On the Relationship Between Inventory Costs and Variety Benefits in Retail Assortments

Garrett van Ryzin • Siddharth Mahajan

Graduate School of Business, Columbia University, Uris Hall, 3022 Broadway, New York, New York 10027
The Fuqua School of Business, Duke University, Box 90120, Durham, North Carolina 27708-0120
gjv1@columbia.edu • mahajan@duke.edu

Consider a category of product variants distinguished by some attribute such as color or flavor. A retailer must construct an *assortment* for the category, i.e., select a subset variants to stock and determine purchase quantities for each offered variant. We analyze this problem using a multinomial logit model to describe the consumer choice process and a newsboy model to represent the retailer's inventory cost. We show that the optimal assortment has a simple structure and provide insights on how various factors affect the optimal level of assortment variety. We also develop a formal definition of the level of fashion in a category using the theory of majorization and examine its implications for category profits.

(*Variety; Inventory; Retailing; Consumer Choice; Assortment; Optimization; Newsboy; Fashion; Majorization; Multinomial Logit*)

1. Introduction and Literature Review

Merchandising is a fundamental responsibility of retail managers. Basic merchandising questions include: What variants should be stocked within any given category of merchandise? How much of each variant should one buy? What is the trade-off between "depth" (large quantities of each stock keeping unit (SKU)) and "breadth" (large numbers of SKUs)? How do factors such as margin and market size affect the assortment decision? And should one manage the buying decision for fashion and basic categories differently?

While product variety has been studied in the economics and marketing literature (see Lancaster (1990) for a survey), such assortment questions typically have not been addressed directly. In particular, this literature focuses primarily on variety at the market level (Shugan (1989)), or variety in terms of product line structuring (Green and Krieger (1985)). In contrast, assortment planning requires detailed decisions on variety and inventory levels for colors, flavors, or sizes within a relatively narrow category of merchandise at the individual store level. Understanding assortment benefits and costs at this microlevel requires more detailed modeling of both the choice process and the inventory cost. An exception is the early work of Baumol and Ide (1956) who consider the joint impact of retail variety on consumer demand and store operating costs. The authors use a highly simplified model in which variety is a scalar quantity and both inventory costs and aggregate demand are deterministic. Nevertheless, their focus on modeling cost-benefit trade-offs in retail variety is very much in the spirit of our work. They also consider several interesting issues such as within-store search costs and customer travel costs, which we do not model directly.

Broadly speaking, our work addresses the costs and benefits of product variety as analyzed from a joint marketing-operations perspective. This topic is receiving increased research attention. de Groote (1994) considers the issue of modeling variety in the context of managing a manufacturing environment. Kekre and Srinivasan (1990) discuss how broadening the product line affects market share and production costs. MacDuffie et al. (1996) use data from the International Motor Vehicle Program to study the effect of product variety on labor productivity and consumer perceived product quality, while Fisher and Ittner (1996) conduct an empirical study on the impact of product variety on automobile assembly operations.

Yet a distinguishing feature of retail variety decisions, as opposed to production variety decisions, is that there are few *direct* costs (set-ups, change-overs, etc.) of variety for a retailer.[1] Despite this fact, retailers in practice do not stock the maximum level of variety possible—e.g., a typical clothing store is not stocked with one piece each of 20,000 different SKUs. Why? First, there are often constraints on display space which obviously limit the variety a retailer can offer. But more subtly, the *indirect* costs of stock-outs and overstocking impose *implicit* costs on variety.

In this paper, we propose and analyze a theoretical model which sheds light on such variety trade-offs in retail assortments. We consider a merchandise category made up of several product variants. While the concept of a "category" of merchandise is not well defined in general, we have in mind a specialized category of merchandise, consisting of alternative variants, offered at identical retail prices and having identical unit costs. The assumption of identical prices and costs is restrictive but realistic in certain cases (e.g., color/size/flavor variety). More generally, it represents a stylized model of reality. Plausible examples fitting these assumptions include different titles in an assortment of music CDs or books; apparel, where variants are different colors or sizes; foods products (ice cream, cereal, soup), where variants are

different flavors; or consumer products (toothpaste, soap, consumer electronics), where variants correspond to different brands with similar features and price points.

A retailer must decide which subset of variants should be offered and how much inventory of each should be stocked. Adding variants to the assortment increases the likelihood that consumers will purchase something from the assortment. However, including more choice alternatives reduces the volume of demand for each variant individually. This thinning—or fragmenting—of total demand increases the variability of demand for each variant, which in turn tends to drive up inventory costs. The main thrust of this paper is to analyze this variety trade-off, and then explore the strategic implications for different categories of merchandise and different operating and competitive environments.

Demand in our model derives from a stochastic choice process in which individual purchase decisions are made according to a multinomial logit (MNL), random utility model. The supply process the retailer uses is modeled as a newsboy problem. The MNL plays a central role in the theory of discrete choice (Anderson et al. 1992), and has also been successfully used in econometric studies to estimate demand for differentiated products (Guadagni and Little (1983)). Likewise, the newsboy model is a fundamental stochastic inventory model, and has played a central role since the early development of inventory theory (Arrow et al. 1958). Therefore, it is important to understand assortment decisions under these two classical models. Indeed, a main contribution of our work is to begin merging these two important streams of research.

Below, we use this model to derive several insights on retail assortments. In doing this, we shall give the model two alternative interpretations. The first is that it represents short-run variety and inventory decisions for a specialized merchandise category within a store (e.g. a buyer's decision for the assortment of men's dress shirts for the fall season). In this case, we are interested in short-run costs and profits. Short-run costs would typically include both direct costs and

---

[1] Of course, clearly there are some direct transactions and information processing cost that are driven by the number of SKUs a retailer carries. However, in most cases these are relatively small costs compared to the cost of goods and other operating expenses.

also variable and/or short-run opportunity costs for the shelf space required to stock each unit.

Alternatively, we will also view the model as a stylized representation of the long-run, strategic assortment decisions of an entire specialty store, in which the assortment is considered to be the entire range of products carried by the store. In this case, the uniform margin assumption can be considered simply as a rough approximation of the typical margins within a given specialty retail segment. For such strategic analysis, one must consider profit and cost over the long run. For example, unit cost should include the (appropriately amitorized) fixed cost of store space, because store size can be increased in the long run (e.g., by leasing more selling space).[2] In this way, the model can provide insights on strategic issues of breadth of assortment and store size. The results in §4 on characterizing a category of merchandise are particularly useful to gain insights about which store formats are best suited to which categories of merchandise.

