Cooperative Advertising and Pricing in a Dynamic Stochastic Supply Chain: Feedback Stackelberg Strategies

Xiuli He
Belk College of Business, The University of North Carolina at Charlotte, Charlotte, North Carolina 28223, xhe8@uncc.edu

Ashutosh Prasad, Suresh P. Sethi
School of Management, The University of Texas at Dallas, Richardson, Texas 75080
sethi@utdallas.edu, aprasad@utdallas.edu

Cooperative (co-op) advertising is an important instrument for aligning manufacturer and retailer decisions in supply chains. In this, the manufacturer announces a co-op advertising policy, i.e., a participation rate that specifies the percentage of the retailer’s advertising expenditure that it will provide. In addition, it also announces the wholesale price. In response, the retailer chooses its optimal advertising and pricing policies. We model this supply chain problem as a stochastic Stackelberg differential game whose dynamics follows Sethi’s stochastic sales-advertising model. We obtain the condition when offering co-op advertising is optimal for the manufacturer. We provide in feedback form the optimal advertising and pricing policies for the manufacturer and the retailer. We contrast the results with the advertising and price decisions of the vertically integrated channel, and suggest a method for coordinating the channel.

Key words: Co-op advertising; sales-advertising dynamics; differential games; sethi model; distribution channel

History: Received: October 2007; Accepted: April 2008, after 1 revision.

1. Introduction
Cooperative (or co-op) advertising is commonly used in supply chains as an incentive by the manufacturer to influence retailer behavior. In co-op advertising programs, the manufacturer contributes a percentage of the retailer’s advertising expenditure to increase the sales of the manufacturer’s product (Dutta et al. 1995, Nagler 2006). Co-op advertising programs can be a significant expense for the manufacturer. As much as 25–40% of local advertisements are cooperatively funded (Dant and Berger 1996). Total expenditure on co-op advertising in 2000 was estimated at $15 billion, compared with $900 million in 1970 (Nagler 2006).

The manufacturer’s decision problem, if offering a co-op advertising program to the retailer, is to determine what percentage of advertising expense to contribute. This is called its participation rate. The participation rate need not be constant but can be a function of the state, which in our case is the proportion of the market aware of the product. If it is assumed that retail price and retail advertising are the primary marketing mix drivers of sales, then the effect of the participation rate on these must be considered when designing a co-op program. That is, it should be expected that the wholesale price and the participation rate of the manufacturer will influence the retail price and the amount of retail advertising by the retailer, and, by backward induction, anticipation of the latter should influence the manufacturer’s decisions. Furthermore, it should be kept in mind that advertising decisions have dynamic effects on awareness and sales, which are also subject to uncertainty. Nevertheless, generating actual decisions is difficult and managerial guidelines do not presently incorporate all of these elements. For the manager, it is also important to understand how costs, advertising effectiveness, and other market and firm specific parameters will affect these decisions.

In attempting to provide managerial guidelines for co-op advertising, we study co-op advertising in the context of a supply chain consisting of a manufacturer selling through a retailer. Following a standard structure, the manufacturer is considered to be the Stackelberg leader and the retailer is the follower. Specifically, the manufacturer announces a wholesale price and the percentage of retail advertising that it will contribute. These can be state dependent, though
we shall show that in equilibrium and under the assumptions of our model they are not. In turn, the retailer decides on its advertising expenditure and the retail price. These can also be state dependent, but in our case only the optimal advertising is and the retail price is not. These determine the sales for the channel, which also is subject to uncertainty. We determine the condition when offering co-op advertising is optimal. We provide explicit solutions and comparative statics. Furthermore, we show that, in the absence of coordinated co-op advertising and pricing strategies, the decentralized channel will always have higher than optimal profit and lower than optimal advertising.

Prior literature on co-op advertising has often used static game models to capture the interaction between the manufacturer and retailer, thus ignoring the dynamic impact of advertising on sales. Dynamic models of co-op advertising have only recently been proposed (e.g., Jørgensen et al. 2000, 2001, 2003; Karray and Zaccour 2005), not least because the Stackelberg game structure with dynamics tends to make the modeling intractable. Often the solution procedure must take recourse to open-loop policies, which are not time consistent, or exclude the pricing decisions, or omit demand uncertainty (see He et al. 2007 for a survey). In contrast, this paper proposes a continuous-time dynamic model of co-op advertising and pricing decisions with uncertainty. In a dynamic stochastic model, we provide, for the first time, explicit feedback Stackelberg solutions of the optimal co-op policy – advertising and pricing decisions. These decisions are time consistent, i.e., neither the manufacturer nor the retailer will have an incentive to revise their policy at any future time. We discuss the literature and contribution in more detail in the next section.

The remainder of the paper is organized as follows. In Section 2 we survey the relevant background literature. Section 3 develops the proposed model, and Section 4 provides the analysis and results. We compare the vertically integrated and decentralized channels analytically and with numerical examples and graphs. In Section 5, we study the use of co-op advertising and revenue sharing contracts. Section 6 provides conclusions and directions for future research. Proofs of all results in the paper are given in Appendix A.

2. Background Literature

This is an interdisciplinary paper related to the literature on supply chain coordination in operations management and the literature on co-op advertising in marketing. In operations management, Cachon (2003) is a good reference on supply chain coordination. In Gerchak and Wang (2004), the retailer sets the shares of the revenue and then the suppliers decide delivery quantities. They show that revenue sharing alone cannot coordinate the assembly system, however, a revenue sharing scheme coupled with a subsidy paid by the retailer to suppliers can. Cachon and Lariviere (2005), in a newsstand setting, study the revenue sharing contracts between a retailer and a manufacturer who sets the wholesale price. Gerchak et al. (2006) study the revenue sharing contracts in a decentralized Stackelberg setting in which the video rental channel and the studio make independent decisions. Consistent with the coordination motif of this literature, we will examine the value of co-op advertising in coordinating the supply chain.

The co-op advertising literature consists of relatively few quantitative studies. We can divide them into two categories: Static models (Berger 1972, Bergen and John 1997, Dant and Berger 1996, Kim and Staelin 1999, Huang et al. 2002) and dynamic models (Jørgensen and Zaccour, 2003, Jørgensen et al. 2000, 2001, 2003, Karray and Zaccour 2005). We proceed to discuss these further.

In the static setting, Berger (1972) appears to be the first paper to analyze co-op advertising. It operationalizes co-op advertising as a wholesale price discount given by the manufacturer to the retailer as an advertising allowance. Sales are a concave function of the level of advertising chosen by the retailer. The conclusion is that both the manufacturer and retailer can be better off from co-op advertising. Dant and Berger (1996) extend the Berger model to study the co-op advertising decisions in franchising systems with demand uncertainty and where there is disagreement between the manufacturer and retailer on anticipated sales. Bergen and John (1997) study the effects of advertising spillover, differentiation across competing retailers, and differentiation across competing manufacturers on the participation rate. Kali (1998) finds that if the co-op advertising subsidy is made conditional on the advertised price being above a specified price threshold, then the channel can be coordinated on price and advertising. Finally, Huang et al. (2002) analyze both manufacturer-as-leader and partnership advertising structures. They consider the impact of local advertising and the national brand name investment on the retailer’s sales.

Huang et al. (2002) justify undertaking a static analysis of co-op advertising by arguing that promotion effects tend to be less long lasting than advertising, and are primarily a driver of short-term sales. Nevertheless, the empirical evidence in studies of concurrent advertising and promotion dynamics by Naik et al. (2005) and concurrent advertising and detailing by Chintagunta and Vílussim (1994) and
Fruchter and Kalish (1998) finds significant support for dynamic promotion effects. In our dynamic model, we use two parameters that can be adjusted to capture how important the present is in relation to the future. One is the finite decay term in the dynamics originating from Vidale and Wolfe (1957) that captures loss of market share due to forgetting and background competition. The other is the firm’s positive discount rate. If this rate is high, then the firm effectively behaves like a myopic agent and if it is low, like a foresighted firm. Setting these two parameters high should therefore resemble a static analysis, yet their presence gives the flexibility to explain dynamic implications. Another difference with the aforementioned papers is that we model the pricing decision in the channel explicitly. But compared with Huang et al. (2002), we have not included the possibility that the manufacturer may separately advertise its product. Some justification for this comes from the study by Nagler (2006, Table 1) who finds that of 2286 brands in the survey, 1470 listed a co-op advertising program but only 599 did national advertising. Ignoring this would not feel tempting to change its decisions depending on the state of the market. Second, we include the price decision as an endogenous decision variable and show how co-op advertising can overcome the double marginalization problem and coordinate the channel. Third, we include uncertainty into the demand dynamics and show that the results are robust to its inclusion. To achieve these goals, we employ the Sethi (1983) advertising model as the dynamics of the optimal control and differential game problems under consideration.

