Supply Chain Coordination and Assortment Planning *

Goker Aydin
Warren H. Hausman

Department of Management Science and Engineering
Stanford University

November 22, 2002

Abstract

In this paper we incorporate issues of assortment planning and consumer choice in a model of delayed differentiation. We model the delayed differentiation environment using Eppen and Schrage’s (1981) depot-warehouse system. The consumer choice model used in this paper follows the multinomial logit model (MNL) as used by van Ryzin and Mahajan (1999). In this setting, our emphasis is on supply chain coordination in assortment planning. In a decentralized supply chain with a single retailer and a single manufacturer, double marginalization and decentralization of inventory management result in a discrepancy between the assortment of products offered by the retailer and the assortment that maximizes total supply chain profits. This raises the question of how to achieve coordination in a decentralized supply chain, given that a coordination mechanism needs to address conflicts due to not only inventory management but also assortment planning. We first show the intuitive result that the optimal assortment of the centralized supply chain consists of a number of end products that have the highest customer appeal. This result can be interpreted as an extension of van Ryzin and Mahajan’s result to a delayed differentiation setting. Next, we show that a centralized supply chain offers at least as many end products as a decentralized one, and we examine how a decentralized supply chain can be coordinated so as to match a centralized supply chain’s performance. We consider a payment scheme that involves a per-product fee paid by the manufacturer to the retailer for every product offered by the retailer in excess of a certain target level. If the wholesale price is below some threshold level, this payment scheme implemented along with delayed differentiation

*This research was supported by Ericsson AB under the Ericsson Supply Chain Research Grant to the Department of Management Science and Engineering at Stanford University.
induces the retailer to offer the supply-chain-optimal number of end products. Furthermore, such a coordination mechanism makes both parties better off. Our results shed some light on whether slotting fees used in the grocery industry may enhance supply chain efficiency or not.

1 Introduction

Choosing from a variety of products that essentially serve the same purpose is part of the shopping routine of consumers. Similarly, manufacturers and retailers have to make routine decisions about what level of product variety to offer their customers. In this paper we incorporate issues of assortment planning and consumer choice in a model of delayed differentiation. This model augments earlier work on delayed differentiation by treating the set of products offered to the customers as an explicit decision. This kind of assortment planning problem arises in situations where a manufacturer markets a product under multiple brand names. For example, a manufacturer that supplies store brands in addition to marketing its own national brand may choose to hold some inventory of the generic product and delay the package labeling. Similarly, manufacturers that sell the same product in different package sizes may hold inventory of the generic product and delay the packaging operation. However, it should be noted that we make the following two assumptions that particularly simplify our model compared to the actual occurrences of the problem:

- The manufacturer does not keep inventories of the generic product; inventory is kept only at the end product level.

- The cost and price parameters are the same for all end products.

In this setting, our emphasis is on supply chain coordination in assortment planning. In a decentralized supply chain with a single retailer and a single manufacturer, double marginalization and decentralization of inventory management result in a discrepancy between the assortment of products offered by the retailer and the assortment that maximizes total supply chain profits. This raises the question of how to achieve coordination in a decentralized supply chain, given that a coordination mechanism needs to address conflicts
due to not only inventory management but also assortment planning. We model the delayed differentiation environment using Eppen and Schrage’s (1981) depot-warehouse system. The consumer choice model used in this paper follows the multinomial logit model (MNL) as used by van Ryzin and Mahajan (1999) who examine the inventory management and assortment planning problem faced by a retailer. We address the following two questions:

- How does the optimal assortment offered by the retailer in a decentralized supply chain differ from that offered by a centralized supply chain?

- Does there exist a coordination mechanism that will induce the retailer to offer the supply-chain-optimal assortment while making both parties better off?

We first show that a centralized supply chain offers at least as many end products as a decentralized one. Next, we consider a payment scheme that involves a per-product fee paid by the manufacturer to the retailer. This payment scheme requires that the manufacturer pay a per-product fee to the retailer for every product the retailer offers in excess of a certain target level. We refer to this payment scheme as ‘incremental fees.’ We show that if the wholesale price is below a threshold level, then incremental fees implemented along with delayed differentiation induce the retailer to offer the supply-chain-optimal assortment. Furthermore, we show that incremental fees guarantee that both the retailer and the manufacturer will be better off, provided that the wholesale price is below the threshold level.

Our results have some implications regarding the usage of slotting fees in the grocery industry. Slotting fees are one-time payments that some retailers require from the manufacturers in order to carry a new product offered by the manufacturer. According to Bloom et. al. (2000), there are two distinct opinions on slotting fees: Proponents argue that slotting fees are tools that enhance channel efficiency, while opponents maintain that slotting fees are anti-competitive and merely a source of extra revenue for retailers. One of the arguments in favor of slotting fees cited by Bloom et. al. is that these fees induce retailers to offer products that would otherwise not make it to the market because of an oversaturation of the
market with product proliferation. Our results partially support this claim: We find that the supply-chain-optimal level of variety is higher than that the retailer would be willing to offer, but a payment scheme that resembles slotting fees may induce the retailer to offer the number of products that is optimal for the entire supply chain, while making both the retailer and the manufacturer better off. However, our results also suggest that the opponents of slotting fees are not entirely off the mark: Depending on the magnitude of the wholesale price relative to production cost and retail price, a payment scheme similar to slotting fees may merely result in improving the retailer’s profits without inducing the retailer to take supply-chain-optimal actions.

In the next section, we discuss the related literature and position our model with respect to existing work on product line selection, delayed differentiation and supply chain coordination. This is followed by a discussion of the centralized supply chain in Section 3. Section 4 describes how the decentralized supply chain differs from the centralized one. In Section 5, we examine how a decentralized supply chain can be coordinated so as to match a centralized supply chain’s performance. The effect of some critical assumptions on our results is discussed in Section 6. Finally, we conclude with a discussion of future research directions in Section 7.

2 Related Literature

There are three main streams of research that are related to the problem being studied in this paper:

1. Product line selection and pricing

2. Delayed differentiation and postponement

3. Supply chain coordination
Product Line Selection and Pricing

The majority of the work in product line selection and pricing is in the marketing literature. The work by Mussa and Rosen (1978) and Moorthy (1984) seem to be the first treatments of the problem. They focus mainly on how cannibalization within a firm’s product line affects the product offerings and the prices charged. Dobson and Kalish (1988) assume that the firm already has a set of products that she has to choose from. All problem parameters are assumed to be known and deterministic. They formulate the product line selection and pricing problem as a mixed integer linear program, and devise heuristics for solving this problem. Having primarily a marketing and economics point of view, most of the work in product line selection and pricing deals with linear production costs. In particular, issues related to inventory management are ignored.

There is also some research on product line selection and pricing in the existence of design and manufacturing costs. Most of the work in this group follows the approach taken in Dobson and Kalish (1988), i.e., modeling the product line selection and pricing problem as a mixed integer linear program, and they enrich this model in order to account for manufacturing complexity associated with product variety. We refer the reader to Yano and Dobson (1998) for a detailed review of this literature. In contrast to this line of research, our model is highly stylized with stochastic demand. In this paper, we do not model the manufacturing complexity associated with product variety, we assume the existence of delayed differentiation capability. Also, we do not address the question of how to price the end products, since we assume the prices and the costs for the products are given, and they are the same for all end products.

Van Ryzin and Mahajan (1999) consider the assortment planning and inventory management problems of a retailer in conjunction. They model a situation where sales volume grows with variety. However, increased variety also results in increased fragmentation of the inventory, leading to higher inventory costs, i.e., a reverse-pooling effect. They find that the optimal assortment consists of a number of products that are most appealing to the consumer population, i.e., a number of products that provide the consumer population with
the highest average utility. The demand model in this paper follows that of van Ryzin and Mahajan. Like van Ryzin and Mahajan, we assume that prices and costs are fixed and the same for all end products. In fact, our result on the form of the optimal assortment for the centralized supply chain is simply an extension of van Ryzin and Mahajan’s main result to another supply process, namely Eppen and Schrage’s depot-warehouse system.

Chen, Eliashberg and Zipkin (1998) consider a product line design and pricing problem using a locational consumer choice model, i.e., a choice model in which each consumer chooses the product that is closest to his or her ideal product in a space of product attributes. The cost associated with offering a product line is a general function of the demand for the products offered to the customers. One particular model they consider involves Poisson arrivals of demands that form a queue where the leadtime for fulfilling a demand is exponentially distributed. This is similar to our supply model in spirit, since the procurement of end products is dependent on common resources in both this model and our model, the common resource in our model being the generic product that is customized into specific end products. They develop a necessary condition for an optimal product line and discuss how the optimal product line can be found using numerical techniques.

**Delayed Differentiation and Postponement**

The literature on delayed differentiation and postponement takes product variety decisions as given. Lee and Tang (1997) consider a situation where the production process consists of a number of stages each of which holds inventory, and there is a single point of differentiation. The authors then consider a number of scenarios motivated from real-life examples, and they analyze the profitability of delayed differentiation. Garg and Tang (1997) consider a problem in which there are multiple points of differentiation. The leadtime of a manufacturing stage is the time it takes the production process to move from one point of differentiation to the one that succeeds it. The authors define postponement of a given stage as reducing the leadtime of that stage by one period while increasing the previous stage’s leadtime by one period. They evaluate the benefits of postponement for two cases, one in which only
finished goods inventory is held, the other in which both finished goods and work in process
inventory are held. Unlike these earlier works, we do not address the question of how to set
the last point of differentiation. Rather, we assume that the last differentiation point is fixed.
However, our work augments earlier literature in that we treat the product variety offered
as an explicit decision. This also allows us to show a benefit of delayed differentiation that
was not identified earlier: Implementing delayed differentiation not only reduces inventory
costs, but also improves revenues by inducing the firm to offer higher variety.

Lee (1996) discusses different ways of modeling inventory problems in product/process
design, and identifies Eppen and Schrage’s depot-warehouse model as one way of modeling
delayed differentiation in build-to-stock environments. In this paper, we use this approach
in modeling delayed differentiation, which simplifies the inventory related aspects of the
problem, and allows us to focus on supply chain coordination issues.

