The Impact of Quality and Variety on Product Assortment Decisions: An Empirical Investigation

Michaela Draganska
Stanford University
Graduate School of Business
Stanford, CA 94305-5015
draganska_michaela@gsb.stanford.edu

Michael Mazzeo
Northwestern University
Kellogg School of Management
Evanston, IL 60208-2001
mazzeo@kellogg.northwestern.edu

August 2003

PRELIMINARY AND INCOMPLETE

Please do not cite or quote

The Impact of Quality and Variety on Product Assortment Decisions: An Empirical Investigation

Abstract

We study the use of variety and quality of a product line as strategic tools, and specifically the link between quality and the composition of product assortments. We observe that individual stores offer assortments such that the same ice cream flavors from brands within the same quality tier do not appear on store shelves at the same time. This suggests that retailers may use flavor selection as a tool to reduce inter-brand competition within quality tiers. Using the ice cream category data, we analyze the assortments offered by stores and the effect of the assortments offered on prices, sales and competition.

Key Words: product line decisions, competitive strategy.

1 Introduction

We study firms' use of variety and quality of their product lines as strategic tools, focusing specifically on the link between quality and the composition of product assortments. Using data on the ice cream category, we document the assortments of brands and flavors offered by stores and analyze the effect of consumer demand and substitution patterns on the variety of products included in retailersÆ assortments. Understanding these effects is of great value to retailers, for whom the product assortment problem is both very complex and very important to profitability. ManufacturersÆ choices regarding what to produce will depend critically on whether retailers find it optimal to offer their particular varieties, given the portfolios manufactured by competitors. In addition, regulators may be concerned about the effects of product variety on utility when analyzing industry mergers.

We use two years of data from five stores, covering 35 flavors and six brands that can be conveniently be divided among three quality tiers, each containing two brands. Two stylized facts emerge from looking at the data. First, across the stores for which we have information, higher quality brands are generally associated with larger assortments. This finding is quite surprising in light of previous research that suggests the opposite correlation would be optimal (Shugan 1989). Therefore, we carefully examine both the demand-side and the supply-side factors that may explain the positive correlation between brand quality and number of product offerings in our context. Second, individual stores offer product assortments such that the same ice cream flavors from brands within the same quality tier rarely appear on store shelves at the same time. This observation suggests that retailers may use flavor selection as a tool to maximize profits by including in their product portfolios brand/flavor combinations that are not close substitutes.

We begin by introducing a simple theoretical framework for analyzing the problem of flavor selection for retailers. We show that, on the revenue side, retailers face a tradeoff between the average utility that each individual product provides to consumers and the extent tow which products appeal to different subsets of the consumer population. Depending on substitution patterns, retailers may find it optimal to avoid offering multiple brands of the same flavor, particularly if the brands are from the same quality tier. In addition, assortment costs or other supply side factors associated with offering particular brand/flavor combinations will influence the product selection decisions of retailers.

Next, we estimate a demand system for the ice cream category. We specify utility at the brand/flavor level for each product in the dataset, and allow the effect of product characteristics to differ across consumers. Estimating such a demand system enables us to further explore the link between quality and variety by evaluating alternative explanations for retailersÆ product assortment decisions. We can evaluate how retailers may use brand/flavor selection to increase profits by introducing products with particular demand

and substitution characteristics.

Our findings also have potentially important implications for manufacturers' competitive strategy. To the extent that demand and substitution factors explain the observed flavor selection decisions by retailers, then a manufacturerÆs strategy of offering unique flavors is right on target. Suppose, for example, that consumer preferences for different flavors of different brands tends to be more variable than for the same flavor produced by these brands. Then, by offering unique flavors, a manufacturer would maximize the chance that a profit maximizing retailer selects a

large subset of its assortment. However, if the explanation is found on the supply side, quantity discounts may be the best tool to give retailers an incentive to carry more flavors from the same brand. Of course, a retailer may not choose to purchase from only one manufacturer within a tier for fear that this may give that manufacturer too much price setting power.

FINDINGS

The remainder of the paper is organized as follows. We start by describing the ice cream market and the data we use for the empirical analysis in Section 2. We present a stylized retailer model in Section 3 and the demand model is derived in Section 4. The results of the empirical analysis are reported in Section 5. Section 6 concludes with a discussion and directions for future research.

2 The Market and Data

Market definition. Ice cream is one of the most popular categories in supermarkets: 92.9% of households in the United States purchase in the category (Marketing Factbook, 1993). In the general category of ice cream, there is a distinction between regular ice cream, frozen yogurt and sorbets, and ice milk[1] with regular ice cream representing more than 60% of total sales. While approximately one-third of all ice cream sales is vanilla, there are literally hundreds of other flavors that have been created over the years. Fruits, nuts, candies, spices, liquors, and syrups are all used to produce a multitude of flavors.

