

**Quantity Discounts for Channel Coordination:  
Transaction and Channel Efficiency**

July, 2003

S. Chan Choi

Rutgers Business School  
Rutgers University  
180 University Ave  
Newark, NJ 07102

[chanchoi@rci.rutgers.edu](mailto:chanchoi@rci.rutgers.edu)

Office: 973-353-5635

Fax: 973-353-1325

## **Quantity Discounts for Channel Coordination: Transaction and Channel Efficiency**

### **Abstract**

Coordinating activities among members of a distribution channel has been a subject of great attention in the past two decades. Among various methods to coordinate independent channel members, a number of studies in the literature suggest quantity discount as a mechanism to achieve incentive-compatible coordination between a manufacturer and his retailer. The rationale behind this coordination mechanism is that a quantity discount schedule can be designed to align the players' interests with the maximum channel gain.

Two separate streams of quantity discount models have been developed in the literature. The operations management literature views quantity discount as a way to minimize the system-wide cost of operation. The manufacturer can offer quantity discount such that the retailer finds it optimal to order a larger quantity that minimizes the total channel's operating cost. The marketing literature employs quantity discount to induce the retailer to lower the retail price than the level she would choose otherwise. The increased market demand more than compensates the reduced margin so that the total channel profit is maximized. In both models, the channel members' profit goals are aligned with the total channel profit, and the resulting channel coordination creates efficiency gains that can be shared between the members.

In this paper, we present a model of quantity discount that combines these two sources of efficiency gains. We briefly review the existing approaches, and then develop a combined model of quantity discount that provides an incentive-compatible mechanism to coordinate the retailer's behavior. The objective for the manufacturer is to find a discount schedule that induces the retailer to choose the channel-optimal retail price *and* order quantity. A three-step approach is developed to solve the complex problem, and a graphical interpretation is provided. A numerical illustration demonstrates its practicality.

**Key Words:** Quantity Discount, Pricing, Channels of Distribution, Channel Cooperation, Wholesaling, Supply Chains

# Quantity Discounts for Channel Coordination: Transaction and Channel Efficiency

## 1. Introduction

Coordinating activities across multiple players in a distribution channel has been a subject of numerous studies in marketing. The channel literature suggests various ways to achieve coordination, including vertical integration, simple contract, implicit understanding, profit sharing, and quantity discount (Jeuland and Shugan 1983). Many of these coordination mechanisms tend to be based on coerced cooperation rather than on incentive compatibility (McGuire and Staelin 1986). When two or more independent companies are involved in a channel, a “coercive solution” that integrates their activities for the channel-level optimality requires that (a) the modeler has a complete knowledge of the channel partners’ costs and capacities, and (b) all involved channel members implement respective actions as prescribed by the solution.

The first requirement becomes increasingly realistic as the companies integrate their information technology on various operational levels (e.g., Wal-Mart and its suppliers; Procter and Gamble and its retailers). However, the second requirement would be difficult to enforce unless (i) there is a binding agreement among the channel members, (ii) the actions are forced and policed, or (iii) the prescribed solution is *incentive compatible*. In the last mechanism, the behavior of a channel member who acts on self-interest motivation is aligned with the objectives of the other members in the channel. It is a more attractive mechanism of channel coordination that benefits the whole system, and it does not require the retailer’s “cooperation.” Quantity discount has been proposed

in two separate research streams as a tool for achieving incentive compatible channel coordination.

Suppose the channel consists of two members: the manufacturer who determines the wholesale price (in the form of a quantity discount schedule) and the retailer who chooses her optimal order quantity *and* the retail price. In the operations management literature, the channel's total transaction cost that includes inventory holding, facility set-up, and order processing costs can be minimized by properly designing the quantity discount schedule so that the retailer orders the channel-optimal economic order quantity (e.g., Monahan 1984; Chakravarty and Martin 1989; Lal and Staelin 1984; Dada and Srikanth 1987; Kim and Hwang 1989). This is called the *transaction efficiency*. In the marketing literature, on the other hand, the manufacturer can offer the retailer a quantity discount, which induces the retailer to choose her price at the channel-optimal level. This eliminates double marginalization, and the increased market demand due to the lower retail price benefits the whole channel (e.g., Jeuland and Shugan 1983; McGuire and Staelin 1986). This is called the *channel efficiency*.

In this paper, we derive a combined model of these two efficiencies in which the manufacturer designs a quantity discount schedule that can induce the retailer to set *both* the order quantity and the retail price at the channel-optimal level. We show that such a discount schedule consists of a quantity discount per retailer's order as well as a periodical franchise fee for profit sharing. A graphical interpretation of the model provides an intuition of its mechanism, and a numerical example shows how it can be implemented in a practical setting.

In the next two sections, we briefly review the two existing quantity discount models of channel coordination: models of channel efficiency and transaction efficiency. Section 4 presents the combined model and discusses its properties via several lemmas and propositions. A numerical example using a linear demand function is presented in Section 5, and the last section concludes the paper by delineating future research topics.

## 2. Quantity Discount Based on Channel Efficiency

In this section, quantity discount models based on channel efficiency are summarized. These models generally assume that the operating costs are fixed regardless of the demand level. First, we define the notations to be used:

$w$  = wholesale price; to be set by the manufacturer

$p$  = retail price; to be set by the retailer

$D(p)$  = retail demand function;  $D'(p) \leq 0$

$c_M$  = variable unit cost of the manufacturer

$c_R$  = variable unit cost of the retailer

Without any coordination between the channel members, the manufacturer and the retailer maximize respective profits by choosing own pricing variables:

$$\Pi_M = (w - c_M)D(p) \tag{1}$$

$$\Pi_R = (p - w - c_R)D(p) \tag{2}$$

Fixed costs are not relevant and are omitted in the model. The manufacturer can be thought of as the Stackelberg leader, and let us denote the resulting Stackelberg equilibrium solution as  $w^*, p^*, \Pi_M^*, \Pi_R^*$ , and  $D^*$ . Note that the retail price is a reaction function of the wholesale price:  $p^* = p^*(w^*)$ .

When the retail price is coordinated in some way, the joint channel profit is maximized:

$$\Pi_J = \Pi_M + \Pi_R = (p - c_M - c_R)D(p). \quad (3)$$

In the joint profit, the wholesale price is cancelled out, since it only serves as a mechanism for splitting the joint profit between the two members. Let  $p^{**}$  denote the optimal retail price in the coordinated case, and  $D^{**}$  the corresponding joint-optimal demand level. It is obvious that the coordinated joint profit is greater than the sum of uncoordinated profits ( $\Pi_J^* \geq \Pi_M^* + \Pi_R^*$ ). This profit gain of the coordinated channel is due to the elimination of double marginalization, and we call it the *channel efficiency gain*. Since  $p$  is the retailer's reaction, however, the manufacturer's problem is to induce the retailer to set  $p = p^{**}$  instead of her myopic  $p^*$ . One way to achieve this is to *force* the retailer to set the price at the joint optimal level using an external influence (e.g., retail price maintenance). If the manufacturer cannot enforce the retail price for legal or practical reasons, however, he needs to use a coordination mechanism that makes it the retailer's best interest to set the retail price at  $p^{**}$ . For this coordination to be sustainable, the retailer needs to be given an incentive compatible profit function.

Note that at any wholesale price  $w$ , the retailer's own profit is maximized at  $p = p^*(w)$ . Thus, the problem is to find the wholesale price  $w^{**}$  that makes  $p^*(w^{**}) = p^{**}$ . This problem is equivalent to aligning the retailer's profit (2) with the joint profit (3) such that the optimal solution that maximizes (2) also maximizes (3). The simplest wholesale price to achieve this goal is to set  $w^{**} = c_M$ : i.e., set the wholesale price at the manufacturer's marginal cost. At this wholesale price, it is easy to see that profit functions (2) and (3) are equivalent, and the retail price that serves the retailer's self-interest is also the channel-optimal price. Since the manufacturer does not have any margin at this wholesale price, however, he has to levy the desired profit through a fixed charge for the planning period (i.e., franchise fee). This is in fact a two-part pricing, which can be thought of as a quantity discount. If the retailer buys less than the joint optimal level  $D^{**}$  for the planning period, she pays a higher average wholesale price.

Now we consider the relationship between  $p^*$  and  $p^{**}$ . The first order condition for the retailer's optimal profit (2) is

$$\frac{\partial \Pi_R}{\partial p} = D(p) + (p - w - c_R)D'(p) = 0, \quad (4)$$

and the second order condition for maximum profit is

$$\frac{\partial^2 \Pi_R}{\partial p^2} = 2D'(p) + (p - w - c_R)D''(p) \leq 0. \quad (5)$$

The total derivative of (4) with respect to  $w$  is

$$2D'(p)\frac{dp}{dw} - D'(p) + (p - w - c_R)D''(p)\frac{dp}{dw} = 0. \quad (6)$$

Solving equation (6) for  $\frac{dp}{dw}$ , we have

$$\frac{dp}{dw} = \frac{D'(p)}{2D'(p) + (p - w - c_R)D''(p)}, \quad (7)$$

where both the numerator and the denominator (which is the same as the second order condition (5)) are negatives. Thus, as  $w$  rises,  $p$  also goes up. In the uncoordinated case, the retail price ( $p^*$ ) is the best reaction for the Stackelberg wholesale price ( $w^*$ ). Since  $w^* \geq w^{**} = c_M$ , we have  $p^*(w^*) \geq p^{**} = p^*(w^{**})$  and  $D^* \leq D^{**}$ . That is, the joint-optimal retail price is lower than the uncoordinated price, and the coordinated demand is greater than the uncoordinated demand.

