainties that may be present, designing coordinating contracts in dynamic settings to improve the performance of the entire channel, and collecting data and performing empirical research to estimate the model parameters, as well as implement the resulting feedback Stackelberg solutions.

REFERENCES

[1] K. Anand, R. Anupindi and Y. Bassok, Strategic inventories in vertical contracts, *Management Science*, **54** (2008), 1792–1804.

[2] A. Bagchi, *Stackelberg Differential Games in Economic Models*, Springer-Verlag, New York, 1994.

[3] T. Başar and A. Haurie, Feedback equilibria in differential games with structural and modal uncertainties, in *Advances in Large Scale Systems* (eds. Jose B. Cruz Jr.), JAE Press Inc., Connecticut, **1** (1984), 163–201.

[4] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed., SIAM, Philadelphia, PA, 1999.

[5] F. M. Bass, A new product growth for model consumer durables, *Chapter: Mathematical Models in Marketing*, **132** (1976), 351–353.

[6] A. Bensoussan, S. Chen, A. Chutani and S. P. Sethi, Feedback Stackelberg-Nash equilibria in mixed leadership games with an application to cooperative advertising, working paper, The University of Texas at Dallas, 2016.

[7] A. Bensoussan, S. Chen and S. P. Sethi, Linear quadratic differential games with mixed leadership: The open-loop solution, *Numerical Algebra, Control and Optimization*, **3** (2013), 95–108.

[8] A. Bensoussan, S. Chen and S. P. Sethi, Feedback Stackelberg solutions of infinite-horizon stochastic differential games, in *Models and Methods in Economics and Management Science* (eds. F. El Ouardighi and K. Kogan), International Series in Operations Research and Management Science 198, Springer, (2014), 3–15.

[9] A. Bensoussan, S. Chen and S. P. Sethi, The maximum principle for global solutions of stochastic Stackelberg differential games, *SIAM Journal on Control and Optimization*, **53** (2015), 1956–1981.

[10] M. Bergen and G. John, Understanding cooperative advertising participation rates in conventional channels, *Journal of Marketing Research*, **34** (1997), 357–369.

[11] P. D. Berger, Vertical cooperative advertising ventures, *Journal of Marketing Research*, **9** (1972), 309–312.

[12] M. Breton, R. Jarrar and G. Zaccour, A note on feedback sequential equilibria in a Lanchester model with empirical application, *Management Science*, **52** (2006), 804–811.

[13] P. Chintagunta and D. Jain, A dynamic model of channel member strategies for marketing expenditures, *Marketing Science*, **11** (1992), 168–188.

[14] K. R. Deal, Optimizing advertising expenditures in a dynamic duopoly, *Operations Research*, **27** (1979), 682–692.

[15] K. R. Deal, S. P. Sethi and G. L. Thompson, A bilinear-quadratic differential game in advertising, in *Control Theory in Mathematical Economics* (eds. P.-T. Liu and J. G. Sutinen), Marcel Dekker, Inc., New York, NY., **47** (1979), 91–109.

[16] N. A. Derzko, S. P. Sethi and G. L. Thompson, Necessary and sufficient conditions for optimal control of quasilinear partial differential systems, *Journal of Optimal Theory and Applications*, **43** (1984), 89–101.

[17] V. S. Desai, Marketing-production decisions under independent and integrated channel structure, *Annals of Operations Research*, **34** (1992), 275–306.

[18] V. S. Desai, Interactions between members of a marketing-production channel under seasonal demand, *European Journal of Operational Research*, **90** (1996), 115–141.

[19] E. Dockner, S. Jørgensen, N. V. Long and G. Sorger, *Differential Games in Economics and Management Science*, Cambridge University Press, 2000.

[20] J. Eliashberg and R. Steinberg, Marketing-production decisions in an industrial channel of distribution, *Management Science*, **33** (1987), 981–1000.

[21] G. M. Erickson, Empirical analysis of closed-loop duopoly advertising strategies, *Management Science*, **38** (1992), 1732–1749.

[22] W. H. Fleming and R. W. Rishel, Deterministic and stochastic optimal control, in *Stochastic Modelling and Applied Probability* (eds. P. W. Glynn and Y. Le Jan), V.1, Springer-Verlag, New York, NY, 1975.

[23] G. E. Fruchter and S. Kalish, Closed-loop advertising strategies in a duopoly, *Management Science*, **43** (1997), 54–63.

[24] J. V. Gray, B. Tomlin and A. V. Roth, Outsourcing to a powerful contract manufacturer: The effect of learning-by-doing, *Production and Operations Management*, **18** (2009), 487–505.

[25] G. J. Gutierrez and X. He, Life-cycle channel coordination issues in launching an innovative durable product, *Production and Operations Management*, **20** (2011), 268–279.

