Pricing and two-tier advertising with one manufacturer and one retailer

Joseph G. Szmerekovsky a,*, Jiang Zhang b

a Department of Management, Marketing, and Finance, College of Business, North Dakota State University, Fargo, ND 58105, United States
b Department of Management, Marketing, and Decision Sciences, School of Business, Adelphi University, Garden City, NY 11530, United States

Received 3 February 2006; accepted 3 October 2007
Available online 10 October 2007

Abstract
This paper considers the pricing decisions and two-tier advertising levels between one manufacturer and one retailer where customer demand depends on the retail price and advertisement by a manufacturer and a retailer. We solve a Stackelberg game with the manufacturer as the leader and the retailer as the follower. With price sensitive customer demand and a linear wholesale contract, we obtain the optimal decisions by the manufacturer and the optimal responses by the retailer. Our results show that cost sharing of local advertising does not work well, it is better for the manufacturer to advertise nationally and offer the retailer a lower wholesale price.

Keywords: Two-tier advertising; Pricing decisions; Stackelberg game; Manufacturer–retailer supply chain

1. Introduction
Manufacturers and retailers have been using advertisements and price reductions to lure customers and boost sales. Recent examples in the car industry are GM’s and Ford’s employee’s discount program for every buyer. It is easy to see that for a particular product, the promotion efforts may come from the manufacturer alone (e.g. Ford launches a national ad campaign), the retailer alone (e.g. a local retailer advertises its grand opening), or both (e.g. Ford and local retailers advertise models for the new year). In addition to their own promotions, manufacturers may pay part or all of the expense of the retailer’s sales promotions. Two-tier advertising is a typical example of this type of joint promotion effort between a manufacturer and a retailer in which the retailer initiates and runs a local advertisement and the manufacturer pays part, if not all, of the cost (Huang and Li, 2001; Huang et al., 2002; Li et al., 2002; Xie and Ai, 2006).

The balance of power between manufacturers and retailers has been the focus of many research studies (Porter, 1974; Chintagunta and Jain, 1992; Kumar, 1996; Kadiyali et al., 2000; Geylani et al., 2007). For example, a powerful manufacturer, such as Proctor & Gamble, will be able to demand certain shelf spaces in a retailer’s stores. In contrast, a powerful retailer, such as Wal-Mart, will be able to squeeze manufacturers’ profit margins and demand extra requirements for their products (e.g. requiring manufacturers to attach RFID tags to their products and to manage their own inventories in Wal-Mart stores). Although powerful manufacturers and retailers can often use their strength to gain concessions from their counterparts and leave them in vulnerable situations with little profits, it is unclear if this type of forceful approach is the most effective way to manage manufacturer–retailer relationships. For example, cooperation between the two parties may prove more effective at producing ongoing benefits for both parties. In this paper, we consider an environment where the manufacturer holds the dominant position in the relationship and discuss the resulting Stackelberg equilibrium regarding advertising by both parties and its implications.
In particular, our paper addresses the following two general questions:

1. In a decentralized system, how effective is two-tier advertising at improving the two parties’ profits?
2. When demand is affected by price, local advertising, and brand advertising, how does the relationship among these parameters influence pricing, advertising, demand and profits for the parties in the system?

To address these questions, this paper considers a single product two-tier advertising model consisting of one manufacturer and one retailer. The demand for the product experienced at the retail outlet is assumed to be a function of the retailer’s local advertising level, the manufacturer’s national brand name investment level, and the retail price charged to the consumers. We assume that the demand function is nonlinear and is known to both parties. Depending on their power in the system, the manufacturer or the retailer may decide on the wholesale price (manufacturer), advertising efforts, and reimbursement rate for local advertising at the retailer outlet. One important focus of our research is analyzing the effects of pricing decisions on a system with two-tier advertising. We study the effect of a linear contract on a decentralized system where the manufacturer is the leader and the retailer is the follower. We then discuss the properties of this non-coordinated system and their managerial implications.

Our results provide several important insights on how two-tier advertising, pricing decisions, and different contracts improve system performance. Our results show that subsidizing local advertising is not an effective contract approach for improving system profits. Rather, what is needed is an increased amount of advertising at the national level by the manufacturer and a lower retail price at the outlet. Interestingly enough, this type of coordinated effort is easy to identify in practice, one example being the aforementioned employee’s discount program in the auto industry.

The remainder of the paper is organized as follows. In Section 2, we review the related literature in sales promotion. Section 3 formulates the basic model. Section 4 develops optimal policies for the manufacturer and the retailer when the manufacturer is the leader in a Stackelberg game. Numerical test cases are studied in Section 5 and a summary of results is provided in Section 6.

2. Literature review

This research is related to the literature rooted in both advertising and pricing techniques. Our paper is particularly inspired by Huang et al. (2002) in which two-tier advertising is studied. They show that the trend of power shifting from manufacturers to retailers gives the latter bargaining power to protect or increase their profits. By comparing results of a leader–follower model with those of a coordinated model, they propose a bargaining technique to generate an equilibrium between members of a supply chain. In that paper, the authors assume the marginal profits are constants and the sale demand is only affected by brand advertising and local advertising. However, “Consumer goods marketers have long recognized a correlation between advertising and pricing strategies” (Farris and Reibstein, 1979) and our model takes price as a driving force for the demand. Yue et al. (2006) extend Huang et al.’s (2002) model and include a manufacturer discount to address the price influence on the demand, but this is done without addressing the issue of constant profit margins.

Two-tier advertising is an interactive scheme in a manufacturer–retailer system where the retailer initiates and runs a local advertisement and the manufacturer pays part of the costs. For example, Small World Toys offers a 2% advertising allowance on total net purchases (Small World Toys, 2007) and Arett Sales, a premier distributor of lawn and garden products in the Northeastern United States, has advertising programs for many products (Arett Sales, 2007). We distinguish between the purpose of national advertising and the purpose of local advertising. A manufacturer uses national advertising to influence potential customers and to help build brand recognition and preference. Whereas a retailer’s local advertising is to bring potential customers to the stage of buying. The sharing of local advertising expenses by the manufacturer motivates the retailer to advertise locally and stimulate sales. A complete discussion about this kind of advertising scheme can be found in Crimmins (1985).