While the above wholesale price  $w^{**} = c_M$  was derived by a simple observation of the profit functions, Jeuland and Shugan (1983) develop a formal model of quantity discount based on an incentive compatible profit design. Suppose the manufacturer would collect a fraction  $k_1$  ( $0 \leq k_1 \leq 1$ ) of the joint channel profit plus a fixed amount  $k_2$ . Then the two profit functions are given by

$$\Pi_M = k_1[(p - c_M - c_R)D(p)] + k_2, \text{ and} \quad (8)$$

$$\Pi_R = (1 - k_1)[(p - c_M - c_R)D(p)] - k_2. \quad (9)$$

The corresponding wholesale price schedule can be found by equating the manufacturer profit functions (8) and (1) and solving for  $w$ :

$$w^{**}(D) = k_1 p + (1 - k_1)c_M - k_1 c_R + k_2 / D, \quad (10)$$



which represents the wholesale price as a function of demand. Substituting (10) into the uncoordinated retailer profit (2) results in the coordinated retailer profit (9). Thus, under the wholesale price schedule (10), the retailer's interest is aligned with the joint interest of the channel (i.e., incentive compatibility). Due to the fixed fee component  $k_2$ , this wholesale price can be thought of as a price discount: i.e., the more the retailer buys during the planning period, the lower is the average wholesale price.

The parameters  $k_1$  and  $k_2$  can be chosen in any combination that is suitable for the industry environment. When  $k_2 = 0$ , equation (10) becomes

$w^{**} = c_M + k_1(p - c_M - c_R)$ : the wholesale price is the manufacturer's cost plus a constant fraction of the retailer's unit profit margin. When  $k_1 = 0$ , we have

$w^{**}(D) = c_M + k_2 / D$ , which is exactly the same wholesale price discussed above. This latter schedule does not require the manufacturer's knowledge of the retailer cost. Profit sharing between the two channel members depends on the values of  $k_1$  and  $k_2$ . For a division of the joint profit between the manufacturer and the retailer, Jeuland and Shugan (1983) suggest a bargaining solution to determine the values of  $k_1$  and  $k_2$ .

Several authors have refined this basic quantity discount model. McGuire and Staelin (1986) derive a modified solution in the form of a two-part tariff that is simpler to implement. Moorthy (1987) shows that any pricing schedule, including a quantity surcharge, that can equate the retailer's marginal cost with his marginal revenue at the channel-optimal demand level  $D^{**}$  will coordinate the channel, as long as the marginal cost is strictly below his marginal revenue at all demand levels  $D < D^{**}$  (i.e., the marginal cost curve "cuts" the marginal revenue curve from below at  $D^{**}$ ). The simplest

such pricing is to charge the retailer a constant price of  $c_M$  (the manufacturer's marginal cost), and to extract the desired profit in the form of a franchise fee:

$$k_1(p^{**} - c_M - c_R)D(p^{**}) + k_2.$$

Ingene and Parry (1995) consider a model with  $N$  independent retailers with different cost and demand structures, where  $N$  is endogenously determined. The retailers are assumed to have exclusive territories so that actions taken by a retailer have no effect on the other retailers' demands. They argue that charging different franchise fees for different retailers requires substantial information, and that a preferred pricing is to use a simple two-part tariff that has a common franchise fee (i.e.,  $k_{2,i} = k_2, \forall i$ ).

Although the franchise fee is not relevant to a retailer's pricing decision, it affects her decision whether to participate in the channel, since she needs a non-negative net profit (i.e.,  $\Pi_i^{**} - k_2 \geq 0$ ) in order to participate. However, this limits the number of participating retailers, and Ingene and Parry show that the resulting "coordinated" channel profit can be lower than the uncoordinated Stackelberg profit that does not have a franchise fee.

### **3. Quantity Discount Based on Transaction Efficiency**

In this section, we summarize coordination models based on transaction efficiency. Working independently, Lal and Staelin (1984), Monahan (1984), and Chakravarty (1984) have shown that quantity discount can be used to coordinate economic order quantity in order to minimize the channel's total operating cost. The total operating cost includes inventory holding, order processing, and facility set-up costs

that are associated with the quantity the retailer orders each time. In this research stream, the demand is usually assumed fixed. Let

$D$  = total yearly number of units demanded at the retail level

$Q$  = order size by the retailer

$Q^*$  = optimal order quantity (EOQ) in the absence of coordination

$Q^{**}$  = optimal order quantity with channel coordination

$H_M, H_R$  = the manufacturer's and the retailer's yearly inventory holding costs per unit

$S_M, S_R$  = the manufacturer's and the retailer's order processing or set-up costs per order

$w^*$  = initial wholesale price in the absence of coordination

$w^{**}$  = optimal wholesale price after discount with channel coordination

$\Pi_M, \Pi_R$  = profits of the manufacturer and the retailer

$C_M, C_R$  = operating costs of the manufacturer and the retailer

Since the demand is constant, we assume that the initial wholesale price  $w^*$  and the retail price  $p^*$  are given exogenously (this assumption will be relaxed in the next section). The retailer's decision is to minimize her total cost by choosing the optimal order size  $Q^*$ . Given yearly demand of  $D$ , the retailer's total operating cost for the planning horizon is a function of her order quantity per order:

$$C_R(Q) = S_R(D/Q) + H_R(Q/2), \quad (11)$$

where the first term is the order cost, and the second the inventory holding cost. This cost function is minimized at the well-known economic order quantity (EOQ)

$$Q^* = (2S_R D / H_R)^{1/2} \text{ and the resulting retailer's cost is } C_R(Q^*) = (2DS_R H_R)^{1/2}.$$

On the other hand, the manufacturer's operating cost, excluding the manufacturing cost, is

$$C_M(Q) = S_M(D/Q) + H_M(Q/2). \quad (12)$$

The manufacturer's set-up cost and inventory holding cost can be derived from an appropriate scenario as to whether he manufactures or orders the product (see Weng 1995). When the retailer chooses  $Q^*$ , the manufacturer's total cost becomes

$$\begin{aligned} C_M(Q^*) &= (S_M D) / (2S_R D / H_R)^{1/2} + H_M (2S_R D / H_R)^{1/2} / 2 \\ &= (H_R S_M + H_M S_R) [D / (2H_R S_R)]^{1/2}, \end{aligned} \quad (13)$$

and the resulting joint cost for the channel is

$$\begin{aligned} J(Q^*) &= C_R(Q^*) + C_M(Q^*) \\ &= [H_R(2S_R + S_M) + H_M S_R] [D / (2H_R S_R)]^{1/2}. \end{aligned} \quad (14)$$

Now consider the case in which the two channel members coordinate to minimize the joint cost for the channel. By definition, the coordination would yield a joint cost no greater than that of the non-cooperative case (14). By adding costs of (11) and (12), we have the joint cost function:

$$J(Q) = \frac{D}{Q}(S_R + S_M) + (H_R + H_M)\frac{Q}{2}, \quad (15)$$

which is minimized at the joint optimal order quantity:

$$Q^{**} = [2(S_R + S_M)D / (H_R + H_M)]^{1/2}. \quad (16)$$

Note that  $Q^{**} \geq Q^*$ , only if  $\frac{S_R + S_M}{H_R + H_M} \geq \frac{S_R}{H_R}$ . That is, the joint optimal order quantity is

larger than that the retailer would choose otherwise, when the ratio of set-up cost to inventory holding cost for the manufacturer is larger than that of the retailer, which is a widely accepted assumption. The resulting joint channel cost at  $Q^{**}$  is

$$J(Q^{**}) = [2D(H_R + H_M)(S_R + S_M)]^{1/2}. \quad (17)$$

Since  $J(Q^{**}) \leq J(Q^*)$  by definition,  $C_R(Q^{**}) + C_M(Q^{**}) \leq C_R(Q^*) + C_M(Q^*)$ . Thus,  $C_R(Q^{**}) - C_R(Q^*) \leq C_M(Q^*) - C_M(Q^{**})$ , which implies that the manufacturer's cost saving is greater than the retailer's cost increase when  $Q^{**}$  is chosen instead of  $Q^*$ . The total channel gain,  $J(Q^*) - J(Q^{**})$ , is called the *transaction efficiency gain*.

However, the retailer would not voluntarily increase her order size to  $Q^{**}$  since it only increases her operating cost. In order to encourage the retailer to order a larger quantity, the manufacturer needs to provide an incentive to cover the incremental operating cost. This can be achieved in the form of a quantity discount, where the manufacturer lowers wholesale price when the retailer orders  $Q^{**}$  instead of  $Q^*$ . Any discount schedule can induce the retailer to order  $Q^{**}$  as long as it is designed such that the retailer's total cost, including the product purchasing cost, is minimized at  $Q^{**}$ . Dada

and Srikanth (1987) provide a graphical analysis that defines the wholesale price range and its profit sharing implications, and it is summarized in the following.