3. The Model
We consider a channel consisting of a manufacturer selling a product to end users through a retailer. The product is in a mature category where sales, expressed as a fraction of the potential market, is influenced through advertising spending and retail price. The manufacturer decides on the wholesale price and implements an advertising support scheme via a participation rate, i.e., for every dollar spent by the retailer, the manufacturer will contribute a certain percentage. Specifically, the manufacturer decides on the wholesale price \( w(t) \) and a co-op participation rate \( \theta(t) \) at time \( t \geq 0 \). The retailer decides on the channel’s total advertising effort level \( u(t) \) and the retail price \( p(t) \), \( t \geq 0 \).

The costs of advertising, i.e., advertising expenditure, are quadratic in the advertising effort \( u(t) \) and the manufacturer’s and retailer’s advertising expenditures at time \( t \) are given by \( \theta u(t)^2 \) and \((1 - \theta)u(t)^2\), respectively. The assumption of a quadratic cost function is common in previous research (e.g., Chintagunta and Jain 1992, Deal 1979, Jorgensen et al. 2000, Prasad and Sethi 2004) and implies increasing marginal cost of advertising effort. Equivalently, the retailer decides on the channel’s total advertising expenditure \( v(t) \), which results in a marginally diminishing advertising effect proportional to \( \sqrt{v(t)} \), \( t \geq 0 \).
We introduce the following additional notation.

**Notation**

- $t$: Time, $t \geq 0$
- $x(t) \in [0, 1]$: Proportion of the market aware of product at time $t$
- $x \in [0, 1]$: Aware proportion, also denotes an arbitrary initial aware proportion
- $u(t) \geq 0$: Advertising effort rate at time $t$
- $U(t) \geq 0$: Advertising effort rate for the integrated channel at time $t$
- $w(t) \geq 0$: Wholesale price at time $t$
- $p(t) \geq 0$: Retail price at time $t$
- $P(t) \geq 0$: Retail price for the integrated channel at time $t$
- $\theta(t) \geq 0$: Co-op advertising participation rate
- $D(p)$: Demand function
- $c \geq 0$: Constant unit production cost for manufacturer
- $\rho > 0$: Advertising effectiveness parameter
- $\delta > 0$: Awareness share decay parameter
- $r > 0$: Discount rate
- $m_M, m_R, m_I$: Gross margins for manufacturer (M), retailer (R), and integrated channel (I)
- $V_M, V_R, V_I$: Value functions for M, R, and I, respectively
- $\alpha_M, \alpha_R, \alpha_I$: Intercepts and slopes of the value functions of M, R and I, respectively

To model the dynamic effect of advertising on sales, we use the Sethi advertising model, which is related to the classical Vidale-Wolfe advertising model. Variants of the Sethi model have been used, for example, by Bass et al. (2005) and references therein, and empirically validated in studies such as Chintagunta and Jain (1995) and Naik et al. (2008).

We use the stochastic formulation of the Sethi model which is given by the Itô equation

$$dx(t) = \left(\rho u(t) \sqrt{1 - x(t)} - \delta x(t)\right)dt + \sigma(x(t))dz(t),$$

$$x(0) = x \in [0, 1], \quad t \geq 0,$$

(1)

where $x(t)$ in this paper represents the awareness share, i.e., the number of aware (or informed) customers expressed as a fraction of the total market at time $t$, $x$ denotes the initial condition, $\rho$ is a response constant, and $\delta$ determines the rate at which potential consumers are lost due to background competition, product obsolescence, forgetting, etc. Because we aim to solve the problem for any initial $x$, we can treat it as arbitrarily chosen. Thus, it will denote as well the variable representing the awareness share. In the second term on the right-hand side, $\sigma(x(t))$ represents a variance term and $z(t)$, $t \geq 0$, represents a standard Wiener process on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We should stress that the admissible class of decisions $u(t)$, $t \geq 0$, are nonnegative stochastic processes, non-anticipative with respect to the Wiener process $z(t)$, $t \geq 0$. This requirement applies also to the other decisions that will be introduced later.

Some elucidation of the properties of Equation (1) is needed. Note that the specification of the dynamics has desirable properties such as concave response with saturation. The awareness share is non-decreasing in advertising and subject to decay. The awareness share dynamics are affected linearly by the advertising effort $u(t)$, which is the square root of the advertising expenditure $u(t)^2$ and is thus a concave function for the advertising expenditure. The rate of awareness share increase is unaffected by the participation rule, which is an internal transfer between the channel members. Advertising affects the proportion of unaware consumers as in Sethi (1983), but sales are generated only by the fraction of aware consumers who are willing to pay the retail price. Price was similarly included in the objective function by Bass et al. (2005) in a simultaneous-move Nash advertising game.

**Remark 3.1.** The awareness share remains bounded within $[0, 1]$ despite the stochastic disturbances, as we assume that the function $\sigma : (0, 1) \rightarrow \mathbb{R}$ is continuous and Lipschitz on every closed subinterval of $(0, 1)$, $u(t) \geq 0$, $t \geq 0$, and $\sigma(0) = \sigma(1) = 0$. This gives a strictly positive drift when the awareness share is 0 and a strictly negative drift when it is 1. Then, from Gihman and Skorohod (1972, p. 149, 157–158), 0 and 1 are the natural boundaries for the solutions of (1) with $x(0) = x \in [0, 1]$, i.e., $x(t) \in (0, 1)$ almost surely for $t > 0$. \[ \square \]

The channel members have a constant and positive discount rate $r$ and play a Stackelberg differential game over an infinite horizon. We regard the manufacturer as the Stackelberg leader and the retailer as the follower. We restrict our attention to feedback Stackelberg solutions where the optimal policy, in general, depends on the current state and time (see Basar and Olsder 1999). However, in an infinite horizon setting with time-independent parameters, we can – with good reason – choose to focus on stationary equilibria. Thus, the feedback policies will not depend explicitly on time $t$. There may be non-stationary equilibria as well, which we do not discover, however.

The sequence of the events is as follows: First, the manufacturer announces the feedback wholesale price policy $w(x) \geq 0$ and the feedback participation rate $\theta(x) \in [0, 1]$. This means that at any time $t \geq 0$, if the
state is \( x(t) \), then at time \( t \), the wholesale price, denoted as \( w(t) \), would be \( w(x(t)) \) and the participation rate, denoted as \( \theta(t) \), would be \( \theta(x(t)) \). Note here a slight abuse of notation in using \( w(x) \) and \( \theta(x) \) as feedback policies, and \( w(t) \) and \( \theta(t) \) as the decisions at time \( t \).

Second, the retailer sets the retail price \( p(t) \) and the advertising effort rate \( u(t) \) as its optimal response to the manufacturer’s announced decisions. The retailer accomplishes this by solving an optimization problem to maximize the present value of its profit stream over the infinite horizon. Given the manufacturer’s announced policies, the retailer’s optimal control problem is given by

\[
V^R(x) = \max_{x\geq 0, u\geq 0} \left\{ \int_0^\infty e^{-\gamma t}(P(x(t),p(t),\theta(t)))dt \right\} \nonumber
\]

subject to (1), where \( E \) denotes the expectation operator.

We will assume that the demand function is downward-sloping and differentiable, and satisfies the usual conditions to ensure an interior solution for price. Moreover, \( 0 \leq D(p) \leq 1 \).

We have denoted the optimal value of the retailer’s discounted total profit at time zero by \( V^R(x) \), clearly indicating that it depends only on the initial value \( x \). Because \( x \) is arbitrary, the function \( V^R(x) \) is defined on the entire domain \([0,1]\). Furthermore, because ours is an infinite horizon problem with stationary parameters, the future at any time \( t \) looks the same as it does at time zero so long as \( x(t) = x \). This means that \( V^R(x) \) will also provide us with the present value of the profit stream associated with an optimal policy on the interval \([t,\infty)\), \( t \geq 0 \), discounted to time \( t \) if \( x(t) = x \) at that time. Thus, \( V^R(x) \) defines the value function for our problem.

The optimal solution of the problem (1), (2) yields the retailer’s feedback retail price \( p(x,w,\theta) \) and advertising effort \( u(x,w,\theta) \) in response to the announced policies \( w(x) \) and \( \theta(x) \). While these responses are simply functions of \( x \), the explicit notation we are using emphasizes their dependence on the announced policies. As an example, if \( u(x,w,\theta) = axw + \theta \), \( w(x) = ax^2 \), and \( \theta(x) = bx \), then \( u(x,w,\theta) \) means \( u(x|w(x),\theta(x)) = ax^3 + bx \). Furthermore, if \( x(t) \) denotes the awareness share at any time \( t \geq 0 \), then the retailer’s advertising effort \( u(t) \) at time \( t \) will be \( u(x(t)|w(t),\theta(t)) = u(x(t)|w(x(t)),\theta(x(t))) = ax(t)^3 + bx(t) \).