Our research is also related to wholesale pricing models in which the retailer can vary the retail price. The contracts in these circumstances are used to reduce double marginalization, which results in a high retail price and a low level of demand. To induce the retailer to sell a higher volume at a lower cost the manufacturer offers a lower wholesale price in the form of a two-part tariff or all-units discount. In a two-part tariff the retailer pays a fixed fee to the manufacturer in order to be able to purchase product at a reduced price and in an all-units discount the retailer must purchase a minimum quantity of product to receive the reduced price. For examples of recent developments in this area see Borger (2000), Kolay et al. (2004), and Raju and Zhang (2005). The work most related to ours is that of Raju and Zhang (2005) as it is the only one which considers the effect of the promotion activity of the retailer. Our model differs from theirs in two ways. First, they consider an environment where there are multiple retailers and a single retailer dominates the market. In contrast we consider the case of only a single retailer. Second, they consider promotional effort to occur only at the retailer and to be discrete: either the retailer engages in promotion at a fixed cost for a fixed benefit or the retailer engages in no pro-
motional activities. In contrast, our model allows for a continuous level of promotion effort to be selected at both the manufacturer and the retailer. This better represents many manufacturers who engage in national promotion of their brand names and accounts for the change in the value of advertising when the profit margin changes. We believe increased flexibility in the level of advertising is of sufficient interest even with only a single retailer. Finally, the principal-agent literature considers the problem of a firm owner who contracts both manufacturing managers and marketing managers to produce and sell his products (Porteus and Whang, 1991). However, these models emphasize the role of a third party, the firm owner, who coordinates the two types of agents whereas our model considers the interaction of the manufacturer and retailer assuming no joint ownership.

Many empirical works (Dhalla, 1978; Jones, 1990; Abraham and Lodish, 1990; Srinivasan et al., 2002, and Raghubir et al., 2004) have studied the relationships between sales promotions, advertising, and pricing, as well as how these factors influence manufacturers, retailers, and consumers. These studies promote the need for our analytical developments by establishing the practical importance of policies which determine these factors.

In the next section, we formalize our own model which allows for varying profit margins, promotion levels and two-tier advertising.

3. Model formulation and notation

In this section, we formalize our model of demand dynamics. Demand is assumed to be a function of the retailer’s local promotion in dollars ($a$), the manufacturer’s brand promotion in dollars ($q$), and the retail sale price paid by the end customer in dollars ($P$). We define the following demand function,

$$D(a, q, P) = (x - \gamma a - \delta q)P^{-\epsilon},$$

where $x$, $\gamma$, and $\delta$ are positive constants and where $\epsilon$ is the price-elasticity index (we assume $\epsilon > 1$ in order to reflect price elasticity).

As explained in Yue et al. (2006), $\gamma$ and $\delta$ are the quasi-advertising elasticity and quasi-investment elasticity, respectively, and $D(a, q, P)$ is a non-decreasing function with respect to $a$ and $q$. When either or both local advertisement and brand name investment tend to infinity, the demand function $D$ takes the well-known multiplicative form and is price sensitive with constant price elasticity $\epsilon$. Here $x$ is the market cap and $\beta$ is the impact of brand name investment and local advertisement on market demand. The quantity $xP^{-\epsilon}$ is referred to as the price adjusted market cap as it equates to the maximum level of demand for any fixed retail price. This quantity will be important for interpreting results later in the paper.

We further assume that all customer demand for the retailer will be satisfied. Practically speaking, the deterministic demand in our model means that, after viewing the contract provided by the manufacturer, the retailer uses $D(a, q, P)$ to choose the quantity of product to order and then takes appropriate actions to satisfy all the demand precisely. Further, since “Coop advertising is intended to generate short-term sales” (Huang et al., 2002), price discount deals are always temporary (Blatterberg et al., 1981; DeNeckere et al., 1996), and the retailer engages in promotion efforts in order to capture a larger share of a stable market, not to create a new market, we study only a one period static model. With a known expected market cap for the period, the assumption that the manufacturer will have the capacity to produce enough to satisfy all of the demand is mild. The assumption that all demand is satisfied is also valid for apparel products and make to order products if temporary backorders are allowed.

Given the level of demand, to determine the profits of the retailer and the manufacturer we make use of the wholesale price in dollars per unit, the manufacturing cost in dollars per unit, and the level of manufacturer reimbursement for local promotion at the retailer. We denote the wholesale price and the manufacturing cost as $W$ and $C$, respectively. Further, for every dollar spent by the retailer for product promotion, we assume that the manufacturer will reimburse him a percentage amount, $t$, where $0 \leq t \leq 1$. This means that if the retailer spends $a$ dollars in promotion, he will get $ta$ dollars back from the manufacturer as an incentive for his promoting efforts. This subsidy could be received in the form of cooperative advertising, channel rebates, dealer hold backs, or other retailer benefits. In general, it is possible for information asymmetry to exist between the manufacturer and the retailer with regard to the amount of money spent on local advertising. In order for the manufacturer to establish the appropriate amount $ta$ for the reimbursement, the retailer must first show proof of the advertising expense $a$. This can be accomplished through a system of receipt checking similar to that used in administering cost plus contracts in project management (see Herroelen et al., 1997). Though such systems can be cumbersome to implement they can also protect the interests of both the manufacturer and the retailer. We assume that such a system is implemented and avoid the concern of information asymmetry in the current work. However, alternative contract designs which address this issue remain an interesting direction for research.

Assuming that the retailer controls the values of the retail price $P$ and the local advertising $a$ and that the manufacturer controls the values of the wholesale price $W$, the brand advertising $q$, and the reimbursement rate $t$, we can now write the retailer’s expected profit $R(P, a)$ and the manufacturer’s expected profit $M(W, q, t)$ as follows:

$$R(P, a) = \int_a^{a_*} \left( xP^{-\epsilon} - \gamma a - \delta q \right)P^{-\epsilon} dP,$$

and

$$M(W, q, t) = \int_q^{q_*} \left( W - C + \beta q \right)q^{-\epsilon} dt.$$
4.1. Optimal policy for the retailer

Stackelberg game. We first address the retailer’s problem.

Proof. Assume that the retailer’s optimal profit is greater than 0. Then, since

\[ R(P,a) = (P - W)(x - \beta a^{-\gamma} q^{-\delta})P^{-\epsilon} - (1 - t)a, \]

\[ M(W,q,t) = (W - C)(x - \beta a^{-\gamma} q^{-\delta})P^{-\epsilon} - q - ta. \]

Using the profits identified above, we next determine the optimal behavior of the retailer and the optimal behavior of the manufacturer assuming that the manufacturer holds the dominant role in the model.

4. Stackelberg equilibrium

This section provides optimal policies for the retailer and the manufacturer when the manufacturer is the leader in a Stackelberg game. We first address the retailer’s problem.

4.1. Optimal policy for the retailer

Consider the problem of finding the Stackelberg equilibrium when in the first stage the manufacturer leads and in the second stage the retailer follows. To that end consider the retailer’s problem. That is, given the manufacturer’s choice of \( q, W, and t \), the retailer selects the amount of local promotion \( a \) and the retail price \( P \) by solving the following optimization problem.

\[
\max_{(P,a)} R(P,a) = \max_{(P,a)} (P - W)(x - \beta a^{-\gamma} q^{-\delta})P^{-\epsilon} - (1 - t)a,
\]

subject to:

\[ P \geq W, \quad a \geq 0. \]

The following theorem establishes the retailer’s optimal policy when the manufacturer is the leader.