Consider the maximum wholesale price the retailer would be interested in paying for an order quantity greater than her initial EOQ,  $Q^*$ . The status quo total cost for the retailer at the initial wholesale price  $w^*$  is  $TC_R(Q^*, w^*) = C_R(Q^*) + w^*D$ , which increases when the order quantity deviates from  $Q^*$ . For any order quantity  $Q$  greater than  $Q^*$  to be acceptable to the retailer, the corresponding wholesale price  $w$  must be decreased to compensate the added cost. The new total retailer cost then becomes  $TC_R(Q, w(Q)) = C_R(Q) + w(Q)D$ . Let  $\bar{w}(Q)$  denote the highest wholesale price that makes the retailer indifferent from ordering  $Q^*$  and any quantity  $Q$ :

$TC_R(Q, \bar{w}(Q)) = TC_R(Q^*, w^*)$ . The corresponding pair  $(Q, \bar{w}(Q))$  represents an indifference (iso-cost) curve for the retailer, and it can be obtained from the last equality as

$$\bar{w}(Q) = \frac{TC_R(Q^*, w^*)}{D} - \frac{S_R}{Q} - \frac{H_R Q}{2D}. \quad (18)$$

From its first and second order derivatives, we find that  $\bar{w}(Q)$  is a concave, decreasing function when  $Q > Q^*$ .<sup>1</sup> At any price-quantity pair on this curve, the manufacturer takes all the transaction efficiency gain.

---

<sup>1</sup> The first order derivative is  $S_R / Q^2 - H_R / (2D)$ . Since we are only interested in the value of  $Q$  such that  $Q > Q^* = (2DS_R / H_R)^{1/2}$ , one can easily verify the first order derivative is negative. The second order derivative is clearly negative.

When the retailer takes all the gain, on the other hand, the manufacturer charges the minimum wholesale price ( $\underline{w}$ ) for a quantity  $Q$  such that  $TC_M(Q, \underline{w}(Q)) = TC_M(Q^*, w^*)$ , where  $TC_M(Q, w(Q)) = C_M(Q) - w(Q)D$ . The corresponding pair  $(Q, \underline{w}(Q))$  represents an indifference (iso-cost) curve for the manufacturer:

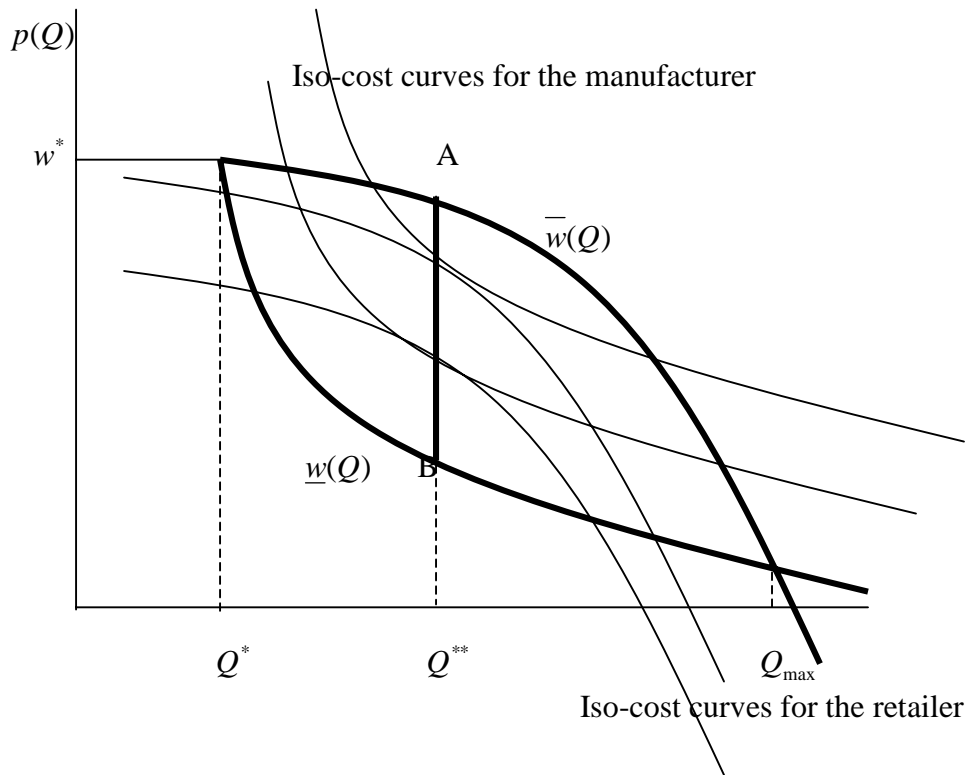
$$\underline{w}(Q) = \frac{S_M}{Q} + \frac{H_M Q}{2D} - \frac{TC_M(Q^*, w^*)}{D}. \quad (19)$$

It is straightforward to verify that  $\underline{w}(Q)$  is a convex, decreasing function due to the assumption that  $\frac{S_R}{H_R} < \frac{S_M}{H_M}$ . These two iso-cost curves meet at  $(Q^*, w^*)$  as depicted in

Figure 1.

In Figure 1, both channel members are better off, if the wholesale price is within the range of  $\underline{w}(Q) < w(Q) < \bar{w}(Q)$  at any given order quantity of  $Q^* < Q < Q_{\max}$ . Thus, any order quantity and wholesale price pair within this region is a feasible price discount. This region is called the *feasible region*. The vertical difference between the two curves represent the amount of the transaction efficiency gain, and by definition, the gain is maximized at  $Q^{**}$ . It is straightforward to show that the joint channel operating cost can be expressed as  $J(Q^{**}) = J(Q^*) - D(\bar{w}(Q^{**}) - \underline{w}(Q^{**}))$ .

**Figure 1. The Feasible Region and Pareto Optimal Price-Quantity Frontier**

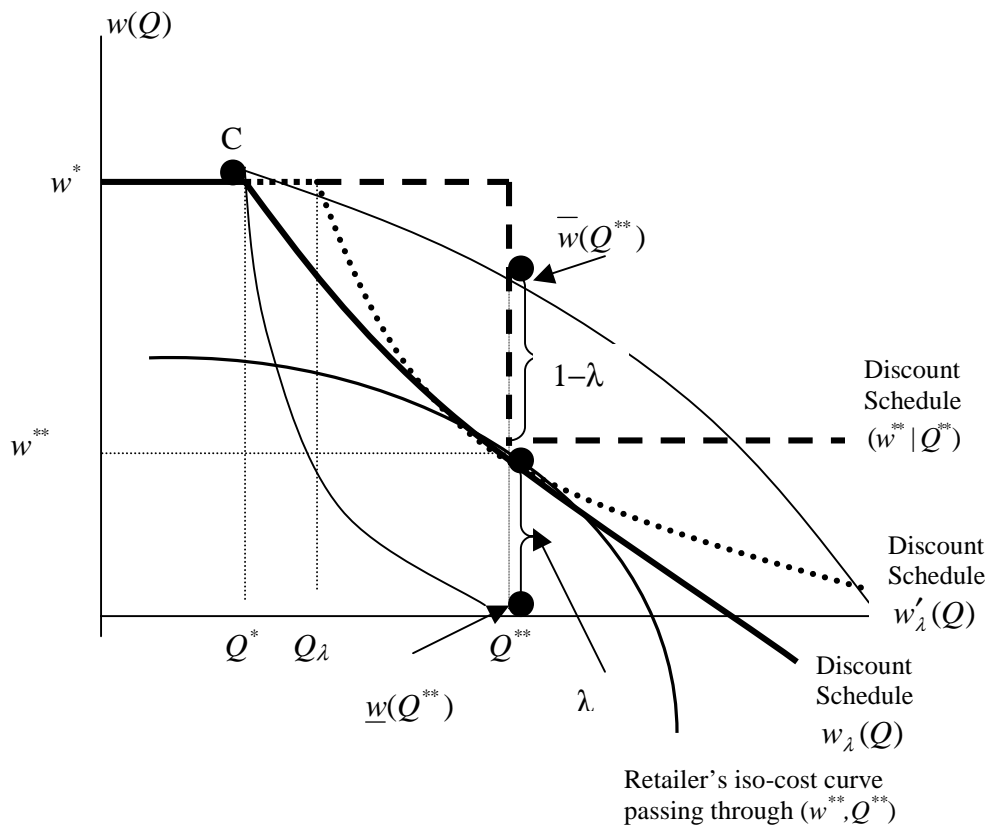


The transaction efficiency gain can be shared between the channel members by setting the wholesale price  $w(Q^{**})$  along the Pareto-optimal frontier between A and B in Figure 1. The higher the wholesale price, the more gain is distributed to the manufacturer. Kohli and Park (1989) propose several methods of dividing the total gain using bargaining mechanisms. Suppose the manufacturer decides to keep a fraction  $\lambda$  ( $0 \leq \lambda \leq 1$ ) of the efficiency gain. The simplest quantity discount schedule is a single-step *all-unit* discount that charges  $w^{**} = \lambda \bar{w}(Q^{**}) + (1 - \lambda) \underline{w}(Q^{**})$  per unit when the retailer orders  $Q^{**}$  or more, and charges  $w^*$  for an order less than that. This discount schedule is depicted as the dashed line in Figure 2 (Schedule  $(w^{**} | Q^{**})$ ). Any multiple-



step variations can be designed as long as the average price lies above the retailer's iso-cost curve passing through the optimal solution, except at  $(w^{**}, Q^{**})$  where the discounted wholesale price touches her iso-cost curve. Note that a straight linear discount between  $(w^*, Q^*)$  and  $(w^{**}, Q^{**})$  may not work especially at a larger value of  $\lambda$ , since it is likely that the discount price line passes below the retailer's iso-cost curve that passes  $(w^{**}, Q^{**})$ . In that case, the retailer would find it even cheaper to deviate from  $Q^{**}$ .