Market structure. Ice cream is one of the few consumer product categories in the U.S. market not dominated by a single company. The top national producer, Kraft (Breyers, Sealtest), had a 15.5% market share in 1993, followed by Dreyers with about 10%, and Häagen Dazs and Ben & Jerry's with about 6% each (Market Share Reporter, Frozen Desserts, 1995). Recently, however, Unilever and Nestle have been pushing for consolidation and fighting for dominance. Unilever bought Breyers and Sealtest from Kraft in 1993, and Ben & Jerry's in 2000. Nestle acquired Häagen Dazs in 1999, and in 2002 proposed to merge its ice cream operations with Dreyers. The U.S. Federal Trade Commission (FTC) has recently sought a preliminary

[1] Ice milk was renamed low-fat ice cream in 1994 due to new nutritional labeling requirements. After the name change, sales increased by 60% in the category.

injunction to block this merger on the grounds that it would "lead to anti-competitive effects ... including less product variety and higher prices".[2]

Data sources. In our empirical analysis we use a store-level scanner panel data set collected by Information Resources Inc. (IRI) in two contagious zip codes of a large Midwestern urban area. The data spans a 104-week period (June 1991-May 1993) and consists of weekly data on price, quantities, features, and displays for all UPCs sold in 5 stores. We focus on the six largest brands offering regular ice cream: Häagen Dazs, Ben & Jerry's, Dreyers[3], Breyers, Sealtest, and Schoeps. Together, they represent about 80% of the market. We supplement the data with information on the product characteristics of the individual flavors obtained from the package labels. For products that were discontinued, we obtained data on close substitutes (e.g., Fieldcrest was used as a proxy for Sealtest, which was discontinued).

Marketing variables. We measure quantity by ounces (and thus aggregate UPCs that belong to the same brand and have the same flavor but vary in size). Market shares are then computed by dividing the quantities sold

[2] Information from the FTC website at www.ftc.gov/opa/2003/03/dreyers.htm. Note that the FTC's concerns relate primarily with Dreyers' super-premium brands (Dreamery, Godiva and Starbucks), which had not yet been introduced at the time of our sample. Nonetheless, our results regarding substitution patterns and flavor selection within and across quality tiers will provide evidence regarding whether the
super-premium category is appropriately defined as a separate market.

[3] Dreyer's ice cream is sold under the brand name Edy's in the Midwest and Eastern United States after Kraft, the makers of Breyers, raised objections in 1985

by each brand by the potential market.[4] Table 1 gives an overview of the characteristics of the brands in the data set. Breyers is the market leader, followed by Dreyers, Sealtest, Schoeps, Häagen Dazs and Ben & Jerry's. It should be noted that this distribution of market shares is based on volume sold as measured by ounces, so brands that are sold in larger sizes have an advantage. Häagen Dazs and Ben & Jerry's have the highest price, followed by Dreyers and Breyers, and Sealtest and Schoeps. There is not much display activity in the category but Sealtest, Breyers, and Dreyers are featured relatively often in retailers' weekly ads (17%, 14%, and 10% of all weeks, respectively).

Quality tiers and product characteristics. In an effort to gain more insight on the competition across firms, we divided the six brands in the data set into three quality tiers: low (Schoeps and Sealtest), medium (Breyers and Dreyers), and high (Ben & Jerry's and Häagen Dazs). Table 2 presents summary statistics for a number of product characteristics for the individual brand-flavor combinations within these quality tiers. While there is variation between the flavors within a brand and between brands within a tier, the differences between the quality tiers are much larger. We see that the super-premium ice creams, Häagen Dazs and Ben & Jerry's have on average twice the calories and the fat of the regular ice creams Sealtest and Schoeps. In

[4] We calculate the potential market using the population size of the two zip codes where the stores are located and multiplying by the average weekly ice cream quantity purchased in grocery stores in the U.S.

terms of calories and fat content, Dreyers and Breyers are closer to Sealtest and Schoeps than to Häagen Dazs and Ben & Jerry's. This is consistent with the industry notion that the higher the butterfat content and the denser the product (less overrun), the higher the quality of the ice cream.[5]

Flavor availability. Table 3 presents the flavors manufactured by each brand. There are 35 flavors in total. The table indicates that there is wide variability in the flavors produced by each of the brands examined here. Of the 35 flavors, only two (Chocolate and Vanilla) are produced by all six firms and just three more (Butter Pecan, Chocolate Chip and Strawberry) are part of the product line of five of the six firms. At the other extreme, more than half - 18 out of 35 - of the flavors in the data set are manufactured by a single brand. We will examine the difference in incentives to offer unique versus common flavors throughout the analysis.

In our theoretical framework, optimal product assortment will hinge on the question of how sales of particular brand-flavor combinations are affected by the availability of other flavors of the same brand versus other brands of the same flavor. Analyzing this question empirically is facilitated by the extensive variability in brand-flavor combinations available among the five stores in the data set and across the 104 weeks for which we have data. Table 4 lists, for each flavor, the number of store-week observations for which the corresponding number of brands of that flavor is offered by the retailer. For

[5] See, for example, Marshall & Arbuckle (2000).

example, there are two brands that produce coffee ice cream, and there are only two store-weeks in which neither of the two brands is available. Both brands are offered in 351 store-weeks and in the remaining 167, just one brand is offered. These differences will allow us to determine the effect on demand for a given brand-flavor combination of the presence of other brands and flavors - an important element of the retailer's decision of which products to stock.

Quality and variety. Table 5 examines the data in a slightly different way by looking at the set of flavors manufactured by each brand. By construction, there are two flavors in the category "produced by five other firms" for all the brands. Of the 18 unique flavors, 14 are produced by the two high quality brands (Ben & Jerry's and Häagen Dazs), with the remaining four spread among the other firms. Breyers, Dreyers, and Schoeps manufacture all eight of the flavors sold by at least three other firms. Sealtest has the most product line overlap among the brands, with only one of its nine flavors sold by two or fewer manufacturers.