**Theorem 1.** If the retailer’s optimal profit is greater than 0, the optimal solution to the retailer’s problem is given by

\[ P^* = \frac{eW}{e - 1}, \]

\[ a^* = \left[ \beta\gamma q^{-\delta} e^{-\epsilon} W / (e - 1) \right]^{1/(\gamma + 1)}. \]

**Proof.** Assume that the retailer’s optimal profit is greater than 0. Then, since \( P \geq W \) it must be the case that \( x - \beta a^{-\gamma} q^{-\delta} > 0 \) for the optimal solution. Consider now the choice of \( P^* \). Taking the partial derivative of \( R(P,a) \) with respect to \( P \), we have

\[
\frac{\partial R(P,a)}{\partial P} = [P^{-\epsilon} - (P - W)eP^{-(\epsilon + 1)}](x - \beta a^{-\gamma} q^{-\delta}) = P^{-\epsilon - 1}[(1 - e)P + eW](x - \beta a^{-\gamma} q^{-\delta}).
\]

Since \( P^{-\epsilon - 1} > 0 \) and \( x - \beta a^{-\gamma} q^{-\delta} > 0 \), the sign of the derivative is determined by \( (1 - e)P + eW \). It follows that \( R(P,a^*) \) is increasing in \( P \) up to \( \frac{eW}{e - 1} \) and decreasing in \( P \) beyond \( \frac{eW}{e - 1} \). Hence, \( P^* = \frac{eW}{e - 1} \).

To derive the optimal local advertising level \( a^* \), we consider the first and second partial derivatives with respect to \( a \) of \( R(P,a) \) at \( P^* = \frac{eW}{e - 1} \).

\[
\frac{\partial^2 R(eW/e-1,a)}{\partial a^2} = a^{-\gamma - 1}(e - 1)^{\gamma + 1} e^{-\epsilon} q^{-\delta} W^{1-\epsilon} \beta \gamma + t - 1,
\]

\[
\frac{\partial^2 R(eW/e-1,a)}{\partial a^2} = -a^{-\gamma - 2}(e - 1)^{\gamma + 1} e^{-\epsilon} q^{-\delta} W^{1-\epsilon} \beta \gamma (\gamma + 1).
\]

With \( \epsilon > 1 \), it is straightforward to observe that \( R(P^*,a) \) is concave in \( a \) and, hence, \( a^* \) is obtained by solving \( \frac{\partial R(P,a)}{\partial a} = 0 \). This completes the proof. \( \square \)

From Theorem 1, the following observations can be made:

1. The optimal retail price depends only on the wholesale price \( W \) and the price elasticity \( e \).
2. The amount of advertising at the retail level is increasing in the linear demand coefficient \( \beta \) and in the reimbursement rate \( t \).
3. The amount of advertising at the retail level is decreasing in \( e, q, \delta, \) and \( W \).
4. The amount of advertising at the retail level is increasing in the quasi-advertising elasticity \( \gamma \) for small values of \( \gamma \) and decreasing in the quasi-advertising elasticity \( \gamma \) for large values of \( \gamma \).
The first observation follows from the multiplicative form of the demand function and tells us that pricing at the retail level can and should be made independent of the chosen level of advertising. For example, the retailer should not attempt to recuperate advertising expenses by increasing the retail price. The second observation is fairly straightforward in that increasing \( \beta \) increases the efficacy of advertising and increasing \( t \) decreases the cost of advertising. The third observation is more interesting. When the price elasticity is high, the retailer will not attempt to compensate by increased advertising but will actually decrease advertising and rely on the reduced \( P^* \) to keep demand high. Further, the more the manufacturer advertises and the greater the quasi-investment elasticity, the less effort the retailer will put into advertising. In this way the retailer uses advertising at the national level as a substitute for his own advertising efforts. Finally, we see that more expensive products will actually receive less advertising expenditures. Just as high retail prices should not be used to make up for advertising expenditures, high advertising expenditures should not be used to compensate for high wholesale prices. The last observation is not surprising. When the quasi-advertising elasticity is small, it is more productive to increase advertising as the quasi-advertising elasticity increases. This is because advertising becomes more effective as the elasticity increases. However, once the quasi-advertising elasticity is large enough, the majority of the benefit of advertising can be obtained with a small level of advertising. Hence, increasing the quasi-advertising elasticity actually results in the optimal level of advertising decreasing.

**Theorem 2.** Given a fixed wholesale price \( W_o \geq C \) the values \( q_0 \) and \( t_0 \) which maximize the manufacturer’s profit are:

\[
t_0 = \begin{cases} 
0, & \text{if } \gamma \geq (e - 2) \text{ or } W_0 < \left( 1 + \frac{1}{c - \gamma} \right) C, \\
\frac{w_o}{(w_o - (e - 1)(w_o - c))}, & \text{otherwise.}
\end{cases}
\]

\[
q_0 = \left[ (e - 1)^{-1} e^{- \gamma} (1 - t_0)^{-1} W_0^{-\gamma} \beta^{-\gamma} (1 + \gamma)^{1 + \gamma} \delta^{-(1 + \gamma)} \right]^{\gamma - 1} \left[ (e - 1)(1 - t_0)(W_0 - C) + t_0 W_0 \delta \right]^{1 - \gamma}.
\]

**Proof.** The proof of **Theorem 2** appears in the Appendix. \( \square \)

From **Theorem 2** the following observations can be made:

1. When the effectiveness of advertising at the retail level is high relative to price elasticity (\( \gamma \geq e - 2 \)) the manufacturer should not subsidize advertising.
2. When the manufacturer’s profit margin is small (\( W_o < (1 + \frac{1}{c - \gamma}) C \)) the manufacturer should not subsidize advertising.
3. The optimal level of advertising expenditure for the manufacturer is increasing in the potential impact of all advertising \( \beta \) and decreasing in the quasi-investment elasticity \( \delta \).

The first observation is intuitive as, when \( \gamma \) is large relative to \( e \), one would expect the retailer to adopt an approach that emphasizes high levels of advertising and high profit margins. Under such circumstances, the retailer is willing to advertise...
work to be done through advertising and effectiveness of his advertising (\(\text{Theorem 1}\)). Hence, as discussed regarding that the potential impact of all advertising (\(\beta\)) increases, the manufacturer will spend more on advertising; but, as the effectiveness of his advertising (\(\delta\)) increases, the manufacturer will spend less on advertising. Essentially, \(\beta\) represents more work to be done through advertising and \(\delta\) represents doing more with less advertising.

Using \(\text{Theorem 2}\), we can identify the optimal values of brand advertising \(g\) and reimbursement rate \(t\) for any \(W = W_0\). Hence, to find the optimal decision for the manufacturer, we need only substitute \((5)\) and \((6)\) into \((3)\) to obtain the manufacturer’s profit as a function of \(t\) and \(W\).

\[
f(W) = (W - C) \left( \frac{e^W}{e - 1} \right)^{-\delta} \\
\quad \times \left[ \alpha - (\beta\gamma^{-\gamma}q_0^\gamma\beta e^{\gamma}W^{(e-\gamma)}(e - 1)^{\gamma(1-\gamma)}(1 - t_0)^{1/(1+\gamma)} \right] - q_0 \\
\quad - t_0[\beta\gamma^{-\gamma}q_0^\gamma\beta e^{\gamma}W^{1-\gamma}(e - 1)^{\gamma-1}(1 - t_0)^{-1/(1+\gamma)}].
\]

In order to identify the value \(W^*\) which maximizes \(f(W)\), we next prove the following theorem which will be used to prove \(\text{Theorem 4}\).