**Figure 2. All-Unit Quantity Discount Schedules**



On the other hand, a simple *continuous* discount schedule can be designed with a price break point at  $Q^*$  (Dada and Srikanth 1987). Consider the following pricing schedule (Schedule  $w_\lambda(Q)$ ):

$$\begin{aligned} w_\lambda(Q) &= \lambda \bar{w}(Q) + (1 - \lambda) \underline{w}(Q), \text{ when } Q > Q^*, \text{ and} \\ &= w^*, \text{ when } Q \leq Q^*. \end{aligned} \quad (20)$$

This schedule divides the total gain according to the predetermined fraction  $\lambda$  at *any* quantity larger than  $Q^*$  (the thick curve in Figure 2). Note that  $w_\lambda(Q^{**}) = w^{**}$ , at which point, the retailer finds that her total cost  $TC_R(Q, w_\lambda)$  is minimized.

A modification of the above schedule can be designed by adding a per-unit fixed margin over the retailer's iso-cost curve (Schedule  $w'_\lambda(Q)$ ):

$$\begin{aligned} w'_\lambda(Q) &= w^*, \text{ when } Q \leq Q_\lambda, \text{ and} \\ &= \underline{w}(Q) + \lambda(\bar{w}(Q^{**}) - \underline{w}(Q^{**})), \text{ for } Q > Q_\lambda. \end{aligned} \quad (21)$$

This schedule is presented as the dotted line in Figure 2. Again, note that  $w'_\lambda(Q^{**}) = w^{**}$ , and the price curve is above the retailer's iso-cost curve elsewhere. The price break point  $Q_\lambda$  can be obtained by equating the two equations in (21). Since a deviation from  $Q^{**}$  becomes more expensive than in Schedule  $w_\lambda(Q)$ , it can provide a stronger incentive for the retailer to choose the channel-optimal order quantity. All these three quantity discount schedules are incentive compatible, since they all make the retailer's interest aligned with that of the total channel.

Note that Schedules  $w_\lambda(Q)$  and  $w'_\lambda(Q)$  require the computation of  $\bar{w}$  and  $\underline{w}$  as functions of  $Q$ , which may not be very attractive in practice. A similar incentive compatibility can be designed in the form of an *incremental* quantity discount that does not require the knowledge of such functions. A simple schedule can be designed such that the seller charges a discounted unit price of  $rw^*$  ( $0 \leq r \leq 1$ ) *only* for the units exceeding the break point  $\hat{Q}$  (Schedule  $(\hat{Q}, r)$ ).

With this schedule, the total wholesale price for a quantity  $Q > \hat{Q}$  is  $w^*(\hat{Q} + r(Q - \hat{Q}))$ , and the average unit price is  $w^*(r + (1 - r)\hat{Q}/Q)$ . Substituting this average unit price into the retailer's total cost, we have

$$TC_R(Q) = w^*[r + (1 - r)\hat{Q}/Q]D + S_R(D/Q) + H_R(Q/2). \quad (22)$$

Its minimum occurs at

$$\tilde{Q} = \sqrt{\{2D\hat{Q}w^*(1 - r) + 2DS_R\}/H_R} = \sqrt{2D\hat{Q}w^*(1 - r)/H_R + Q^{*2}},$$

where  $Q^*$  is the retailer's initial EOQ. The manufacturer designs the incremental quantity discount schedule by solving the equation system:  $Q^{**} = \tilde{Q}$  and

$w^{**} = w^*(r + (1 - r)\hat{Q}/Q^{**})$ , whose solution is

$$\begin{aligned} r &= (H_R Q^{*2} + 2w^{**} Q^{**} D - H_R Q^{**2}) / (2w^* Q^{**} D), \\ \hat{Q} &= Q^{**} (w^{**} - rw^*) / w^* (1 - r). \end{aligned} \quad (23)$$

Any desired price-quantity combination within the feasible region thus can be achieved by employing an incremental quantity discount schedule (23). Its graphical property can be shown to be similar to that of Schedule  $w'_\lambda(Q)$  in Figure 2, which follows from the

fact that any incremental discount can be converted into an all-unit discount using an average price function (Dolan 1987).

#### **4. Quantity Discount Based on Both Channel and Transaction Efficiency**

In the previous two sections, we have seen two separate sources of efficiency when the channel members coordinate their actions. In both cases, by providing a wholesale price quantity discount, the manufacturer can induce the retailer to choose a decision (order quantity or retail price) that is optimal for the whole channel. However, there are substantial differences in these two forms of quantity discount. In the transaction efficiency models, quantity discount is based on the *quantity per order*. The yearly demand is usually assumed fixed in which case the retail price is irrelevant. The discount types are all-unit and incremental quantity discount schedules, which do not have the fixed (or franchise) fee component. In the channel efficiency model, on the other hand, the discount is based on the *yearly (or for a planning horizon) demand*. In order to redistribute the efficiency gain, a periodic franchise fee is employed. In a combined model, however, the quantity discount schedule is compounded by the fact that the manufacturer needs to induce the retailer to not only set the retail price for the channel-optimal demand level but also choose the jointly optimal order quantity, which in turn depends on the demand.

When we combine these two sources of efficiency, it is intuitive that the solution would combine an all-unit or incremental discount with a periodic franchise fee. Weng (1995) shows that this combined quantity discount is sufficient for channel coordination, but his analysis implicitly assumes that the Pareto-optimal solution is a line segment as in

Figure 1, which is generally not the case when the demand is elastic as will be shown below.

Without coordination, the manufacturer's problem is to maximize his yearly profit with respect to the wholesale price  $w$ :

$$\Pi_M(w) = (w - c)D(p) - S_M D(p)/Q - H_M Q/2. \quad (24)$$

The manufacturer's own replenishment cost  $S_M$  includes order processing and his own ordering (or set-up, if he makes the product) costs per retailer's order. His inventory holding cost  $H_M$  also depends on whether he buys or makes the product.<sup>2</sup> Note that these operating costs for the manufacturer are functions of the order quantity chosen by the *retailer*.

Similarly, the retailer's problem is to maximize her yearly profit with respect to two variables--the retail price  $p$  and the order quantity  $Q$ :

$$\Pi_R(p, Q) = (p - w)D(p) - S_R D(p)/Q - H_R Q/2. \quad (25)$$

When there is no channel coordination, we again assume that the two channel members play a manufacturer-Stackelberg game. Conditional on the wholesale price  $w$ , the retailer maximizes her own profit by solving her first order conditions:

$$\frac{\partial \Pi_R}{\partial p} = D(p) + (p - w)D'(p) - \frac{S_R D'(p)}{Q} = 0 \quad \text{and} \quad (26)$$

---

<sup>2</sup> In the operations management literature, it is well known that the manufacturer's optimal lot size is an integer multiple of the retailer's order quantity  $Q$ . Both  $S_M$  and  $H_M$  are functions of this multiple coefficient ( $m$ ). (See Weng 1995 and Joglekar 1988.) We assume the seller always sets the optimal value of  $m$ .

$$\frac{\partial \Pi_R}{\partial Q} = -\frac{S_R D(p)}{Q^2} - \frac{H_R}{2} = 0. \quad (27)$$

In Appendix, we show that the Hessian matrix of  $\Pi_R$  is negative definite when the retailer profit is concave in price. The retailer's problem can be simplified by noting that equation (27) produces a unique value of  $Q^*$  (i.e., EOQ) for any given  $p$ . Thus, equations (26) and (27) can be simplified by substituting  $Q^*(p)$  in the retailer profit (25):

$$\Pi_R(p) = (p - w)D(p) - [2S_R H_R D(p)]^{1/2}. \quad (28)$$

The first order condition for the maximum retailer profit is

$$\frac{\partial \Pi_R}{\partial p} = D(p) + (p - w)D'(p) - D'(p) \sqrt{\frac{H_R S_R}{2D(p)}} = 0, \quad (29)$$

Its solution is a set of the retailer's reaction functions to the manufacturer's wholesale price:  $p^*(w)$  and  $Q^*(w)$ .