The manufacturer anticipates the retailer’s reaction functions and incorporates them into its optimal control problem, and solves for its wholesale price policy \( w(x) \) and the participation rate policy \( \theta(x) \). Therefore, the manufacturer’s problem can be stated as

\[
V^M(x) = \max_{w(t),\theta(t)} \left\{ \int_0^\infty e^{-\gamma t}(P(x(t),w(t),\theta(t)))dt \right\} \nonumber
\]

\[
dx(t) = \left( \rho u(t) \sqrt{1 - x(t) - \delta x(t)} \right) dt + \sigma(x(t)) \, dz(t), \nonumber
\]

where \( V^M(x) \) is the manufacturer’s value function.

The solution to problem (3) yields the equilibrium feedback policies \( w^*(x) \) and \( \theta^*(x) \) for the manufacturer.

Once we have the solution \( w^*(x) \) and \( \theta^*(x) \), we can then express the retailer’s feedback price and advertising effort as \( p^*(x) = p(x|w^*,\theta^*) \) and \( u^*(x) = u(x|w^*,\theta^*) \). We would like to stress here that the policies \( w^*(x) \), \( \theta^*(x) \), \( p^*(x) \) and \( u^*(x) \) constitute a feedback Stackelberg equilibrium, which is time consistent. Using these policies in the state Equation (1) results in the stochastic awareness process \( x^*(t) \), \( t \geq 0 \), and the respective decisions will be \( w^*(x^*(t)) \), \( \theta^*(x^*(t)) \), \( p^*(x^*(t)) \) and \( u^*(x^*(t)) \) at time \( t \geq 0 \).

We will obtain these policies explicitly in this paper. Thus, this paper represents a major advance over the bulk of the literature on Stackelberg differential game models in advertising and pricing, where open-loop policies are obtained, which are in general not time consistent.

Before we proceed to analysis in the next section, we also formulate the problem for the vertically integrated channel. Here the pricing and advertising decisions are made by a centralized decision maker in order to maximize the present value of the total channel profit. By denoting the retail price as \( P(t) \) and the joint advertising effort by \( U(t) \) at time \( t \), the problem can be written as

\[
V^I(x) = \max_{p(t),U(t)|t\geq 0} \left\{ \int_0^\infty e^{-\gamma t}(P(t) - c)D(P(t),U(t))dt \right\} \nonumber
\]

\[
dx(t) = \left( \rho U(t) \sqrt{1 - x(t) - \delta x(t)} \right) dt + \sigma(x(t)) \, dz(t), \nonumber
\]

\[
x(0) = x \in [0,1], t \geq 0. \nonumber
\]

We now proceed to the analysis of the decentralized and integrated channels.

4. Analysis

We first consider the case of the integrated channel in which there is a central decision maker making the advertising and pricing decisions. This provides the benchmark against which the decentralized channel results can be compared.

4.1. The Integrated Channel

The analysis can be broken into two parts that provide the optimal price and optimal advertising, respectively. First, observe that \( P \) occurs only in the
integrand and not in the dynamics of (4). Thus, we can maximize the channel profit by first maximizing the integrand of (4) with respect to $P$, holding $U$ fixed, and then solve the resulting optimal control problem to obtain the optimal $U$ (see, e.g., Sethi and Thompson 2000, §7.1.1). The optimal price $P^*$ satisfies the first-order condition (FOC) of maximizing $(P - c)D(P)$ with respect to $P$. Thus, the optimal pricing decision can be obtained independent of the advertising decisions. Specifically, the optimal price $P^*$ is constant over time, and it is the solution of the equation $(P^* - c)D'(P^*) + D(P^*) = 0$. Let $m_1 \equiv (P^* - c)D(P^*)$ denote the integrated channel’s optimal gross margin per unit share of awareness. Note that this is different from the standard margin $(P^* - c)$ per unit sale. Using the implicit function theorem, it can be shown that $\partial P^*/\partial c > 0$ and $\partial m_1/\partial c < 0$. The former occurs because marginal revenue equals marginal cost at the optimal demand, and when marginal cost increases, the firm sells fewer units by charging a higher price so that the marginal revenue is higher. But overall, due to cost increases, the gross margin is reduced.

We next solve the optimal control problem for the advertising decision. Here the objective function in (4) is rewritten with the maximized retail price.

**Proposition 1:** For the integrated channel, we have the following results.
(a) The integrated channel’s optimal profit is linear in $x$, i.e., $V^I = a_1 + b_1x$, where $a_1$ and $b_1$ are positive constants given in terms of the system parameters as follows:

$$a_1 = \frac{\beta_1^2 \rho^2}{4r}, \quad b_1 = \frac{2m_1 \sqrt{r + \delta}}{\rho^2 m_1 + (r + \delta)}.$$  

Comparative statics are $\partial a_1/\partial m_1 > 0, \partial b_1/\partial m_1 > 0, \partial b_1/\partial \rho < 0, \partial b_1/\partial r < 0$.

(b) The optimal feedback advertising policy is

$$U^I(x) = \frac{\beta_1 \rho \sqrt{1 - x}}{2} = \sqrt{r}a_1(1 - x).$$

Comparative statics are $\partial U^I/\partial m_1 > 0, \partial U^I/\partial \rho > 0, \partial U^I/\partial r < 0, \partial U^I/\partial \delta < 0$. □

The solutions are explicit, and the comparative statics on the decision variables and the value function appear to be intuitive. When the decay parameter is higher or the effectiveness of advertising lower, it is optimal to have less advertising. Advertising is higher when the firm is less myopic, i.e., it discounts the future less. This is due to the carryover effect of advertising. To obtain the effect of demand function parameters on the results, which will have an impact via the gross margins, we can employ, for example, the linear demand function $D(P) = 1 - \eta P$, where $\eta$ captures price sensitivity and $0 \leq P \leq 1/\eta$. In this case, $P^* = (1 - \eta)/2\eta$ and $m_1 = (1 - \eta^2)/4\eta$, so $\partial m_1/\partial \eta < 0$. When demand is more sensitive to price, the optimal price is lower and the optimal advertising should also be lower, because the returns from advertising have decreased.

The optimal advertising effort is proportional to $\sqrt{1 - x}$, i.e., the optimal advertising budget is proportional to $(1 - x)$. In other words, when the awareness share is higher, there should be less advertising than when it is lower. Thus, the benefit of advertising is greatest at the beginning, over the steepest part of the sales-advertising response function when the awareness share is low, rather than when it is closer to saturation.

4.2. The Decentralized Channel

The results for the Stackelberg differential game are obtained as described in the previous section. We present them in the following two propositions. Proposition 2 will specify the retailer’s best response and Proposition 3 will characterize the Stackelberg equilibrium for the supply chain.

**Proposition 2:** For any given policies $w(x)$ and $\theta(x)$ of the manufacturer, the retailer’s reaction price is independent of $\theta(x)$ and can be expressed as $p(x|w) = \hat{p}(w(x))$, where the function $\hat{p}(w)$ solves

$$D(\hat{p}) + (\hat{p} - w)D'(\hat{p}) = 0,$$

and its advertising reaction policy is independent of $w(x)$, and can be expressed as

$$u(x|\theta) = \frac{V^R p \sqrt{1 - x}}{2(1 - \theta(x))}.$$  □

We insert these reaction functions into the maximization problem for the manufacturer and proceed further. It turns out that, depending upon the parameter values, there can be two types of equilibria. In the first, which we term as the no co-op equilibrium, the manufacturer does not provide any co-op advertising program to the retailer, whereas in the second, it does. We state these results in the following proposition.