Finally, we note that in a recent paper, Smith and Agrawal (1996) consider a problem that is very similar to ours. However, the work differs in some important respects. First, rather than using the MNL, they propose their own model of demand substitution. They also consider other cost components and impose a variety of constraints. The resulting model is therefore more realistic for decision support purposes; however, it is analytically intractable, and the authors must resort to numerical solution of a nonlinear, integer program. In contrast, our primary aim is to obtain theoretical insights rather than to provide a decision support tool. As a result, we use a parsimonious model that builds on established choice and inventory models. To keep the model simple, we do not include any direct costs or constraints on variety decisions, though direct variety cost in particular would not be hard to incorporate. Indeed, we believe it is an interesting observation that, even in the *absence* of any direct variety costs or constraints, there still exists a significant trade-off between inventory costs and va-

---

[2] We are assuming here that there are approximately constant returns to scale in a store's fixed costs in this case.

riety benefits. Our work, in this sense, provides a theoretical complement to the work of Smith and Agrawal (1996), and there are interesting extensions along both lines of research.

The remainder of the paper is organized as follows: In §2, we formulate several versions of our assortment model for a single merchandise category. In §3, we find a simple characterization of the optimal subset of variants. In §3.2, we investigate how the scale of the business, gross profit margins and the attractiveness of outside alternatives affect the optimal level of variety offered. In §4, we compare categories similar in cost structure and volume of business, but having different profiles of customer preference (utility). This leads naturally to a definition of a fashionable merchandise category based on the theory of majorization (Marshall and Olkin 1979). The implications of our results for store operating strategies are discussed in §4.2.

2. Model Formulation

2.1. Variants, Prices, and Costs
The set of possible variants is denoted $N = \{1, 2, \ldots, n\}$, and we let $S \subseteq N$ denote the subset of variants stocked by the store. Each variant is offered at an identical retail price, p, and has an identical unit cost, c. As mentioned, we can allow different prices and costs, provided the ratio p/c is the same for all variants. However, for clarity of exposition we will retain the assumption of identical prices and costs.

We assume that p is exogenously determined. Endogenizing the price decision would be a desirable extension, but it complicates the problem considerably. Indeed, even joint optimization of inventory and price for a *single* variant is complex (see Hempenius (1970), Karlin and Carr (1962), Mills (1959), and Whitin (1955)). Yet there are certainly realistic cases in which a retailer's pricing flexibility is quite limited, in particular when price competition is very high and/or manufacturers exert control over retail prices (manufacturer suggested retail price (MSRP)).

2.2. The Customer Choice Process
Each potential customer considers the subset S of variants offered by the store. A customer may choose

to purchase a variant in S or she may choose not to purchase at all. This choice decision is modeled using the multinomial logit (MNL) model (see Luce (1959) and Luce and Suppes (1965)). For a recent monograph on the MNL, see Anderson et al. (1992). The MNL is a generalization of the classical model of utility maximizing consumers, and is widely used. See for example Guadagni and Little (1983) for application in marketing or Ben-Akiva and Lerman (1985) for application in estimating travel demand. Each individual in the population associates a utility U_j with the variants $j \in S$. In addition, there is a no purchase option, denoted $j = 0$, with associated utility U_0. An individual chooses the variant with the highest utility among the set of available choices, $\{U_j : j \in S \cup \{0\}\}$.

Note that increasing the choice set S makes it more likely that a given consumer chooses to purchase, because a consumer chooses to purchase only if $\max_{j \in S} \{U_j\} \geq U_0$. On the other hand, the variants are choice alternatives in the sense that if the set S is enlarged by adding a variant i, a customer will switch from i to j if $U_i > U_j$. Thus, the selection of the set S affects the choice decision of each consumer and hence the total demand for each variant.

The key difference between the MNL and classical utility maximization is that the utility U_j is assumed to be *unobservable*, so that an individual's choice is uncertain. Moreover, utility values can vary from individual to individual due to heterogeneity of preference among customers (the value of a size small shirt depends on the purchaser), and these differences among customers may also be unobservable a priori.[3]

In the MNL, these characteristics are modeled as random components in the utility values. The utility U_j is decomposed into two parts: one part, denoted u_j, represents the nominal (expected) utility; the other part, denoted ξ_j, is a zero-mean random variable representing the difference between an individual's actual utility, U_j, and the nominal utility, u_j. Thus, $U_j = u_j + \xi_j$. We note that while the actual utility U_j

[3] The interpretation of random utility as being unobservable heterogeneity in a population of consumers was formalized by Manski (1977); Luce (1959) and Luce and Suppes (1965) considered the randomness as being primarily due to an individual's inherent inconsistency in making choice decisions.

assigned to variant j differs from individual to individual, the nominal utility u_j for variants is the same for all individuals. The random component, ξ_j, is modeled as a Gumbel (extreme value of Type I, see, for example, Evans et al. 1993) random variable with distribution $P(\xi_j \leq x) = \exp(-e^{-((x/\mu)+\gamma)})$ with mean zero and variance $\mu^2 \pi^2 / 6$ (γ is Euler's constant, $\gamma \approx 0.5772$). Here μ is a positive constant, with a higher value of μ corresponding to a higher degree of heterogeneity among the population. The assumption of zero mean is without loss of generality, while the Gumbel distribution is used primarily because it is closed under maximization (Gumbel 1958). The assumption of the Gumbel distribution in the MNL model, while restrictive, leads to a simple form of the choice probabilities as shown below in Equation (1).

Given the choices S and the no-purchase option, a consumer chooses the option with the largest utility. Let $q_j = P(U_j = \max\{U_i : i \in S \cup \{0\}\})$ denote the probability that variant j has the maximum utility. A standard result of the MNL yields (see Anderson et al. 1992, p. 39):

$$q_j = \frac{v_j}{\sum_{i \in S} v_i + v_0} \quad (1)$$

where, for $j = 0, \ldots, n$, we define

$$v_j = e^{u_j / \mu}. \quad (2)$$

We refer to the quantities v_j as *preferences*. Note that v_j is increasing in the nominal utility u_j, so a high value of v_j corresponds to variants with higher expected utility. Also, as mentioned above, (1) reflects the fact that variants are choice alternatives, in the sense that the choice probabilities depend on S. For example, if we add a variant $l \notin S$ to S, the denominator in (1) increases by v_l; hence, the probability that a customer selects one of the original variants decreases. However, the overall probability that a customer selects *something* from the set S, $\sum_{j \in S} q_j$, increases. Below, we work directly with the value v_j and let v denote the vector $(v_0, v_1, \ldots, v_n)$.