**Lemma 1:**  $p^*(w)$  is an increasing function, and  $Q^*(w)$  is a decreasing function of  $w$ .

**Proof:** See Appendix.

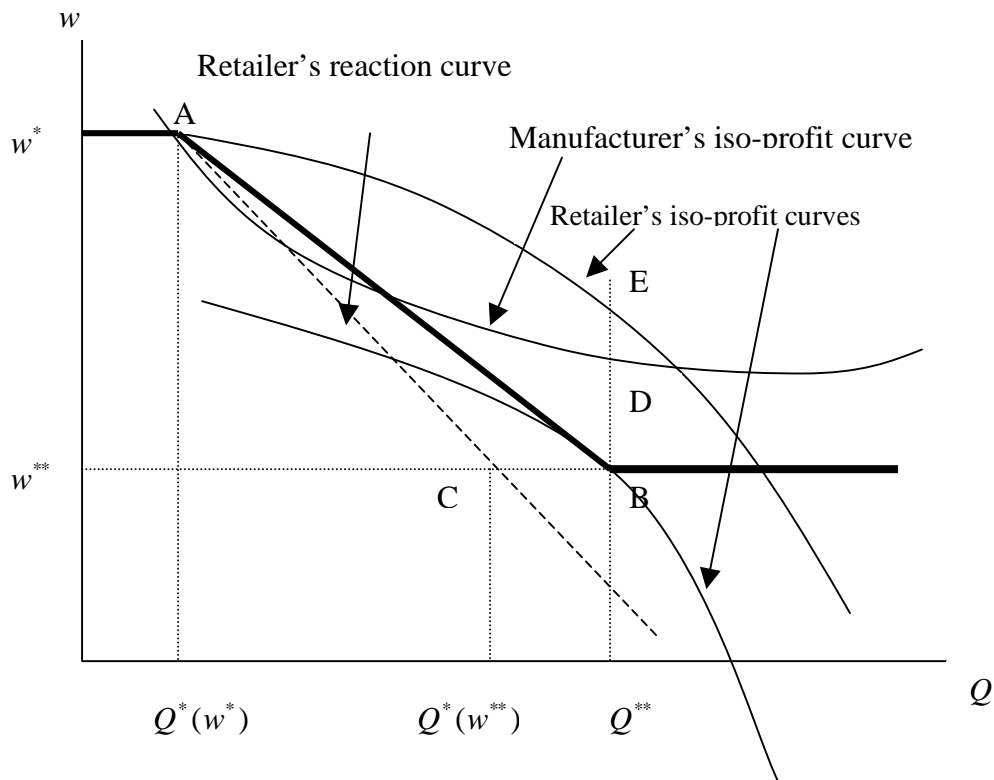
**Corollary 1:** Demand  $D^*(w)$  is a decreasing function of  $w$ .

In particular, the relationship between  $w$  and the corresponding  $Q^*(w)$  is depicted as a dashed line in Figure 3, and is termed as the *retailer's reaction curve*.

The manufacturer, who plays the Stackelberg leader, can find the optimal wholesale price by substituting these retailer reaction functions--  $p^*(w)$  and  $Q^*(w)$ --into his profit function (24) and solving the resulting first order condition for  $w$ :

$\frac{\partial \Pi_M(w, p^*(w), Q^*(w))}{\partial w} = 0$ . Let us denote the resulting uncoordinated wholesale price as  $w^*$ . All other related equilibrium quantities for the uncoordinated channel outcome can be obtained by substitution; let us denote them as  $p^*, Q^*, D^*, \Pi_R^*$ , and  $\Pi_M^*$ . The resulting total channel profit is  $\Pi_J^* = \Pi_R^* + \Pi_M^*$ . The wholesale price-order quantity pair  $(w^*, Q^*)$  of this uncoordinated channel is marked as point A in Figure 3.

**Figure 3. Uncoordinated and Coordinated Solutions and Quantity Discount**



Now let us consider the case in which the channel members use some mechanism (in our case, quantity discount) to coordinate the retail price and the order quantity in

order to maximize the joint profit  $\Pi_J = \Pi_R + \Pi_M$ . The corresponding optimal joint profit  $\Pi_J^{**}$  is by definition no less than  $\Pi_J^*$ . In this problem, the wholesale price  $w$  cancels out, and the channel determines only the retail price and the order quantity. After simplification, the joint profit can be rewritten as

$$\Pi_J = (p - c)D(p) - S_J D(p) / Q(p) - H_J Q(p) / 2, \quad (30)$$

where  $S_J = S_M + S_R$  and  $H_J = H_M + H_R$ . The last two terms can be separated, and they produce the joint optimal order quantity  $Q^{**}(p) = [2S_J D(p) / H_J]^{1/2}$ , which is a function of the retail price. Comparing  $Q^{**}$  with  $Q^*$ , it is easy to see that for any given  $p$ ,  $Q^{**}(p) > Q^*(p)$  as long as  $S_M / H_M > S_R / H_R$ . As in Section 3, this condition requires that the manufacturer incur a relatively higher set up cost and lower inventory holding cost than the retailer does, which is reasonable and customary in the operations management literature. It is straightforward to obtain the resulting joint operating cost:  $C_J^{**} = [2S_J H_J D(p)]^{1/2}$ .

**Lemma 2:** Suppose  $S_M / H_M > S_R / H_R$ . Then  $Q^{**}(p) > Q^*(p)$  for any given  $p$ .

(Proof is obvious, and is omitted.)

Substituting  $Q(p)$  in (30) with  $Q^{**}(p)$ , equation (30) can be rewritten as a function of only the retail price:

$$\Pi_J = (p - c)D(p) - [2S_J H_J D(p)]^{1/2}. \quad (31)$$

Its first order condition for the maximum  $\Pi_J$  is



$$\frac{\partial \Pi_J}{\partial p} = D(p) + (p - c)D'(p) - D'(p)\sqrt{\frac{H_J S_J}{2D(p)}} = 0. \quad (32)$$

Note that, without the operating costs (the third term), this condition is equivalent to the first order condition when only the channel efficiency is considered (Section 2). Since we know  $D'(p) < 0$ , the optimal price for the joint profit (31) must be higher than that for the joint profit (3). That is, when the operating costs are considered, the channel-optimal retail price  $p^{**}$  is higher than that without the operational costs.

**Proposition 1:** The channel-optimal retail price  $p^{**}$  is higher when operating costs are considered than when they are not.

**Proof:** See the above discussion

On the other hand, from the discussions in the channel efficiency case of Section 2, it is intuitive that the channel-optimal retail price  $p^{**}$  must be lower than the uncoordinated retail price  $p^*$ , because it eliminates double marginalization in the channel. The following proposition confirms that this remains true in the current case of combined efficiency:

**Proposition 2:** The channel-optimal retail price  $p^{**}$  (i.e., the solution for equation 32) is lower than the uncoordinated retail price  $p^*(w^*)$  (i.e., the solution for equation 29).

**Proof:** See Appendix

Once the channel-optimal retail price  $p^{**}$  is found, the corresponding market demand  $D^{**} = D(p^{**})$  and the channel-optimal order quantity  $Q^{**} = (2S_J D^{**} / H_J)^{1/2}$  are uniquely determined. Now the manufacturer's channel-coordination quantity

discount problem becomes one of designing the wholesale price  $w^{**}$  as a quantity discount that induces the retailer to choose the channel optimal solution: i.e.,  $(p^{**}, Q^{**})$ . Compared to the quantity discount schedules derived in the previous sections, however, making the retailer's profit incentive compatible with the channel profit in this combined case is much more complicated due to the elastic demand. Moreover, the wholesale price must be designed such that it influences *both* the retail price and the optimal order quantity.

Our approach to solving this problem is to divide it into the following three steps that are sequentially solved:

**Step 1:** Find the optimal wholesale price  $w^{**}$  such that it induces the retailer to choose her retail price  $p^{**}$  as the best reaction

**Step 2:** Design a quantity discount schedule to induce the retailer to choose an order quantity  $Q^{**}$  as the best reaction

**Step 3:** Find a proper level of franchise fee in order to share the efficiency gain between the channel members

In the following, these steps are elaborated in detail:

### *Step 1*

In solving the first step problem, the retailer's first order condition (29) is solved to generate the retailer's price reaction function  $p = p^*(w)$ . By Lemma 1, we know that there is a unique  $p$  for any positive, real  $w$ . Therefore, once the demand function is specified, it is straightforward to find an optimal value of  $w^{**}$ , which produces the

optimal retailer reaction such that  $p^*(w^{**}) = p^{**}$ . This retail price determines the demand  $D^{**} = D(p^{**})$  and the corresponding channel-optimal order quantity  $Q^{**}$ . This solution pair  $(w^{**}, Q^{**})$  is marked in Figure 3 as Point B. The following proposition shows that the wholesale price must decrease for channel coordination.

**Proposition 3:** The channel-coordinating wholesale price  $w^{**}$  is lower than the uncoordinated wholesale price  $w^*$ .

**Proof:** We saw from Proposition 2 that  $p^{**} < p^*(w^*)$ , and from Lemma 1 that  $\frac{dp}{dw} > 0$ .

Therefore, if there exists a wholesale price  $w^{**}$  that results in the retailer's retail price  $p^{**}$ , then  $w^{**} < w^*$ . QED

### *Step 2*

However, given the channel-optimal wholesale price  $w^{**}$  and the corresponding demand  $D^{**} = D(p^*(w^{**}))$ , the retailer will choose the order quantity based on her own EOQ,  $Q^*(D^{**})$  (via equation 27), which follows her best reaction curve in Figure 3. The following proposition shows that the retailer's best reaction order quantity is less than the joint optimal order quantity.

**Proposition 4:** At  $w^{**}$ , the retailer's best reaction order quantity is less than the channel-optimal order quantity: i.e.,  $Q^*(w^{**}) \leq Q^{**}$ .

**Proof:** It directly follows from Lemmas 1 and 2.

**Corollary 2:** At the channel-optimal wholesale price  $w^{**}$ , the myopic retailer profit  $\Pi_R(w^{**}, Q^*(w^{**}))$  is greater than the coordinated profit  $\Pi_R(w^{**}, Q^{**})$ .