**Theorem 3.** The optimal advertisement subsidy \(t^*\) is 0.

**Proof.** Notice first that if \(\gamma \geq e - 2\) then the result follows from \(\text{Theorem 2}\). Otherwise, \(\gamma < e - 2\) and \(\text{Theorem 2}\) still gives \(t^* = 0\) if \(W^* < \left(1 + \frac{1+\gamma}{e-\gamma}\right)C\). Hence, if we can show that \(W^*\) satisfies this condition then we will have completed the proof.

When \(W \geq \left(1 + \frac{1+\gamma}{e-\gamma}\right)C\), as given by \(\text{Theorem 2}\), substituting the optimal values for \(t_0\) and \(q_0\) into \(f(W)\) results in

\[
f(W) = g W^{\gamma-\gamma}(W - C) + k W^{\gamma-\gamma} \left( (1 - e)C + (e - \gamma - 1)W \right)^{\gamma-\gamma},
\]

where \(g\) and \(k\) are the following constants determined by \(x\), \(\beta\), \(C\), \(e\), \(\gamma\) and \(\delta\).

\[
g = a \left( e - 1 \right)^{\gamma} e^{-\gamma}, \\
k = - \left( e - 1 \right)^{\gamma} e^{-\gamma} \beta^{\gamma-\gamma} \delta^{\gamma-\gamma} \delta^{-\gamma}(1 + \gamma + \delta).
\]

Taking the derivative of \(f(W)\), we obtain:

\[
f'(W) = g W^{\gamma-\gamma}(eC - (e - 1)W) + k \left( \frac{e - 1}{1 + \gamma + \delta} \right) W^{(e-\gamma)\gamma} \times (1 - e)C + (e - \gamma - 1)W^{\gamma-\gamma}(eC - (e - \gamma - 1)W).
\]

To identify the behavior of \(f(W)\), we observe that when \(W > \left(1 + \frac{1+\gamma}{e-\gamma}\right)C\), \(W^*\) becomes positive if and only if

\[
-\frac{(eC - (e - 1)W)}{(eC - (e - \gamma - 1)W)} < \frac{k(e - 1)}{g(1 + \gamma + \delta)} W^{\gamma-\gamma}(1 - e)C + (e - \gamma - 1)W^{\gamma-\gamma}.
\]

We next compute the derivative of the left-hand and right-hand sides of \((8)\). They are as follows:

\[
\text{(LHS)} = -e^{\gamma}C \\
\text{(RHS)} = \frac{k(e - 1)^2(\gamma + \delta) W^{(e-\gamma)\gamma} \times (1 - e)C + (e - \gamma - 1)W^{\gamma-\gamma}(eC - (e - \gamma - 1)W)}{g(1 + \gamma + \delta)^2(eC - (e - \gamma - 1)W)^2}
\]

It follows that when \(W > \left(1 + \frac{1+\gamma}{e-\gamma}\right)C\), the left-hand side is decreasing and, since \(k < 0\) and \((eC - (e - \gamma - 1)W) < 0\), the right-hand side is increasing. Hence, \(f(W)\) is either increasing first then decreasing or decreasing first then increasing on \(\left(\left(1 + \frac{1+\gamma}{e-\gamma}\right)C, \infty\right]\). Using \((7)\), \(f(W)\) still approaches 0 as \(W\) approaches \(\infty\) and \(f(W)\) will not achieve its maximum on \(\left(\left(1 + \frac{1+\gamma}{e-\gamma}\right)C, \infty\right]\) for any of the three cases. This means that when \(\gamma < e - 2\), it is the case that \(W^* < \left(1 + \frac{1+\gamma}{e-\gamma}\right)C\). So, \(t^* = 0\).  \(\Box\)
From Theorem 3, we have the following observation:

- The manufacturer’s optimal solution always results in no percentage reimbursement for the retailer’s local advertising expenditures (i.e., \( t_0 = 0 \) for \( W^* \)).

Theorem 2 has already provided two cases when this occurs: (1) the local quasi-advertising elasticity \( \gamma \) is large relative to the retail price elasticity \( e \) and (2) the wholesale price \( W \) is sufficiently small. Hence, Theorem 3 handles the remaining case where advertising at the retailer is not overly effective relative to price elasticity and the wholesale price is large. Put another way, Theorem 3 provides a “NO” answer to the following question: “When the influence of local advertising is not large relative to the influence of price, will it ever be optimal for the manufacturer to charge a high wholesale price and balance it by subsidizing local advertising?” Stated in this way, it is not surprising that we are able to obtain the “NO” response because the large wholesale price will surely result to a large retail price. This will result to a small price adjusted market cap which will limit not only the potential market size but also the potential impact of advertising.

Further, Theorem 1 predicts that the large wholesale price induces lower local advertising. Combining the limited potential adjustment market cap which will limit not only the potential market size but also the potential impact of advertising. Further, Theorem 1 predicts that the large wholesale price induces lower local advertising. Combining the limited potential for advertising with the low level of local advertising results in a limited potential for any effort to recover market share through subsidies.

It follows that if the manufacturer wants to increase demand at the retailer, he should either increase brand promotion or lower his wholesale price; but, not offer to share local promotion costs. Other models (Huang et al., 2002; Yue et al., 2006) which do not allow for profit margins to vary (i.e., \( W \) and \( P \) are fixed) do not exhibit this property. This suggests that the manufacturer’s ability to influence the retail price by manipulating the wholesale price is more advantageous than his ability to subsidize local advertising.

In order to assist in finding \( W^* \), the next theorem identifies the behavior of \( f(W) \) given its derivative \( f'(W) \).

**Theorem 4.** (a) If \( f'((1 + \frac{\delta}{(e-1)(\gamma + \delta)})C) \leq 0 \) then either \( W = C \) maximizes \( f(W) \) over \([C, \infty)\) or there exist numbers \( W_1 \) and \( W_2 \) such that \( C < W_1 < W_2 \leq (1 + \frac{\delta}{(e-1)(\gamma + \delta)})C \), \( f'(W) \leq 0 \) over \([C, W_1]\), \( f'(W) > 0 \) over \([W_1, W_2]\), \( f'(W) \leq 0 \) over \([W_2, (1 + \frac{\delta}{(e-1)(\gamma + \delta)})C]\), and \( W = W_2 \) maximizes \( f(W) \) over \([C, \infty)\).

(b) Otherwise, \( f'((1 + \frac{\delta}{(e-1)(\gamma + \delta)})C) > 0 \) and \( f(W) \) has exactly one mode over \([(1 + \frac{\delta}{(e-1)(\gamma + \delta)})C, (1 + \frac{1+\delta}{\epsilon}(\gamma + \delta))C]\) at \( W = W_3 \) and either \( W = W_3 \) or \( W = C \) maximizes \( f(W) \) over \([C, \infty)\).