The MNL possesses a somewhat restrictive property known as the *independence from irrelevant alternatives* (IIA) property. Roughly, IIA states that the ratio of choice probabilities for alternatives i and j is indepen-

dent of the choice set S containing the alternatives. The IIA property is not realistic if the choice set contains alternatives that can be grouped such that alternatives within a group are more similar than alternatives outside the group, because adding a new alternative reduces the probability of choosing similar alternatives more than dissimilar ones. Therefore, the MNL model should be restricted to choice sets containing alternatives that are, in some sense, "equally dissimilar" (e.g., different colors or different sizes, but not different color-size combinations). Despite these well-known limitations, however, the MNL model remains widely used and therefore serves as a useful starting point to understand inventory under consumer choice.

We next impose two assumptions concerning the customer choice process, which we call *static choice* assumptions. (These same static choice assumptions are made in the work of Smith and Agrawal 1996.)

Assumption (A1) *Customers choose based only on knowledge of the set S, and they have no knowledge of the inventory status of the variants in S.*

Assumption (A2) *If a customer selects a variant in S and the store does not have it in stock, the customer does not undertake a second choice, and the sale is lost.*

Note that as a result of Assumptions A1 and A2, variants are *not* substitutes in the sense that a customer will dynamically substitute one variant for another if their first choice is out of stock. Rather, it is only a customer's initial choice that is influenced by the set of alternatives that are offered.

While these are perhaps not the most realistic assumptions, they provide a reasonable starting point and, importantly, greatly simplify an otherwise quite complex demand process. In particular, assuming that a customer's choice is made from stock on hand introduces a complicated dependency between customer choice decisions and inventory status. (We address a problem in which customers choose dynamically based on the on-hand stock in Mahajan and van Ryzin (1998).)

However, there are retail environments where Assumptions A1 and A2 may reasonably approximate certain types of customer behavior. For example, A1

certainly holds for customers of a catalog retailer who do not know the inventory status of items prior to ordering. In a store setting, if customers choose based on inspection of "floor models," the status of the on-hand inventory is also not readily discernable. As another justification of A1, Smith and Agrawal (1996) argue that in many retail settings opportunity costs of shortages are quite high, and hence optimal inventory decisions hedge strongly against stock outs (see §2.4). In such situations, stock outs are relatively rare events, and hence most customers are indeed choosing from the full assortment.[4]

As for Assumption A2, it may be plausible if one views customers as relatively uninformed *prior* to visiting the store. Upon inspecting the choice set S (e.g., the store's floor models), they learn about the variants offered and identify one that they like best. If the store then turns out to be out of stock, the newly-informed customer elects to go elsewhere to obtain his or her preferred variant rather than settling for another alternative. Thus, the very act of inspecting S changes the information a customer has about his or her possible choices, making the first and second choice fundamentally different.

Another setting in which Assumptions A1 and A2 are plausible approximations is when one views the customer's initial choice as a store-visit decision. That is, S is viewed as a store's strategic variety decision (e.g., the collection of brands they routinely offer), and customers are viewed as making a choice whether or not to visit a store (e.g., select a variant in S) based only on their knowledge of S. (In this case, the "no-purchase" utility should be interpreted as the maximum utility among all other available alternatives, including visiting other stores or not shopping.) Since customers are not physically in the store when they make their store-visit decision, they are unaware of the current inventory status of individual brands (A1). Upon visiting the store, they learn if their preferred brand is in stock; if it is not, they go elsewhere or decide not to purchase at all (A2).

Of course, each of these explanations, in turn, im-

---

[4] However, it is not clear, a priori, whether or not these rare stock-out events truly have a negligible effect on inventory and variety decisions.

plies other behavioral assumptions that may be equally unpalatable. For example, the uninformed-customer explanation assumes strong preference once a favorite variant is identified. The store-visit explanation assumes no learning about long-run inventory availability, and so forth. Nevertheless, Assumptions A1 and A2 are partially defensible, and, as we show below, they lead to intuitively satisfying insights on retail assortments.

2.3. Two Models of the Aggregate Demand

The above description details how an individual customer makes his or her decision. Aggregate demand, however, is determined by the collection of decisions made by all customers. We consider two models of aggregate demand. In the *independent population* model, aggregate demand is the result of a series of independent choices from a heterogeneous population of consumers. In the *trend-following population* model, aggregate demand is the result of a series of dependent choices from a homogeneous population. Each model is appropriate in different retail settings, as we discuss below. (See Marvel and Peck 1995 for a similar pair of demand models.)

2.3.1. The Independent Population Model. In the independent population model, each customer assigns utilities to the variants in the subset S offered by the store based on *independent* samples of the MNL model described above. The utility U_{ij} that customer i assigns to variant j is given by $U_{ij} = u_j + \xi_{ij}$ where the $\{\xi_{ij}; i \geq 1\}$ are i.i.d. random variables. Hence, the different customers make i.i.d. choices among the variants offered by the retailer according to the MNL probabilities q_j defined in (1). Since customers are heterogeneous and independent, the observation of one customer's choice reveals no additional information about the choice of subsequent customers.

The mean number of customers making choice decisions per unit time is λ. Let the number of customers selecting variant j per unit time be denoted Y_j, and note that by Assumptions A1 and A2, $EY_j = \lambda q_j$. We assume Y_j is normally distributed with a standard deviation that is a power function of the mean; that is, the standard deviation is $\sigma(\lambda q_j)^\beta$, where $\sigma > 0$ and $0 \leq \beta < 1$. The assumption that $\sigma > 0$ and $0 \leq \beta < 1$

implies that the coefficient of variation is decreasing in λ. If the coefficient of variation is constant, the assortment profit is simply proportional to λ. We do not consider the case where the coefficient of variation is increasing in volume; however, this does not seem to be a very natural case. We assume that $\lambda^{1-\beta} q_j \gg \sigma q_j^\beta$, so that the probability of having negative demand is small. A natural special case of this model is when the total volume of customers visiting the store is Poisson with mean λ. Then the total demand for variant j, Y_j, is also Poisson with mean λq_j, because the MNL results in a thinning of the aggregate Poisson demand. In this case, a normal approximation to the Poisson distribution yields $\sigma = 1$ and $\beta = \frac{1}{2}$. This approximation is justified if $EY_j \gg 1$.

The independent purchase model is useful for basic product categories, in which aggregate consumer preference is relatively stable and the primary uncertainty is over individual preferences for color, size, or flavor that inherently vary from one customer to the next.

2.3.2. The Trend-Following Population Model. The trend-following population model assumes an opposite extreme. A fixed number λ of customers make choice decisions. Each customer has *identical* valuations of the utilities for the variants, and this uniform set of utilities is determined by a *single* sample of the MNL model. So, $\{\xi_{ij} = \xi_{1j}, \forall i > 1\}$. As a result, each customer makes the same choice. Hence, once the outcome of one customer's choice is observed, the decision of subsequent customers is perfectly predictable. However, these common utilities are not observable to the retailer prior to making the assortment decision, and the retailer therefore has an incentive to hedge against this uncertainty by stocking more than one variant.