Supply Chain Coordination

There is a growing and already rich literature on supply chain coordination. For reviews of
the literature on supply chain coordination and contracting from different perspectives, see
and Agrawal (1998). The premise of supply chain coordination is that decentralized decision-
making and ownership in a supply chain result in inferior performance of the chain compared
to one where ownership and decision-making are centralized. In this paper, we focus on the
effect of decentralization on the assortment of products offered to the customers. To the
best of our knowledge, the only work that has considered the inefficiency introduced to the
product line selection activity as a result of decentralized decision-making is that of Villas-
Boas (1998). He focuses on how the assortments offered by the centralized and decentralized
supply chains are different, and does not address in detail the issue of how a decentralized
supply chain can be coordinated. The problem faced by the centralized channel of Villas-
Boas corresponds to the product line selection and pricing problem modeled in Moorthy
(1984). Villas-Boas considers the decentralized case in which the manufacturer decides what
product line to offer to the retailer, and the retailer decides which of these products to carry in stock. The manufacturer charges the retailer some transfer price for each product, and the retailer is free to set the price the end customers will pay. In this scenario, the manufacturer acts as the Stackelberg leader in offering a product line to the retailer. The author finds the equilibrium solution for the optimal product line and the optimal transfer and retail prices. In this model all demand is met, so there are no inventory issues, and the only source of channel misalignment is double marginalization. We model a problem with inventory costs, and focus on how to achieve coordination. However, we do not deal with the pricing aspect of the problem. Also, in our decentralized model, we assume that there is a given set of potential end products that the retailer chooses from, and we do not allow for gaming in forming this set of potential end products.

3 The Centralized Supply Chain

In this section, we investigate the assortment planning and inventory management problem of the centralized supply chain. In the centralized supply chain, a single decision-maker that will be referred to as ‘the firm’ both manufactures the end products and sells the assortment to the end customers.

The Inventory Model

The firm has $I$ potential end products each of which could be included in the assortment. All these end products serve the same general purpose for the end customer, so the end customer chooses one of these end products in case she decides to make a purchase. Each end product is obtained by customizing a generic product. The customization takes $l$ periods for all end products, and the leadtime for procuring the generic product is $L$ periods. We assume that either there is an outside supplier that has an ample stock of the generic product or the firm has ample capacity for the production of the generic product. Inventory is held only at the end product level. There is no inventory of work in process except those items that are in manufacture or in transit. In particular, we assume that no inventory of the generic product
is kept. This assumption, although applicable in some cases, makes our model a special case of the more general problem in which the firm can hold inventory of the generic product. We assume a periodic-review model, and the timing of the events is as follows:

1. In the beginning of every period, the firm observes the inventory positions of the end products, and allocates the arriving inventory of the generic product to the end products, i.e., the firm begins customizing the arriving inventory of the generic product into end products.

2. In the beginning of every period, the firm places an order for the generic product in accordance with a base-stock policy.

3. The demand for end products is realized, and unmet demand is backlogged.

This model of delayed differentiation is simply an interpretation of the depot-warehouse system discussed in Eppen and Schrage (1981). The inventories of the end products in our model correspond to the inventory held at the warehouses in Eppen and Schrage, and the generic product corresponds to the inventory that arrives at the depot from the outside supplier.

We assume that the price and cost parameters are the same across all end products. We define the following notation:

\[ p = \text{unit selling price of the end products} \]
\[ c = \text{unit cost of the end products} \]
\[ b = \text{penalty cost per unit of backlogged demand for the end products} \]
\[ h = \text{unit inventory holding cost per period for the end products} \]

We assume that the holding cost incurred for an item during the manufacturing process is included in the unit cost. The assumption of identical price and cost parameters is clearly a restrictive one. This assumption would probably hold, for example, when the end products
are different flavors of a food product or different colors of a garment, etc. However, this assumption makes our model inappropriate for addressing quality or size differences in the assortment.

**The Consumer Choice Process and Demand Model**

We follow the ‘independent population’ model of van Ryzin and Mahajan (1999) closely, in which the purchasing decisions of individuals in the customer population are independent. Each arriving customer decides whether to purchase or not, and, if so, which end product to purchase. We let \( q_i(S) \) be the probability that an arriving customer chooses end product \( i \) given that the assortment being offered is \( S \subseteq \{1, \ldots, I\} \). We assume that the demand for end product \( i \) is normal with mean and variance \( \lambda q_i(S) \). Furthermore, we assume that the demand for end products are independent from each other. In effect, this is equivalent to a normal approximation of the Poisson arrival with rate \( \lambda \) of customers, each of which buys end product \( i \) with probability \( q_i(S) \). All results would extend to the case where demand for end product \( i \) is normally distributed with a mean of \( \lambda q_i(S) \) and variance of \( (\lambda q_i(S))^{2\beta} \) for \( 0 < \beta < 1 \). In order to model the choice process of a customer, we use the MNL model following van Ryzin and Mahajan.

The MNL model is a utility-based consumer choice model in which the utility of a consumer for end product \( i \), denoted by \( U_i \), is given by

\[
U_i = \alpha_i + \epsilon_i
\]

where \( \alpha_i \) is a fixed term and \( \epsilon_i \) is a random error term that has a Gumbel distribution with mean zero and shape parameter \( \mu \). In this consumer choice model, \( \alpha_i \) can be interpreted as the average utility of a customer for end product \( i \), and this is assumed to be the same for all customers. The random error term \( \epsilon_i \) can be interpreted as modeling the heterogeneity in the consumer population. \(^1\) We assume that, for any given customer, \( \epsilon_i \)'s are independent

---

\(^1\)In marketing literature, \( \epsilon_i \) is interpreted as modeling measurement error, and heterogeneity is modeled
across end products. Furthermore, for end product $i$, $\epsilon_i$ is assumed to be independent across customers.

The utility of a customer for not purchasing is denoted by $U_0$ and is given by $U_0 = \alpha_0 + \epsilon_0$, where $\alpha_0$ is a fixed term and $\epsilon_0$ is a random error term with a Gumbel distribution of mean zero and shape parameter $\mu$. Throughout the paper, when we refer to a customer choosing end product 0, we mean that the customer chooses not to purchase.

Each arriving customer makes a purchasing decision so as to maximize her utility. It follows from a standard result of the MNL model that

$$q_i(S) = \frac{e^{\alpha_i}}{\sum_{j \in S \cup \{0\}} e^{\alpha_j}}.$$  \hspace{1cm} (1)

Following the approach in van Ryzin and Mahajan, we let $v_j = e^{\alpha_j}$ for $j = 0, \ldots, I$, and we call $v_j$ the preference of the customer population for end product $j$. The higher the value of $\alpha_j$ is, the higher the value of $v_j$, i.e., the higher the average utility from end product $j$ is, the higher the customer population’s preference for end product $j$. Now, we can write $q_i(S) = \frac{v_i}{\sum_{j \in S \cup \{0\}} v_j}$ for $i \in S \cup \{0\}$. The key aspect of this choice model is that the more end products there are in assortment $S$, the smaller the probability that an arriving customer will not purchase any of the end products, i.e., as $S$ gets larger, $q_0(S)$ decreases.

Detailed reviews of the use of consumer choice models in managing product variety can be found in Lancaster (1998) or Mahajan and van Ryzin (1998). Here, we will only highlight some important limitations of the demand model used in this paper. First, the arrival rate is not a function of the variety offered by the manufacturer, and it would be an interesting extension to assume otherwise. Also, we assume that unmet demand is backlogged, the

by estimating a separate $\alpha_i$ for every individual. Therefore, MNL model as used here suffers from the so-called independence from irrelevant alternatives (IIA) property, i.e., any new alternative will affect the choice probabilities of existing alternatives to the same extent regardless of the degree of similarity between the existing alternatives and the new alternative. For a more detailed discussion of IIA, see Guadagni and Little (1983).
implication being that a customer may choose to purchase an end product that is out of stock. This assumption would hold when the customer makes the purchasing decision before observing the inventory status, for example, when buying from a catalogue. However, it is admittedly a restrictive assumption and only a stylized version of reality. Finally, if a customer decides to purchase, he or she purchases only one unit of the item.

The Optimal Assortment

As a preliminary, we first consider the inventory management problem faced by a firm whose product line is already determined. Let $S \subseteq \{1, \ldots, I\}$ denote the assortment offered by the firm. As mentioned in Section 3, the demand for end product $i \in S$ in a given period is normally distributed with mean $\lambda q_i(S)$ and standard deviation $\sqrt{\lambda q_i(S)}$ where $q_i(S) = \sum_{j \in S \cup \{0\}} v_j$ for $i \in S \cup \{0\}$. Furthermore, the demands for the end products are independent across time and independent from each other. Therefore, when the assortment is already determined, the inventory management problem reduces to the one in Eppen and Schrage.

In every period, the manufacturer needs to allocate the arriving amount of the generic product to the end products. Eppen and Schrage show that the optimal allocation is one for which the probability of a stockout $l$ periods later is equal for all end products. However, in order to obtain this result, they assume that such an allocation is feasible in any given period. The same ‘allocation assumption’ is made in this paper as well. Eppen and Schrage demonstrate that when the demand for end products have independent and identical normal distributions with mean $\mu$ and standard deviation $\sigma$, the probability that the allocation assumption holds is bounded below by $1 - \sum_{j \in S} \Phi \left( -\frac{\mu}{\sigma} \right)$ where $\Phi$ is the standard normal cdf. They then demonstrate that the probability that the allocation assumption holds is quite high when $\frac{\mu}{\sigma}$ is reasonably high. In our model, the demands for the end products are not identically distributed. The following proposition specifies the lower bound on the probability that the allocation assumption will be satisfied in our model.
**Proposition 1** Assume that the firm offers assortment \( S \subseteq \{1, \ldots, I\} \). Then, the probability that the allocation assumption holds is bounded below by

\[
1 - \sum_{j \in S} \Phi \left( -\frac{\sqrt{\lambda \sum_{i \in S} \sqrt{q_i(S)}}}{\Psi_j} \right)
\]

where

\[
\Psi_j^2 = \left( \frac{\sum_{i \in S} \sqrt{q_i(S)}}{q_j(S)} \right)^2 - 2 \frac{\sum_{i \in S} \sqrt{q_i(S)}}{\sqrt{q_j(S)}} + 2 \frac{\sum_{i \in S} q_i(S)}{q_j(S)}
\]

and \( \Phi \) is the standard normal cdf.