**Proposition 3 (Feedback Stackelberg Equilibrium):** For any given set of parameters, there exists a unique feedback Stackelberg equilibrium $(w^*(x), p^*(x), \theta^*(x), u^*(t))$ for the game, given by $w^*(x) = w^*$ satisfying $(w^* - c)\frac{d[D(p(w^*))]}{dw^*} + D'(p(w^*)) = 0$ and $p^*(x) = p^* = \hat{p}(w^*)$, with the function $\hat{p}(w)$ as defined in Proposition 2, and with $\Delta = \frac{m_1^2 \rho - m^2 \rho^2}{(r + \delta)^2 + \rho^2 m_2 + (r + \delta)}$, $m_M \equiv (w^* - c)D(p^*)$, $m_R \equiv (p^* - w^*)$ follows:

First, we discuss the optimal prices obtained in Proposition 3. The retail price is higher in a decen-
Retailer’s profit $V^R$

$V^R(x) = a_R + \beta_R x$

Manufacturer’s profit $V^M$

$V^M(x) = a_M + \beta_M x$

Coefficients of profit functions, $\alpha_R, \beta_R, \alpha_M, \beta_M$ obtained from:

\[
\begin{align*}
\beta_R &= \frac{2m_R}{\sqrt{(r + \delta)^2 + r^2 + (r + \delta)}} \\
\beta_M &= \frac{2m_M}{2(r + \delta) + \beta_R r^2} \\
\alpha_R &= \frac{\beta_R r^2}{4r} \\
\alpha_M &= \frac{\beta_R \beta_M r^2}{2r}
\end{align*}
\]

Participation rate $\theta^*(x) = 0$

Advertising effort $u^*(x) = \frac{\rho \beta_R \sqrt{1 - x}}{2} = \sqrt{r a_R (1 - x)}$

(a) if $\Delta \leq 0$

No Co-op Equilibrium

(b) if $\Delta > 0$

Co-op Equilibrium

\[
\begin{align*}
\beta_R &= \frac{m_R}{r + \delta} - \frac{\beta_R (\beta_R + 2\beta_M)r^2}{8(r + \delta)} \\
\beta_M &= \frac{m_M}{r + \delta} - \frac{(\beta_R + 2\beta_M)^2 r^2}{16(r + \delta)} \\
\alpha_R &= \frac{\beta_R (\beta_R + 2\beta_M)r^2}{8r} \\
\alpha_M &= \frac{(\beta_R + 2\beta_M)^2 r^2}{16r}
\end{align*}
\]

EXAMPLE: With the demand function $D(p) = 1 - \eta P$, we see that $p^* = (3 + \eta c)/4\eta$, $w^* = (\eta c + 1)/2\eta$, $m_R = (1 - \eta c)^2/16\eta$, $m_M = (1 - \eta c)^2/8\eta$, and $m_M + m_R = 3(1 - \eta c^2).$

Thus, $p^* > P^* = (1 + \eta c)/2\eta$, because $\eta c < \eta P^* \leq 1$. □

Next, we discuss the no co-op solution appearing in the left column (a) of the table in Proposition 3, which happens when $\Delta \leq 0$ The salient points of the no co-op solution, which is completely explicit, are, first, that advertising is proportional to $\sqrt{a_R}$, and a comparison with the integrated channel case shows that it is necessarily suboptimal for the channel. Second, when the margin of the manufacturer is sufficiently smaller than that of the retailer, then the condition under which the no co-op equilibrium holds is satisfied. Thus, the manufacturer should not offer a co-op advertising incentive when the retailer has sufficient channel power to swing the split of channel margins in its direction. It is useful that the analysis provides a result with no co-op advertising, because in practice co-op advertising is not observed all the time. While noting that the condition for no co-op advertising specified in the proposition also applies when $m_M$ and $m_R$ are exogenously specified, we can verify that there are demand functions from which the condition or its negation emerge endogenously.

EXAMPLES: The linear demand function and the isoelastic demand function provide the verification being sought, since they generate opposite implications for whether there should be co-op advertising or not in equilibrium. Note that these are the most commonly used demand functions (e.g., Petruzzi and Dada 1999). With the linear demand function $D(p) = 1 - \eta p$, $\eta \in (0, 1/p)$, it can be derived that $\hat{w}(w) = (1 + \eta w)/2\eta$, $\hat{w}(w) - w = (1 - \eta w)/2\eta$, and further that $w^* = (1 + \eta c)/2\eta, w^* - c = (1 - \eta c)/2\eta$. Thus $m_R/m_M = (1 - \eta w^*^)/(1 - \eta c)$. With the isoelastic demand function $D(p) = p^{1 - \gamma}, \gamma > 1$, and $c > 1$, where $\gamma$ is the price elasticity of demand, we get $\hat{w}(w) = w(\gamma - 1), \hat{w}(w) - w = w/(\gamma - 1)$, and further that $w^* = c/\gamma(\gamma - 1), w^* - c = c/\gamma(\gamma - 1)$. Thus $m_R/m_M = w^*/c$. Noting further that $w^* > c$ in both cases, it is clear that in the linear demand case $m_R < m_M$, which means that Proposition 3(a) never applies and co-op advertising is optimal. On the other hand, with isoelastic demand, we have $m_R > m_M$, which implies that Proposition 3(a) applies for low values of the elasticity $\gamma$, and thereby rules out the
optimality of co-op advertising. For high values of \( \beta \), co-op advertising is optimal. \( \square \)

Finally, we consider the case where the condition (a) for no co-op advertising does not hold. Then, the right column condition (b) holds, i.e., \( \Delta > 0 \). The parameters \( z_R, \beta_R, z_M, \beta_M \) are obtained from the simultaneous solution of the four implicit equations. Because it is the value of \( \Delta \) that determines whether the manufacturer will share in advertising or not, we could term \( \Delta \) as the differential power of the manufacturer over the retailer. Note that \( m_M \geq m_R \) is sufficient for \( \Delta > 0 \), but not necessary.

It is now possible to see that the advertising in the decentralized channel is lower than in the vertically integrated channel. The ratio of the advertising expenditures \( u^*(x)/U^*(x) \) is given by \( \sqrt{z_M/z_I} \) if \( \Delta > 0 \) and \( \sqrt{z_R/z_I} \) if \( \Delta \leq 0 \). Because the integrated channel must have at least as high a value function as the decentralized channel for any \( x \), including \( x = 0 \), we have \( z_I \geq z_M + z_R \). And further, because the manufacturer cannot extract all the surplus from the decentralized channel, \( z_I > z_M \) and \( z_I > z_R \). Therefore, the advertising in the decentralized channel is lower than the channel optimal rate. This extends the result obtained by Huang et al. (2002) in the static case. We may also note that the ratio is constant, depending only on the problem parameters and not on the current awareness share.

Figure 1 shows how we can interpret the optimal advertising policy results such as \( U^*(x) = \sqrt{z_I(1-x)} \), obtained in Proposition 1 for the integrated case, in the usual marginal benefit equals marginal cost argument. For the integrated channel, we see by differentiating Equation (A1) with respect to \( U \) that the marginal cost of advertising is \( 2U \) and its marginal benefit is \( V_I' \rho \sqrt{1-x} \). To see it more clearly, let us see what happens when we apply a constant control \( U \) in a small interval \([t, t+\delta t]\), when \( x(t) = x \). From the Itô differential Equation (1), we can see that \( E(x(t+\delta t)) = x + (\rho U \sqrt{1-x}) \delta t \), because the expectation of the stochastic term is zero. But the marginal value (or the shadow price) of the state is given by \( V_I^\alpha \). Thus, the expected benefit is \( V_I'[E(x(t+\delta t))] - x = V_I' \rho U \sqrt{1-x} \delta t \). Moreover, the total advertising cost in the small interval is \( U^2 \delta t \). The difference of these two terms gives the expected profit \( [V_I' \rho U \sqrt{1-x} - U^2] \delta t \) resulting from our action in the interval \([t, t+\delta t]\). We can maximize this profit by taking its derivative with respect to \( U \) and set it equal to zero. But this gives precisely the FOC (A2), which can now be seen as the condition of equating marginal revenue and marginal cost of advertising. Further analysis reveals that the marginal revenue in the integrated case can also be expressed as \( MR_I = V_I' \rho \sqrt{1-x} = \beta_I \rho \sqrt{1-x} = 2 \sqrt{z_I(1-x)} \).

For the decentralized channel as a whole, the instantaneous advertising expense rate is \( u^2 \), and so the marginal cost is \( 2u \). We can disaggregate this for the retailer and manufacturer as follows: The marginal costs are \( 2(1-\theta)u \) and \( 2u \rho \), respectively. Furthermore, the condition showing the equality between the marginal cost \( 2(1-\theta)u \) and the marginal benefit \( V_R^\alpha \rho \sqrt{1-x} \) for the retailer is given by (A12).