The demand for variant j, denoted Y_j, under this model is (a scaled) Bernoulli random variable with probability mass function,

$$P(Y_j = y) = \begin{cases} q_j & y = \lambda, \\ 1 - q_j & y = 0, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

This models an extreme case of trend-following behavior. For example, it is appropriate as a stylized model of color or style variety for trendy apparel.

Next, we analyze inventory cost under both these models of aggregate demand.

2.4. The Supply Processes

We consider the supply model. We do not consider constraints on the number of units stocked. Such constraints may affect the structure of the optimal assortment in the short run and are a worthy topic of research. On the other hand, over a long time horizon such constraints can be relaxed at a cost, either by reallocating store space among categories or by expanding the selling space of the store. This is the perspective we take below.

The supply model is analyzed under both the independent and trend-following demand models. Both cases yield qualitatively similar profit functions but provide different insights into the assortment decision.

2.4.1. Supply Cost Under the Independent Population Model.

We assume without loss of generality that the sales season lasts one unit of time. The number of customers selecting variant j, Y_j, is normally distributed with mean λq_j and standard deviation $\sigma(\lambda q_j)^\beta$ as described above. For each unit not sold, we assume the loss to the retailer is the cost c. For each unit short, the opportunity cost to the retailer is the loss in margin, $p - c$. We assume that there is no salvage value, though this assumption can easily be relaxed.

Let x_j denote the number of units of variant j stocked by the store and $x = (x_1, \ldots, x_n)$. The expected profit made by the retailer on the jth variant is then $E[p \min \{x_j, Y_j\} - cx_j]$, so the maximum expected profit given S and v, denoted $\pi_I(S, v)$, is

$$\pi_I(S, v) = \max_{x \geq 0} \sum_{j \in S} E[p \min \{x_j, Y_j\} - cx_j]. \quad (4)$$

The optimal profit depends on S and v through their effect on the choice probabilities q_j, which in turn determine the demand Y_j.

Under the assumption of a normal distribution, the optimal procurement level, denoted x_j^*, is given by

$$x_j^* = \lambda q_j + z\sigma(\lambda q_j)^\beta, \quad (5)$$

where

$$z = \Phi^{-1}\left(1 - \frac{c}{p}\right), \quad (6)$$

and $\Phi(z)$ denotes the c.d.f. of a standard normal random variable. (We let $\phi(z) = (1/\sqrt{2\pi})e^{-z^2/2}$ denote the standard normal density function.) Define $x^* = (x_1^*, \ldots, x_n^*)$ to be the optimal vector of stocking levels, where we adopt the convention that $x_j^* = 0$, $j \notin S$. Substituting this expression for x_j^*, in (4) we find

$$\pi_I(S, v) = (p - c)\lambda \sum_{j \in S} q_j - \frac{p\sigma\lambda^\beta e^{-z^2/2}}{\sqrt{2\pi}} \sum_{j \in S} q_j^\beta, \quad (7)$$

where we have used the fact that for a standard normal random variable Z, $E(Z - z)^+ = \phi(z) - z(1 - \Phi(z))$.

2.4.2. Supply Cost Under the Trend-Following Population Model.

In this case, a constant number λ of customers visit the store and each makes identical utility valuations. The valuations are unknown to the retailer at the time the assortment decisions are made, and the distribution of demand for variant j, Y_j, is given by (3), as described above.

The expected profit from variant j given x_j is again $E[p \min\{x_j, Y_j\} - cx_j]$. One can easily determine that the optimal expected profit, denoted $\pi_T(S, v)$, is

$$\pi_T(S, v) = \sum_{j \in S} (pq_j - c)^+ \lambda, \quad (8)$$

using the fact that the optimal procurement level is $x_j^* = 0$ if $pq_j < c$ and $x_j^* = \lambda$ if $pq_j \geq c$. Note that either demand model can also be interpreted as the single-period cost in a periodic review system with lost sales as in the work of Smith and Agrawal (1996).

3. The Optimal Assortment Problem

With these different profit functions in hand, we can formulate the optimal assortment selection problem. The problem admits a two level hierarchy; namely, at the lower level the retailer selects the optimal stocking levels given S and v, yielding the profits (7) or (8). At a higher level, the retailer chooses the best set of variants S by solving $\max_{S \subseteq N} \pi(S, v)$, where the appropriate profit function is used. Let S^* denote an

optimal solution to this problem. Then the corresponding pair (S^*, x^*) defines an *optimal assortment* for the store.

3.1. The Structure of the Optimal Assortment

Without loss of generality, let the variants be ordered according to decreasing values of v_j, so that $v_1 \geq v_2 \geq \cdots \geq v_n$. To see that the optimal subset S^* is not trivial, consider the case where $n = 8$ and $v_1 = v_2 = 10$ while $v_8 = 1$. It turns out that the set $\{1, 8\}$ can be better than the set $\{1, 2\}$. That is, it may be better to pair a popular variant with a very unpopular variant rather than offering two popular variants. One or the other could dominate depending on the cost structure and the vector of preferences of the variants, v.

To see why, note that the variant that gets paired with Variant 1 will add some incremental sales. On the other hand, its presence will divert demand from Variant 1. In the extreme case, introducing a second variant provides a negligible increase in overall sales and only results in splitting total demand between the two variants. The resulting increase in variability requires more total inventory and results in lower expected profits. Indeed, from example like this one can show that the most profitable assortment of k variants need not consist of the most popular k variants.

Despite such counterintuitive examples, the optimal set S^* turns out to be structurally quite simple. To show this, however, we need the following lemma (proof in Appendix):

Lemma 1. *Let $0 \leq \delta \leq v_1$ and define the following functions:*

$$f(\delta) = \sum_{j \in S} v_j + \delta + v_0 \qquad (9)$$

$$g_I(\delta) = (p - c)\lambda \left(\sum_{j \in S} v_j + \delta \right) - p \frac{\sigma(\lambda)^\beta e^{-z^2/2}}{\sqrt{2\pi}}$$

$$\times \left(\sum_{j \in S} v_j^\beta + \delta^\beta \right) \left(\sum_{j \in S} v_j + \delta + v_0 \right)^{1-\beta} \qquad (10)$$

$$g_T(\delta) = \lambda \sum_{j \in S} (pv_j - cf(\delta))^+ + (p\delta - cf(\delta))^+. \qquad (11)$$

Then the functions $h_I(\delta) = g_I(\delta)/f(\delta)$ and $h_T(\delta) = g_T(\delta)/f(\delta)$, are both quasi-convex in δ.