Table 5 indicates that, at the manufacturer level, the size of product assortments are roughly equal across the quality tiers, with Sealtest having the smallest number of products available among the brands. In contrast, Table 6 presents the data at the retail level, demonstrating a more pronounced difference among the three quality tiers. The table lists the average number of flavors (per week) each store offers in each brand over the two years for which we have information. Retailers offer more flavors of brands in the highest quality tier, with the lowest quality tier having the fewest flavors offered by retailers. By focusing attention on the vertically disintegrated retailers, we obtain a result that stands in contrast with Shugan's (1989) prediction that high-quality brands will offer fewer varieties than lower-quality brands.

The raw data can also provide some initial evidence regarding potential connections between quality tiers, brands and flavor selection by retailers. As an example, table 7 focuses on six individual flavors that are manufactured by Breyers, Sealtest and Schoeps. Again, stores vary in the number of weeks in which they offer 0, 1, 2 or 3 of these brand-flavor combinations. Of particular interest are the store-week observations where two of the brands are offered for the given flavor. The table breaks out those cases, demonstrating that for most flavors the two brands sold are in different quality tiers (that is, either Schoeps and Breyers or Sealtest and Breyers, rather than Schoeps and Sealtest). We will be able to examine the importance of quality differentiation in flavor selection using these tier distinctions as our empirical analysis proceeds.

Assortment composition. We computed a series of statistics to measure the variety of the assortments for the retailers in our data set. The first measure is a simple count of products offered averaged over all the weeks in the data set for each store. The second variety measure goes beyond the length of product lines to incorporate some information about their composition.

For each store/week, we compute the following statistic:

$$\left(\frac{\text{\# products offered}}{\text{\# possible products}}\right) \times \left(\frac{\text{\# flavors offered}}{\text{\# possible flavors}}\right) \times \left(\frac{\text{\# brands offered}}{\text{\# possible brands}}\right) \tag{1}$$

A third category assortment variety was suggested by Hoch, Bradlow & Wansink (1999). To compute their statistic, we pair each of the products offered with each other product (if a retailer's assortment contains N products, there are $N(N-1)/2$ relevant pairs). For each pair, there are two potential similarities among the products, as they can be the same flavor or the same brand. The pairs are then grouped by the number of characteristics in common - n1 is the count of all pairs with neither the same brand nor flavor and n2 is the count of pairs where flavor or brand is the same. The variety statistic is then:

$$\sqrt{m} \times n_1 + \sqrt{m} \times n_2,$$

where m equals the number of differences between the products(so, $m = 2$ for n_1 and $m = 1$ for n_2).

Table 8 displays the three variety measures, we compute an average across the weeks for each store and an overall average for all the store/weeks. The table shows a considerable difference in the variety of assortments, with a large spread between the lowest and highest variety stores. In addition, we find that the various statistics to measure variety yield similar results; however, the measure that incorporate product characteristics do add some nuance beyond the simple line-length measure.

Summary and discussion. The data presented in this section highlight several factors that make the ice cream category a good context for studying portfolio assortment decisions. Foremost among these is the considerable variety in the flavors product by the six manufacturers in the category. Since there is also variety in the assortments retailers offer week-to-week, we will be able to estimate a demand system that incorporates how consumersÆ choices are affected by the options that are available to them. This is particularly critical for retailers deciding on their product assortments because including or excluding a particular brand/flavor combination will affect profits differently depending on how consumers substitute among the various products. Finally, the organization of the category into three, two-brand quality tiers allows us to examine possible connections between assortments andproduct quality ù an important issue in the marketing and economics literatures.

3 Stylized Retailer Model

A rigorous theoretical analysis of optimal brand and flavor selection is exceedingly complex, though the retailer's problem is reasonably straightforward. We begin with the assumption that retailers maximize profits for the ice cream category at their store and the number and preferences of consumers are exogenous with respect to ice cream category decisions. Then, we work with the following profit function for each period (with the t subscripts suppressed):

$$\Pi = \sum_{jf} [p_j D_{jf} * I_{jf}] - C(J, F) \tag{2}$$

where the summation is over all flavors available to a retailer (i.e., the ones that the manufacturers produce), p_j is now the fixed price that the retailer charges for (all flavors in) brand j, D_{jf} is the demand for brand-flavor jf, and I_{jf} is a dummy variable indicating whether the retailer has chosen to include the particular brand/flavor combination in its product assortment. We leave the cost function unspecified, except to suggest that the cost to offering additional products potentially increases in both the number of flavors offered and the number of brands offered. Otherwise, it would be trivially optimal for the retailer to offer all the brand/flavor combinations that are manufactured.

A small but growing literature has started the process of solving this difficult problem. Earlier work, such as Costjens & Doyle (1981) and Bultez & Naert (1988) began by addressing the related question of how to optimally allocate an exogenous amount of shelf space, once the set of products to be offered had been determined. Dobson & Kalish (1988) examine the complementary problem of optimal pricing once the set of products is established. A recent paper by Chong, Ho & Tang (2001) considers the endogenous determination of the category assortment - and is, in fact, motivated by observation of the ice cream category. This paper, and several of the others in this literature, address the inherent complexities by developing heuristics and algorithms to guide practitioners, rather than deriving robust predictions from the optimization problem.