Although more complicated to show, the equality between marginal benefit and marginal cost also holds for the manufacturer. To see this, let us solve the manufacturer’s problem (3) subject to the retailer’s response (A13). This can be done by either using the Lagrange multiplier method or by substituting for \( \theta \) from (A13) into (3). If we did the latter, then instead of the HJB Equation (A15), we would get the equation

\[
rV^M = \max_{u \geq 0} \left[ m_M(x) - u^2 + \frac{V_R^\alpha \rho \sqrt{1-x}}{2} + V_x^M \rho \sqrt{1-x} \right].
\]

The first order condition for \( u \) is

\[
-2u + \frac{V_R^\alpha \rho \sqrt{1-x}}{2} + V_x^M \rho \sqrt{1-x} = 0,
\]
which, by using (A13) and adding $\theta u$ on both sides, can be written as

$$V_x^M \rho \sqrt{1 - x} - u(1 - \theta) = 2\theta u. \quad (5)$$

Here we can see that without constraint (A13), the manufacturer’s marginal benefit would be simply $V_x^M \rho \sqrt{1 - x}$. The decrease in the marginal benefit by $u(1 - \theta)$ is because of constraint (A13). If we had used the Lagrange multiplier approach, then it can be shown that $u(1 - \theta)$ will be replaced by $2\lambda(1 - \theta)$ with $\lambda = u/2$ as the value of the Lagrange multiplier, which we know represents the shadow price associated with the constraint (A13).

Finally, by summing (5) and (A12), we obtain

$$(V_x^M + V_x^R) \rho \sqrt{1 - x} - u(1 - \theta) = 2u.$$

The right-hand side is the total marginal cost of advertising for the decentralized channel and the left-hand side represents the marginal revenue $MR_D$ of the decentralized channel, which can be written as

$$MR_D = \begin{cases} 2\sqrt{r \delta} (1 - x) & \text{in the no co-op solution}, \\ 2\sqrt{r \delta} (1 - x) & \text{in the co-op solution}. \end{cases}$$

We can now see that the optimal advertising effort derived in Proposition 3 for the decentralized channel can be obtained from $MC_D = 2u$.

We have seen already that $x_L > x_M$ and $x_L > x_R$, and hence Figure 1 shows that under-advertising, leading to lower sales and lower channel profit, takes place in the decentralized solution in both co-op and no co-op equilibria. Moreover, this under-advertising comes about from the application of the marginal cost equals the margin cost argument two times — once by the retailer and once by the manufacturer. This effectively generalizes the inefficiency caused by double marginalization for pricing in static supply chains to advertising decisions in dynamic stochastic supply chains.

### 4.3. Comparative Statics in the Decentralized Channel

The comparative statics on the advertising decision in the no co-op solution follow from direct observations, and they are

$$\frac{\partial u^*}{\partial m_R} > 0, \quad \frac{\partial u^*}{\partial m_M} = 0, \quad \frac{\partial u^*}{\partial \rho} > 0, \quad \frac{\partial u^*}{\partial \theta} < 0, \quad \frac{\partial u^*}{\partial \delta} < 0. \quad (6)$$

These are as in the integrated channel case, except for the clarification that advertising effort does not depend on the manufacturer’s gross margin, but only on the retailer’s. The comparative statics for advertising in the co-op case are likely identical, but the intractability of the algebraic expressions presents a hurdle to confirming this.

For the comparative statics on the participation rate with respect to the gross margins of the firms, we employ numerical analysis. For this we assume the margins to be given. This analysis is fairly straightforward because we have the formula for the optimal participation rate in Proposition 3. In Figures 2 and 3, we use parameters $r = 0.05$, $\delta = 1$ and $\rho = 2$. We fix the retailer’s margin to $m_R = 0.10$ when plotting the relationship between the manufacturer’s margin $m_M$ and the participation rate $\theta$ in Figure 2, and the manufacturer’s margin to $m_M = 0.4$ when plotting $m_R$ against $\theta$ in Figure 3. Figure 2 shows that as the manufacturer’s margin increases, the participation rate increases. In contrast, Figure 3 shows that as the retailer’s margin increases, the participation decreases.

The result appears to be consistent with previous literature (e.g., Berger 1972) that has found that the participation rate increases with the manufacturer’s
margin, and this is supported by the empirical studies as well, such as by Nagler (2006). The argument is that the manufacturer has an incentive to motivate the retailer to increase advertising as it has a higher value for sales. However, a substantiation of this result may not hold in general. An investigation of the conditions under which it would hold is left as a topic of further research.

When the margins are not given, they need to be derived in terms of the problem parameters such as \( r, \delta, \) and \( \rho \) for any given demand functions. While we do not obtain comparative statics in terms of these parameters, these can be easily obtained numerically.

The analysis in the next section proceeds with the case that there is co-op advertising in equilibrium.

### 4.4. Evolution Process of Awareness Share

We next examine the awareness share processes analytically for both the integrated and the decentralized channels. Inserting the values of the advertising effort into the state equation, we obtain,

\[
dx(t) = (A(1-x(t)) - \delta x(t))dt + \sigma(x(t))dz(t),
\]

\[
x(0) = x \in [0,1], \quad t \geq 0,
\]

where \( A = \rho \sqrt{r_I} \) for the integrated channel, and \( A = \rho \sqrt{r_M} \) for the decentralized channel.

To characterize the evolution processes, a specification of the disturbance function is required. Following Prasad and Sethi (2004), we use \( \sigma(x) = \sigma \sqrt{1-x} \), where \( \sigma \) is a constant. This specification has the properties discussed in Remark 3.1, and it ensures the awareness share to remain bounded within \([0,1]\) despite the stochastic disturbances. With this specification, we get the following result.

**Proposition 4:** The density of the stationary distribution of the awareness share is given by the Beta density, i.e.,

\[
f(y) = \frac{\Gamma\left(\frac{24}{\sigma^2} + \frac{1}{\sigma^2} \right)}{\Gamma\left(\frac{24}{\sigma^2}\right) \Gamma\left(\frac{24}{\sigma^2} \right)} y^{\frac{24}{\sigma^2} - 1} (1-y)^{\frac{24}{\sigma^2} - 1},
\]

where \( \Gamma(s) = \int_0^\infty q^{s-1}e^{-q}dq \), \( s > 0 \) is the gamma function. Furthermore, the long-run equilibrium awareness shares, denoted by \( \bar{x}_I \) for the integrated channel and \( \bar{x}_D \) for the decentralized supply chain, are given by

\[
\bar{x}_I = \frac{\rho \sqrt{r_I}}{\rho \sqrt{r_I} + \delta}, \quad \bar{x}_D = \frac{\rho \sqrt{r_M}}{\rho \sqrt{r_M} + \delta}.
\]

Clearly, as a consequence of \( x_I > x_M \), we get \( \bar{x}_I > \bar{x}_D \). What this means is that, due to higher advertising in the integrated channel, the average awareness about the product is higher in this case in the long run than when it is sold through a decentralized supply chain, ceteris paribus.

Illustrative examples can be obtained for different parameter values. For the graphs in Figures 4–7, we used the linear demand function and parameter values \( c = 0 \) and \( \eta = 0.25 \). This generates \( m_I = 1 \), \( m_M = 0.5 \) and \( m_R = 0.25 \). Next, the values \( r = 0.05 \), \( \delta = 1 \), \( \sigma = 0.25 \), and advertising effectiveness \( \rho = 2 \) are used. In practice, a decision calculus approach could be followed to obtain the parameter values. The solutions of the unknown coefficients are \( x_I = 7.3 \), \( x_M = 3.29 \), \( x_R = 1.39 \), \( \beta_I = 0.6 \), \( \beta_M = 0.32 \), \( \beta_R = 0.17 \). Therefore, \( \theta^* = 0.58 \).

From this information, the value functions are plotted. In Figure 4, it can be seen that the value function for the integrated channel is over 1.5 times that of the decentralized channel. This is higher than \( m_I/(m_M + m_R) = 1.33 \) times, which would be the result given by a static analysis. The additional inefficiency is, of course, due to lower than the centralized optimal level of advertising as shown in Figure 1 and as can be seen from Figure 5. Channel coordination is thus of greater value when co-op advertising is used. Another thing to note is that the manufacturer enjoys the first mover advantage, obtaining over twice as much profit as the retailer.

Next, we graph the sample paths of awareness shares and advertising efforts over time, to make the
Figure 6 Sample Paths of Awareness Shares for Integrated and Decentralized Channels

Figure 7 Sample Paths of Optimal Advertising Efforts for Integrated and Decentralized Channels

effect of the stochastic term much clearer. To do this, note that the SDE $dx(t) = a(x)dt + b(x)d\omega(t)$ can be numerically approximated by $x(t + \Delta) = x(t) + a(x(t))\Delta + b(x(t))\sqrt{\Delta}\sigma(t)$. The $\{\sigma(t)\}$ are i.i.d. Normal with mean 0 and variance 1, generated using Excel’s random number generator, and using the time step $\Delta = 0.01$.