The function $h_I(\delta)$ (similarly for $h_T(\delta)$) represents the profit associated with adding a variant with preference δ to the existing set S. When $\delta = 0$, the function reduces to $\pi_I(S, v)$ and when $\delta = v_j$ for $j \notin S$, it equals $\pi_I(S \cup \{j\}, v)$. Because the function is quasi-convex, it follows that if we want to maximize $h_I(\delta)$ on the closed interval $[0, \max \{v_j: j \notin S\}]$, the maximum is achieved at the end points of the interval. So the profit is maximized by either not adding any more variants or by adding the variant with highest preference among those not included in S. The same holds for the trend following demand case. This property allows us to prove our main result on the structure of the optimal assortment.

Theorem 1. *Let $A_i = \{1, \ldots, i\}$ for $1 \leq i \leq n$. Then for each of the assortment problems defined above, there exists an $S^* \in \{A_1, \ldots, A_n\}$ that maximizes store profits.*

Proof. The proof is by construction and is the same for both cases of the assortment problem. Let S^* be an optimal subset with cardinality m, and let the objects in S^* be denoted v_j^* with $v_1^* \geq v_2^* \geq \cdots \geq v_m^* > 0$. If $S^* = A_m$ the theorem holds trivially. If $S^* \neq A_m$, then there exists a $v_j \notin S^*$ such that $v_j > v_m^*$. However, from the quasi convexity of the functions $h(\delta)$ in Lemma 1, it must also be true that we can either remove v_m^* or exchange it for $v_j > v_m^*$ without decreasing profits. Redefine S^* to be this new set and repeat the procedure. Eventually, one arrives at an optimal set $S^* \in \{A_1, \ldots, A_n\}$. □

In words, the optimal assortment can be restricted to one of n possible types. We simply choose the best i variants, where $1 \leq i \leq n$.

To gain some intuition about the result, reconsider our earlier comparison of the sets $\{1, 2\}$ and $\{1, 8\}$. We can now say the following: If $\{1, 8\}$ is indeed better than $\{1, 2\}$, then it must be true that $\{1\}$ is the optimal subset for the merchandise category. That is, the only reason that pairing a popular item with a very unpopular item is preferred is because it is in some sense "closer" to offering only the popular variant. More generally, while the most profitable assortment of k variants need not consist of the most popular k vari-

ants, the optimal assortment does consists of the most popular k variants *for some* value of k.

3.2. Implications

The simple structure of the optimal assortment makes it easy to define the level of variety a store offers. That is, more variety corresponds directly to a higher index i among the possible subsets $\{A_1, \ldots, A_n\}$. This raises interesting questions about what affects the level of variety offered in an optimal assortment. The answers to these questions are resolved by the following theorem (proof in Appendix):

Theorem 2. *For all $n > i \geq 1$,*

a) $\pi(A_{i+1}, v) > \pi(A_i, v)$ *for sufficiently high selling price p.*

b) $\pi(A_{i+1}, v) < \pi(A_i, v)$ *for sufficiently low no-purchase preference v_0.*

c) $\pi(A_{i+1}, v) > \pi(A_i, v)$ *(independent population cases only) for sufficiently high store volume λ.*

Part (a) states that high margins create an incentive to stock higher levels of variety. This is intuitive, since as margins increase the risk of lost sales dominates the risk of overstocking. Therefore, a wider variety is offered to minimize the likelihood of customers not purchasing. The result suggests that as margins rise in a category, existing retailers will have an incentive to broaden the range of merchandise they carry.[5]

Part (b) considers the effect of no-purchase utility on variety. A high no-purchase utility could correspond to a product category that is somewhat frivolous (e.g., souvenirs, toys, or jewelry), in which not purchasing is a common outcome. Alternatively, a high no-purchase utility can represent the existence of many attractive outside alternatives, including other stores in close proximity carrying similar merchandise. In either case, as the no-purchase utility declines, the prospect of losing a purchase to an external option decreases while the threat of within-assortment cannibalization increases. Therefore, it is in the retailer's interest to

decrease the breadth of the assortment. For example, this result predicts that stores in less competitive retail environments will tend to offer lower variety than similar stores in more competitive retail environments.

Part (c) implies that, in the independent demand case, as the volume of business increases, high variety becomes increasingly more profitable, and a store will carry all variants for a sufficiently high volume. That is, there are *scale economies* in offering variety and, as traffic grows, a store not only stocks more units of each variant, but also tends to stock more variants.

The reason higher variety becomes more desirable in the independent population case is due to the risk pooling inherent in a large number of independent purchase decisions. As the volume of purchase decisions goes up, the relative amount of overstocking and understocking error goes down and the cost of having fragmented purchase decisions become relatively smaller. Hence, more variety becomes profitable. "Super stores" are plausible examples of this sort of scale effect. Indeed, the results suggest a reason why the super store format is economically viable. There may be a natural, positive feedback in this format; the large variety attracts a high volume of traffic which in turn allows the super store to profitably offer a large variety of merchandise.

But this scale economy only exists for the independent purchase case. In the trend-following case, (8) shows that profit is directly proportional to λ. That is, volume has no effect on the relative profitability of different levels of assortment variety. Hence, there are no scale economies; a large store simply places bigger bets and ends up making proportionately bigger losses.

While this is clearly a stylized model, the general conclusions appear consistent with retailing practices. For example, consider the categories of merchandise offered by the super stores. They tend to be those that experience highly fragmented but largely independent demand, e.g. Borders (books, magazines), Home Depot (hardware, home remodeling), Staples (office supplies), and Toys "R" Us (children's toys). In contrast, small stores, independent boutique clothing stores for example, tend to thrive with merchandise

---

[5] We note that Shugan (1989) shows that it is possible for lower levels of variety to be optimal for higher priced categories when demand is highly price sensitive. The reason is that the reduced purchase volume from high prices can be enough to offset the positive effect of higher margins. Part (a) assumes store volume and purchase probabilities are fixed.

categories that have strong trend-following customers. Here, store scale is not an impediment to profitability. Indeed, in the trend-following case, competitive advantage derives from having intimate knowledge of local markets and the buying flexibility to respond to that information. This is precisely the advantage small, independent stores have.

4. Defining Fashion Using Majorization Ordering

Thus far, our analysis has been restricted to a single category. We next consider comparisons among different merchandise categories. We show below that the "evenness" of the preference vector v provides a natural measure of "fashion" and this notion can be made precise using the theory of majorization (Marshall and Olkin 1979).