The proof of Proposition 1 is included in the Appendix.

An examination of the lower bound shows that the higher \( \lambda \), the higher the lower bound is. Now, using the approach of Eppen and Schrage, the optimal order-up-to level for the generic product, denoted by \( y^* \), is

\[
y^* = (L + l + 1) \lambda \sum_{i \in S} q_i(S) + \sigma_e \Phi^{-1} \left( \frac{b}{b + h} \right) \sum_{i \in S} \sqrt{q_i(S)}
\]

where

\[
\sigma_e^2 = L \lambda \sum_{i \in S} q_i(S) + (l + 1) \lambda \left( \sum_{i \in S} \sqrt{q_i(S)} \right)^2
\]

and \( \Phi \) is the standard normal cdf.

Note that the optimal order-up-to level for the generic product is equal to the expected sales of all end products over \( L + l + 1 \) periods plus a safety stock.

Let \( G(S) \) be the optimal expected inventory holding and shortage cost per period when the assortment offered is \( S \subseteq \{1, \ldots, I\} \). Using the results of Eppen and Schrage, we find
that

\[ G(S) = (b + h) \sigma \phi \left( \frac{b}{b + h} \right) \]

where \( \phi \) is the standard normal pdf. Throughout the rest of the paper, we let \( \phi \left( \frac{b}{b+h} \right) = \phi^* \).

The expected sales in a period is \( \lambda \sum_{i \in S} q_i(S) \). (Recall that we assume unmet demand is backlogged.) Since the contribution per item sold is \( p - c \), the expected gross profit per period is \( \lambda (p - c) \sum_{i \in S} q_i(S) \). Combining this with the inventory cost \( G(S) \), the net profit of the centralized supply chain per period for assortment \( S \subseteq \{1, \ldots, I\} \) is given as

\[
\Pi^c(S) = \lambda (p - c) \sum_{i \in S} q_i(S) - (b + h) \sigma \phi^*.
\]

(3)

The firm’s problem is to choose the assortment that maximizes profits:

\[
\max_{S \subseteq \{1, \ldots, I\}} \Pi^c(S).
\]

Let us first consider the problem faced by a firm who is currently offering assortment \( S \subseteq \{1, \ldots, I\} \), and is considering an expansion of the current assortment, i.e., the firm has to decide whether or not to offer one more end product, and, if so, which end product to choose out of a given set of candidates. It turns out, as one would expect, that the firm should either expand its product line with the candidate that has the highest customer appeal or not expand the product line at all. The following proposition formalizes this result.

**Proposition 2** Assume the firm is expanding an assortment \( S \subseteq \{1, \ldots, I\} \) with a new end product. Let \( \delta \) be the preference of a customer for the new end product. The profit of the manufacturer after expansion is quasi-convex in \( \delta \).

The proof of Proposition 2 is provided in the Appendix. This is an extension of the analogous result in van Ryzin and Mahajan: with \( L = 0 \), the problem would reduce to a case considered by van Ryzin and Mahajan. Once we know Proposition 2 holds, the following
result regarding the form of the optimal product line follows:

**Theorem 1** Assume the firm has $I$ end products to choose from in order to compose its assortment, and the end products are indexed so that $v_1 > v_2 > \ldots > v_I$. The optimal product line consists of the first $k$ end products for some $k \in \{1, \ldots, I\}$.

In plain language, Theorem 1 says that the optimal product line consists of a number of end products with the highest customer appeal. The proof of the theorem exploits the property established in Proposition 2 and is the same as the proof of Theorem 1 of van Ryzin and Mahajan, so we do not reproduce it here.

4 The Decentralized Supply Chain

In this section we consider a situation in which the manufacturer and the retailer are separate entities. We make the following assumptions regarding the structure of the decentralized supply chain:

- The retailer holds inventories of the end products and sells to the end customers.
- The retailer uses an order-up-to policy for replenishing its inventory of the end products.
- The manufacturer works in a build-to-order fashion. Therefore, the leadtime observed by the retailer is $L + l$ periods for all end products, i.e., the manufacturer’s delayed differentiation capability is of no use in the decentralized supply chain in the absence of any coordination.
- In addition to the previously defined price and cost parameters, we now let $w$ be the wholesale price per unit of end product, charged by the manufacturer to the retailer. Again, $w$ is assumed to be fixed and the same for all end products.
• The retailer selects the product line from a given set of $I$ end products. \footnote{It would be interesting to allow an extension in which the set of potential products is decided by a strategic manufacturer who foresees the profit-maximizing behavior of the retailer. In Section 6 we give a numerical example of how the results would change if this game-theoretic extension were allowed.}

Let $\Pi^d_r(S)$, $\Pi^d_m(S)$ and $\Pi^d_{sc}(S)$ denote the expected profit per period, respectively, for the retailer, the manufacturer and the supply chain when the decentralized supply chain offers assortment $S \subseteq \{1, \ldots, I\}$. From standard inventory theory, it is known that this infinite horizon problem decomposes into a separate problem for each end product. The optimal expected inventory holding and shortage cost per period is given by

$$(b + h) \sigma_d \phi^*,$$

where

$$\sigma^2_d = (L + l + 1) \lambda \left( \sum_{i \in S} \sqrt{q_i(S)} \right)^2$$  

This inventory holding and shortage cost is borne by the retailer. Therefore, the profit functions in the decentralized supply chain are given by

$$\Pi^d_r(S) = \lambda(p - w) \sum_{i \in S} q_i(S) - (b + h) \sigma_d \phi^*$$ \hspace{1cm} (5)  

$$\Pi^d_m(S) = \lambda(w - c) \sum_{i \in S} q_i(S)$$ \hspace{1cm} (6)  

$$\Pi^d_{sc}(S) = \lambda(p - c) \sum_{i \in S} q_i(S) - (b + h) \sigma_d \phi^*.$$ \hspace{1cm} (7)  

Since the retailer makes the assortment planning decision by considering its own profits only, the profit maximization problem of the retailer is

$$\max_{S \subseteq \{1, \ldots, I\}} \Pi^d_r(S).$$
Note that the optimization problem above is a special case of the optimization problem of the centralized supply chain. ($\Pi_r^d(S)$ is a restriction of $\Pi_r^c(S)$ to $L := 0$ and $l := L + l$.) Therefore, the optimal assortment has the same form as the one obtained in Section 3 for the centralized supply chain.

We assume for the rest of the paper that the end products are indexed so that $v_1 > v_2 > \ldots > v_l$. Due to the form of the optimal assortment, we need only consider assortments that contain end products 1 through $k$ for some $k \in \{1, \ldots, I\}$. Throughout the rest of the paper, we let $A_k$ denote the assortment of end products 1 through $k$.

The next proposition will be useful for proving our main result regarding how coordination affects the optimal product line offered to the end customers.

**Proposition 3** $\Pi^c(A_k) - \Pi_r^d(A_k)$ is increasing in $k$.

The proof of Proposition 3 is provided in the Appendix. Note that the profit functions $\Pi^c$ and $\Pi_r^d$ are maximized in the assortment planning problems of the centralized and decentralized supply chains, respectively.

**Theorem 2** Let $k_c$ and $k_d$ be the optimal number of end products offered by the centralized and decentralized supply chains, respectively. Then, $k_c \geq k_d$.

**Proof of Theorem 2:**

The proof is by contradiction. Suppose $k_d > k_c$. Then, from Proposition 3, it follows that 

$$\left(\Pi^c(A_{k_d}) - \Pi_r^d(A_{k_d})\right) - \left(\Pi^c(A_{k_c}) - \Pi_r^d(A_{k_c})\right) > 0.$$ 

Regrouping the terms, we can write

$$\left(\Pi^c(A_{k_d}) - \Pi^c(A_{k_c})\right) - \left(\Pi_r^d(A_{k_d}) - \Pi_r^d(A_{k_c})\right) > 0.$$ 

However, this leads to a contradiction, since $A_{k_d}$ maximizes $\Pi_r^d$ and $A_{k_c}$ maximizes $\Pi_r^c$. Therefore, we cannot have $k_d > k_c$, which concludes the proof. \qed
This result is not surprising. The reasons for the centralized supply chain offering at least as many end products as the decentralized supply chain are twofold. First, the profit margin per sale of the firm in the centralized supply chain is larger than the profit margin per sale of the retailer who makes the assortment planning decision in the decentralized supply chain, i.e., the double-marginalization effect. Second, the centralized supply chain enjoys reduced inventory costs due to the pooling effect created by delayed differentiation. This savings in inventory cost is absent in a decentralized supply chain. Therefore, the centralized supply chain suffers less from fragmentation of inventory due to increased variety, which induces the centralized supply chain to offer higher variety. Finally, we note that customers have more variety to choose from in the centralized supply chain.

5 Coordinating a Decentralized Supply Chain

As described in the previous section, there are two sources of channel misalignment in the decentralized model:

1. the double-marginalization effect;

2. lack of delayed differentiation.

Implementing Delayed Differentiation

A first step in coordinating the supply chain would be to implement delayed differentiation. In this context, the implementation of delayed differentiation can have two physical interpretations:

1. In the beginning of every period, the manufacturer obtains full knowledge of the inventory levels of the end products at the retailer. The manufacturer then places an order for the generic product in accordance with an order-up-to policy, and decides how to allocate the arriving order of the generic product to the end products in order to minimize the inventory costs.
2. In the beginning of every period, the retailer places an order for the generic product. When this order of the generic product arrives $L$ periods later, the retailer decides how the manufacturer should allocate the arriving inventory of the generic product to specific end products.