[26] X. He, A. Prasad, S. P. Sethi and G. J. Gutierrez, A survey of differential game models in supply and marketing channels, *Journal of Systems Science and Systems Engineering*, **16** (2007), 385–413.

[27] X. He and S. P. Sethi, Dynamic slotting and pricing decisions in a durable product supply chain, *Journal of Optimization Theory and Applications*, **137** (2008), 363–379.

[28] X. He, A. Prasad and S. P. Sethi, Cooperative advertising and pricing in a dynamic stochastic supply chain: Feedback Stackelberg strategies, *Production and Operations Management*, **18** (2009), 78–94.

[29] Z. Huang, S. X. Li and V. Mahajan, An analysis of manufacturer-retailer supply chain coordination in cooperative advertising, *Decision Sciences*, **33** (2002), 469–494.

[30] R. Jarrar, G. Martin-Herran and G. Zaccour, Markov perfect equilibrium advertising strategies of Lanchester duopoly model: A technical note, *Management Science*, **50** (2004), 995–1000.

[31] S. Jørgensen, S. P. Sigue and G. Zaccour, Dynamic cooperative advertising in a channel, *Journal of Retailing*, **76** (2000), 71–92.

[32] S. Jørgensen, S. P. Sigue and G. Zaccour, Stackelberg leadership in a marketing channel, *International Game Theory Review*, **3** (2001), 13–26.

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

[34] S. Jørgensen, S. Taboubi and G. Zaccour, Incentives for retailer promotion in a marketing channel, *Annals of the International Society of dynamic Games*, **8** (2006), 365–378.

[35] S. Karray and G. Zaccour, A differential game of advertising for national and store brands, in *Dynamic Games: Theory and Applicatoins* (eds. A. Haurie and G. Zaccour), Springer, New York, NY, **10** (2005), 213–229.

[36] K. Kogan and C. S. Tapiero, *Supply Chain Games: Operations Management and Risk Valuation*, Springer, New York, NY, 2007.

[37] K. Kogan and C. S. Tapiero, *Co-investment in Supply Chain Infrastructure*, working paper, Bar Ilan University, Israel, 2007.

[38] H. L. Lee, V. Padmanabhan, T. A. Taylor and S. Whang, Price protection in the personal computer industry, *Management Science*, **46** (2000), 467–482.

[39] G. Leitmann, On generalized Stackelberg strategies, *J. Optimization Theory and Applications*, **26** (1978), 637–643.

[40] T. Li, S. P. Sethi and X. He, Dynamic pricing, production, and channel coordination with stochastic learning, *Production and Operations Management*, **24** (2015), 857–882.

[41] C. T. Linh and Y. Hong, Channel coordination through a revenue sharing contract in a two-period newsboy problem, *European Journal of Operational Research*, **198** (2009), 822–829.

[42] J. D. C. Little, Aggregate advertising models: The state of the art, *Operations Research*, **27** (1979), 629–667.

[43] X. Lu, J. S. Song and A. Regan, Rebate, returns and price protection policies in channel coordination, *IIE Transactions*, **39** (2007), 111–124.

[44] G. Martin-Herran and S. Taboubi, Incentive strategies for shelf-space allocation in duopolies, in *Dynamic Games Theory and Applications* (eds. A. Haurie and G. Zaccour), Springer, New York, NY, **10** (2005), 231–253.

[45] M. Nerlove and K. J. Arrow, Optimal advertising policy under dynamic conditions, *Mathematical Models in Marketing*, **132** (1976), 167–168.

[46] G. P. Papavassilopoulos and J. B. Cruz, Jr, Nonclassical control problems and Stackelberg games, *IEEE Transactions on Automatic Control*, **24** (1979), 155–166.

[47] D. Pekelman, Simultaneous price production in channels, *Marketing Science*, **7** (1974), 335–355.

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

[49] S. P. Sethi, Deterministic and stochastic optimization of a dynamic advertising model, *Optimal Control Applications and Methods*, **4** (1983), 179–184.

[50] S. P. Sethi and G. L. Thompson, *Optimal Control Theory: Applications to Management Science and Economics*, 2nd edition, Springer, New York, 2000.

[51] M. Simaan and J. B. Cruz, Jr, On the Stackelberg strategy in nonzero-sum games, *J. Optimization Theory and Applications*, **11** (1973), 533–555.

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

[53] H. V. Stackelberg, *The Theory of the Market Economy*, translated by Peacock A.T., William Hodge and Co., London, 1952.

[54] T. A. Taylor, Channel coordination under price protection, midlife returns, and end-of-life returns in dynamic markets, *Management Science*, **47** (2001), 1220–1234.