**Proof.** The proof of Theorem 4 appears in the Appendix. □

Whenever \( f(W) \) achieves one and only one local maximum on an interval that local maximum can be found using any of a number of line search techniques. Hence, using Theorem 4, it is easy to obtain the manufacturer’s optimal wholesale price \( W^* \). A procedure for doing this is presented below.

1. Compute \( f'((1 + \frac{\delta}{(e-1)(\gamma + \delta)})C) \). If \( f'((1 + \frac{\delta}{(e-1)(\gamma + \delta)})C) \leq 0 \), then go to 2. Otherwise, go to 4.
2. Search \([C, (1 + \frac{\delta}{(e-1)(\gamma + \delta)})C]\) for a point \( W' \) where \( f'(W') > 0 \). If no such \( W' \) exists, set \( W^* = C \) and go to 5. Otherwise, go to 3.
3. Use a line search to find \( W_2 \) which maximizes \( f(W) \) on the interval \([W', (1 + \frac{\delta}{(e-1)(\gamma + \delta)})C]\). If \( f(W_2) > 0 \), then set \( W^* = W_2 \). Otherwise, set \( W^* = C \). Go to 5.
4. Search \([(1 + \frac{1+\delta}{\epsilon}(\gamma + \delta))C, (1 + \frac{1+\delta}{\epsilon})C]\) for the mode \( W_3 \). If \( f(W_3) > 0 \), set \( W^* = W_3 \). Otherwise, set \( W^* = C \). Go to 5.
5. Return \( W^* \) which maximizes \( f(W) \).

Despite the lack of a closed form solution for \( W^* \), Theorem 4 still provides a compelling insight into the manufacturer’s optimal behavior.

- The optimal wholesale price will never exceed \((1 + \frac{1+\delta}{\epsilon})C\).

This suggests that when the retail price elasticity \( e \) is large relative to the local quasi-advertising elasticity \( \gamma \), the optimal wholesale price \( W^* \) will be close to the manufacturing cost \( C \). In other words, when price elasticity has a more pronounced effect on demand than promotion, the wholesale price should be kept low so that the retailer chooses a lower retail price. This results in a lower profit margin, but a higher volume of demand. Hence, a manufacturer should be careful not to over emphasize the profit margin at the expense of market volume.

In light of the absence of a closed form solution for \( W^* \), to better understand the relationship between the retailer’s and manufacturer’s policies, we provide numerical results in Section 5.
5. Insights with a set of example problems

Because no closed form solution could be obtained for the optimal value of the wholesale price in the Stackelberg equilibrium, to better understand the structure of the optimal solution we examine the performance over a set of test cases. Specifically, we consider cases with \( a = 1000, b = 500, \) and \( C = 1 \). Note that the cost can be set to 1 without loss of generality and the values for \( a \) and \( \beta \) can be interpreted as follows: when \( a = q = 1 \), exactly one-half of the total possible market for a fixed retail price will be serviced. To select a range of values for the price elasticity \( \epsilon \), we note that practical estimates of price elasticity range from 0.7 to 2.8 (see Schiller, 1997, p. 113; Frank and Bernanke, 2007, p. 101). Since our results apply only when \( \epsilon > 1 \), we test values of \( \epsilon = 1.1, 2, \) and 3. For the range of advertising elasticities, practical estimates range from 0.0% to 0.8% where the percentage is a linear increase in demand per 1% increase in advertising expenditures (see PNG, 2001, p. 75). This notion of advertising elasticity differs from ours in that it provides only a linear approximation of the effect of advertising on demand given a base advertising level. To equate 0.5% price elasticity to our model, we use \( a = q = 1 \) as the base levels of advertising which equates to demand of 500. Then, if we increase advertising 100% at the manufacturer (\( q = 2 \)), we should experience a 50% increase in demand to 750. This requires that \( 500 \times 2^\delta = 250 \) or that \( \delta = 1.0 \). Hence, for our range of parameters, an advertising elasticity of 1.0 should be typical and we test values of \( \gamma = 0.5, 1.0, \) and 1.5 and \( \delta = 0.5, 1.0, \) and 1.5.

For each case, we find the Stackelberg equilibrium when the manufacturer is the leader. Results for the test cases appear in Tables 1–3 and Figs. 1–5. Table 1 compares the optimal wholesale price \( W^o \) to a heuristic solution of \( \frac{e^\epsilon}{e^\epsilon - 1} C \) by reporting \( \left( \frac{W^h}{W^o} \right) - W^o \) for each case. Tables 2 and 3 show the effectiveness of this heuristic for the manufacturer and the retailer by reporting \( \left( \frac{M(W^h)}{M(W^o)} \right) - M(W^o) \) and \( \left( \frac{R(W^h)}{R(W^o)} \right) - R(W^o) \), respectively. Here \( M(W) \) and \( R(W) \) represent the manufacturer’s and the retailer’s profits when the wholesale price is \( W \). Figs. 1 and 2 plot the optimal levels of advertising for the manufacturer and the retailer, Fig. 3 plots the optimal level of demand and Figs. 4 and 5 plot the optimal profits for the manufacturer and the retailer, respectively. In all cases with \( \gamma = 1.0, \) the values follow the trends represented by the values with \( \gamma = 0.5 \) and \( \gamma = 1.5 \). Therefore, for brevity, we do not plot the values in Figs. 1–5 for cases with \( \gamma = 1.0 \).

From the tables and figures, we state the following results:

1. The optimal wholesale price is approximately \( \frac{e^\epsilon}{e^\epsilon - 1} C \) with the gap becoming smaller as the quasi-advertising elasticity \( \gamma \) and the quasi-investment elasticity \( \delta \) increase. By comparison with Eq. (5), the retail price is approximately \( \left( \frac{1}{e^\epsilon} \right) C \) (Table 1).

<table>
<thead>
<tr>
<th>( \gamma )</th>
<th>( \delta )</th>
<th>( e ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>1.5</td>
<td>0.5</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>( \gamma )</th>
<th>( \delta )</th>
<th>( e ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0</td>
</tr>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0</td>
</tr>
<tr>
<td>1.5</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.0</td>
</tr>
</tbody>
</table>
2. A wholesale price of \( \frac{e}{C_0} \) is approximately optimal for the manufacturer but can be detrimental to the retailer’s profit when the price elasticity \( e \) is not small (Tables 2 and 3).

3. The optimal level of advertising for the manufacturer is decreasing in the quasi-advertising elasticity \( c \) and the quasi-investment elasticity \( d \) (Fig. 1).

4. The optimal level of advertising for the manufacturer is the lowest when the price elasticity \( e \) is small. When \( e \) is equal to 2 or 3, the difference in the manufacturer advertising \( q^* \) is negligible (Fig. 1).

5. The optimal level of advertising for the retailer is decreasing in the price elasticity \( e \). No clear relationship exists with the quasi-advertising elasticity \( \gamma \) and the quasi-investment elasticity \( \delta \) (Fig. 2).