Figure 6 shows a sample path from an initial starting point $x = 0.1$. Because the equilibrium awareness shares for both channel structures are higher than the initial value, there is a rise over time towards the long-run equilibrium awareness shares $\bar{x}_I = 0.55$ and $\bar{x}_D = 0.45$. A comparison of Figures 6 and 7 together will show that when the awareness share decreases below its mean value, then the advertising is higher, and vice versa. When the stochastic disturbances move the awareness share away from the mean, the optimal advertising changes in a way that tends to return the awareness share back to the mean. Thus the advertising acts to reduce the effects of stochastic disturbances. Figure 7 also shows that the advertising effort at each instant of time is also lower in the decentralized case than in the integrated case.

5. Revenue Sharing Contracts and Co-op Advertising

We consider the case in which the co-op advertising program is combined with a revenue sharing contract. We want to examine whether the combination of these two schemes can coordinate the channel, i.e., achieve the same profit level in the vertically integrated channel. We assume that revenue sharing is applied to all units. The firms’ profit functions for any constant sharing rule $\theta$ are

$$V^R = E \int_0^\infty e^{-rt} \{((1-\theta)p(t) - w(t))D(p(t))x(t) - (1-\theta)uw^2(t)\}dt$$

$$V^M = E \int_0^\infty e^{-rt} \{[\theta p(t) + w(t) - c]D(p(t))x(t) - \theta uw^2(t)\}dt$$

**Proposition 5:** A revenue sharing and co-op advertising contract where the manufacturer receives a constant fraction $\theta$ of the retail revenue and contributes the same fraction towards co-op advertising will coordinate the channel. In equilibrium,

$$w^*(x) = c(1-\theta), p^*(x) = P^*, u^*(x) = U^*(x),$$

$$V^R = (1-\theta)V^I$$

We make a few observations. First, the retailer shares the supply chain’s profit by sharing both its cost and revenue. Second, the manufacturer’s wholesale price is less than its production cost, which implies that it loses money in selling the product to the retailer, but makes profit by participating in revenue sharing. Third, the revenue sharing contract can arbitrarily split the profit between the manufacturer and the retailer. These three observations carry over in the dynamics case, the results obtained in Cachon and Lariviere (2005) in a one-period newsvendor model.

There are other contracts by which the channel can be coordinated. For example, a two-part tariff consisting of a wholesale price per unit and a fixed fee is possible. The wholesale price is set to the marginal cost and the participation rate is set to zero, which will make the retailer’s problem equivalent to the integrated channel problem. The fixed fee extracts the maximum surplus and hence the manufacturer sets it equal to the integrated channel profit leaving the retailer with no surplus. Instead of this, another method is to have a bargained split of channel revenues so that it is possible for both parties to have some surplus.
6. Conclusions

Co-op advertising is a widely used marketing tool that affects advertising and pricing policies throughout the supply chain. A co-op advertising plan specifies a participation rate, which is defined as the percentage of the retailer’s advertising expense on the manufacturer’s product that is contributed by the manufacturer to the retailer. In this paper we provide a theoretical analysis of co-op advertising plans in a dynamic stochastic supply chain.

We first develop a model that has the following salient features. Unlike much of the work in the co-op literature, this is a dynamic model. Modeling dynamics is particularly important when studying advertising, because the carryover of advertising and promotion effects is well documented. We employ the dynamics of the Sethi model, which is analytically tractable and has been validated in empirical studies. In further contrast to the literature, we incorporate uncertainty in awareness share and derive optimal feedback policies for price and advertising by the manufacturer and retailer.

The sequence of events is that the manufacturer announces a participation rate policy and wholesale price policy, and subsequently the retailer determines its selling price and local advertising policy. These decisions constitute a feedback strategy and are thus time consistent.

There are two types of equilibria identified in the analysis that may occur depending on the parameter values. One corresponds to no co-op advertising (Proposition 3(a)) and the other to positive co-op advertising (Proposition 3(b)). Thus, we note that it is not always optimal for the manufacturer to offer a co-op advertising program. In the case when the retailer’s margin is lower than the manufacturer’s margin, or close to it, then the manufacturer will offer a co-op program. Otherwise, when the retailer’s margin is significantly higher than the manufacturer’s, then the manufacturer will not offer a co-op program.

We discussed how the shape of the demand curve is relevant to this conclusion, with co-op advertising being likely when there is linear demand and less likely with isoelastic demand.

We find that the optimal participation rate, the wholesale price, and the retail price are constants that depend only on the model parameters (Proposition 2–3). The optimal advertising expenditure on the other hand depends on the awareness share and follows an inverse law (Proposition 3). The results are robust against uncertainty. With a specification of the random noise, we are able to show that the awareness process has a Beta distribution (Proposition 4).

We solve the model for a vertically integrated channel (Proposition 1). Comparing its results to those for the decentralized channel, we find that in the absence of co-op advertising, the decentralized channel has higher than optimal prices and lower than optimal advertising. The higher prices can be explained by the standard double marginalization argument. Under-advertising, on the other hand, has a related but more involved explanation. Whereas wholesale price by itself cannot correct for these problems, we demonstrate that a revenue and advertising sharing contract allows the channel to achieve the coordinated outcome (Proposition 5). Thus, we show that for the manufacturer, decision making that jointly optimizes co-op advertising and price has an important benefit.

A few limitations and future extensions of the model should be noted. We assume the manufacturer to be the Stackelberg leader, but there are practical examples of large retailers such as Walmart that have the channel power to dictate terms to the manufacturer. In that case, the retailer as Stackelberg leader should be studied. Another extension is suggested by the survey of Dutta et al. (1995) of over two thousand co-op advertising plans, which finds that these plans specify not only a participation rate but also an accrual rate. For example, a 50% participation rate capped by a 3% accrual rate means that the manufacturer will compensate 50% of the retailer’s local advertising expenses up to 3% of the purchases made by the retailer. A possible reason for the use of accrual rates is to limit the maximum liability of the manufacturer in the absence of full information about the retailer and market conditions. Thus, incorporation of the accrual rate in our model presents a topic of future research.

It is also important to consider the impact of competition between manufacturers, although tractability can become an issue there. Dutta et al. (1995) find that participation rates are higher when there are fewer competitors, which seems surprising; but they argue that manufacturers with monopoly power can use co-op advertising to resolve coordination problems. Missing in their data is competitive intensity among retailers. Both manufacturer and retailer level competition deserve further attention in analytical and empirical models.

Acknowledgments

This article is dedicated to the memory of Lev Semenovich Pontryagin (1908–1988) and presented as a plenary talk at the International Conference on “Differential Equations and Topology” in commemoration of the Centennial Anniversary of L.S. Pontryagin, Lomonosov Moscow State University, Moscow, Russia, June 17–22, 2008.

This paper won the PICMET (Portland International Center for Management of Engineering and Technology) 2008 Outstanding Student Paper Award and was

The authors also thank Hyo duk Shin and seminar participants at UT Dallas, UT El Paso, UNC Charlotte, National Institute of Industrial Engineers (India), Indian School of Business (India), Indian Institute of Technology Delhi (India), Southern Methodist University, INRIA (France), Bina Nusantara University (Indonesia), Singapore Management University, and one can see that the linear value function satisfies (A6). By inserting it, we get

\[ r x_t + r \beta_t x = m_t x + \frac{\beta_t^2 \rho^2 (1 - x)}{4} - \beta_t \delta x. \]  

Comparing the coefficients of like powers yields

\[ x_t = \beta_t^2 \rho^2 / 4r, \quad (r + \delta) \beta_t = m_t - \beta_t^2 \rho^2 / 4. \]  

(A8)

Because it makes sense to assume \( \beta_t \) to be positive, its solution is

\[ \beta_t = \sqrt{(r + \delta)^2 + \rho^2 m_t - (r + \delta)} \frac{\rho^2/2}{2m_t} \]

\( \sqrt{(r + \delta)^2 + \rho^2 m_t + (r + \delta)} \).  

\( \Box \)

Proof of Proposition 2. The analysis proceeds by backward induction, solving the retailer’s problem first. The retail price \( p \) only occurs in the integrand, which allows us to first derive the best response retail price \( p(x|w, \theta) \) by maximizing the integrand in (2) with respect to \( p \). It is obvious that the result will be independent of \( x \) and \( \theta(x) \). We can thus denote by \( \hat{p}(w) \) the retailer’s best response retail price. It satisfies the FOC \( D(\hat{p}) + (\hat{p} - w)D'(\hat{p}) = 0 \). We can now write the retailer’s margin as \( m_R = (\hat{p}(w) - w)D'(\hat{p}(w)) \).