Both interpretations result in the same model where the inventory holding and shortage cost per period is given by Eppen and Schrage’s depot-warehouse model as discussed in Section 3. In both interpretations we require that the retailer continues to bear the inventory holding and shortage cost. In Section 6 we discuss how the results would change if delayed differentiation were implemented along with consignment. Let $\Pi_{r}^{dd}(S)$, $\Pi_{m}^{dd}(S)$ and $\Pi_{sc}^{dd}(S)$ denote the expected profit per period, respectively, for the retailer, the manufacturer and the supply chain when the decentralized supply chain with delayed differentiation offers assortment $S \subseteq \{1, \ldots, I\}$. These profit functions are given by

\[
\Pi_{r}^{dd}(S) = \lambda(p - w) \sum_{i \in S} q_{i}(S) - (b + h) \sigma_{c} \phi^{*} \tag{8}
\]

\[
\Pi_{m}^{dd}(S) = \lambda(w - c) \sum_{i \in S} q_{i}(S) \tag{9}
\]

\[
\Pi_{sc}^{dd}(S) = \lambda(p - c) \sum_{i \in S} q_{i}(S) - (b + h) \sigma_{c} \phi^{*}. \tag{10}
\]

with $\sigma_{c}$ defined in (2). Clearly, delayed differentiation in itself would not necessarily achieve supply chain coordination, since delayed differentiation does not address the double marginalization issue. Nevertheless, the implementation of delayed differentiation would result in both parties being not worse off at the very least. To see why this is true, observe that, with delayed differentiation, the retailer benefits from pooling effects and incurs reduced inventory costs. Consequently, the retailer will offer at least as many end products under delayed differentiation as in the decentralized supply chain, thereby making the manufacturer possibly better off. This result is summarized in the following corollary.

**Corollary 1** Let $k_{dd}$ be the optimal number of end products offered by the retailer in the
decentralized supply chain with delayed differentiation. Then:

1. \( k_c \geq k_{dd} \geq k_d \),

2. \( \Pi_r^{dd}(A_{k_{dd}}) \geq \Pi_r^d(A_{k_d}) \) and \( \Pi_m^{dd}(A_{k_{dd}}) \geq \Pi_m^d(A_{k_d}) \).

**Proof of Corollary 1:**

**Proof of 1:** The arguments in the proof of Proposition 3 can be used to show that, both \( \Pi_c^c(A_k) - \Pi_r^{dd}(A_k) \) and \( \Pi_r^{dd}(A_k) - \Pi_r^d(A_k) \) are increasing in \( k \). Then, the same argument in the proof of Theorem 2 applies, and we conclude that \( k_c \geq k_{dd} \geq k_d \).

**Proof of 2:** Since \( \sigma_c \leq \sigma_d \), we know that, for a given \( k \in \{1, \ldots, I\} \), \( \Pi_r^{dd}(A_k) \geq \Pi_r^d(A_k) \). Then, \( \Pi_r^{dd}(A_{k_{dd}}) \geq \Pi_r^d(A_{k_d}) \) follows from the fact that assortments \( A_{k_{dd}} \) and \( A_{k_d} \) maximize \( \Pi_r^{dd} \) and \( \Pi_r^d \), respectively. \( \Pi_m^{dd}(A_{k_{dd}}) \geq \Pi_m^d(A_{k_d}) \) follows from the facts that \( k_{dd} \geq k_d \) and \( \sum_{i=1}^k q_i(A_k) \) is increasing in \( k \), which was shown in the proof of Proposition 3. ■

The implication of the corollary is that delayed differentiation can be implemented without putting neither party at a disadvantage. However, in order to coordinate the supply chain, more needs to be done in order to address the double-marginalization issue.

**Implementing the Incremental Fee Scheme**

The next question we address is the following: Can we implement a coordination scheme where, in addition to implementing delayed differentiation, the manufacturer pays the retailer \( K \) per period, per end product offered by the retailer in excess of a certain target level, i.e., an *incremental fee scheme* is implemented, so that

1. the retailer will offer end products 1 through \( k_c \),

2. both the retailer and the manufacturer will be better off compared to implementing delayed differentiation only?

We will consider the case where the target number of end products is set as \( k_{dd} \), i.e., the manufacturer pays the retailer \( K \) per period for each end product offered in excess of \( k_{dd} \).
Also, we will impose a restriction on the retailer’s choice by requiring the retailer to offer end products 1 through $j$ for some $j \in \{1, \ldots, I\}$. In Section 6, we discuss why this requirement is necessary, and how results would change without this limitation. With incremental fees and delayed differentiation in place, the profit of a retailer that offers end products 1 through $j$ is given by

$$
\begin{cases}
\Pi_{r}^{dd} (A_j) + K (j - k_{dd}), j \geq k_{dd} \\
\Pi_{r}^{dd} (A_j), j < k_{dd}.
\end{cases}
$$

It is easy to observe that the retailer will never choose to offer less than $k_{dd}$ end products when incremental fees are implemented along with delayed differentiation. (This follows from the fact that assortment $A_{k_{dd}}$ maximizes $\Pi_{r}^{dd}$.) Given $p$ and $c$, when $w$ is sufficiently small, e.g., $w = c$, there will be no coordination problem once delayed differentiation is implemented, since we will have $k_{dd} = k_c$, i.e., if the difference between the retailer’s profit margin and the supply chain’s profit margin is sufficiently small, then the retailer will choose to offer the supply-chain-optimal assortment once delayed differentiation is implemented. Let $\tilde{w}$ be such that $k_{dd} = k_c$ for $w \in [c, \tilde{w}]$. (The existence of such a $\tilde{w}$ is guaranteed, since we know that $k_{dd} = k_c$ for $w = c$.) However, for $w$ sufficiently large, we will have $k_c > k_{dd}$, in which case the retailer needs to be given an incentive in order to offer the supply-chain-optimal assortment. For such sufficiently large $w$, incremental fees will induce the retailer to offer $k_c$ end products if the following set of inequalities hold:

$$
\Pi_{r}^{dd} (A_{k_c}) + K (k_c - k_{dd}) \geq \Pi_{r}^{dd} (A_k) + K (k - k_{dd})
$$

Alternatively, we can write the above set of inequalities as

$$
K \leq K(w) = \min_{k_c < k \leq I} \left\{ \frac{\Pi_{r}^{dd} (A_{k_c}) - \Pi_{r}^{dd} (A_k)}{k - k_c} \right\}, \quad (11)
$$

$$
K \geq K(w) = \max_{k_{dd} \leq k < k_c} \left\{ \frac{\Pi_{r}^{dd} (A_k) - \Pi_{r}^{dd} (A_{k_c})}{k_c - k} \right\}. \quad (12)
$$
The first inequality requires that $K$ is not so large that the retailer will offer more than $k_c$ end products. Likewise, the second inequality requires that $K$ is not so small that the retailer will offer less than $k_c$ end products. In order for such an incremental fee to exist at a given wholesale price $w$, we need $\overline{K}(w) \geq K(w) \geq 0$; then any $K \in [K(w), \overline{K}(w)]$ would induce the retailer to offer end products 1 through $k_c$. The following proposition characterizes the behavior of $\overline{K}(w)$ and $\underline{K}(w)$ with respect to $w$, when all other parameters, e.g., $p$ and $c$, are held constant.

**Proposition 4** Let $\bar{w}$ be such that $k_{dd} = k_c$ for $w \in [c, \bar{w}]$. Then:

1. $\overline{K}(w) \geq 0$ and $\underline{K}(w) \geq 0$ for $w \in (\bar{w}, p]$,

2. $\overline{K}(w) - \underline{K}(w)$ is decreasing in $w$ for $w \in (\bar{w}, p]$.

See the Appendix for the proof of Proposition 4. An example of how $\overline{K}(w)$ and $\underline{K}(w)$ change with $w$ is depicted in Figure 1. In this example $\lambda = 1$, $p = 70$, $c = 20$, $I = 10$, $b = 10$, $h = 1$ and $L = l = 2$. Given the set of preference values for products one through ten listed in Table 1, it turns out that $k_c = 4$. For this example, we have $k_c = k_{dd}$ when the wholesale price is less than roughly 30, i.e., $\bar{w} \approx 30$. Recall that $w \leq \bar{w}$ is the region where coordination is trivially achieved once delayed differentiation is implemented. (In this region we plot $\overline{K}(w)$ as zero. Note that $\underline{K}(w)$ is actually undefined in this region.) For any wholesale price larger than 30, the implementation of delayed differentiation is not enough to induce the retailer to offer the supply-chain-optimal assortment.

From Proposition 4, it is easy to see that, given $c$ and $p$, an incremental fee that will induce the retailer to choose the supply-chain-optimal number of end products can be found, provided that the wholesale price being charged is below some threshold level. The following theorem formalizes this result.

**Theorem 3** There exist $K \geq 0$ and $w^* \in [\bar{w}, p]$ such that if $w \leq w^*$ and the retailer receives $K$ per period per end product offered in excess of $k_{dd}$, then the retailer’s optimal assortment will be end products 1 through $k_c$. 
Proof of Theorem 3:
First, note that for any \( w \leq \tilde{w} \), the supply chain with delayed differentiation is trivially coordinated, since \( k_{dd} = k_c \) for \( w \leq \tilde{w} \) by definition of \( \tilde{w} \). Now, if \( w > \tilde{w} \), we may have two separate cases: If we have \( \lim_{w \to \tilde{w}^+} K(w) \geq \lim_{w \to \tilde{w}^+} K(w) \), then part 2 of Proposition 4 implies that we will have \( K(w) \geq K(w) \) for \( w < w^* \in [\tilde{w}, p] \), which in turn implies that such a supply chain can be coordinated by some incremental fee at all wholesale prices less than \( w^* \). On the other hand, if \( \lim_{w \to \tilde{w}^+} K(w) < \lim_{w \to \tilde{w}^+} K(w) \), then such a supply chain cannot be coordinated for any \( w > \tilde{w} \), since part 2 of Proposition 4 implies that we will have \( K(w) < K(w) \) for any \( w > \tilde{w} \). For such a supply chain, the only coordination that exists is the trivial case that arises when \( w \leq \tilde{w} \). 

Going back to our example in Figure 1, we see that \( K(w) \geq K(w) \) for wholesale prices less than roughly 53. Consequently, if the wholesale price being used is above 30 but below 53, this supply chain can be coordinated by the use of delayed differentiation along with an incremental fee scheme. In this region, unlike the region where the wholesale price is less than 30, delayed differentiation in itself would not be enough for inducing the retailer to offer the supply-chain-optimal variety. On the other hand, if the wholesale price being used in the supply chain is above 53, then there exists no incremental fee that will achieve coordination when implemented along with delayed differentiation.