[55] J. T. Teng and G. L. Thompson, Oligopoly models for optimal advertising when production costs obey a learning curve, *Management Science*, **29** (1983), 1087–1101.

[56] M. L. Vidale and H. B. Wolfe, An operations research study of sales response to advertising, *Operations Research*, **5** (1957), 370–381.

[57] H. Von Stackelberg, *Marktform und Gleichgewicht*, Springer, Vienna, 1934. (An English translation appeared in *The Theory of the Market Economy*, Oxford University Press, Oxford, England, 1952.)

Table 1: Notations

$*$	Optimal/ equilibrium levels	$\hat{c}_i$	Unit advertising cost
i	Denotes the ith player	f	Production function, $\dot{x}$
A , $\hat{A}$	Advertising level	h_i	Unit inventory/ backlog cost
B ,$\hat{B}$	Local advertising	h_i^+	Unit inventory holding cost
C_i	Cost of production/advertising	h_i^-	Unit backlog cost
D	Demand, Revenue rate	m_i	Margin
G, G_i	Goodwill	p, p_i	Retail price
$H_i, \bar{H}_i$	Current -value Hamiltonians	q	Manufacturer's share of revenue
I_i	Inventory level	r_1, r_2	Effectiveness of advertising
J_i	Objective functional	t_1^d, t_2^d	Time parameters
K	Infrastructure capital	t_s, t_f	Start&end of promotional period
L_i	Labor force	u	Leader's control variable
M	Market size	$v,$	Follower's control variable
N	Number of firms	v^0	Optimal response of the follower
Q_i	Production rate, Processing rate	w	Wholesale price
$\bar{Q}_i$	Capacity limit	x	State variable; Sales rate
S	Shelf space	α	External market influence
S_i	Salvage value, Unit salvage value	α_i	Demand parameters
T	Planning horizon	β	Internal market influence
r	Optimal response of the follower	δ	Decay rate
V_i	Value function	$\theta, \hat{\theta}$	M's share of R's advertising cost
X	Cumulative sales	λ_i	Adjoint variable, Shadow price
$a,$	Market potential/Advertising effectiveness	π_i	Instantaneous profit rate
a_i, b_i, d, e_i	Problem parameters	ρ	Discount rate
a_l, a_s, b_l, b_s	Advertising effectiveness	ϕ, ψ	Adjoint variable, Shadow price
b	Price sensitivity/Advertising effectiveness	ω_i	Coefficient of incentive strategy
c_i	Unit production/advertising cost	$\mathcal{U}, \mathcal{V}$	Feasible set of controls
q_i	Order quantity	Λ, γ_i	Learning efficiency

Table 2: Summary of Model Descriptions

Section	Dynamics	L's decisions	F's decisions	Solution[1]
2.2.1	Seasonal	Production rate, Price	Price	OLSE
2.2.2	Seasonal	Production rate, Price	Price	OLSE
2.2.3	Seasonal	Production rate, Price	Price	OLSE
2.2.4	General	Price, Production rate	Price	OLSE
2.2.5	General	Price	Order quantity	OLSE
2.2.6	General	Price	Order quantity	OLSE
2.2.7	General	Price	Price	OLSE
2.2.8	General	Labor, investment	Labor, investment	OLSE
2.2.9	Bass-type	Price	Price	OLSE
2.2.10	Bass type	Price	Price	OLSE
2.3.1	NA dynamics	Participation rate, Ad. effort	Ad. effort	FSE
2.3.2	NA dynamics	Ad. effort, Price	Ad. effort, Price	FSE
2.3.3	NA dynamics	Participation rate, Ad. effort	Ad. effort, Price	FSE
2.3.4	NA dynamics	Ad. effort, (coop), Price	Ad. effort, Price	FSE
2.3.5	NA dynamics	Ad. effort, Incentive	Shelf-space	FSE
2.3.6	NA dynamics	Ad. effort	Ad. effort	FSE
2.3.7	Lanchester type	Ad. effort	Ad. effort	FSE
2.3.8	Sethi 1983	Participation rate, Price	Ad. effort, Price	FSE
2.3.9	Sethi 1983	Participation&Ad. rate	Participation&Ad. rate	FSE
2.5.1	Inventory	Price	Order Quantity, Price	OLSE, FSE
2.5.2	Production Cost	Price	Order quantity	FSE
2.5.3	Inv., Prod. Cost	Price, Production quantity	Price, Order quantity	FSE

[1]The symbol OLSE = Open-loop Stackelberg equilibrium and FSE = Feedback Stackelberg equilibrium.

Received January 2016; revised June 2016.

E-mail address: `tli1@scu.edu`
E-mail address: `sethi@utdallas.edu`