4.1. Characterizing a Fashionable Merchandise Category

For a vector $x \in \Re^n$, let $[i]$ denote a permutation of the indices $\{1, 2, \ldots, n\}$ satisfying $x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$. We then have the following definition of the partial order based on majorization:

Definition 1. For $x, y \in \Re^n$, x is said to be majorized by y, $x < y$ (y majorizes x), if $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$, and for all $k = 1, \ldots, n - 1$, $\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}$.

Intuitively, a nonnegative vector y that majorizes x tends to have more of its "mass" concentrated in a few components. In our problem, majorization provides an appropriate measure for the degree of fragmentation in consumer preference within a given category of merchandise. Indeed, we propose the following definition:

Definition 2. A merchandise category v is said to be more *fashionable* than w if $(v_1, \ldots, v_n) < (w_1, \ldots, w_n)$.

To compare two categories v and w using Definition 2, we must make the assumption that they have the same number of variants n, and that $\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} w_i$. However, these assumptions are not restrictive. Indeed, one can add an arbitrary number of variants j with preference values $v_j = 0$ without altering the problem. Also, one can scale all values of v or w (and

the no-purchase preference, v_0 or w_0) by an arbitrary multiplier without affecting the resulting choice probabilities and thus ensure that $\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} w_i$, so the second assumption is not restrictive either. Finally, we emphasize that majorization only produces a partial ordering on vectors, so even after these adjustments it is entirely possible that two vectors v and w cannot be ordered according to Definition 2.

Definition 2 says that, for fashion categories, preferences are more evenly spread out across variants. Alternatively, in the trend-following demand case, the degree of fragmentation of the retailer's prior information of consumer preferences for variants is higher for the fashion category than for the basic category. Note also that one can interpret v and w either as two categories at the same point in time or one category observed at two different points in time. In the latter case, one can then meaningfully talk of a category becoming "more fashionable" over time. Pashigian's (1988) study of bed sheets (white vs. fancy) is an example of such a category. He shows that the fraction of sales of white sheets has declined over time, while the fraction of fancy sheets, which includes all other colors and patterns, has risen. Such behavior corresponds, roughly, to an increase in fashion over time according to our definition.

Note that it is common to think of "trendy" and "fashionable" as being synonymous. However, in our terminology, *fashion* corresponds to fragmentation in preference (or prior information on preference), while *trendy* corresponds to correlated purchase behavior among consumers.

We require the following standard result (see Marshall and Olkin 1979):

Lemma 2. *If $g: \Re \to \Re$ is a convex function and $x < y$ then, $\sum_{i=1}^{n} g(x_i) \leq \sum_{i=1}^{n} g(y_i)$.*

We next show that the optimal profit obtained from two merchandise categories is intimately related to the majorization ordering.

Theorem 3. *Consider two merchandise categories, v and w, that have identical cost structures (p and c), demand volumes, λ, and equal no purchase utilities ($v_0 = w_0$). Define r and l such that $\pi(A_r, v) = \max_{S \subseteq N} \pi(S, v)$ and $\pi(A_l, w) = \max_{S \subseteq N} \pi(S, w)$, where $\pi(\cdot)$ can be*

either of the assortment profit functions, (7) *or* (8). *Then if* $v < w$, $\pi(A_r, v) \leq \pi(A_l, w)$.

Proof. We first prove the result for the independent demand function (7) and then for the function (8). The proofs proceed as follows: First, we find r such that A_r is the optimal subset for the fashion category v. We then construct a feasible solution set for the basic category w that yields at least as much profit as the optimal profit for the fashion category v.

From the definition of majorization, we can find a $t \leq r$ such that, $\sum_{j=1}^{t-1} w_j + \delta w_t = \sum_{j=1}^{r} v_j$, where $0 < \delta \leq 1$. Define the convex function,

$$g(x) = \frac{(p - c)\lambda}{L} x - \frac{p\sigma(\lambda)^\beta e^{-z^2/2}}{\sqrt{2\pi} L^\beta} x^\beta,$$

where $L = \sum_{j=1}^{r} v_j + v_0$, $0 < \beta < 1$, and let w' be the r-dimensional vector such that $w' = (w_1, w_2, \ldots, w_{t-1}, \delta w_t, 0, \ldots, 0)$. Note that $\sum_{j=1}^{r} w'_j = \sum_{j=1}^{r} v_j$. Since $(v_1, \ldots, v_r) \prec w'$, using Lemma 2, it follows that

$$\sum_{j=1}^{r} g(v_j) \leq \sum_{j=1}^{r} g(w'_j). \tag{12}$$

Now since $g(0) = 0$, the right hand side of (12) is simply $h_I(\delta w_t)$, where, $h_I(\cdot)$ is the quasiconvex function defined in Lemma 1 with subset of variants $S = A_{t-1}$. Using (12) and the quasiconvexity of $h_I(\cdot)$, it follows that

$$\pi_I(A_r, v) = \sum_{j=1}^{r} g(v_j) \leq \sum_{j=1}^{r} g(w'_j)$$

$$\leq \max \{\pi_I(A_{t-1}, w), \pi_I(A_t, w)\} \leq \pi_I(A_l, w).$$

This completes the proof for the independent population case.

For the trend-following population case (profit function (8)), we define the convex function $g^T(x) = (\lambda/L)(px - cL)^+$, where $L = \sum_{j=1}^{r} v_j + v_0$, as defined above. Replacing $g(\cdot)$ by $g^T(\cdot)$, and $h_I(\cdot)$ by $h_T(\cdot)$ in the argument above, we conclude that $\pi_T(A_r, v) \leq \pi_T(A_l, w)$. □

In contrast to the previous assumptions on v and w, the conditions equal λ and $v_0 = w_0$ are restrictive. Essentially, these two assumptions serve to equalize the relative demand ("market potential") of each category. This follows since if we have two assortments, S and T with $\sum_{i \in S} v_i = \sum_{i \in T} w_i$ (equally attractive) and they face the same mean number of customers λ, then the mean number of customers who decide to purchase is also the same. That is, $\lambda(\sum_{i \in S} v_i)/(\sum_{i \in S} v_i + v_0) = \lambda(\sum_{i \in T} w_i)/(\sum_{i \in T} w_i + w_0)$. This normalization is necessary to isolate the effect of fashion on profits because demand volume strongly affects a category's profits. One should therefore view Theorem 3 as primarily a theoretical comparison. The profits of two categories are affected, in general, by a combination of volume, gross margin and fashion effects. In a real-world comparison of two categories, $v_0 = w_0$ is equivalent to assuming that the outside choice alternatives for each category are equally attractive. The assumption of equal λ may approximate a case where two categories are stocked at the same store and λ is viewed as a measure of store traffic.