One way to interpret this result is to say that the supply chain can be coordinated with a contract that specifies a wholesale price and a per-product, per-period incremental fee from the manufacturer to the retailer. The key aspect of the result summarized in Theorem 3 is that there exist wholesale prices other than the manufacturing cost \( c \) at which the supply chain can be coordinated with an incremental fee, i.e., an incremental fee scheme may achieve coordination even in the existence of double marginalization. (The only time this is not true is when \( \tilde{w} = w^* = c \), which is very unlikely to happen.) Another interesting aspect of this result is that the ability of incremental fees to coordinate the supply chain depends critically on what wholesale price is being used in the supply chain. Supply chains in which the
manufacturer receives a relatively lower portion of the profit margin are more likely to be coordinated by an incremental fee.

The Effects of Coordination on Profits

We can now examine, given a wholesale price that is being used in the supply chain, whether delayed differentiation in conjunction with an incremental fee scheme may achieve supply chain coordination while making both parties better off. It is easy to see that the retailer will be better off; the implementation of delayed differentiation already made the retailer better off as shown in Corollary 1, and the addition of a per-product fee from the manufacturer to the retailer cannot hurt the retailer. In order for the manufacturer to be better off as a result of coordination, we need $K$ to satisfy the following inequality:

$$K \leq K_m(w) = \frac{\Pi_{m}^{dd}(A_{k_c}) - \Pi_{m}^{dd}(A_{k_{dd}})}{k_c - k_{dd}}. \quad (13)$$

If this inequality holds, then the manufacturer would prefer paying the retailer $K$ for every end product offered in excess of $k_{dd}$ in order to have the retailer offer $k_c$ end products instead of $k_{dd}$, which would be offered in the absence of incremental fees. Now, in order for an incremental fee to exist that coordinates the supply chain while making both parties better off, there needs to exist a $K$ such that $K \leq K_m(w)$ (so that the manufacturer will be better off) and $\overline{K}(w) \geq K \geq \underline{K}(w)$ (so that the retailer will offer supply-chain-optimal variety.) It turns out that some of the $K$ values that satisfy $\overline{K}(w) \geq K \geq \underline{K}(w)$ will also satisfy $K \leq K_m(w)$. Consequently, as long as we can find an incremental fee that induces the retailer to offer the supply-chain-optimal variety, we will be able to find a coordination mechanism that will make both parties better off. The following theorem summarizes this result:

**Theorem 4** There exist $K \geq 0$ and $w^* \in [\bar{w}, p]$ such that if $w \leq w^*$ and the retailer receives $K$ per period per end product offered in excess of $k_{dd}$, then:
1. The retailer’s optimal assortment will be end products 1 through $k_e$.

2. Both the retailer and the manufacturer will be better off with this coordination scheme than in the decentralized supply chain with delayed differentiation only.

See the Appendix for the proof of Theorem 4. This result is rather surprising in that it shows that a manufacturer may find it beneficial to use a payment scheme similar to slotting fees even when the retailer is not particularly more powerful than the manufacturer. In particular, our result supports the idea that slotting fees may actually improve the manufacturer's profit by luring the retailer into offering a high variety level that would be suboptimal for the retailer in the absence of slotting fees. However, it must be cautioned that the effect of slotting fees on a supply chain is critically dependent on the wholesale price being used in the supply chain. Furthermore, in this model we do not address the controversial issue of how slotting fees affect competing manufacturers, and whether or not slotting fees may be anti-competitive.

It is usually asserted that slotting fees may achieve supply chain efficiency by serving as a signal of the manufacturer’s confidence in high demand for the prospective product. Lariviere and Padmanabhan (1997) investigate this issue in the context of introducing a single new product. They consider a setting where the manufacturer may be one of two types, the type of the manufacturer determining what the demand for the prospective product will be. The retailer does not know the type of the manufacturer and there is a fixed cost that the retailer will incur if she decides to carry the product. Lariviere and Padmanabhan find that unless there is both information asymmetry about the type of the manufacturer and a fixed cost for carrying the new product, the manufacturer will not choose to pay a slotting fee. Unlike Lariviere and Padmanabhan, we model a situation where there is cannibalization within a product line, and the manufacturer and the retailer share the same level of uncertainty about the demand for a prospective product. Our result suggests that even when there are no issues regarding demand signaling, slotting fees may still arise in the interest of both parties in the supply chain, if the retailer bears additional burden related to higher variety, which is modeled in the context of inventory costs in this paper.
6 Discussion of Some Critical Assumptions

In this section we discuss how the results would change if some critical assumptions were modified.

Let us first consider what would happen if the coordination scheme did not require that the retailer offer end products 1 through \( j \) for some \( j \in \{1,\ldots,I\} \). (Recall that products are numbered in descending order of preference values, and \( A_j \) denotes the assortment containing products 1 through \( j \).) The form of the optimal policy described in Theorem 1 does not imply that the best way to choose \( j \) products is to select the \( j \) products with the highest preference values, i.e., although the optimal solution will be assortment \( A_j \) for some \( j \in \{1,\ldots,I\} \), the optimal assortment of \( j \) products for some \( j \in \{1,\ldots,I\} \) is not necessarily \( A_j \). (See van Ryzin and Mahajan (1999) for a numerical example.) If we did not require the retailer to choose assortment \( A_j \) for some \( j \in \{1,\ldots,I\} \), then we might run into situations where an incremental fee scheme induces the retailer to offer \( k_c \) end products, but the retailer chooses an assortment of \( k_c \) end products other than \( A_{k_c} \). Consider the example depicted in Table 2.

In this example, \( k_{idd} = 1 \) and \( k_c = 2 \). By requiring the retailer to offer assortment \( A_j \) for some \( j \in \{1,\ldots,I\} \) and using an incremental fee scheme with \( K = 0.21 \), the coordination scheme induces the retailer to offer assortment \( A_2 \), which is the supply chain optimal assortment. If the retailer was not limited to assortments \( A_j \), then the retailer could have chosen to carry end products 1 and 5, which is an improvement over assortment \( A_2 \) for the retailer. The coordination scheme discussed in the previous section avoids such a possibility by limiting the retailer to offering \( A_j \) for some \( j \in \{1,\ldots,I\} \). Note that, even with such a limitation, the retailer is better off in the coordinated supply chain than in the decentralized one.

One may wonder how the results would change if we considered a payment scheme in which the manufacturer pays the retailer a per-product, per-period fee for all the products offered by the retailer. In such a case, the result summarized in Theorem 3 would continue to hold, i.e., we could still find a per-product, per-period fee that induces the retailer to offer \( k_c \) end products. However, such a payment scheme would no longer guarantee that the manufacturer will be better off. Note that with such a payment scheme, inequality (13)
would change to

\[
K \leq K_m(w) = \frac{\Pi_m^{dd}(A_{k_c}) - \Pi_m^{dd}(A_{k_{dd}})}{k_c}.
\]

so that the manufacturer would prefer paying the retailer \( K \) for every end product offered in order to have the retailer offer \( k_c \) end products instead of \( k_{dd} \). In Figure 2, \( \overline{K}(w) \) and \( \underline{K}(w) \) and the updated \( K_m(w) \) are plotted for the same example depicted in Figure 1. In this example there exists a range of wholesale prices where \( \overline{K}(w) \geq K \geq \underline{K}(w) \) can be satisfied without satisfying \( K \leq K_m(w) \), i.e., there exist wholesale prices at which the retailer cannot be induced to offer \( k_c \) end products without making the manufacturer worse off.

One critical assumption in our model is the assumption that there exists a given set of \( I \) end products from which the retailer makes her product line selection. Another approach would be to assume that the manufacturer has a set of \( I \) potential end products, and the manufacturer decides which of these \( I \) products to let the retailer choose from, i.e., the manufacturer acts as the Stackelberg leader in deciding what end products to offer to the retailer. In such a case, the manufacturer would not necessarily make all \( I \) products available to the retailer. Consider the example depicted in Table 3. In this example, the manufacturer decides which of the eight potential end products to make available to the retailer. If the manufacturer lets the retailer choose from all eight end products, the retailer will offer end products one through four in her assortment. However, if the manufacturer decides not to make product two available, then the retailer will offer end products 1, 3, 4, 5, and 6 in her assortment, which provides the manufacturer with a profit greater than the assortment of end products one through four. Therefore, with the data depicted in Table 3, the manufacturer will never make all products available to the retailer. Note that, for a centralized supply chain, it is never beneficial to limit the assortment planning to a reduced set of end products. Therefore, if the retailer does not have access to all the potential end products, the decentralized supply chain will not perform as efficiently as the centralized supply chain, and the manufacturer needs to be given some incentive for making the entire set of products
available to the retailer.

Another critical assumption in our model is the assumption that delayed differentiation is implemented without consignment. If delayed differentiation were implemented with consignment, i.e., if the manufacturer incurred the inventory holding and shortage cost instead of the retailer, then the retailer would choose to offer all end products available to her. This follows from the fact that sales volume increases with variety, and the only cost that is keeping the retailer from offering all available variety is the inventory cost. However, such an outcome is not necessarily desirable, unless the supply-chain-optimal solution happens to be to offer all end products.

7 Conclusion

In this paper we first considered the assortment planning problem faced by a firm that both manufactures and sells end products, and uses delayed differentiation. We followed van Ryzin and Mahajan (1999) in modeling the benefits from variety and we used Eppen and Schrage’s (1981) depot-warehouse system to model delayed differentiation. We showed that the optimal product line of the firm has a simple form: Offer a number of end products with the highest customer appeal. This first part of the paper is an extension of a result in van Ryzin and Mahajan (1999) to a delayed differentiation setting. Building on the results of this problem, we considered a simple supply chain coordination model: In the decentralized supply chain, a retailer keeps inventory of the end products and chooses what products to offer to customers. The manufacturer in the supply chain has the capability to practice delayed differentiation but builds to order in the decentralized scenario. We showed that the centralized chain offers at least as many products as the decentralized one, and that delayed differentiation makes both parties better off. We then considered the effects of coupling delayed differentiation with a payment scheme that requires the manufacturer to pay a fee to the retailer per end product the retailer carries in excess of a certain target level. We showed that if the wholesale price in the decentralized supply chain is below some threshold level, then such a payment scheme along with delayed differentiation will induce the retailer
to offer supply-chain-optimal number of end products. Furthermore, provided that such a
payment scheme exists, it is guaranteed to make both parties better off.