Next we solve the retailer’s best advertising strategy \( u(x|w, \theta) \). In view of the above, we can rewrite the retailer’s problem (1), (2) as follows:

\[ V^R(x) = \max_{u \geq 0} \left\{ \frac{E}{u} \int_0^\infty e^{-\nu t} (m_R x(t) - (1 - \theta(x(t))) u(t)^2) dt \right\}, \]

\[ dx(t) = \left[ \rho u(t) \sqrt{1 - x(t) - \delta x(t)} \right] dt + \sigma(x(t)) dz(t), \]

\[ x(0) = x \in [0, 1], \quad t \geq 0. \]

(A10)

The Hamilton–Jacobi–Bellman (HJB) equation is

\[ r V^R = \max_{a \geq 0} \left\{ m_R x - (1 - \theta(x)) a^2 + V_x^R \left( \rho u \sqrt{1 - x - \delta x} + \frac{\sigma(x)^2 V_{xx}^R}{2} \right) \right\}. \]

(A11)

The FOC for a maximum is

\[ -2u(1 - \theta(x)) + V^R_x \rho \sqrt{1 - x} = 0, \]

which gives

\[ u(x|w, \theta) = \frac{V^R_x \rho \sqrt{1 - x}}{2(1 - \theta(x))}. \]  

(A13)

Note that we have kept the dependence of this response on \( w \), because \( V^R_x \) would depend on \( w \) in general. Substituting \( u(x|w, \theta) \) for \( u \) into (A11), we get

\[ r V^R = m_R x + \frac{(V^R_x)^2 \rho^2 (1 - x)}{4(1 - \theta(x))} - V^R_x \frac{\sigma(x)^2 V_{xx}^R}{2}. \]

(A14)
Proof of Proposition 3. We substitute \( \dot{p}(w) \) and \( u(x,w,\theta) \) into the manufacturer’s problem (3) and obtain the following problem:

\[
V^M(x) = \max_{0 \leq \theta \leq 1} \left\{ E \int_0^\infty e^{-rt} \left[ (w(t) - c)D(\dot{p}(w(t)))x(t) + \frac{\theta(V^M_x)^2\rho^2(1-x(t))}{4(1-\theta)^2} \right] dt \right\}
\]

\[dx(t) = \left( \frac{VR^2(1-x(t))}{2(1-\theta)} - \delta x(t) \right) dt + \sigma(x(t))dz(t),\]

\[x(0) = x \in [0,1], t \geq 0.\]

Note that the wholesale price does not enter the dynamics of awareness share, and can therefore be solved for separately. The optimal wholesale price \( w^* \) satisfies the FOC \((w^* - c)D'(\dot{p}(w^*)) + D' (\ddot{p}(w^*)) = 0\). Let \( m_M = (w^* - c)D'(\dot{p}(w^*)) \) denote the manufacturer’s optimal margin. We can rewrite the manufacturer’s value function as

\[
V^M(x) = \max_{0 \leq \theta \leq 1} \left\{ E \int_0^\infty e^{-rt} \left[ m_M x(t) - \frac{\theta(V^M_x)^2\rho^2(1-x(t))}{4(1-\theta)^2} \right] dt \right\}.
\]

The Hamilton–Jacobi–Bellman (HJB) equation is

\[
rV^M = \max_{0 \leq \theta \leq 1} \left\{ m_M x - \frac{\theta(V^M_x)^2\rho^2(1-x)}{4(1-\theta)^2} + \frac{V^M_x}{2} \left( \frac{VR^2(1-x)}{2(1-\theta)} - \delta x \right) + \frac{\sigma(x)^2 V^M_{xx}}{2} \right\} \tag{A15}
\]

In maximizing the right-hand side of (A15), whether \( \theta = 0, \theta \in (0,1) \) or \( \theta = 1 \) depends, respectively, on whether the derivative of the right-hand side of (A15), which is

\[
-(V^M_x)^2\rho^2(1-x) \left[ \frac{(1-\theta)^2 + 2\theta(1-\theta)}{(1-\theta)^4} \right] + \frac{V^M_x V^R_x \rho^2(1-x)}{2(1-\theta)^2}
\]

is less than, equal to, or greater than zero. From the expression, \( \theta = 1 \) can be ruled out, and we are left with the result that

\[
\theta(x) = \max \left\{ 0, \frac{2V^M_x - V^R_x}{2V^M_x + V^R_x} \right\} \tag{A16}
\]

Next, we investigate the two cases where the participation rate is zero (a) or positive (b), and determine the condition required for \( \theta = 0 \) to be optimal.

Case (a): No co-op advertising, \( \theta = 0 \).

Inserting \( \theta = 0 \) into (A13) gives

\[
u^*(x) = \frac{1}{2} \rho (VR_x \sqrt{1-x}). \tag{A17}
\]

Inserting \( \theta = 0 \) into (A15) and (A14), we have

\[
rV^M = m_M x + \frac{V^M_x^2 \rho^2(1-x)}{2} - \frac{V^M_x}{2} \frac{\sigma(x)^2 V^M_{xx}}{2}, \tag{A18}
\]

\[
rV^R = m_R x + \frac{(VR_x)^2 \rho^2(1-x)}{4} - \frac{VR_x}{2} \frac{\sigma(x)^2 V^R_{xx}}{2}. \tag{A19}
\]

Let \( V^M = \beta_M x \) and \( V^R = \beta R x \). Then, \( V^M_x = \beta_M \) and \( V^R_x = \beta R). Substituting these into (A18) and (A19) and equating like powers of \( x \), we can express all the unknowns in terms of \( \beta R \), which itself can be explicitly solved. That is,

\[
\beta_R = \frac{2m_R}{\sqrt{(r+\delta)^2 + \rho^2 m_R + (r+\delta)}}, \tag{A20}
\]

\[
\beta_M = \frac{2m_M}{(r+\delta) + \beta R \rho^2}, \tag{A21}
\]

\[
x_M = \frac{\beta_R \rho^2 m_R^2}{4r}, \tag{A22}
\]

Using (A20) in (A17), we can write \( u^*(x) = \sqrt{r/x_R(1-x)} \). Finally, we can derive the required condition from (A16), which is \( 2V^M_x \leq V^R_x \), for no co-op advertising (\( \theta = 0 \)) in the equilibrium. This is given by \( 2\beta_M \leq \beta R \), or

\[
\frac{4m_M}{2(r+\delta) + \frac{2m_R \rho^2}{\sqrt{(r+\delta)^2 + \rho^2 m_R + (r+\delta)}}} \leq \frac{2m_R}{\sqrt{(r+\delta)^2 + \rho^2 m_R + (r+\delta)}}. \tag{A21}
\]

After a few steps of algebra, this yields the required condition

\[
\Delta = \frac{m_M}{\sqrt{(r+\delta)^2 + \rho^2 m_R}} - \frac{m_R}{\sqrt{(r+\delta)^2 + \rho^2 m_R + (r+\delta)}} \leq 0. \tag{A22}
\]

Next, we obtain the solution when \( \Delta > 0 \). Case (b): Co-op advertising, \( \theta > 0 \).

Using the expression from (A16), with \( \theta(x) > 0 \), into (A13) gives

\[
u^*(x) = \frac{1}{4} \rho (VR_x + 2V^M_x) \sqrt{1-x}. \tag{A23}
\]
Inserting (A16) into (A15) and (A14), we have
\[
\begin{align*}
rV^M &= m_M x - \frac{\rho^2 (1-x)}{16} \left[ 4(V^M_x)^2 - (V^R_x)^2 \right] \\
&\quad + \frac{V^M_x \rho^2 (1-x) [2V^M_x + V^R_x]}{4} \\
&\quad - V^M_x \delta x + \frac{\sigma(x)^2 V^M_x}{2}, \\
\end{align*}
\]

(A24)

\[
\begin{align*}
rV^R &= m_R x - \frac{(V^R_x)^2 \rho^2 (1-x)}{4} \times \frac{2V^M_x + V^R_x}{2V^R_x} \\
&\quad - V^R_x \delta x + \frac{\sigma(x)^2 V^R_x}{2}. \\
\end{align*}
\]

(A25)

Let \( V^M = \alpha_M + \beta_M x \) and \( V^R = \alpha_R + \beta_R x \). Then, \( V^M_x = \beta_M \) and \( V^R_x = \beta_R \). Substituting these into (A24) and (A25) and equating like powers of \( x \), we have
\[
\alpha_R = \frac{\beta_R (\beta_R + 2\beta_M) \rho^2}{8r}, \tag{A26}
\]
\[
(\rho + \delta) \beta_R = m_R - \frac{\beta_R (\beta_R + 2\beta_M) \rho^2}{8}, \tag{A27}
\]
\[
\alpha_M = \frac{(\beta_R + 2\beta_M)^2 \rho^2}{16r}, \tag{A28}
\]
\[
(\rho + \delta) \beta_M = m_M - \frac{(\beta_R + 2\beta_M)^2 \rho^2}{16}. \tag{A29}
\]

Using (A28) in (A23), we can write \( u^+(x) = \sqrt{r x_M (1-x)} \).