According to the above theorem, if one merchandise category is more fashionable than another, then, all other things being equal, the optimal profit of the fashion category will be lower than that of the more basic category. The intuitive reason for this is that the risk of inventory overage and underage is higher due to the higher fragmentation of consumer purchase decisions in the fashion category. Thus, even under optimal variety and stocking decisions, the fashion category is less profitable at a given price.

However, one might question whether basic categories, under similar demand volumes, could reasonably be expected to be more profitable in the long run. Indeed, one would expect the market to eventually compensate retailers for the added costs of fashion categories by allowing higher market prices. That is, retailers recognizing the higher profitability of a basic category would have an incentive to add the category to their store. The resulting increase in the number of retailers offering this category would, in turn, tend to increase price competition until the basic category becomes less profitable and further new entrants are discouraged.

A simple approximation of this effect is obtained by freeing up the price variable and asking the question: At what prices do two categories have equal profits?

The following corollary shows that these equalizing prices are ordered if the categories are ordered according to our definition of fashion (proof in Appendix):

COROLLARY 1. *Assume $v < w$, all other parameters except price, p, are identical and that prices are adjusted to satisfy the equilibrium profit condition,*

$$\max_{S \subseteq N} \pi(S, v) = \max_{S \subseteq N} \pi(S, w) = \pi_e,$$

where $\pi(\cdot)$ is either of the assortment profit functions (7) or (8) and π_e is an arbitrary, nonnegative equilibrium profit level. Let the equilibrium prices be denoted p_v and p_w, respectively, for categories v and w. Then $p_v \geq p_w$.

Clearly, the equilibrium profit hypothesis behind this result represents a highly simplistic view of retail markets. In reality, factors such as location, assortment, store image and customer service enable a retail store to differentiate itself from its competitors. In addition, categories may not have the same costs for space, fixtures, signage, etc. Also, customers may shop stores to purchase a "basket" of products from different categories based on a "shopping list" (see, for example, Bell et al. (1997)), and stores may differentiate themselves based on the total basket they offer.

Despite these myriad limitations, Corollary 1 does provide a simple and intuitively appealing explanation of the fact (see Pashigian (1988) for empirical evidence) that fashion goods tend to have higher margins than basic goods. In short, Corollary 1 simply suggests that higher gross margins may serve to compensate retailers for the increased inventory risks induced by the highly fragmented purchase choices of the fashion category. (Lazear (1986) proposes an alternative theory for this effect based on differences in the variability of reservation prices for each category.)

4.2. Implications
The equilibrium price results of Corollary 1 and the scale economy results of Theorem 2 have interesting implications for a retailer's operating strategy. Corollary 1 suggests that equilibrium margins tend to be higher for fashion categories. High margins, in turn, may justify changes in the way goods are supplied to stores. In particular, they may justify using fast—and potentially expensive—logistics processes (e.g., air

freight) to replenish stocks of popular variants in season rather than stocking to forecasts.

While this may be a viable strategy to manage a fashion category with dependent (trend-following) purchase behavior, Theorem 2 suggests that there are scale economies to offering variety in the independent (non-trend-following) population case. As a result, fast replenishment may not be a viable logistics strategy for this type of category because competing retailers can mitigate fashion risks using large-scale store formats (or centralized warehouses) without resorting to expensive logistics options.

Dvorak and Paasschen (1996), writing in *McKinsey Quarterly*, describe several retailer's logistical strategies that are consistent with these conclusions. The authors describe the operations of one (anonymous) "fast-to-market" high fashion retailer as follows:

> High fashion is a high-risk business . . . [Most] retailers usually air-ship only those items that have unexpectedly run out. But this retailer air-ships all items that have sold better than expected during tests . . . None of this speed is cheap, but the expense is more than covered by higher sales and fewer mark-downs.

For a retailer of casual cotton clothes, which are moderately priced fashion items and correspond more closely to the independent (non-trend-following) case, they describe a very different strategy:

> What is important is to make sure stores are always stocked with the right color, size and design . . . Production lead times are long . . . [and] price constraints also rule out shipping goods by air . . . The thrust of the logistics strategy is therefore to achieve not speed, but a smooth, seamless transition from one wave of goods to the next . . . To cope with variations in demand—for specific colors, sizes, or designs—a second strand of the retailer's strategy introduces flexibility. Regional warehouses, located close enough to stores to allow cost-effective, frequent replenishment, provide buffer stocks ready to fill gaps on the store's shelves.

This description is consistent with the scale sensitive characteristics of the independent population case.

Appendix
Proof of Lemma 1. We use the following result from Mangasarian (1969): The function $g(\cdot)/f(\cdot)$ is quasiconvex on X if (i) $g(\cdot)$ is convex and $f(\cdot) > 0$ for all $v \in X$ and (ii) $f(\cdot)$ is linear on X. The function $f(\cdot)$ defined in (9) is linear in δ. It is easy to show that the functions $g_I(\cdot)$ and $g_T(\cdot)$ defined by (10) and (11), respectively, are

convex in δ. Thus, the functions $h(\cdot)$ of the lemma satisfy the above conditions.

Proof of Theorem 2. From (1), we have

$$q_j = \frac{v_j}{\sum_{j \in S} v_j + v_0}.$$

Define $q_j = t_j$, when $S = A_i$ and $q_j = w_j$, when $S = A_{i+1}$.

Then from (7), in the independent demand case, $\pi_I(A_{i+1}, v) \geq \pi_I(A_i, v)$ if and only if,

$$(p - c)\lambda \left[\sum_{j=1}^{i+1} w_j - \sum_{j=1}^{i} t_j \right] \geq \frac{p\sigma\lambda^\beta e^{-(z^2/2)}}{\sqrt{2\pi}} \left[\sum_{j=1}^{i+1} w_j^\beta - \sum_{j=1}^{i} t_j^\beta \right].$$

Because $\sum_{j=1}^{i+1} w_j > \sum_{j=1}^{i} t_j$, this yields

$$\left(1 - \frac{c}{p} \right) \frac{\sqrt{2\pi}}{\sigma} \lambda^{1-\beta} e^{z^2/2} \geq \frac{\sum_{j=1}^{i+1} w_j^\beta - \sum_{j=1}^{i} t_j^\beta}{\sum_{j=1}^{i+1} w_j - \sum_{j=1}^{i} t_j}. \quad (13)$$

Similarly, for the trend-following demand case, we have by (8) that $\pi_T(A_{i+1}, v) \geq \pi_T(A_i, v)$ implies

$$\sum_{j=1}^{i+1} (pw_j - c)^+ \geq \sum_{j=1}^{i} (pt_j - c)^+. \quad (14)$$

To prove part (a), we note that by (6), $z = \Phi^{-1}(1 - (c/p))$ is an increasing function of p. So the left hand side of (13), is an increasing function of p, while the right hand side is independent of p. Thus, for sufficiently high p, A_{i+1} is better than A_i. For (14), note that for sufficiently high p, the inequality becomes $p(\sum_{j=1}^{i+1} w_j - \sum_{j=1}^{i} t_j) \geq c$, which is satisfied for large enough p since $\sum_{j=1}^{i+1} w_j > \sum_{j=1}^{i} t_j$. This completes the proof of part (a).