### Table 3

<table>
<thead>
<tr>
<th>( \gamma )</th>
<th>( \delta )</th>
<th>( e ) (%)</th>
<th>( 1.1 )</th>
<th>( 2.0 )</th>
<th>( 3.0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>-0.5</td>
<td>-4.1</td>
<td>-6.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.2</td>
<td>-1.5</td>
<td>-2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>-0.1</td>
<td>-1.0</td>
<td>-1.4</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>0.5</td>
<td>-0.4</td>
<td>-4.1</td>
<td>-6.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.2</td>
<td>-1.5</td>
<td>-2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>-0.1</td>
<td>-1.0</td>
<td>-1.4</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>0.5</td>
<td>-0.3</td>
<td>-3.6</td>
<td>-6.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.2</td>
<td>-1.5</td>
<td>-2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>-0.1</td>
<td>-1.0</td>
<td>-2.7</td>
<td></td>
</tr>
</tbody>
</table>

### Fig. 1

Manufacturer’s advertising by \( \delta \) for various \( \gamma \).

### Fig. 2

Retailer’s advertising by \( \delta \) for various \( \gamma \).
6. The optimal level of advertising for the manufacturer is significantly less than that for the retailer when the price elasticity $e$ is small or the quasi-investment elasticity $\delta$ and quasi-advertising elasticity $\gamma$ are large. Otherwise, the manufacturer advertises more than the retailer (Figs. 1 and 2).

7. The optimal level of demand is increasing in the quasi-advertising elasticity $\gamma$, decreasing in the quasi-investment elasticity $\delta$, and increasing in the price elasticity $e$. The demand is more sensitive to $e$ than to $\gamma$ and $\delta$, but the demand’s sensitivity to $\gamma$ and $\delta$ is increasing as $e$ increases (Fig. 3).
8. The manufacturer’s profit is highest for the moderate level of price elasticity $e$ and it is increasing in the quasi-advertising elasticity $\gamma$ and the quasi-investment elasticity $\delta$ (Fig. 4).

9. The retailer’s profit is decreasing in the price elasticity $e$ and is less sensitive to changes in the quasi-advertising elasticity $\gamma$ and the quasi-investment elasticity $\delta$ (Fig. 5).

10. The retailer makes more profit than the manufacturer. This gap becomes smaller when the quasi-advertising elasticity $\gamma$, the quasi-investment elasticity $\delta$, and the price elasticity $e$ are all large (Figs. 4 and 5).

The first result indicates that the optimal wholesale price is slightly less than $W^h$. More importantly, the second result indicates that this heuristic is essentially optimal from the manufacturer’s standpoint, but not from the retailer’s. In particular, unless price elasticity is low ($e = 1.1$), a manufacturer who chooses $W^h$ will cut the retailer’s profits at least 1.0% and as much as 6.8%. For the retailers who are known to have slim profit margins (for example discounted retailers such as Wal-Mart and Target), these reductions are significant. This suggests that if price elasticity is low, a manufacturer may use the heuristic value $W^h$ without concern, as both parties’ profits will be insensitive to the difference in wholesale price. However, when price elasticity is moderate to high, the larger wholesale price translates into a larger retail price which significantly reduces the retailer’s profits. This occurs at the retailer not the manufacturer. Because the retailer’s profit margin per unit is larger than the manufacturer’s, the retailer’s profit is more sensitive to changes in demand. Further, though high levels of advertising elasticity can reduce the effect of price elasticity, the retailer will still experience a significant loss. Hence, the retailer has an incentive to encourage the manufacturer to select an optimal rather than a heuristic wholesale price. Interestingly enough, lowering the wholesale price from $W^h$ to $W^*$ will have little impact on the manufacturer but significantly help the retailer.

Result 3 suggests that when local advertising is effective the manufacturer pushes advertising responsibilities on to the retailer. However, when national advertising is effective, the manufacturer chooses to spend less to achieve similar results. This is consistent with the lack of a relationship with advertising elasticities in Result 5. As the leader, the manufacturer manipulates advertising expenditures so that the retailer can not adjust to higher or lower advertising elasticities. Regarding price elasticity and advertising levels, Results 4 and 5 indicate that low price elasticity is exploited through savings on advertising by the manufacturer, while the retailer advertises more to exploit larger profit margins. For larger price elasticities, the manufacturer’s advertising expenditure remains more or less stable while the retailer’s expenditure decreases as profit margins decrease. Again, this demonstrates the manufacturer’s ability to push advertising responsibilities onto the retailer. Result 6 confirms this as the retailer will spend more on advertising whenever price elasticity is low or national advertising elasticity is high.

In Result 7, the effects of the local quasi-advertising elasticity $\gamma$ and the retail price elasticity $e$ on demand are not surprising. For, a higher value of $\gamma$ suggests more effective local advertising and a higher value of $e$ suggests a low profit margin and high volume strategy. In contrast, the fact that larger values of the quasi-investment elasticity $\delta$ result in lower levels of demand reflects the previously discussed tendency of the manufacturer to avoid exploiting the effectiveness of his own advertising and, instead, push advertising responsibilities onto the retailer. This also explains the profit patterns of Results 8–10. When $\gamma$ and $\delta$ are large the manufacturer can achieve his own advertising goals with little expense and rely on the retailer to make up the difference. In this way, increases in advertising elasticities benefit the manufacturer but not the retailer. Finally, when price elasticity is low, even though the manufacturer is the leader, the retailer’s ability to control the retail price allows the retailer to experience significantly more profit than the manufacturer. As the price elasticity increases this advantage is lost and neither the manufacturer nor the retailer does particularly well. This results in the manufacturer doing best with moderate levels of price elasticity, as he loses out to the retailer’s control of the retail price under low price elasticity and struggles along with the retailer under high price elasticity.

6. Conclusion and future research

We have studied pricing and two-tier advertising in a system consisting of one manufacturer and one retailer. For the model where the manufacturer dominates and acts as the leader in a Stackelberg game, we have provided a closed form solution for the retailer’s problem and an algorithm for identifying the manufacturer’s optimal policy. We have shown that it is not optimal for the manufacturer to subsidize advertising at the retail level. Rather, when advertising at the retail level is inefficient, the manufacturer should increase advertising at the national level. But, when advertising is efficient at the retail level, the manufacturer should push advertising responsibility on to the retailer without subsidizing it. Further, when price elasticity is high, the manufacturer should offer a low wholesale price and not seek to increase demand by subsidizing advertising. In general, a heuristic value for the wholesale price, which uses only price elasticity and manufacturing cost, was seen to work well for the manufacturer, but to be detrimental to the retailer except when price elasticity was low.

It would be interesting to see if the suboptimality of cost sharing persists in more general supply chain environments. Situations where the retailer is the leader, where there is more than one retailer and/or supplier, and where demand is
uncertain should be explored in the future to determine whether the results remain valid under these more complex conditions.

Acknowledgements

The authors would like to thank the anonymous referees for all constructive comments and suggestions that have greatly improved the quality of the paper.