This best-reaction order quantity is marked as Point C in Figure 3, which is located on the left side of  $Q^{**}$  due to Proposition 4. In order to induce the retailer to choose  $Q^{**}$ , the manufacturer needs to levy a *surcharge* for any quantity lower than  $Q^{**}$ . This surcharge has to be greater than the retailer's incremental gain from choosing a lower quantity, yet it must be low enough to provide the retailer a profit no less than the uncoordinated Stackelberg solution  $\Pi_R(w^*, Q^*(w^*))$ . Otherwise, the retailer will choose to pay the surcharge and order a smaller quantity. This surcharge can be viewed as a quantity discount that “waives” the surcharge so that the retailer finds it more profitable to order  $Q^{**}$ . In other words, any order quantity greater than  $Q^*(w^*)$  will cost the retailer more, and she needs to be compensated for the incremental cost in the form of quantity discount.

For the quantity discount to be plausible, we need to show that ordering  $Q^{**}$  at  $w^{**}$  is indeed more profitable to the retailer than ordering  $Q^*$  at  $w^*$ . From the retailer's point of view, her profit before coordination is  $\Pi_R(w^*, Q^*(w^*))$ . We are interested in the discounted wholesale price ( $\bar{w}$ ) that makes the retailer indifferent between ordering a quantity  $Q$  and  $Q^*(w^*)$ . The iso-profit curve for the retailer can be obtained from solving the following equation with respect to  $\bar{w}$ :

$$\Pi_R(w^*, Q^*(w^*)) = (p^*(\bar{w}) - \bar{w})D(p^*(\bar{w})) - S_R D(p^*(\bar{w})) / Q - H_R Q / 2, \quad (33)$$

where the left-hand side of the equation is a constant—the uncoordinated retailer profit. Figure 3 shows two iso-profit curves for the retailer: the top one for the uncoordinated profit level, and the bottom one for the coordinated profit level. Note that the closer the retailer’s iso-profit curve is located toward the origin, the higher retailer profit it represents. Because our demand function is a generic form, we cannot derive a closed form equation of the iso-profit curve  $\bar{w}(Q)$ . However, by definition, the base-line curve passes the uncoordinated solution  $(w^*, Q^*(w^*))$ , at which point it crosses the retailer’s best reaction curve as shown in Figure 3. In addition, since it represents the wholesale price discount for an increased order quantity, it is a decreasing function of  $Q$ .

Figure 3 depicts that the coordinated solution B lies under the retailer’s iso-profit curve. This implies that without any other side payment, the coordinated retailer profit  $\Pi_R(w^{**}, Q^{**})$  is greater than the uncoordinated one  $\Pi_R(w^*, Q^*(w^*))$ . Chakravarty and Martin (1989) show that, when an elastic demand is introduced, the Pareto-optimal (i.e., neither player is worse off) price-quantity coordinate is not a line segment but a point on the manufacturer’s iso-profit curve (the “bottom” of the feasible region). If we relax the Pareto-optimality constraint (i.e. the manufacturer is allowed to be worse off before the franchise fee), the channel-optimal solution would go below the manufacturer’s iso-profit curve, which in turn is below the retailer’s base-line iso-profit curve. Hence,

$\Pi_R(w^{**}, Q^{**}) \geq \Pi_R(w^*, Q^*)$ . Based on numerous simulations, we believe this is generally the case, although we were not able to prove it under the generic demand function and general cost parameters. Instead, the following lemma and proposition provide a *sufficient* condition for this profit relation:

**Lemma 3:** When the wholesale price changes, the retail price changes by less than half of the wholesale price change.

**Proof:** See Appendix

**Proposition 5:** Let  $m$  denote the retailer's margin (i.e.,  $m = p - w$ ). Further, let

$x = S_J / S_R$  and  $y = H_J / H_R$  and  $R = (x + y) / 2\sqrt{xy}$ . A sufficient condition for

$\Pi_R(w^{**}, Q^{**}) \geq \Pi_R(w^*, Q^*(w^*))$  is that  $m^* (\sqrt{D^{**}} + \sqrt{D^*}) \geq (R - 1)\sqrt{2S_R H_R}$ .

**Proof:** See Appendix

This condition can be loosely translated to state that the retailer's operating costs are not too large compared to the incremental retailer revenue. In addition, it helps if the ratios of the set-up to inventory-holding costs between the two channel members are not very asymmetric.<sup>3</sup> Under this sufficient condition, the retailer profits at the three points (A, B, and C) in Figure 3 have the following relationship:

$$\Pi_R(w^*, Q^*) \leq \Pi_R(w^{**}, Q^{**}) \leq \Pi_R(w^*, Q^*(w^{**})).$$

As discussed briefly above, the manufacturer's pricing goal now is to let the retailer buy the product at wholesale price  $w^{**}$  when she orders  $Q^{**}$ . Therefore, an all-unit quantity discount schedule can be designed by connecting the two points A and B in the same way as discussed in the previous section. That is, any smooth or step function can induce the retailer to choose the channel-optimal solution as long as the discount line stays above the retailer's iso-profit curve that passes Point B. All quantity discount schedules discussed in Section 3 can be applied here. However, due to the complexity of

---

<sup>3</sup> When the inventory holding costs and set up costs are arbitrarily high to violate this condition, our numerical trials found that the profits become negatives.

the inverse functions, one needs to resort to a numerical approach to determine the discount schedule.<sup>4</sup> Section 5 provides an illustration of a numerical solution.

### *Step 3*

In the previous two steps, the channel-optimal solution could be achieved by an incentive compatible wholesale price quantity discount. However, the resulting wholesale price is determined without considering the manufacturer's or the retailer's individual profit implication. Instead, it maximizes the total channel profit. The discussion in Step 2 shows that the manufacturer's profit is likely to be lower than that of the uncoordinated solution. Therefore, the next step is to find a mechanism to share the efficiency gain between the channel members.

The manufacturer's indifference curve is also shown in Figure 3. At  $Q^*$ , it has a negative slope, steeper than the retailer's indifference curve. However, unlike the case in Section 3, its slope will eventually become positive at a larger value of  $Q$  due to the operating cost (see Appendix for a proof). The quantity discount schedule derived above will let the retailer choose to order  $Q^{**}$  at a marginal wholesale price  $w^{**}$  (point B), which lies below the manufacturer's iso-profit curve. In order to recover the loss from the coordination, and possibly claim a portion of the efficiency gain, the manufacturer needs to charge a periodic franchise fee. The minimum franchise fee (per unit) is represented by the line segment between B and D, in which case the retailer takes all the efficiency gain. The maximum fee is the line segment between B and E, where the

---

<sup>4</sup> Some authors had to rely on numerical solutions for transaction efficiency models when some assumptions are relaxed. See for example Chakravarty and Martin (1989) and Kim and Hwang (1989)

manufacturer takes all the efficiency gain. Depending on the bargaining solution (e.g., Kohli and Park 1989), the franchise fee can be set between these two extremes. However, we note that this franchise fee was derived under the assumption that the quantity discount in Step 2 is implemented as an all-unit discount. When an incremental discount schedule is used instead, the franchise fee needs to be adjusted for the *additional* revenue the manufacturer gains from the higher prices at lower quantity breaks.

We have shown that channel coordination requires both per-order quantity discount as well as a periodic franchise fee, which is considered as another form of quantity discount. Due to the generic form of the demand function, however, we do not have a closed form solution for the quantity discount schedules. In the next section, we provide a numerical illustration using a simple linear demand and specific values for the cost parameters.

## 5. An Illustration with a Linear Demand

Suppose the demand function takes the simplest linear form:  $D(p) = 1 - bp$ . In this section we assume the following specific parameter values:<sup>5</sup>

$$S_R = 0.2, S_M = 0.5, H_R = 0.3, H_M = 0.1, c = 0.3, b = 0.2.$$

When there is no coordination, the retailer maximizes her profit function with respect to the retail price (and the corresponding order quantity) conditional on the wholesale price. A closed form solution of equation (29) for the reaction function involves imaginary numbers outside the relevant variable domain and is difficult to

---

<sup>5</sup> Our simulation using a number of parameter value combinations resulted in virtually the same implications.



derive in a simple form. Instead, we have used a quadratic approximation of the reaction function as follows:

$$p^*(w) = 2.6258 + 0.5125w + 0.0038w^2. \quad (36)$$

The corresponding economic order quantity is, after simplification,

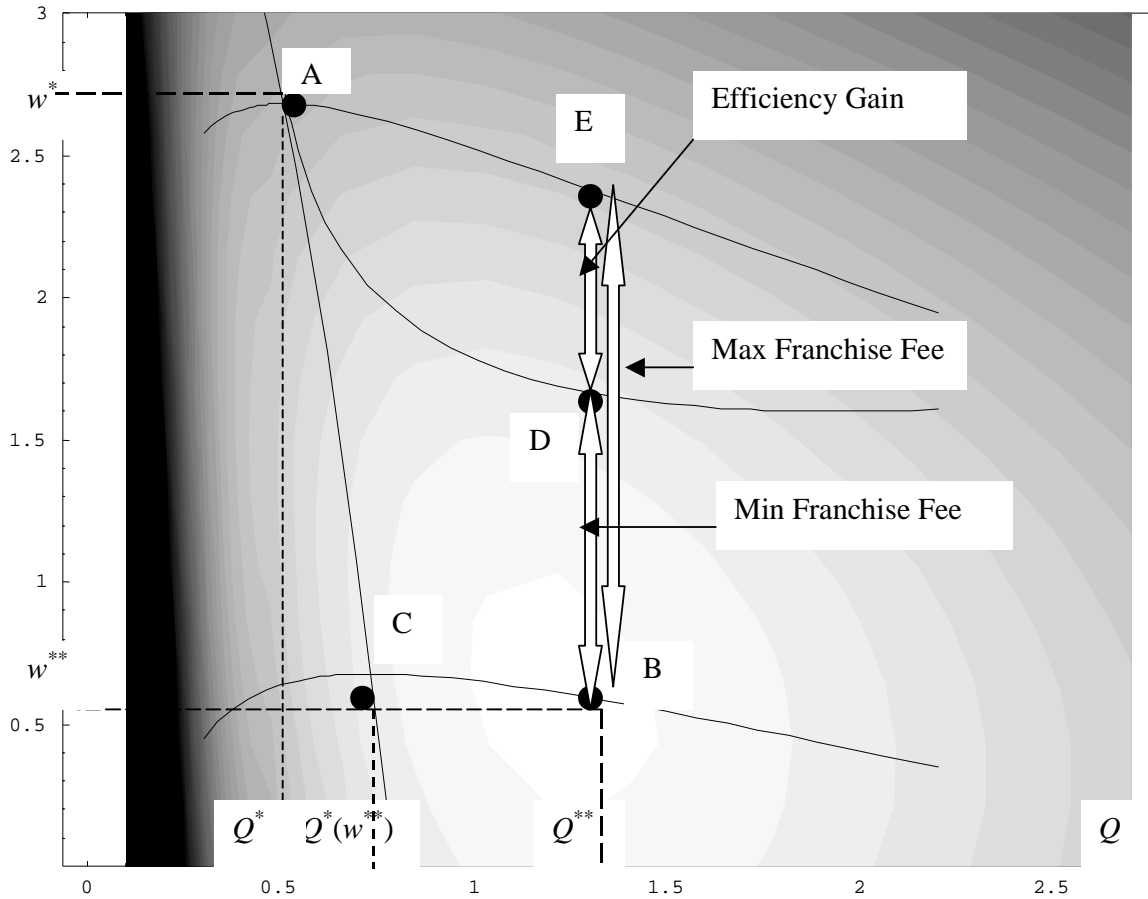
$$Q^*(w) = 0.0318\sqrt{-(w - 4.4838)(w + 139.656)}. \quad (37)$$

Substituting these quantities into the manufacturer's profit function (24), and solving for the optimal value of  $w$  for profit maximization results in  $w^* = 2.6816$ . The corresponding uncoordinated order quantity is  $Q^* = 0.5093$ . Point A in Figure 4 marks this  $(w^*, Q^*)$  pair. All other related Stackelberg equilibrium quantities are summarized in Table 1.

**Table 1. Uncoordinated Stackelberg Equilibrium vs. Coordinated solution**

	Uncoordinated Stackelberg	Coordinated Solution
$w$	2.6816	0.6133
$m$	1.3458	2.3283
$p$	4.0274	2.9416
$D$	0.1945	0.4117
$Q$	0.5093	1.2004
$\Pi_M$	0.2468	-0.1025
$\Pi_R$	0.1090	0.7099
$\Pi_J$	0.3558	0.6073

**Figure 4. A Numerical Example of Quantity Discount**



On the other hand, when the channel members coordinate actions for joint profit maximization (equation 30), the optimal retail price is  $p^{**} = 2.9416$  and the corresponding order quantity  $Q^{**} = 1.2004$ . In order to attain this retail price, the manufacturer needs to utilize the retailer's reaction function (36) to determine his wholesale price:  $w^{**} = 0.6133$ . The joint optimal solution and the related quantities are also presented in Table 1. The background of Figure 4 shows the contour graph of the joint profit function, the maximum of which occurs at point B by definition. Note in the table that, as a result of coordination, the order quantity is more than doubled, and the joint channel profit is almost doubled as well.

However, at this demand level  $D^{**} = 0.4117$ , the selfish retailer's optimal EOQ is  $Q^*(w^{**}) = 0.7409$  (Point C in Figure 4) which follows her best reaction function (37), instead of the channel-optimal quantity  $Q^{**} = 1.2004$ . The retailer's profit at this best reaction order quantity is greater than the coordinated one:  $\Pi_R(w^{**}, Q^*(w^{**})) = 0.7363$  whereas  $\Pi_R(w^{**}, Q^{**}(w^{**})) = 0.7099$ . In order to keep the retailer's order quantity at  $Q^{**}$ , the manufacturer needs to charge per-order surcharge for any order less than  $Q^{**}$ .

From the initial point of the uncoordinated solution  $(w^*, Q^*)$ , the pricing can be viewed as a quantity discount. The simplest discount schedule can be designed with a single break such that the manufacturer charges  $w^*$  if the order quantity is less than  $Q^{**}$ , and  $w^{**}$  for  $Q^{**}$  or higher. Then it is the retailer's best interest to choose an order quantity  $Q^{**}$ . On the other hand, a linear discount schedule (all-unit) can be designed such that the unit price is  $w^*$  for any quantity less than  $Q^*$ , and after that, the price decreases by the rate of  $\frac{w^* - w^{**}}{Q^{**} - Q^*(w^*)}$  up to  $Q^{**}$ . The wholesale price  $w^{**}$  remains the

same for order quantities greater than  $Q^{**}$ . In our example, the wholesale price is  $w = 2.6816$  if the retailer orders less than or equal to  $Q = 0.5093$  units,  $w = 2.6816 - 2.9928(Q - 0.5093)$  between order quantity  $Q = 0.5093$  and 1.2004, and  $w = 0.6133$  thereafter. Any  $(w, Q)$  combination under this schedule provides a higher retailer profit than the uncoordinated solution, since the schedule is below the retailer's status quo iso-profit curve. Using either quantity discount schedule, the retailer profit is

maximized at the maximum discount offered at  $Q^{**}$ , where the joint profit is also maximized.

Table 1 shows that, the manufacturer profit can even go negative, because the manufacturer cannot recover his total cost at this wholesale price.<sup>6</sup> Therefore, he has to rely on the periodic franchise fee for a positive profit. The franchise fee can be as high as the retailer profit difference between the coordinated and uncoordinated cases: Maximum per-unit franchise fee =  $\Pi_R^{**} - \Pi_R^* = 0.6009$ . In this case, the manufacturer takes all the efficiency gain. The manufacturer's profit under no coordination  $\Pi_M^*(w^*, Q^*)$  becomes the bottom-line profit for setting the minimum franchise fee: Minimum per-unit franchise fee =  $\Pi_M^* - \Pi_M^{**} = 0.3493$ . Here, the retailer takes all the efficiency gain, which amounts to the difference between the two limit franchise fees: Total efficiency gain=0.2516 (one can obtain the same total gain by taking the difference between the joint profits at the bottom of Table 1). Multiplied by the demand level  $D^{**} = 0.4117$ , the total maximum and minimum franchise fees are 0.2474 and 0.1438, respectively. Figure 4 also depicts these franchise fees. The level of franchise fee between these two limits determines the division of the total efficiency gain between the channel members.

## 6. Conclusion

In this paper, we investigated an incentive-compatible quantity discount mechanism that coordinates actions between the channel members when both channel and transaction efficiencies are considered. One remarkable feature of this mechanism is

---

<sup>6</sup> It assumes an all-unit quantity discount. If an incremental quantity discount were used, the manufacturer profit would be higher because of the higher price for lower quantity breaks.

that it does not require the retailer to “cooperate” for the benefit of the whole channel. While cooperation may leave an incentive for a channel member to deviate from the prescribed action, an incentive-compatible coordination is a self-regulating mechanism.

In our quantity discount model, the manufacturer designs the wholesale pricing schedule to induce the retailer to set up the retail price as well as the order quantity to the levels that maximize the total channel profit. The channel efficiency stems from the increased market demand, while the transaction efficiency comes from minimizing the operating costs. When both efficiencies are combined, the model becomes substantially complicated, and closed-form solutions are difficult to derive. However, we were able to provide basic intuitions of the coordination mechanism using a graphical interpretation. Our numerical illustration shows that such a discount schedule can be practically designed once the demand and cost parameters are known.

On the other hand, our model is limited by the assumption of a deterministic demand function. In practice, the demand can be stochastic, and the inventory policy can be substantially complicated with various assumptions such as stock-out, back order, safety stock, and lead time. Generalizing our basic model by including these features would be a fruitful future research topic. Also, we considered only a single retailer case. Extending the model into multiple retailers with different demand and cost structure would be a challenging but rewarding research topic that increases practicality of our current model.