The four equations (A26)–(A29) determine the solutions for the four unknowns, \( \alpha_R, \beta_R, \alpha_M, \) and \( \beta_M \). From (A27) and (A29), we can obtain
\[
\beta_R^3 + \frac{2m_M \beta_R^2 + 8m_R \beta_R}{r + \delta} - \frac{8m_M^2}{(r + \delta) \rho^2} = 0. \tag{A30}
\]

If we denote \( a_1 = \frac{2m_M}{r + \delta} \), \( a_2 = \frac{8m_R}{r + \delta} \), and \( a_3 = -\frac{8m_M^2}{(r + \delta) \rho^2} \), then \( a_1 > 0 \), \( a_2 > 0 \) and \( a_3 > 0 \). From Descarte’s Rule of Signs, there exists a unique, positive real root. The two remaining roots may be both imaginary or both real and negative. Because this is a cubic equation, a complete solution can be obtained. Using Mathematica or following Spiegel (1968), we can write down the three roots
\[
\begin{align*}
\beta_R(1) &= S + T - \frac{1}{3} a_1, \\
\beta_R(2) &= -\frac{1}{2} (S + T) - \frac{1}{3} a_1 + \frac{\sqrt{3}}{2} i(S - T), \\
\beta_R(3) &= -\frac{1}{2} (S + T) - \frac{1}{3} a_1 - \frac{\sqrt{3}}{2} i(S - T),
\end{align*}
\]

where
\[
\begin{align*}
Q &= \frac{3a_2 - a_1^2}{9}, \\
R &= \frac{9a_1 a_2 - 27a_3 - 2a_1^3}{54}, \\
S &= \sqrt[3]{R + \sqrt{Q^3 + R^2}}, \\
T &= \sqrt[3]{R - \sqrt{Q^3 + R^2}}.
\end{align*}
\]

Next, we identify the positive root in each of three cases.

Case 1 \((Q > 0)\): We have \( S > 0 > T \) and \( Q^3 + R^2 > 0 \). There is one positive root and two imaginary roots. The positive root is \( \beta_R = S + T - \frac{1}{3} a_1 \).

Case 2 \((Q < 0)\) and \( Q^3 + R^2 > 0\): There are three real roots with one positive. The positive root is \( \beta_R = S + T - \frac{1}{3} a_1 \).

Case 3 \((Q < 0)\) and \( Q^3 + R^2 < 0\): \( S \) and \( T \) are both imaginary. We have three real roots with one positive. While subcases can be given to identify the positive root, for our purposes, it is enough to identify it numerically.

Finally, we can conclude that \( 2\beta_M - \beta_R > 0 \) so that \( \theta^* > 0 \), since if this were not the case, then \( \theta^* \) would be zero and we would once again be in Case (a).

**Proof of Proposition 4.** An important property of the solution \( x(t) \) of an Itô stochastic differential equation
\[
\frac{dx}{dt} = a(x(t), t) dt + b(x(t), t) dz(t), x(0) = x
\]

is that it is a Markov process (Cyganowski et al. 2002, §8.3). The transition probability of this Markov process has a density \( p(t; y; 0, x) \) for going from awareness share \( y \) at time \( 0 \) to awareness share \( x \) at time \( t > 0 \), which satisfies the Fokker-Planck equation
\[
\frac{\partial p}{\partial t} + \frac{\partial}{\partial y} (ap) - \frac{1}{2} \frac{\partial^2}{\partial y^2} (b^2 p) = 0, \quad p(t; y; 0, x) = \delta(y - x).
\]

For our problem, we shall first obtain and then attempt to solve the Fokker-Planck equation. The stochastic differential equation can then be expressed as
\[
\begin{align*}
\frac{dx(t)}{dt} &= (A - (A + \delta)x(t) dt \\
&\quad + \sigma \sqrt{x(t)(1 - x(t))} dz(t), \\
x(0) &= x, \quad t \geq 0,
\end{align*}
\]

(A31)

The corresponding Fokker-Planck equation is given by
\[
\frac{\partial p}{\partial t} + \frac{\partial}{\partial y} ((A - (A + \delta)y)p) - \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma^2 y(1 - y)p) = 0.
\]

(A32)
which simplifies to
\[
\frac{\partial p}{\partial t} + \frac{\sigma^2 y(y-1) \partial^2 p}{2} + ((2\sigma^2 - (A + \delta))y + A - \sigma^2) \frac{\partial p}{\partial y} + (\sigma^2 - (A + \delta))p = 0.
\]

(A33)

For its density \( f(y) \), we can set \( \frac{\partial p}{\partial y} = 0 \) and obtain the second-order ordinary differential equation
\[
\frac{\sigma^2 y(y-1) d^2 f}{dy^2} + ((2\sigma^2 - (A + \delta))y + A - \sigma^2) \frac{df}{dy} + (\sigma^2 - (A + \delta))f = 0.
\]

(A34)

A slight rearrangement makes it clear that this is the hypergeometric equation
\[
y(y-1) \frac{d^2 f}{dy^2} + \left( \frac{4 - 2(A + \delta)}{\sigma^2} y - \frac{2A}{\sigma^2} \right) \frac{df}{dy} + \left( 2 - \frac{2(A + \delta)}{\sigma^2} \right) f = 0.
\]

(A35)

The solution is the Beta density
\[
f(y) = \frac{\Gamma \left( \frac{2A}{\sigma^2} + \frac{2\delta}{\sigma^2} \right)}{\Gamma \left( \frac{2A}{\sigma^2} \right) \Gamma \left( \frac{2\delta}{\sigma^2} \right)} y^{\frac{2A}{\sigma^2} - 1} (1 - y)^{\frac{2\delta}{\sigma^2} - 1}.
\]

(A36)

The long-run equilibrium awareness shares are obtained from the mean of this density function, which is \( \bar{x}_i/(A + \delta) \). If we denote this by \( \bar{x}_I \) for the integrated channel and \( \bar{x}_D \) for the decentralized supply chain, we have
\[
\bar{x}_I = \frac{\rho \sqrt{\tau_I}}{\rho \sqrt{\tau_I} + \delta}, \quad \bar{x}_D = \frac{\rho \sqrt{\tau_M}}{\rho \sqrt{\tau_M} + \delta}.
\]

(A37)

Proof of Proposition 5. For any given \( w(x) \) and \( \theta \), the retailer’s retail price response \( p(x|w, \theta) = \hat{p}(w, \theta) \) satisfies the FOC
\[
[(1 - \theta)\hat{p}(w, \theta) - w]D'(\hat{p}(w, \theta)) + D(\hat{p}(w, \theta)) = 0.
\]

(A38)

Its advertising response is
\[
u(x|w, \theta) = \frac{V^R \rho \sqrt{(1 - x)}}{2(1 - \theta)}.
\]

(A39)

Taking (A38) into consideration, the manufacturer’s decision \( w \) satisfies the FOC
\[
\frac{\partial p(w, \theta) - c}{D'(\hat{p}(w, \theta))} + D(\hat{p}(w, \theta)) \frac{d\hat{p}(w, \theta)}{dw} - \{(1 - \theta)\hat{p}(w, \theta) - w\}D'(\hat{p}(w, \theta)) + (1 - \theta)D(\hat{p}(w, \theta))
\]
\[\times \frac{d\hat{p}(w, \theta)}{dw} = 0.
\]

(A40)

Using (A38) and simplifying, we get
\[
\frac{\partial p(w, \theta) - c}{D'(\hat{p}(w, \theta))} + D(\hat{p}(w, \theta)) = 0. \quad (A41)
\]

Comparing this condition to the corresponding one for the integrated channel and using (A38), we can conclude that \( w^* \) satisfies
\[
(p^* - c) = - \frac{D(p^*)}{D'(p^*)} = \frac{(1 - \theta)p^* - w^*}{1 - \theta}, \quad (A42)
\]

which gives the prices
\[
w^*(x) = (1 - \theta)c, \quad p^*(x) = \hat{p}(w^*(x), \theta) = P^*, \quad (A43)
\]

which are both constants independent of \( x \). With these prices, the retailer’s problem reduces to the problem of the integrated channel except for the constant multiplier \( (1 - \theta) \). Thus, from (A39), we have
\[
u^*(x) = u^*(x|w^*, \theta) = \frac{V^I \rho \sqrt{(1 - x)}}{2(1 - \theta)} = u^*(x). \quad (A44)
\]

Finally, it is easy to see that \( V^R = (1 - \theta)V^I \) and \( V^M = \theta V^I \). □

References