A paper by Aydin & Ryan (2000) is closest to the spirit of the analysis that we are pursuing here. They work with a slightly simpler problem, to the extent that product differentiation is assumed to be vertical and substitution patterns between products are constrained to be proportional to market shares. Using this set up, they are able to establish some simple, intuitive results: for example, that a retailerÆs optimal product line should consist of the products that have the highest potential margins and that the profitability of adding more products decreases with the number of products produced. It is worth noting, however, that these authors are not able to make predictions about the relationship between optimal product selection and demand elasticities, given their assumptions. Their model is also simplified by the fact that it does not constrain prices for all flavors within a given brand to be equal.

To illustrate some of the important features of the product assortment problem, we present and discuss a simplified example. Suppose that the retailer in question can choose between two manufacturers (A and B), each of which sells two flavors, with one flavor offered by both brands: AX, AY, BY, and BZ will indicate the four possible brand/flavor combinations. The retailer then has the choice among 15 different options for which portfolio of products to offer:

0 products	1 products	2 products	3 products	4 products
$\emptyset$	AX	AX AY	AX AY BY	AX AY BY BZ
	AY	AX BY	AX AY BZ	
	BY	AX BZ	AY BY BZ	
	BZ	AY BY		
		AY BZ		
		BY BZ		

Clearly, the retailer would offer all four products if there were no cost to offering additional brands and flavors. To the extent that there is a cost, $C(J, F)$, the retailer needs to select the set of flavors that maximizes the revenue produced less such costs. In perhaps the simplest example, if it turns out that costs are such that offering more than one flavor is prohibitive, then the flavor with the highest demand will be chosen.

The important tradeoffs in this problem can be illustrated by considering whether to offer a second flavor, and if so which among the other three flavors should be the second flavor. When the retailer sells product AX only, its profits are simply:

$$\Pi_{AX} = p_A D_{AX|AX} - C(1, 1), \tag{3}$$

where $D_{AX|AX}$ signifies the quantity of flavor AX demanded when AX is one product sold. Essentially, $D_{AX|AX}$ is the entire measure of shoppers at that retailer for which:

$$U_i = d^i_{AX} + x'_{AX}\beta^i + \alpha^i p_A + \xi_{AX} + \epsilon^i_{AX} > 0 \tag{4}$$

The following section provides more details about this utility function and its estimation. So long as the $C(1,1)$ term is not greater than the revenue Generated, it is profitable for the retailer to offer AX.

Now, let us consider the second flavor. The simplest case is for product AY, since the prices of the two products will be the same. The retailer will be willing to add AY to its product assortment so long as:

$$p_A(D_{AX|AX,AY} + D_{AY|AX,AY}) - C(1,2) > p_A D_{AX|AX} - C(1,1), \tag{5}$$

which can be re-written as:

$$p_A(D_{AX|AX,AY} + D_{AY|AX,AY} - D_{AX|AX}) > C(1,2) - C(1,1) \tag{6}$$

When the second flavor is offered, there will be some shifting from AX to AY for those consumers for whom $U(AY) > U(AX)$.Even if we assume that product AX has a higher mean utility, there will be a certain share of the population that prefers an alternative product, like AY. However, such a demand shift is revenue neutral for the retailer as long as we assume that the retailer keeps prices fixed. Therefore, any additional revenue earned when AY is also offered comes from the consumers who did not previously purchase ice cream when AX was the only flavor available. These are the measure of i for which $U(AX) < 0$ and $U(AY) > 0$.

The situation is a bit more complex to evaluate when the prices are different for the products that are added. For example, it would be optimal to add a flavor of a different brand - say BY - if the following holds:

$$p_A D_{AX|AX,BY} + p_B D_{BY|AX,BY} - C(2,2) > p_A D_{AX|AX} - C(1,1) \qquad (7)$$

which can be re-written as:

$$p_A(D_{AX|AX,AY} - D_{AX|AX}) + p_B D_{BY|AX,BY} > C(2,2) - C(1,1) \qquad (8)$$

Now, the weighting by the different prices makes things a bit more difficult. The first term on the left-hand side is for sure a negative number, since there will be consumers of AX lost now that BY is also available - specifically, those consumers for which $U(BY) > U(AX)$ and $U(AX) > 0$. Those consumers lost must be more than made up for by the consumers purchasing product BY. They are the consumers for which $U(BY) > U(AX)$ and $U(BY) > 0$. The total number of consumers purchasing the product ought to be larger to the extent that there are more consumers for which only one of the products has utility greater than zero. These new quantities then need to be weighted by the relative prices and the total additional profits (assuming the left-hand side is positive) must exceed the cost of offering an additional flavor that is a different brand. [6]

From this example, it is clear that substitution patterns take on additional relevance when two or more products are offered. It is not certain, for example, that the alternative with the second highest average utility will be chosen as the second product that the retailer stocks. Suppose that consumers are distributed as follows in terms of their utility functions - such that there are two possible consumer types with these preference orderings for products:

Consumer Type	Type 1	Type 2
Share of All Consumers	75%	25%
Favorite Product	AY	BZ
2nd Favorite Product	BY	AY
3rd Favorite Product	AX	BY
Least Favorite Product	BZ	AX

If the retailer offers only one product, AY would likely be the choice since it generates the greatest utility from consumers of type 1 and may even generate some interest from consumer type 2 (since it is their second favorite flavor, there may be several type 2 consumers for whom $U(AY) > 0$).