The models used in this paper have many simplifying assumptions. Consequently, there
exist several possible extensions to the problem considered here:

- Firstly, the assumption of equal price and cost parameters for the end products could
  be relaxed in order to allow for product lines in which the end products differ from
each other in quality, size, etc.

- It was assumed in this paper that there is a single point of differentiation in the
  production process. This assumption could be modified to allow for multiple points
  of differentiation. Each differentiation point would then correspond to an attribute
  over which the end products vary. The question would then be what levels of these
  attributes to allow the customers to choose from.

- The arrival rate of customers can be modeled to be a function of the variety offered by
  the seller.

- One can assume a capacity constraint on the production of the generic product or
  shelf-space constraints on the inventories of the end products.

- In this paper it was assumed that the customers made their purchasing choice without
  knowledge of the inventory levels of the end products. This assumption could be
  relaxed to allow the customers to first observe the inventory status and then make a
  purchasing decision. Furthermore, we assumed that unmet demand is backlogged, i.e.,
  if the end product chosen by the customer is out of stock, the customer is willing to
  wait. One could alternatively use a model in which customers buy, for example, their
  second choice, if their first choice is not in stock.

- One could use models other than Eppen and Schrage’s depot-warehouse system to
  model delayed differentiation. In particular, it would be interesting to consider a case
  where the manufacturer holds inventory of the generic product.
• When modeling coordination issues, we assumed that there exists a given set of potential products from which the retailer chooses. One could also consider a model where the manufacturer acts as the Stackelberg leader in choosing the set of potential end products to offer to the retailer.

• In order to gain a better understanding of the effects of slotting fees, it would be desirable to use a model that takes into account external competition from other manufacturers.

8 References


Appendix

Proof of Proposition 1:

Assume, without loss of generality, that $S = \{1, \ldots, k\}$. Eppen and Schrage show that when there are $k$ warehouses, i.e., when there are $k$ end products in the product line, and the demand distributions are stationary over time, the probability that the allocation assumption holds is bounded below by

$$1 - \sum_{j=1}^{k} P \left\{ \left[ \sum_{i=1}^{k} D_{i,t-1} - \sum_{i=1, i \neq j}^{k} D_{i,t+L-1} - D_{j,t+L-1} \left( 1 - \frac{\sum_{i=1}^{k} \sigma_i}{\sigma_j} \right) \right] < 0 \right\} \quad (14)$$

where $D_{i,t}$ is the demand for end product $i$ in period $t$ and $\sigma_i$ is the standard deviation of $D_{i,t}$. In our model, $D_{i,t}$ is normally distributed with mean $\lambda q_i$ and standard deviation $\sqrt{\lambda q_i}$. Moreover, the demands are independent across both time and end products. Therefore, in our model, for $j \in \{1, \ldots, k\}$,

$$\sum_{i=1}^{k} D_{i,t-1} - \sum_{i=1, i \neq j}^{k} D_{i,t+L-1} - D_{j,t+L-1} \left( 1 - \frac{\sum_{i=1}^{k} \sigma_i}{\sigma_j} \right)$$

is distributed normally with mean $\lambda \sqrt{q_j} \left( \sqrt{q_1} + \ldots + \sqrt{q_k} \right)$ and variance

$$\lambda q_j \left[ \frac{\left( \sqrt{q_1} + \ldots + \sqrt{q_k} \right)^2}{q_j} - 2 \left( \frac{\sqrt{q_1} + \ldots + \sqrt{q_k}}{\sqrt{q_j}} \right) + 2 \frac{q_1 + \ldots + q_k}{q_j} \right].$$

Letting

$$\Psi_j^2 = \frac{\left( \sqrt{q_1} + \ldots + \sqrt{q_k} \right)^2}{q_j} - 2 \left( \frac{\sqrt{q_1} + \ldots + \sqrt{q_k}}{\sqrt{q_j}} \right) + 2 \frac{q_1 + \ldots + q_k}{q_j}$$

we can write the probability that $\sum_{i=1}^{k} D_{i,t-1} - \sum_{i=1, i \neq j}^{k} D_{i,t+L-1} - D_{j,t+L-1} \left( 1 - \frac{\sum_{i=1}^{k} \sigma_i}{\sigma_j} \right) < 0$
as
\[
\Phi \left( -\frac{\sqrt{\lambda} \left( \sqrt{q_1} + \ldots + \sqrt{q_k} \right)}{\Psi_j} \right)
\]
where \(\Phi\) is the standard normal cdf. Substituting this into Equation (14) yields the desired result. □

Proof of Proposition 2:

Without loss of generality, assume that \(S = \{1, \ldots, k\}\). Letting \(\sum_{i=1}^k v_i = V\) and \(\sum_{i=1}^k \sqrt{v_i} = X\) and substituting from Equations (1) and (3), we can write the profit after expansion as
\[
\lambda(p - c) \frac{V + \delta}{v_0 + V + \delta} - (b + h) \phi^* \left[ L\lambda \frac{V + \delta}{v_0 + V + \delta} + (l + 1)\lambda \frac{(X + \sqrt{\delta})^2}{v_0 + V + \delta} \right]^{\frac{1}{2}}.
\]
By rearranging the terms, we obtain
\[
\frac{1}{v_0 + V + \delta} \left\{ \lambda(p - c)(V + \delta) - (b + h) \phi^* \left[ L\lambda(V + \delta) + (l + 1)\lambda \left( X + \sqrt{\delta} \right)^2 \right]^{\frac{1}{2}} (v_0 + V + \delta)^{\frac{1}{2}} \right\}.
\]
Since \(v_0 + V + \delta\) is linear in \(\delta\), if we can show that the term in curly brackets above is convex in \(\delta\), then we can conclude the profit after expansion is quasi-convex in \(\delta\). (See the reference to Mangasarian in van Ryzin and Mahajan (1999), Lemma 1.) In order to show the convexity of the term in curly brackets, it is enough to show that
\[
\left[ L\lambda(V + \delta) + (l + 1)\lambda \left( X + \sqrt{\delta} \right)^2 \right]^{\frac{1}{2}} (v_0 + V + \delta)^{\frac{1}{2}}
\]
is concave. Letting \(M = L\lambda(V + \delta) + (l + 1)\lambda \left( X + \sqrt{\delta} \right)^2\), what we need to show reduces to showing
\[
\frac{d^2 \left( \sqrt{M(v_0 + V + \delta)} \right)}{d\delta^2} < 0.
\]
Taking the second derivative and rearranging the terms yields

$$
\frac{d^2}{d\delta^2} \left( \sqrt{M(v_0 + V + \delta)} \right) = \frac{1}{2} M^{-\frac{1}{2}} \frac{d^2 M}{d\delta^2} (v_0 + V + \delta)^{\frac{1}{2}} - \frac{1}{4} M^{-\frac{3}{2}} (v_0 + V + \delta)^{-\frac{3}{2}} \left[ \left( \frac{d M}{d\delta} \right)^2 (v_0 + V + \delta)^2 - 2M(v_0 + V + \delta) \frac{d M}{d\delta} + M^2 \right].
$$

It is easy to show that $\frac{d^2 M}{d\delta^2} < 0$. Moreover,

$$
\left( \frac{d M}{d\delta} \right)^2 (v_0 + V + \delta)^2 - 2M(v_0 + V + \delta) \frac{d M}{d\delta} + M^2 = \left( M - \frac{d M}{d\delta} (v_0 + V + \delta) \right)^2 > 0.
$$

Therefore, we obtain $\frac{d^2}{d\delta^2} \left( \sqrt{M(v_0 + V + \delta)} \right) < 0$, concluding the proof. □

**Proof of Proposition 3:**

Throughout the proof, recall that the end products are indexed so that $v_1 > v_2 > \ldots > v_k > 0$. Substituting for $\Pi^e(A_k)$ and $\Pi^d(A_k)$ from Equations (3) and (5), we can write the difference of the two profit functions as

$$
\lambda (w - c) \sum_{i=1}^{k} q_i(A_k) + (b + h) \phi^*(\sigma_d - \sigma_c).
$$

First, we note that $\sum_{i=1}^{k} q_i(A_k) = \sum_{i=0}^{m} v_i - \sum_{i=0}^{m} v_i$, which is clearly increasing in $k$. Therefore, we only need to show that $\sigma_d - \sigma_c$ is increasing in $k$. Substituting for $\sigma_c$ and $\sigma_d$ from Equations (2) and (4) yields

$$
\sigma_d - \sigma_c = \sqrt{(L + l + 1) \lambda \sum_{i=1}^{k} \sqrt{q_i(A_k)}} - \left[ L \lambda \sum_{i=1}^{k} q_i(A_k) + (l + 1) \lambda \left( \sum_{i=1}^{k} \sqrt{q_i(A_k)} \right)^2 \right]^{\frac{1}{2}}.
$$

Rearranging terms, we can write the expression above as

$$
\sum_{i=1}^{k} \sqrt{q_i(A_k)} \left\{ \sqrt{(L + l + 1) \lambda} - \sqrt{(l + 1) \lambda} \left( \frac{L}{l + 1} \frac{\sum_{i=1}^{k} q_i(A_k)}{\left( \sum_{i=1}^{k} \sqrt{q_i(A_k)} \right)^2 + 1} \right)^{\frac{1}{2}} \right\}.
$$
Since $\sigma_d - \sigma_c \geq 0$, the term in curly brackets is always non-negative, and in order to show $\sigma_d - \sigma_c$ is increasing in $k$, all we need to show is that $\sum_{i=1}^{k} \sqrt{q_i (A_k)}$ is increasing in $k$ and $\frac{\sum_{i=1}^{k} q_i (A_k)}{\left( \sum_{i=1}^{k} \sqrt{q_i (A_k)} \right)^2}$ is decreasing in $k$. Let us first show that $\frac{\sum_{i=1}^{k} q_i (A_k)}{\left( \sum_{i=1}^{k} \sqrt{q_i (A_k)} \right)^2}$ is decreasing in $k$. In order to simplify the notation, let $X = \sum_{i=1}^{k} \sqrt{v_i}$ and $V = \sum_{i=1}^{k} v_i$. Substituting from Equation (1) yields

$$\frac{\sum_{i=1}^{k} q_i (A_k)}{\left( \sum_{i=1}^{k} \sqrt{q_i (A_k)} \right)^2} = \frac{\sum_{i=1}^{k} v_i}{\left( \sum_{i=1}^{k} \sqrt{v_i} \right)^2} = \frac{V}{X^2}.$$