## References

- Chakravarty, Amiya K.** (1984) "Joint, Inventory Replenishment with Discounts Based on Invoice Value," *Management Science*, 30 (9), 1105-1112.
- Chakravarty, Amiya K. and G.E. Martin** (1989) "Discount Pricing Policies for Inventories Subject to Declining Demand," *Naval Research Logistics*, 36 (1), 89-102.
- Dada, Maqbool and K. N. Srikanth** (1987), "Pricing Policies for Quantity Discounts," *Management Science*, 33, 10, 1247-1252.
- Dolan, Robert J.** (1987), "Quantity Discounts: Managerial Issues and Research Opportunities," *Marketing Science*, 6, 1, 1-22.
- Ingene, Charles A. and Mark E. Parry** (1995), "Coordination and Manufacturer Profit Maximization: The Multiple Retailer Channel," *Journal of Retailing*, 71, 2, 129-151.
- Jeuland, Abel, P. and Steven M. Shugan** (1983), "Managing Channel Profits," *Marketing Science*, 2, 3, 239-272.
- Joglekar, Prafulla N.** (1988), "Comments on 'A Quantity Discount Pricing Model to Increase Vendor Profits,'" *Management Science*, 34, 11, 1391-1398.
- Kim, Kap Hwan and Hark Hwang** (1989), "Simultaneous Improvement of Supplier's Profit and Buyer's Cost by Utilizing Quantity Discount," *Journal of the Operations Research Society*, 40, 3, 255-265.
- Kohli, Rajeev and Heungsoo Park** (1989), "A Cooperative Game Theory Model of Quality Discounts," *Management Science*, 35, 6, 693-707.
- Lal, Rajiv and Richard Staelin** (1984), "An Approach for Developing an Optimal Discount Policy," *Management Science*, 30, 12, 1524-1539.
- McGuire, Timothy W. and Richard Staelin** (1986), "Channel Efficiency, Incentive Compatibility, Transfer Pricing, and Market Structure: An Equilibrium Analysis of Channel Relationships," in *Research in Marketing: Distribution Channels and Institutions* (L.P. Bucklin and J.M. Carman eds.), V 8, Greenwich, CT: JAI Press, 181-223.
- Monahan, James P.** (1984), "A Quantity Discount Pricing Model to Increase Vendor Profits," *Management Science*, 30, 6, 720-726.
- Moorthy, Sridhar** (1987), "Managing Channel Profits: Comment," *Marketing Science*; 6, 4, 375-379.
- Weng, Z. Kevin** (1995), "Channel Coordination and Quantity Discounts," *Management Science*, 41, 9, 1509-1522.

## Appendix

### Proof of Concavity of the Retailer Profit Function (25)

The Hessian matrix of  $\Pi_R$  is

$$\begin{bmatrix} 2D' + \frac{(pQ - S_R - wQ)D''}{Q}, & \frac{S_R D'}{Q^2} \\ \frac{S_R D'}{Q^2}, & -\frac{2S_R D}{Q^3} \end{bmatrix}. \quad (\text{A1})$$

Unless  $D''$  is large positive, we have the diagonal elements all negatives. The determinant of the Hessian matrix is

$$-S_R [S_R D'^2 + D\{4QD' + 2((p-w)Q - S_R)D''\}] / Q^4. \quad (\text{A2})$$

It is easy to see that the term in the square bracket in the numerator is the second order condition of  $\Pi_R$  with respect to  $p$ , which is negative if the retailer profit is concave in price. Thus, under this condition, the determinant of the Hessian matrix is positive.

Since its principal minors alternate in sign starting from negative, the Hessian is negative definite, and the retailer profit function (25) is concave. QED

### Proof of Lemma 1: Comparative Statics of Optimal Retail Price and Order Quantity in Wholesale Price

Total differentiation of the first order conditions (26) and (27) with respect to  $w$  results in

$$-D'(p) + \frac{S_R D'(p)}{Q^2} \frac{dQ}{dw} + (2D'(p) - \frac{S_R D''(p)}{Q} + (p-w)D''(p)) \frac{dp}{dw} = 0, \text{ and} \quad (\text{A3})$$

$$-\frac{2S_R D(p)}{Q^3} \frac{dQ}{dw} + \frac{S_R D'(p)}{Q^2} \frac{dp}{dw} = 0. \quad (\text{A4})$$

Simultaneously solving (A1) and (A2) for  $\frac{dp}{dw}$  and  $\frac{dQ}{dw}$ , we obtain

$$\frac{dp}{dw} = \frac{2QD(p)D'(p)}{D'(p)[4QD(p) + S_R D'(p)] + 2[(p-w)Q - S_R]D(p)D''(p)}, \quad (\text{A5})$$

$$\frac{dQ}{dw} = \frac{Q^2 D'(p)^2}{D'(p)[4QD(p) + S_R D'(p)] + 2[(p-w)Q - S_R]D(p)D''(p)}. \quad (\text{A6})$$

Note that the denominators are negative of the numerator of (A2), hence are negative in sign under the conditions stated above. Therefore, it is clear that the retail price rises and the order quantity decreases as the wholesale price rises as long as the demand function is not highly convex (i.e., large positive  $D''(p)$ ) and the retailer's order cost is not very

large. Therefore,  $\frac{dp}{dw} > 0$  and  $\frac{dq}{dw} < 0$ . QED

In a special case of a linear demand, we have  $D''(p) = 0$ , they become

$$\frac{dp}{dw} = \frac{2QD(p)}{4QD(p) + S_R D'(p)}, \quad (\text{A7})$$

$$\frac{dQ}{dw} = \frac{Q^2 D'(p)}{4QD(p) + S_R D'(p)}. \quad (\text{A8})$$

To ensure vertical strategic complementarity (i.e., the retail price goes up when the wholesale price rises), the condition is to have  $S_R < \frac{-4QD(p)}{D'(p)}$ : i.e., the retailer's order set up cost is not too high. Under this condition, we have  $\frac{dp}{dw} > 0$  and  $\frac{dq}{dw} < 0$ . QED

## Proof of Proposition 2

Compare the retailer's profit maximization condition (29) (repeated here):

$$\frac{\partial \Pi_R}{\partial p} = D(p) + (p-w)D'(p) - D'(p)\sqrt{\frac{H_R S_R}{2D(p)}} = 0, \quad (\text{A9})$$

with the first-order condition of the joint profit maximization (32) (repeated here):

$$\frac{\partial \Pi_J}{\partial p} = D(p) + (p-c)D'(p) - D'(p)\sqrt{\frac{S_J H_J}{2D(p)}} = 0. \quad (\text{A10})$$



Note that the only difference between these two equations is the last term and the unit cost. In our previous discussion, we assumed that  $S_J H_J > S_R H_R$ . Since  $D'(p) < 0$ , the solution to (A9) is greater than the solution to (A10) at any  $w \geq c$ , including  $w^*$ . QED

### Proof of Lemma 3

Rearranging the denominator of the derivative in (A5), we have:

$$\frac{dp}{dw} = \frac{2QD(p)D'(p)}{4QD(p)D'(p) + S_R D'(p)^2 + 2[(p-w)Q - S_R]D(p)D''(p)}. \quad (\text{A11})$$

The first term of the denominator is two times the numerator, and the remaining terms are nonnegatives. Therefore,  $\frac{dp}{dw} \leq \frac{1}{2}$ . QED

### Proof of Proposition 5

By definition,

$$\Pi_R^* \equiv \Pi_R(p^*, Q^*) = (p^* - w^*)D^* - \sqrt{2S_R H_R D^*}, \text{ and} \quad (\text{A12})$$

$$\Pi_R^{**} \equiv \Pi_R(p^{**}, Q^{**}) = (p^{**} - w^{**})D^{**} - \sqrt{S_J H_J D^{**} / 2(H_R / H_J + S_R / S_J)} \quad (\text{A13})$$

First, we derive the incremental operating cost when the wholesale price decreases to  $w^{**}$ . Using the definitions of  $x$  and  $y$ , the last term of equation (A13) can be rewritten as

$\sqrt{2S_R H_R D^{**}} (x+y) / 2\sqrt{xy}$ . We have defined  $R = (x+y) / 2\sqrt{xy}$ , and it is easy to verify that  $R \geq 1$  for any real positive values of  $x$  and  $y$ . The more different  $x$  and  $y$  are, the larger is the value of  $R$ . Thus, for  $\Pi_R^{**} \geq \Pi_R^*$ , we have the following inequalities:

$$\begin{aligned} (m^* D^{**} - m^* D^*) &\geq m^* (D^{**} - D^*) \\ &\geq \sqrt{2S_R H_R D^{**}} R - \sqrt{2S_R H_R D^*} = (R-1)\sqrt{2S_R H_R} (\sqrt{D^{**}} - \sqrt{D^*}), \end{aligned}$$

which can be reduced to  $m^* (\sqrt{D^{**}} + \sqrt{D^*}) \geq (R-1)\sqrt{2S_R H_R}$ . QED

### Proof of Decreasing Iso-Profit Curve for the Manufacturer

Totally differentiating the manufacturer's profit (24) with respect to  $Q$ , we have

$$\frac{d\Pi_M}{dQ} = (1/2Q^2)[2D(S_M + Q^2 \frac{dw}{dQ}) - Q(H_M Q - 2((w-c)Q - S_M) \frac{dw}{dQ} D'p')]. \quad (A14)$$

Solving  $\frac{d\Pi_M}{dQ} = 0$  for  $\frac{dw}{dQ}$  yields

$$\frac{dw}{dQ} = \frac{H_M Q^2 - 2S_M D(p(w))}{2Q^2 D(p(w)) + 2Q((w-c)Q - S_M) D'(p(w)) p'(w)}. \quad (A15)$$

The denominator, which is the first order condition of maximum manufacturer profit, is zero by definition at the uncoordinated order quantity  $Q^*$ , and is negative for  $Q > Q^*$ . The numerator is in the same form as the first order condition of the retailer's EOQ (i.e., minimum operating cost). We know that  $H_R Q^2 - 2S_R D$  is zero at  $Q^* = \sqrt{2S_R D / H_R}$ . Since we assume that  $S_M / H_M > S_R / H_R$ , the numerator is negative within the range of  $Q^* < Q < \sqrt{2S_M D / H_M}$ , and positive beyond that. Therefore, the manufacturer's iso-profit curve has  $w$  as a decreasing function of  $Q$  within this range and then increasing there after. QED