When considering the choice of a second product, the retailer may well find the product BZ as the optimal choice - even though most consumers find it to be their least favorite option. The product BY likely has the

[6] Note that price differences may also occur when the two products offered are the same brand, as the profit maximizing price may increase to the extent that more consumers will see their highest utility alternative increase when additional products are offered.

second highest average utility; however, few consumers would switch from product AY to product BY if both were offered. In contrast, product BZ will attract many type 2 consumers, some of whom would be willing to switch from product AY and others that previously did not purchase. In addition, the retailer may be able to raise its price for product AY since it would no longer be trying to sell this product to type 2 consumers. Offering a second flavor would be profitable for the retailer provided that the revenue from these two sources exceeds the additional costs associated with the extra product BZ.

Therefore, there are two critical factors necessary for the retailers to consider in optimal flavor selection. The first is, naturally, overall utility – retailers will be able to sell more ice cream at a higher price to the extent that consumers demand those products more. This is the intuition that is provided by the analysis of Aydin and Ryan, when they conclude that higher margin products are those most likely to be offered. Since we do not have any available data on costs, we will only be able to examine utility differences among flavors.

Beyond that, optimal product selection will depend critically on the extent to which utility for the products with lower demand are correlated with the utility for the most desired products. If product B has a lower average utility than product A, and U(B) ¿ U(A) for every consumer, then introducing product B in addition to product A would add nothing to the retailerÆs profits. On the other hand, if a product C exists such that there is some measure of consumers that have U(C) ¿ U(A), then it may be profitable to introduce product C along with product A. Note that in this case, the retailer would profit more from offering product C, even if it had a lower average utility than product B. To the extent that consumersÆ utility for products are more highly correlated among products within the same quality tier, or among same-flavor or same-brand products, we would expect the retailers' optimal product portfolio to exhibit greater levels of variability.

Note that factoring in optimal product selection may complicate demand estimation and calculation of elasticities. Our identification strategy takes advantage of the considerable variation in the portfolio of products offered in each store-week observation. However, there may be different implications for demand estimation if these portfolios are generated based on retailer profit maximization as opposed to being selected randomly, which we will assume to start. We will analyze how the product portfolios offer might differ from what we would expect if retailers were maximizing profits using our price elasticity estimates.

Finally, the analysis of retailer behavior will have implications for manufacturers. This will be particularly true with regard to the choice made by firms over which flavors to manufacture. Prior literature has demonstrated an advantage in having longer product lines – the analysis here will focus on the importance of the composition of product lines in terms of flavors offered. Clearly, manufacturers will want to produce products that retailers prefer to stock on their shelves, so to the extent that retailers choose based on demand substitution patterns, offering unique flavors may be more profitable for the manufacturers. It is interesting to note that - returning to Table 5 - while the number of flavors offered by firms in the various quality tiers is similar, it is striking that most of the unique flavors are offered by the high quality manufacturers. We intend to use the retailer model of flavor selection as a sort of demand model with respect to the supply choices made by manufacturers.

To summarize, we will use our estimates to evaluate a series of questions regarding the role of flavor portfolio competition, such as:

- How do the flavors offered by retailers rank in terms of average utilities (i.e., evaluating Aydin & Ryan's (2000) hypothesis)?

- What are the terms of the tradeoff between overall utility and heterogeneity in consumersÆ preferences in determining the profitability of retailers portfolio selection? In other words, can we find evidence that the higher utility products that are not offered as regularly generate little additional demand beyond those products that are offered (and vice versa)?

- How important/distinct are the various quality tiers in this market? Are the substitution patterns of consumers variable within and across quality tiers? If so, does this help explain why the relationship between quality and variety in our data does not conform to the prediction of Shugan's model?

- What do the utility function estimates and product line assortment decisions of retailers imply for the ice cream manufacturers? Under what circumstances should firms at each quality level produce new or unique flavors?

4 Demand Model

Let $j = 1, \ldots, J$ denote brands and $f = 1, \ldots, F$ flavors. Let F_{jt} denote the flavors offered by brand j in period t. We assume that at each consumption occasion, consumer i selects one brand-flavor combination that maximizes his/her utility (Nevo 2001, Villas-Boas 2001) or decides to not consume anything, in which case his/her utility is $U^i_{0t} = \epsilon^i_{0t}$.

The utility of a brand-flavor combination jf is given by

$$U^i_{jft} = d_{jf} + x'_{jft}\beta^i + \alpha^i p_{jt} + \xi_{jft} + \epsilon^i_{jft}, \tag{9}$$

where d_{jf} are fixed effects capturing the intrinsic preference for a given brand-flavor combination, x_{jft} is a vector of observed product characteristics, p_{jt} denotes the price of brand j in period t. The random coefficients β^i and α^i allow us to capture the heterogeneity of consumer responses to price and observed characteristics such as feature and display. We model this response as:

$$[\beta^i, \alpha^i]' = [\beta, \alpha]' + \Upsilon \nu^i, \tag{10}$$

where $\nu^i \sim N(0, I)$ is a consumer-specific standard-normal random vector and Υ is a matrix to be estimated from the data.