Then:

$$\frac{\sum_{i=1}^{k+1} q_i (A_{k+1})}{\left( \sum_{i=1}^{k+1} \sqrt{q_i (A_{k+1})} \right)^2} - \frac{\sum_{i=1}^{k} q_i (A_k)}{\left( \sum_{i=1}^{k} \sqrt{q_i (A_k)} \right)^2} = \frac{V + v_{k+1}}{X^2 + 2X \sqrt{v_{k+1}} + v_{k+1}} - \frac{V}{X^2},$$

and all we need to show is $(V + v_{k+1}) X^2 - V \left( X^2 + 2X \sqrt{v_{k+1}} + v_{k+1} \right) < 0$. Now:

$$(V + v_{k+1}) X^2 - V \left( X^2 + 2X \sqrt{v_{k+1}} + v_{k+1} \right) = 2 \sum_{i=1}^{k} \sum_{j=i}^{k} \sqrt{v_i} \sqrt{v_j} \sqrt{v_{k+1}} - 2 \sum_{i=1}^{k} \sum_{j=1}^{k} v_i \sqrt{v_j} \sqrt{v_{k+1}}$$

$$< 0$$

where the equality follows from $X = \sum_{i=1}^{k} \sqrt{v_i}$ and the inequality follows from $v_1 > v_2 > \ldots > v_k$.

Next, we consider $\sum_{i=1}^{k} \sqrt{q_i (A_k)}$. Substituting from Equation (1) yields

$$\sum_{i=1}^{k} \sqrt{q_i (A_k)} = \frac{\sum_{i=1}^{k} \sqrt{v_i}}{\sqrt{\sum_{i=0}^{k} v_i}}.$$

Therefore,

$$\sum_{i=1}^{k+1} \sqrt{q_i (A_{k+1})} - \sum_{i=1}^{k} \sqrt{q_i (A_k)} = \frac{X + \sqrt{v_{k+1}}}{\sqrt{v_0 + X + v_{k+1}}} - \frac{X}{\sqrt{v_0 + X}},$$

and in order to show $\sum_{i=1}^{k} \sqrt{q_i (A_k)}$ is increasing in $k$, it suffices to show that
\[(X + \sqrt{v_{k+1}})^2 (v_0 + V) - X^2(v_0 + V + v_{k+1}) > 0.\]

Now:

\[
(X + \sqrt{v_{k+1}})^2 (v_0 + V) - X^2(v_0 + V + v_{k+1}) = (2X\sqrt{v_{k+1}} + v_{k+1})(v_0 + V) - X^2v_{k+1}
\]

\[
= (2X\sqrt{v_{k+1}} + v_{k+1})v_0 + 2\sqrt{v_{k+1}}XV - 2 \sum_{i=1}^{k} \sum_{j>i}^{k} \sqrt{v_i\sqrt{v_j}}v_{k+1}
\]

\[
> 2 \sum_{i=1}^{k} \sum_{j>i}^{k} v_i\sqrt{v_j}\sqrt{v_{k+1}} - 2 \sum_{i=1}^{k} \sum_{j>i}^{k} \sqrt{v_i\sqrt{v_j}}v_{k+1}
\]

\[
> 0
\]

where the next to last inequality follows from \(2\sqrt{v_{k+1}}XV > 2 \sum_{i=1}^{k} \sum_{j>i}^{k} v_i\sqrt{v_j}\sqrt{v_{k+1}}\) and the last inequality follows from \(v_1 > v_2 > \ldots > v_k\). This concludes the proof. \(\blacksquare\)

**Proof of Proposition 4:**

**Proof of 1:** To see why \(K(w) \geq 0\) for \(w \in (\bar{w}, p]\), first note that, using Equations (8), (9) and (10), we can write

\[
\Pi^d (A_{k_c}) - \Pi^d (A_k) = \Pi^d (A_{k_c}) - \Pi^d (A_k)
\]

\[
+ \Pi^d (A_k) - \Pi^d (A_{k_c}).
\]

Now, \(\Pi^d (A_{k_c}) - \Pi^d (A_k) \geq 0\), since assortment \(A_{k_c}\) maximizes \(\Pi^d\). Also, for \(k > k_c\), \(\Pi^d (A_k) - \Pi^d (A_{k_c}) \geq 0\) follows from the fact that \(\sum_{i=1}^{k} q_i (A_k)\) is increasing in \(k\), which was shown in the proof of Proposition 3. Therefore, \(\Pi^d (A_{k_c}) - \Pi^d (A_k) > 0\) for \(k > k_c\), and hence \(K(w) \geq 0\).

As for why \(K(w) \geq 0\), note that by definition of \(\bar{w}\), \(k_c \neq k_d\) for \(w \in (\bar{w}, p]\). Therefore, \(\Pi^d (A_{k_d}) - \Pi^d (A_{k_c}) \geq 0\), from which we conclude that \(K(w) \geq 0\) for \(w \in (\bar{w}, p]\).
Proof of 2: Let us first define some notation that will be helpful in the proof. Let

\[
\overline{K}_k(w) = \frac{\Pi^{dd}_r (A_{k_c}) - \Pi^{dd}_r (A_k)}{k - k_c}
\]

so that \( \overline{K}(w) = \min_{k_c < k \leq I} \{ \overline{K}_k(w) \} \). Similarly, define

\[
\underline{K}_k(w) = \frac{\Pi^{dd}_r (A_k) - \Pi^{dd}_r (A_{k_c})}{k_c - k}
\]

so that \( \underline{K}(w) = \max_{k_{dd} \leq k < k_c} \{ \underline{K}_k(w) \} \). Define \( \Psi(k) \) for \( k_c < k \leq I \) such that

\[
(w - c)\Psi(k) = \frac{\Pi^{dd}_m (A_k) - \Pi^{dd}_m (A_{k_c})}{k - k_c}
\]

Similarly, define \( \Phi(k) \) for \( k_{dd} \leq k < k_c \) such that

\[
(w - c)\Phi(k) = \frac{\Pi^{dd}_m (A_{k_c}) - \Pi^{dd}_m (A_k)}{k_c - k}
\]

Finally, define \( \Delta(k) \) as \( \Pi^{dd}_sc (A_{k_c}) - \Pi^{dd}_sc (A_k) \).

It follows from the definitions above and Equations (8), (9) and (10), that \( \overline{K}_k(w) = \frac{\Delta(k)}{k - k_c} + (w - c)\Psi(k) \) for \( k_c < k \leq I \) and \( \underline{K}_k(w) = -\frac{\Delta(k)}{k_c - k} + (w - c)\Phi(k) \) for \( k_{dd} \leq k < k_c \). We now establish some properties of \( \Psi(k) \) and \( \Phi(k) \) that will be useful in the proof. From the definition of \( \Pi^{dd}_m (A_k) \), it is easy to check that

\[
\Psi(k) = \lambda v_0 \frac{\sum_{i=k+1}^{k_c} v_i}{(k - k_c) \sum_{i=0}^{k} v_i \sum_{i=0}^{k_c} v_i},
\]

for \( k_c < k \leq I \). Clearly, \( \sum_{i=0}^{k} v_i \) is increasing in \( k \). Since, \( v_1 \geq v_2 \geq \ldots \geq v_I \), \( \frac{\sum_{i=k_{c+1}}^{k_c} v_i}{k - k_c} \) is decreasing in \( k \). (Note that \( \sum_{i=k_{c+1}}^{k_c} v_i \) is the average value of \( v_{k+1} \) through \( v_k \), and \( v \)'s are in descending order.) Therefore, \( \Psi(k) \) is decreasing in \( k \). Similarly,

\[
\Phi(k) = \lambda v_0 \frac{\sum_{i=k_{c+1}}^{k_c} v_i}{(k_c - k) \sum_{i=0}^{k} v_i \sum_{i=0}^{k_c} v_i},
\]
for $k_{dd} \leq k < k_c$ and $\Phi(k)$ is decreasing in $k$ as well. Moreover, for any $m > k_c$ and $n < k_c$, $\Psi(m) \leq \Phi(n)$ follows from the facts that $\sum_{i=0}^{k} v_i$ is increasing in $k$ and and $\frac{\sum_{i=k-n+1}^{k} v_i}{m-k_c} \leq \sum_{i=k-n+1}^{k} v_i$.

Now we are ready to show that $\overline{K}(w) - \underline{K}(w)$ is decreasing in $w$ for $w \in (\bar{w}, p]$. Note that $\overline{K}(w)$ and $\underline{K}(w)$ are not necessarily continuous and differentiable in $w \in (\bar{w}, p]$. However, at $w \in (\bar{w}, p]$ where $\overline{K}(w)$ is continuous and differentiable,

$$\frac{\partial \overline{K}(w)}{\partial w} \leq \max_{k_c < k \leq I} \{\Psi(k)\},$$

since $\overline{K}(w) = \min_{k_c < k \leq I} \{\overline{K}_k(w)\}$ and $\frac{\partial \overline{K}_k(w)}{\partial w} = \Psi(k)$. (To see why $\frac{\partial \overline{K}_k(w)}{\partial w} = \Psi(k)$, note that $\overline{K}_k(w) = \frac{\Delta(k)}{k-k_c} + (w-c) \Psi(k)$ for $k_c < k \leq I$, and $\Delta(k)$ and $k_c$ do not depend on $w$.)