Appendix

Proof of Theorem 2. We first establish the correctness of $t_0$ for a given value of $W_0$. To that end, substituting (5) and (6) into (3), we write the manufacturer’s profit in the following form:

$$M_{\text{lead}}(q,t) = -q + xq\frac{\gamma x^t}{C_0} - 1 + yq\frac{\gamma x^t}{C_0} + z,$$

where $x$, $y$ and $z$ are the following constants determined by $x$, $\beta$, $C$, $e$, $\gamma$, and $W_0$.

$$x = -(e - 1)\frac{\gamma x}{C_0}W_0^\frac{\gamma x}{C_0}b\gamma x_1^t,$$

$$y = -(e - 1)\frac{\gamma x}{C_0}W_0^\frac{\gamma x}{C_0}(W_0 - C)b\gamma x_1^t,$$

$$z = (e - 1)^2\frac{\gamma x}{C_0}e^{-eW_0^\frac{\gamma x}{C_0}t}.$$

To determine the optimal value for $t$ we first take the partial derivative of (9) with respect to $t$.

$$\frac{\partial M_{\text{lead}}(q,t)}{\partial t} = \left[\frac{1}{1 + \gamma}\right]q\frac{\gamma x}{C_0}(1 - t)^{-\gamma x}((1 + \gamma)x - \gamma y + \gamma(y - x)t).$$

Note that for any $q > 0$ and $0 \leq t \leq 1$, we have $\left[\frac{1}{1 + \gamma}\right]q\frac{\gamma x}{C_0}(1 - t)^{-\gamma x} \geq 0$. It follows that the sign of the partial derivative is determined by $(1 + \gamma)x - \gamma y + \gamma(y - x)t$ which is a linear function of $t$. Taking $t = 0$ and substituting (10) and (11) give the following expression for the constant term:

$$(1 + \gamma)x - \gamma y = (e - 1)^2\frac{\gamma x}{C_0}W_0^\frac{\gamma x}{C_0}b\gamma x_1^t((e - 1)(W_0 - C) - (1 + \gamma)W_0).$$

The sign of the constant term is determined by the expression $(e - 1)(W_0 - C) - (1 + \gamma)W_0$. It is easy to show that the expression is negative whenever $\frac{c_2}{\gamma}W_0 < C$. If $\gamma \geq e - 2$, then this will always be the case since $C > 0$. Otherwise, this will only be the case if $W_0 < \frac{\gamma}{\gamma - 1}C$.

Next, we substitute (10) and (11) into the coefficient of $t$.

$$\gamma(y - x) = (e - 1)^2\frac{\gamma x}{C_0}W_0^\frac{\gamma x}{C_0}b\gamma x_1^t(\gamma W_0 - (e - 1)(W_0 - C))$$

The sign of the coefficient is then determined by the expression $\gamma W_0 - (e - 1)(W_0 - C)$. It is easy to show that the expression is positive whenever $\frac{c_2}{\gamma}W_0 < C$. If $\gamma \geq e - 1$, then this will always be the case since $C > 0$. Otherwise, this will only be the case if $W_0 < \frac{\gamma}{\gamma - 1}C$.

Notice that if the coefficient and the constant term are both negative, then it will be optimal to take $t = 0$ since $M_{\text{lead}}(q,t)$ will be a decreasing function of $t$ over $[0,1]$. This occurs when $\gamma \in [e - 2, e - 1]$ and $W_0 \in \left[\frac{c_1}{\gamma - 1}C, \frac{c_2}{\gamma - 1}C\right]$. Similarly, if the constant and the coefficient are both positive, then it will be optimal to take $t = 1$ since $M_{\text{lead}}(q,t)$ will be an increasing function of $t$ over $[0,1]$. This occurs when $\gamma \leq e - 2$, $W_0 > \frac{c_1}{\gamma - 1}C$, and either $\gamma \geq e - 1$ or $W_0 < \frac{c_1}{\gamma - 1}C$. Clearly enough, if the first two conditions are satisfied then the third will not be satisfied as $\gamma \leq e - 2 < e - 1$ and $W_0 > \frac{c_1}{\gamma - 1}C > \frac{c_1}{\gamma - 1}C$. So, this case never occurs.

For the constant term to be positive and the coefficient to be negative, we must have $\gamma < e - 2$ and $W_0 \geq \frac{c_1}{\gamma - 1}C$. When this is the case, the objective achieves a global maximum over $[0,1]$ which can be identified by solving $(1 + \gamma)x - \gamma y + \gamma(y - x)t = 0$. Hence, the optimal value of $t$ will be:

$$\frac{\gamma y - (1 + \gamma)x}{\gamma(y - x)} = 1 + \frac{W_0}{\gamma W_0 - (e - 1)(W_0 - C)}.$$

It is straightforward to show that since $\gamma < e - 2$ and $W_0 \geq \frac{c_1}{\gamma - 1}C > \frac{c_1}{\gamma - 1}C$, the righthand side of (14) is less than 1 and greater than or equal to 0.

Finally, for the constant term to be negative and the coefficient to be positive, we must have $\gamma \geq e - 2$ or $W_0 < \frac{c_1}{\gamma - 1}C$, and $\gamma \geq e - 1$ or $W_0 < \frac{c_1}{\gamma - 1}C$. When this is the case, the objective achieves a global maximum over $[0,1]$ at either 0 or 1. However, since either $W_0 < \frac{c_1}{\gamma - 1}C$ or $\gamma > (e - 1)\frac{c_1}{\gamma - 1}C$, the the zero identified by (14) will be greater than 1. Hence, $M_{\text{lead}}(q,t)$ will be a decreasing function of $t$ over $[0,1]$ and the optimal value for $t$ will be 0.
Given $W_0$ and $t_0$, $q_0$ can be derived using straightforward calculus.

\[
\frac{\partial M_{\text{lead}}(q,t_0)}{\partial q} = \left(\frac{\delta}{1 + \gamma}\right) q^{-(1-\gamma-\delta)/(1+\gamma)} (W_0 - C) \left(\frac{eW_0}{e-1}\right)^{-e} \beta^{(1+\gamma)/\gamma} q^{-(1-\gamma-\delta)/(1+\gamma)} e^{\gamma/(1+\gamma)}
\]
\[
\times W_0^{(-\gamma)/(1+\gamma)} (e - 1)^{1/(1+\gamma)} (1 - t_0) \gamma/(1+\gamma) - \left(\frac{\delta}{1 + \gamma}\right) q^{-(1-\gamma-\delta)/(1+\gamma)} t_0 \beta^{(1+\gamma)/\gamma} \gamma/(1+\gamma)
\]
\[
\times e^{-\gamma/(1+\gamma)} W_0^{(-\gamma)/(1+\gamma)} (e - 1)^{1/(1+\gamma)} (1 - t_0)^{-1/(1+\gamma)}.
\]