The term ξ_{jft} explicitly acknowledges the existence of time-varying attributes such as national advertising and shelf space that affect consumer utility but are unobserved by the researcher (Berry 1994, Besanko, Gupta & Jain 1998). ϵ^h_{jft} is an idiosyncratic demand shock, i.i.d. extreme value distributed.

If there are too many brand-flavor combinations, we may focus on the main effects of brand and flavor by setting $d_{jf} = d_j + d_f$, where d_j denotes the brand effect and d_f denotes the flavor effect (Fader & Hardie 1996). An additional benefit of this formulation is that it enables us to evaluate new product-brand/flavor offerings as long as they are a combination of existing ones. For example, if Dreyers were contemplating the launch of cherry/vanilla flavor, the utility of this new offering can be computed by adding up the brand effect for Dreyers and the flavor effect for cherry/vanilla. Since brand and flavor effects always appear as a sum, we need to normalize one brand or flavor effect.

The market share of a brand-flavor combination in week t is given by:

$$S_{jft} = \int \frac{\exp(d_{jf} + x'_{jft}\beta + \alpha p_{jt} + \xi_{jft} + [x_{jft}, p_{jt}]'(\Upsilon\nu))}{1 + \sum_{j=1}^{J} \sum_{f \in F_{jt}} \exp(d_{jf} + x'_{jft}\beta + \alpha p_{jt} + \xi_{jft} + [x_{jft}, p_{jt}]'(\Upsilon\nu)} dF(\nu). \tag{11}$$

The market share for the full random coefficients model defined above cannot be expressed in closed form. If consumer heterogeneity enters only through the idiosyncratic error term ϵ^i_{jft}, then the market share can be written as:

$$S_{jft} = \frac{\exp(d_{jf} + x'_{jft}\beta + \alpha p_{jt} + \xi_{jft})}{1 + \sum_{j=1}^{J} \sum_{f \in F_{jt}} \exp(d_{jf} + x'_{jft}\beta + \alpha p_{jt} + \xi_{jft})}. \tag{12}$$

To obtain additional insights into consumers' choice process, we use a nested logit model and compare different nesting structures (which in turn imply different decision process). Similar to the random coefficients model presented above, this model also allows for some correlation in tastes across products (McFadden 1980). We follow Cardell (1997) and Berry (1994) in our exposition of the nested logit model and assume a variance components structure.[7] All products are grouped in mutually exclusive sets, with the outside good being the only member of group 0. We explore two different possibilities: (1) all flavors are grouped by brand and consumers select flavors conditional on brand choice; (2) all brands, which offer a particular flavor form a group, and consumers first choose the flavor then the brand.

In the case of flavor choice conditional on brand choice, the idiosyncratic error term is given by

$$\epsilon^i_{jt} + (1-\sigma)\epsilon^i_{jft}, \ \sigma \in [0, 1), \tag{13}$$

where the parameter σ captures the within group (nest) correlation. If it equals zero, then the multinomial logit model obtains. The share for a brand-

[7] Cardell (1997) shows that ϵ^i_{jt} has a distribution function that depends on σ with the property that, if ϵ^i_{jft} is extreme-value distributed, then $\epsilon^i_{jt} + (1-\sigma)\epsilon^i_{jft}$ is also extreme-value distributed.

flavor combination in this case can be written as

$$S_{jft} = \int \frac{\exp\left[(d_{jf} + x'_{jft}\beta + \alpha p_{jt} + \xi_{jft} + [x_{jft}, p_{jt}]'(\Upsilon\nu))/(1-\sigma)\right]}{\{A_{jt}\}^{\sigma}\left(1 + \sum_{j'}\{A_{j't}\}^{1-\sigma}\right)} dF(\nu), \tag{14}$$

where

$$A_{jt} = \sum_{f' \in F_{jt}} \exp\left[(d_{jf'} + x'_{jf't}\beta + \alpha p_{jt} + \xi_{jf't} + [x_{jf't}, p_{jt}]'(\Upsilon\nu))/(1-\sigma)\right].$$

In the case where the nests are defined by flavors, i.e., consumers choose first the flavor they wish to consume and then the brand the idiosyncratic term is given by

$$\epsilon^{i}_{ft} + (1-\sigma)\epsilon^{i}_{jft},\ \sigma \in [0,1). \tag{15}$$

To simplify notation, we define

$$B_{ft} = \sum_{j' \in J_{ft}} \exp\left[\exp(d_{j'f} + x'_{j'ft}\beta + \alpha p_{j't} + \xi_{j'ft} + [x_{j'ft}, p_{j't}]'(\Upsilon\nu))/(1-\sigma)\right],$$

where J_{f_t} is the set of brands that offer flavor f in week t. The market share of a brand-flavor combination is then

$$S_{jft} = \int \frac{\exp\left[(d_{jf} + x'_{jft}\beta + \alpha p_{jt} + \xi_{jft} + [x_{jft}, p_{jt}]'(\Upsilon\nu))/(1-\sigma)\right]}{\{B_{ft}\}^{\sigma}\left(1 + \sum_{f'}\{B_{f't}\}^{1-\sigma}\right)} dF(\nu). \tag{16}$$

5 Estimation Results

In this section we report the results for the estimation of the demand system described in Section 4. We start by estimating a multinomial logit model where consumer heterogeneity enters only through the idiosyncratic shock. This model formulation allows for a closed-form expression of the market shares as a function of the marketing variables and product characteristics and can be estimated using standard OLS and 2SLS methods. Adding heterogeneity in the response parameters yields the full random coefficients model, where the market shares cannot be expressed in closed form. We follow Berry (1994), Berry, Levinsohn & Pakes (1995), and Nevo (2001) and apply a contraction mapping procedure in order to obtain our parameter estimates.