Similarly, at $w \in (\bar{w}, p]$ where $\underline{K}(w)$ is continuous and differentiable, $$\frac{\partial \underline{K}(w)}{\partial w} \geq \min_{k_{dd} \leq k < k_c} \{\Phi(k)\},$$

since $\underline{K}(w) = \max_{0 < k < k_c} \{\underline{K}_k(w)\}$ and $\frac{\partial \underline{K}_k(w)}{\partial w} = \Phi(k)$. Therefore, at $w \in (\bar{w}, p]$, if both $\overline{K}(w)$ and $\underline{K}(w)$ are continuous and differentiable, then

$$\frac{\partial K(w)}{\partial w} \geq \min_{k_{dd} \leq k < k_c} \{\Phi(k)\} \geq \max_{k_c < k \leq I} \{\Psi(k)\} \geq \frac{\partial \overline{K}(w)}{\partial w}$$

since $\Psi(m) < \Phi(n)$ for any $m > k_c$ and $n < k_c$ as shown earlier. Now, take any $(a, b] \subseteq (\bar{w}, p]$. If both $\overline{K}(w)$ and $\underline{K}(w)$ are continuous and differentiable in $(a, b]$, then $\overline{K}(w) - \underline{K}(w)$ is decreasing over $(a, b]$, since $\frac{\partial K(w)}{\partial w} \geq \frac{\partial \overline{K}(w)}{\partial w}$.

If $\underline{K}(w)$ is discontinuous at $b$, then the jump in $K(w)$ is upwards, causing $\overline{K}(w) - \underline{K}(w)$ to decrease. (To see why this is true, note that a discontinuity in $K(w)$ will arise at $w = b$ only if $\arg \max_{k_{dd} \leq k < k_c} \{K_k(w)\}$ changes at $b$. Moreover, $k_{dd}$ is non-increasing in $w$, which implies that the domain $k_{dd} \leq k < k_c$ over which $K_k(w)$ is maximized is not going to get smaller as $w$ increases.) Similarly, if $\overline{K}(w)$ is discontinuous at $b$, then the jump in $\underline{K}(w)$ is downwards, causing $\overline{K}(w) - \underline{K}(w)$ to decrease. The entire interval $(\bar{w}, p]$ can be divided into a number of such $(a, b]$ intervals. Thus, $\overline{K}(w) - \underline{K}(w)$ is decreasing in $w$ for $w \in (\bar{w}, p]$. ■
**Proof of Theorem 4:**

**Proof of 1:** This was already shown in Theorem 3.

**Proof of 2:** In this proof we will use definitions and results from the proof of Proposition 4. Inequality (13) must hold in order for the manufacturer to be better off, i.e., we need

\[ K \leq K_m(w) = \frac{\Pi_m^d(A_{k_c}) - \Pi_m^d(A_{k_{dd}})}{k_c - k_{dd}}. \]

But then \( K_m(w) = (w - c)\Phi(k_{dd}). \) Since \( \Delta(k) > 0 \) and \( \Phi(k) \) is decreasing in \( k \), we have

\[ K_m(w) = (w - c)\Phi(k_{dd}) \geq \max_{k_{dd} \leq k < k_c} \left\{ -\frac{\Delta(k)}{k_c - k} + (w - c)\Phi(k) \right\} = K(w). \]

Therefore, whenever \( w \) is such that \( K(w) \leq \overline{K}(w) \), there exists \( K \) that satisfy both \( K(w) \leq K \leq \overline{K}(w) \) and \( K \leq K_m(w) \). This concludes the proof. \( \blacksquare \)
<table>
<thead>
<tr>
<th>Product #</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (no buy)</td>
<td>67.768</td>
</tr>
<tr>
<td>1</td>
<td>94.885</td>
</tr>
<tr>
<td>2</td>
<td>117.164</td>
</tr>
<tr>
<td>3</td>
<td>127.962</td>
</tr>
<tr>
<td>4</td>
<td>165.169</td>
</tr>
<tr>
<td>5</td>
<td>168.678</td>
</tr>
<tr>
<td>6</td>
<td>171.919</td>
</tr>
<tr>
<td>7</td>
<td>173.863</td>
</tr>
<tr>
<td>8</td>
<td>197.9439</td>
</tr>
<tr>
<td>9</td>
<td>202.7657</td>
</tr>
<tr>
<td>10</td>
<td>226.3618</td>
</tr>
</tbody>
</table>
### TABLE 2

<table>
<thead>
<tr>
<th>λ</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p - w$</td>
<td>8</td>
</tr>
<tr>
<td>$w - c$</td>
<td>15</td>
</tr>
<tr>
<td>$(b+h) \phi^*$</td>
<td>1</td>
</tr>
<tr>
<td>$v$</td>
<td>5</td>
</tr>
<tr>
<td>$L$</td>
<td>1</td>
</tr>
<tr>
<td>$l$</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product #</th>
<th>$v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

#### Retailer Profit Under Delayed Differentiation

<table>
<thead>
<tr>
<th>Assortment</th>
<th>Retailer Revenue</th>
<th>Inventory Cost</th>
<th>Retailer Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>7.273</td>
<td>1.651</td>
<td>5.621</td>
</tr>
<tr>
<td>{1,2}</td>
<td>7.600</td>
<td>2.179</td>
<td>5.421</td>
</tr>
<tr>
<td>{1,2,3}</td>
<td>7.667</td>
<td>2.552</td>
<td>5.115</td>
</tr>
<tr>
<td>{1,...,4}</td>
<td>7.692</td>
<td>2.831</td>
<td>4.861</td>
</tr>
<tr>
<td>{1,...,5}</td>
<td>7.695</td>
<td>2.938</td>
<td>4.757</td>
</tr>
</tbody>
</table>

#### Supply Chain Profit Under Delayed Differentiation

<table>
<thead>
<tr>
<th>Assortment</th>
<th>SC Revenue</th>
<th>Inventory Cost</th>
<th>SC Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>20.909</td>
<td>1.651</td>
<td>19.258</td>
</tr>
<tr>
<td>{1,2}</td>
<td>21.850</td>
<td>2.179</td>
<td>19.671</td>
</tr>
<tr>
<td>{1,2,3}</td>
<td>22.042</td>
<td>2.552</td>
<td>19.490</td>
</tr>
<tr>
<td>{1,...,4}</td>
<td>22.115</td>
<td>2.831</td>
<td>19.284</td>
</tr>
<tr>
<td>{1,...,5}</td>
<td>22.122</td>
<td>2.938</td>
<td>19.184</td>
</tr>
</tbody>
</table>

#### Retailer Profit Under Delayed Differentiation and Incremental Fee Scheme with $K = 0.21$

<table>
<thead>
<tr>
<th>Assortment</th>
<th>Retailer Revenue</th>
<th>Inventory Cost</th>
<th>Retailer Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>7.273</td>
<td>1.651</td>
<td>5.621</td>
</tr>
<tr>
<td>{1,2}</td>
<td>7.810</td>
<td>2.179</td>
<td>5.631</td>
</tr>
<tr>
<td>{1,2,3}</td>
<td>8.087</td>
<td>2.552</td>
<td>5.535</td>
</tr>
<tr>
<td>{1,...,4}</td>
<td>8.322</td>
<td>2.831</td>
<td>5.491</td>
</tr>
<tr>
<td>{1,...,5}</td>
<td>8.535</td>
<td>2.938</td>
<td>5.597</td>
</tr>
<tr>
<td>{1,5}</td>
<td>7.496</td>
<td>1.799</td>
<td>5.696</td>
</tr>
</tbody>
</table>
Figure 2
### TABLE 3

<table>
<thead>
<tr>
<th>λ</th>
<th>1</th>
<th>Product #</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>p - w</td>
<td>20</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>w - c</td>
<td>10</td>
<td>2</td>
<td>38</td>
</tr>
<tr>
<td>(b + h) φ</td>
<td>0.5</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>v α</td>
<td>8</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>L</td>
<td>1</td>
<td>5</td>
<td>24</td>
</tr>
<tr>
<td>l</td>
<td>1</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

Scenario 1: Manufacturer offers all eight products

<table>
<thead>
<tr>
<th>Assortment</th>
<th>Retailer Revenue</th>
<th>Inventory Cost</th>
<th>Retailer Profit</th>
<th>Manufacturer Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>16.596</td>
<td>0.789</td>
<td>15.807</td>
<td>8.298</td>
</tr>
<tr>
<td>{1,2}</td>
<td>18.118</td>
<td>1.064</td>
<td>17.054</td>
<td>9.059</td>
</tr>
<tr>
<td>{1,2,3}</td>
<td>18.689</td>
<td>1.279</td>
<td>17.410</td>
<td>9.344</td>
</tr>
<tr>
<td>{1,....4}</td>
<td>18.912</td>
<td>1.454</td>
<td>17.458</td>
<td>9.456</td>
</tr>
<tr>
<td>{1,....5}</td>
<td>19.064</td>
<td>1.611</td>
<td>17.453</td>
<td>9.532</td>
</tr>
<tr>
<td>{1,....6}</td>
<td>19.175</td>
<td>1.755</td>
<td>17.421</td>
<td>9.588</td>
</tr>
<tr>
<td>{1,....7}</td>
<td>19.252</td>
<td>1.885</td>
<td>17.367</td>
<td>9.626</td>
</tr>
<tr>
<td>{1,....8}</td>
<td>19.289</td>
<td>1.993</td>
<td>17.296</td>
<td>9.644</td>
</tr>
</tbody>
</table>

Scenario 2: Manufacturer removes Product # 2 from the offering

<table>
<thead>
<tr>
<th>Assortment</th>
<th>Retailer Revenue</th>
<th>Inventory Cost</th>
<th>Retailer Profit</th>
<th>Manufacturer Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>16.596</td>
<td>0.789</td>
<td>15.807</td>
<td>8.298</td>
</tr>
<tr>
<td>{1,3}</td>
<td>18.095</td>
<td>1.063</td>
<td>17.032</td>
<td>9.048</td>
</tr>
<tr>
<td>{1,3,4}</td>
<td>18.532</td>
<td>1.268</td>
<td>17.264</td>
<td>9.266</td>
</tr>
<tr>
<td>{1,3,4,5}</td>
<td>18.797</td>
<td>1.446</td>
<td>17.351</td>
<td>9.398</td>
</tr>
<tr>
<td>{1,3,4,5,6}</td>
<td>18.974</td>
<td>1.605</td>
<td>17.369</td>
<td>9.487</td>
</tr>
<tr>
<td>{1,3,4,....7}</td>
<td>19.091</td>
<td>1.748</td>
<td>17.343</td>
<td>9.545</td>
</tr>
<tr>
<td>{1,3,4,....8}</td>
<td>19.144</td>
<td>1.865</td>
<td>17.279</td>
<td>9.572</td>
</tr>
</tbody>
</table>