For proving part (b), note that $q_j \gtrsim v_j/\sum_{j \in S} v_j$ as $v_0 \gtrsim 0$, and hence $\sum_{j=1}^{i+1} w_j \gtrsim \sum_{j=1}^{i} t_j$. As a result, the denominator of the right hand side of both (13) tends to zero, while one can easily show the numerators tend to a positive constant, which establishes the result. Reversing the inequality in (14) and using the fact that $\sum_{j=1}^{i+1} w_j \gtrsim \sum_{j=1}^{i} t_j$ as $v_0 \gtrsim 0$ the claim holds in the trend-following case as well.

Part (c) is true because the left hand side of both (13) is increasing in λ.

Proof of Corollary 1. We need to show that $\pi(S, v)$ is an increasing function of p. Hence, $\pi(A_r, v) = \max_{S \in A_1 \ldots A_n} \pi(S, v)$, is also an increasing function of p. Combining this property with Theorem 3, it then follows that the equilibrium prices satisfy the stated condition.

For the independent demand case, differentiating $\pi_I(S, v)$ with respect to p, we obtain the following condition for $\pi_I(S, v)$ to be an increasing function of p,

$$\frac{\sqrt{2\pi}\lambda^{1-\beta} \sum_{j \in S} q_j}{\sigma \sum_{j \in S} q_j^\beta} \geq \left[1 - \frac{c\Phi^{-1}\left(1 - \frac{c}{p} \right) \left[\Phi^{-1}\left(1 - \frac{c}{p} \right) \right]'}{p} \right] e^{-z^2/2}$$

where prime denotes the derivative with respect to p. Under our assumption that $\lambda^{1-\beta} q_j \gg \sigma q_j^\beta$, the left hand side is greater than 1,

while the right hand side is less than 1, since $[\Phi^{-1}(1 - (c/p))]' \geq 0$, so the above condition is always satisfied. For the trend following demand case, it is easily seen that $\pi_T(S, v)$ is an increasing function of p.

References

Anderson, S. P., A. de Palma, J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* The MIT Press, Cambridge, MA.

Arrow, K. J., S. Karlin, H. Scarf, eds. 1958. *Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, CA.

Baumol, W. J., E. A. Ide. 1956. Variety in retailing. *Management Sci.* **3** 93–101.

Bell, D. R., T. Ho, C. Tang. 1997. Determining where to shop: Fixed and variable costs of shopping. Working Paper, The Anderson School, UCLA, Los Angeles, CA.

Ben-Akiva, M., S. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Predict Travel Demand.* The MIT Press, Cambridge, MA.

de Groote, X. 1994. Flexibility and marketing/manufacturing coordination. *Internat. J. Production Econom.* **36** 153–167.

Dvorak, R. E., F. van Paasschen. 1996. Retail logistics: One size doesn't fit all. *McKinsey Quart.* **2** 120–129.

Evans, M., N. Hastings, B. Peacock. 1993. *Statistical Distributions.* John Wiley, New York.

Fisher, M. L., C. D. Ittner. 1996. The impact of product variety on automobile assembly operations. Working Paper, Operations and Information Management, The Wharton School, Philadelphia, PA.

Green, P., A. Krieger. 1985. Models and heuristics for product line selection. *Marketing Sci.* **4,** 1–19.

Guadagni, P. M., J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **3** 203–238.

Gumbel, E. J. 1958. *Statistics of Extremes.* Columbia University Press, New York.

Hempenius, A. L. 1970. *Monopoly with Random Demand.* Rotterdam University Press, Rotterdam, The Netherlands.

Karlin, S., C. R. Carr. 1962. Prices and optimal inventory policies. Arrow, K. J., S. Karlin, and H. Scarf eds. *Studies in Applied Probability and Management Science.* Stanford University Press, Stanford, California.

Kekre, S., K. Srinivasan. 1990. Broader product line: A necessity to achieve success? *Management Sci.* **36** 1216–1231.

Lancaster, K. 1990. The economics of product variety: A survey. *Marketing Sci.* **9** 189–210.

Lariviere, M., E. L. Porteus. 1996. The ice cream stocking problem. Presented in Session SC06, INFORMS, Washington D.C., May 5, 1996.

Lazear, E. P. 1986. Retail pricing and clearance sales. *Amer. Econom. Rev.* **76** 14–32.

Luce, R. 1959. *Individual Choice Behaviour: A Theoretical Analysis.* Wiley, New York.

——, P. Suppes. 1965. Preference, utility and subjective probability.

R. Luce, R. Bush, and E. Galanter, eds. *Handbook of Mathematical Psychology.* Vol. 3, Wiley, New York.

MacDuffie, J. P., K. Sethuraman, M. L. Fisher. 1996. Product variety and manufacturing performance: Evidence from the international automotive assembly plant study. *Management Sci.* **42** 350–370.

Mahajan, S., G. van Ryzin. 1998. Stocking retail assortments under dynamic consumer substitution. To appear, *Oper. Res.*

Mangasarian, O. L. 1969. *Nonlinear Programming.* McGraw-Hill, New York.

Manski, C. 1977. The structure of random utility models. *Theory and Decisions* **8** 229–254.

Marshall, A. W., I. Olkin. 1979. *Inequalities: Theory of Majorization and its Application.* Academic Press, New York.

Marvel, H. P., J. Peck. 1995. Demand uncertainty and returns policies. *Internat. Econom. Rev.* **36** 691–714.

Mills, E. S. 1959. Uncertainty and price theory. *Quart. J. Econom.* **73** 117–130.

Pashigian, P. B. 1988. Demand uncertainty and sales: A study of fashion and markdown pricing. *Amer. Econom. Rev.* **78** 936–953.

Shugan, S. M. 1989. Product assortment in triopoly. *Management Sci.* **35** 304–320.

Smith, S. A., N. Agrawal. 1996. Management of multi-item retail inventory systems with demand substitution. Working Paper, Decision and Information Sciences, Santa Clara University, Santa Clara, CA.

Whitin, T. M. 1955. Inventory control and price theory. *Management Sci.* **2** 61–68.