Our model explicitly acknowledges the presence of unobserved (to the researcher) factors that may affect demand such as national

advertising and shelf space allocation. Ignoring these unobserved attributes has been shown to lead to understated estimates of the price response parameter (Besanko et al. 1998, Villas-Boas & Winer 1999). For this reason we need to instrument for prices. The instruments should ideally be highly correlated with prices and uncorrelated with the demand shocks. We use a number of cost factors such as federal interest rates, prices of cream, milk, and weekly wages in the dairy industry. We get overidentifying restrictions by interacting our original instruments with a full set of brand dummies as in Villas-Boas (2001).

To assess the importance of accounting for the endogeneity of prices, we first compared our parameter estimates obtained using OLS to the ones obtained after instrumenting for prices. The OLS price estimates across all model specifications were much lower in absolute value than the 2SLS estimates, which is consistent with previous findings in the empirical literature.[8]

Tables 9 and 10 report the 2SLS estimation results. We estimate four models: a multinomial logit (Standard Logit), a nested logit with nests defined by brands (Nested Logit BF), a nested logit with nests defined by flavors (Nested Logit FB), and a full random coefficients logit model. With respect to the fixed effects, we explore two specifications: (1) fixed effects for all brand-flavor combinations (Table 9), and (2) hedonic specification with separate brand and flavor fixed effects (Table 10).

Looking at the results reported in Table 9, we see that the estimated brand-flavor demand constants are negative, which is easily explained by the large market share of the outside alternative which serves as a base for the estimation. All marketing variables are significant and have the expected signs:

The effect of price on demand is negative, whereas feature advertising and display positively affect demand across the different models. None of the estimated standard deviations is significant, however. The model with separate brand and flavor effects (Table 10) produces intuitive results, except

[8] The results of the OLS estimation are not reported to conserve space but are available from the authors upon request.

for the coefficient on feature, which is only significant and positive for the nested logit BF specification. In this model, the standard deviation of price is significant, suggesting that there may be a fair amount of unobserved heterogeneity in price response. The estimated brand coefficients are consistent with our notion that there are three different quality tiers in the ice cream category: Häagen Dazs and Ben & Jerry have the highest values, 4.70 and 3.67, respectively, followed by Dreyers (1.11) and Breyers (1.80), Sealtest (0.49) and Schoeps (0). The most attractive flavor is vanilla (-3.82), the least attractive is honey vanilla (-6.19)[9].

An interesting question is to see how the perceived quality captured by the estimated brand-flavor fixed effects relates to objective quality measures in the ice cream category such as calories or fat content. To this end, we regress the estimated brand-flavor constants on a number of product characteristics. As evident from Table 11, the calories from fat and vanilla flavor positively affect perceived quality. Since butterfat content is the a key determinant of ice cream quality, it appears that consumers valuations are in line with industry quality definitions. The total calories and sugar content are highly collinear with fat calories, which explains the relatively high standard errors on the estimated coefficients and the ensuing lack of statistical significance. The presence of nuts, fruit, or chocolate seems to be a purely horizontal characteristic, not related to perceived quality. On the other hand, vanilla has a positive effect, which can be explained by the fact that vanilla is a versatile

[9]It has since been discontinued.

flavor that can be not only enjoyed on its own but also in combination with a variety of desserts such as fruit or cake. Versatility can thus be viewed as a quality-related characteristic of an ice cream brand-flavor combination.

The manufacturers seem to also know which attributes are valued by consumers and charge them accordingly as a hedonic regression of average prices on product characteristics reveals. As evident from Table 12, prices are significantly related to calories and sugar content. The effect of calories from fat is not significant but this is due to the high correlation of this variable with the total calories measure.

Comparing the SSE's of the different model specifications, it appears that consumers first select a brand, and then choose among the flavors within this brand. This conclusion is corroborated by the fact that σ is statistically significant from zero, thus allowing us to reject the hypothesis of equivalence between the nested logit and the multinomial logit model. The parameter σ can be also viewed as measuring the similarity of the alternatives within a nest. The value of 0.725(0.036) indicates that flavors within a brand are perceived as fairly similar.

Within the individual nests the nested logit model, however, shares the unappealing IIA property of the standard logit model. To obtain more flexible substitution patterns, we need to look at the full random coefficients model. We use this model to obtain a matrix of own and cross-price elasticities.

The matrix of cross-price elasticities for all 79 products is too large to include (and interpret), so we report a number of summary measures to get a feel for the substitution patterns. Specifically, we define a measure of substitution between flavors within a brand and different brands offering the same flavor as the ratio of the average cross-price elasticity within a brand and the average cross-price elasticity across brands (for a given flavor). Table 13 displays this ratio for the five flavors manufactured by at least five firms. Interestingly, the middle quality tier has the lowest average ratio. This may result because these brands are relatively close substitutes with both the low and high quality brands. It is also worth noting that the flavor with the highest average ratio is chocolate chip, which could be argued has several close substitutes in other flavors such as Cookies & Cream, Cookie Dough, and Chocolate Chocolate Chip.