Taking the second derivative of $M_{\text{lead}}(W,q,t)$ with respect to $q$, we obtain:

\[
\frac{\partial^2 M_{\text{lead}}(q,t_0)}{\partial q^2} = \left(\frac{-\delta - \delta \gamma - \delta^2}{(1 + \gamma)^2}\right) q^{-2(1-2\gamma-\delta)/(1+\gamma)} (W_0 - C) \left(\frac{eW_0}{e-1}\right)^{-e} \beta^{(1+\gamma)/\gamma} q^{-(1-\gamma-\delta)/(1+\gamma)} e^{\gamma/(1+\gamma)}
\]
\[
\times W_0^{(-\gamma)/(1+\gamma)} (e - 1)^{1/(1+\gamma)} (1 - t_0) \gamma/(1+\gamma) - \left(\frac{\delta + \delta \gamma + \delta^2}{(1 + \gamma)^2}\right) q^{-2(1-2\gamma-\delta)/(1+\gamma)} t_0 \beta^{(1+\gamma)/\gamma} \gamma/(1+\gamma)
\]
\[
\times e^{-\gamma/(1+\gamma)} W_0^{(-\gamma)/(1+\gamma)} (e - 1)^{1/(1+\gamma)} (1 - t_0)^{-1/(1+\gamma)}.
\]

It is straightforward to observe that $\frac{\partial M_{\text{lead}}(q,t_0)}{\partial q} \leq 0$ for all $W_0 > C$ and $0 \leq t_0 \leq 1$. Hence, we can set $M_{\text{lead}}(q,t_0) = 0$ to obtain the optimal value for $q$ given any $W_0$ and $t_0$. This results in

\[
q_0 = (e - 1)^{1/(1+\gamma)} e^{\gamma/(1+\gamma)} (1 - t_0)^{-1/(1+\gamma)} W_0^{(-\gamma)/(1+\gamma)} \beta^{(1+\gamma)/\gamma} (1 + \gamma) \gamma/(1+\gamma) \times ((e - 1)(1 - t_0)(W_0 - C) + t_0 W_0) \gamma/(1+\gamma).
\]

This completes the proof of Theorem 2. \[ \square \]

**Proof of Theorem 4.** By Theorem 3, we know that $\tau^* = 0$. Hence, substituting $t_0 = 0$ and $q_0$ as given by Theorem 2 into $f(W)$, we obtain

\[
f(W) = g W^{-\tau}(W - C) + h W^{-\tau}(W - C)^{1/\gamma},
\]

where $g$ is as in (8) and $h$ is the following constant determined by $\alpha$, $\beta$, $C$, $e$, $\gamma$, and $\delta$.

\[
h = -(e - 1)^{1/(1+\gamma)} e^{\gamma/(1+\gamma)} \beta^{1/\gamma} \gamma/(1+\gamma) (1 + \gamma)^{-1/\gamma} \gamma/(1+\gamma) \delta/(1 + \gamma + \delta).
\]

Taking the derivative of $f(W)$, we obtain:

\[
f'(W) = g W^{-1-\tau}(eC - (e - 1)W) + h(1 + \gamma + \delta)^{-1} W^{-1-\tau}(W - C)^{1/\gamma} \gamma/(1+\gamma),
\]

To identify the behavior of $f(W)$, observe that when $W < (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C$, $f'(W)$ will be positive if and only if

\[
\frac{eC - (e - 1)W}{(e + \gamma)C - (e - 1)W} > -h \frac{g}{e}(1 + \gamma + \delta)^{-1} W^{-1-\tau}(W - C)^{1/\gamma}.
\]

When $W > (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C$, the inequality (17) guarantees that the derivative will be negative. Observe that straightforward calculus establishes that the left-hand side of (17) is concave over $[C, (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C]$, decreasing everywhere but $(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C$ where it approaches negative infinite from the left and positive infinite from the right, and has value $1/(1 + \gamma)$ at $W = C$. Straightforward calculus also establishes that the right-hand side of (17) is decreasing and convex over $\left(C, (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C\right]$, increasing over $\left(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C, \infty\right)$, and tends to $\infty$ as $W$ approaches $C$ from the right. We are now ready to establish parts (a) and (b) of Theorem 4.

Assume that $f''((1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C) \leq 0$. It follows that the inequality (17) does not hold at $(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C$ and $f(W)$ is decreasing at $(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C$. Since the left-hand side is increasing and the right-hand side is decreasing over $\left[1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C, (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C\right]$, the inequality will not hold anywhere over $\left(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C, (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C\right]$ and $f(W)$ will be decreasing over $\left[1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C, (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C\right]$. Further, on the interval $\left(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C, \infty\right)$, $f'(W)$ is negative up to some point and then positive up to $\infty$ since the right-hand side is decreasing and the left-hand side is increasing over this interval. Since $f(W)$ is decreasing on $\left(1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C, (1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C\right]$, no point over which $f(W)$ is decreasing can be optimal. Further, $f(W)$ approaches $0$ from below as $W$ approaches $\infty$ so that no point over which $f(W)$ is increasing can be optimal since $f(C) = 0$. It then follows that either $W = C$ maximizes $f(W)$ over $[C, \infty)$ or $f(W)$ achieves another local maximum over $\left(C, 1 + \frac{\delta}{\gamma(e - 1)/(1+\gamma)})C\right]$. If $f(W)$ achieves another local maximum, then there can be only one such local maximum. This is
because the left-hand side is concave and the right-hand side is convex so that they can only intersect at most two points on the interval \( \left( C, \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \right) \). For the assumed local maximum to exist, they must intersect at exactly two points in which case \( f(W) \) is decreasing, then increasing, and finally decreasing over the interval \( \left( C, \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \right) \). This provides the existence of the two points \( W_1 \) and \( W_2 \) as described in the theorem of which \( W_2 \) must be the local maximum which then maximizes \( f(W) \) over \( [C, \infty] \) if \( f(W_2) \geq 0 \). Otherwise, \( W = C \) with \( f(C) = 0 \) maximizes \( f(W) \) over \( [C, \infty] \). This completes the proof of part (a).

Assuming \( f'\left( \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \right) > 0 \), the inequality (17) holds at \( \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \) and \( f(W) \) is increasing at \( \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \). Since the left-hand side is 0 at \( \left( 1 + \frac{1}{e-1} \right) C \) and the right-hand side is strictly positive, it follows that \( f(W) \) switches from increasing to decreasing over \( \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \) \( (1 + \frac{1}{e-1})C \]. Recalling, as in the proof of part (a), that (17) can only hold at equality at at most two points over the interval \( \left( C, \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \right) \) and that \( f'(C) < 0 \), it follows that once \( f(W) \) begins decreasing, it continues decreasing for all \( W < \left( 1 + \frac{1}{e-1} \right) C \). Hence, there exists exactly one local maximum for \( f(W) \) over \( \left( 1 + \frac{\delta}{\left( e-1 + \frac{1}{e-1} \right)} \right) C \) \( (1 + \frac{1}{e-1})C \] which establishes the existence of \( W_3 \). Again, as in the proof of part (a), \( f(W) \) is not maximized for any point greater than \( (1 + \frac{1}{e-1})C \) and either \( W = C \) or \( W = W_3 \) maximizes \( f(W) \) over \( [C, \infty] \). This completes the proof of part (b). This completes the proof of Theorem 4. \( \square \)

References