Looking at Table 14, we see that price response varies by quality tier. This finding is in contrast with one of Shugan's key assumptions and explains why we do not observe the pattern of quality-variety interaction his model predicts. Note, however, that logit elasticities are higher for higher-priced products. To relax this assumption, we propose to let the price response vary by quality tier. Including interaction terms in the demand model reveals that the price sensitivity is by far lowest for the high quality brands (a price coefficient of -17), followed by the coefficients for the medium tier (-59), and he low-quality tier (-76). Looking at the corresponding price elasticities, we see that the lowest quality tier still has the highest price sensitivity but now the high quality-tier consumers seem to be more price sensitive than the middle tier ones.

Finally, we want to use our demand estimates to begin to explore the retailers decision regarding category assortments. Recall from our stylized retailer model in Section 3 the intuition that substitution among products at the demand level was as important (if not more) as overall utility of products in determining which products retailers ought to carry. Using our empirical results, we computed a baseline utility measure for each of the 79 products that retailers could potentially offer and divided them into quartiles. Then, for a typical week, we compared the products that stores actually offered against these rankings. The results in Table 15 indicate that, while skewed slightly toward the highest utility products, retailers fill out their assortments from flavors all throughout the utility ranking. While it may be initially surprising that retailers do not necessarily offer the most popular products, in the context of our stylized model popular flavors that are similar to other flavors would be passed over. We will look in greater detail at similarities among products to establish the remainder of this connection.

6 Discussion and Conclusions

In this paper, we have begun to explore the role of product line composition in the behavior of consumers, retailers and manufacturers. Our analysis of data from the ice cream category suggests that consumers base their purchase decisions on brand first, then on flavor. Nonetheless, retailers appear to stock flavors in such a way as to avoid having multiple brands of the same flavor – particularly for brands in the same quality tier. Retailers also tend to offer more flavors of higher quality brands, even though the manufacturers produce roughly the same

number of flavors in all quality levels. In future work, we will attempt to explain these potentially conflicting facts in terms of either costs of providing variety, differing retailer margins, or departures from profit-maximizing behavior.

We will also place greater emphasis on the implications of the consumer and retailer results for the product-line decisions made by manufacturers. While previous research has established benefits associated with increased line length, our estimates allow us to evaluate the relative benefits of offering unique flavors versus more popular varieties. In doing so, we plan to more explicitly consider the connections between product quality and variety in terms of both incentives to take advantage of consumer demand and the role of competition in horizontal and vertical dimensions.

Table 1: Means of variables in the data set (standard deviations in parentheses).

brand	Häagen Dazs	Ben & Jerry's	Dreyers	Breyers	Sealtest	Schoeps
# ounces sold	4302.49	2252.09	4760.00	10398.15	3178.21	1853.75
	(3516.76)	(2111.29)	(6928.80)	(8508.59)	(3470.71)	(3437.18)
market share (in %)	1.96	1.02	2.16	4.73	1.44	0.84
	(1.60)	(0.96)	(3.15)	(3.87)	(1.58)	(1.56)
price per ounce	0.15	0.16	0.06	0.06	0.05	0.05
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
display	0.01	0.01	0.03	0.01	0.01	0.02
	(0.09)	(0.09)	(0.18)	(0.09)	(0.10)	(0.15)
feature	0.08	0.04	0.10	0.14	0.17	0.11
	(0.26)	(0.20)	(0.30)	(0.35)	(0.38)	(0.31)

Table 2: Product characteristics for brands in the data set. Means reported with standard deviations in parentheses.

brand	Häagen Dazs	Ben & Jerry's	Dreyers	Breyers	Sealtest	Schoeps
calories	290.00	278.33	163.57	158.00	136.67	140.77
	(27.33)	(22.90)	(17.81)	(14.74)	(8.66)	(11.15)
fat calories	170.63	151.83	77.86	83.33	57.78	60.77
	(22.05)	(16.52)	(13.11)	(11.13)	(10.93)	(10.38)
sugar	22.13	23.75	14.21	15.13	12.33	15.54
	(1.71)	(3.14)	(1.37)	(1.30)	(0.50)	(1.51)
chocolate	0.31	0.50	0.43	0.33	0.22	0.31
fruit	0.06	0.17	0.07	0.07	0.11	0.08
nuts	0.19	0.17	0.21	0.27	0.11	0.08
pure flavor	0.25	0.17	0.21	0.27	0.33	0.31
vanilla	0.06	0.08	0.07	0.07	0.11	0.08

Table 3: Flavors offered by the brands in the data set.

Flavor	Häagen Dazs	Ben & Jerry's	Dreyers	Breyers	Sealtest	Schoeps
Butter Almond				X		
Butter Pecan	X		X	X	X	X
Cherry Garcia		X				
Chocolate	X	X	X	X	X	X
Chocolate Chip		X	X	X	X	X
Cookie Dough		X	X			X
Chocolate Chocolate Chip	X					
Chocolate Fudge	X		X			
Chocolate Fudge Brownie		X	X			
Chocolate Peanut Butter	X					
Chunky Monkey		X				
Coffee	X			X		
Coffee Heath Crunch		X				
Cookies & Cream	X		X	X		X
French Vanilla			X	X	X	X
Heath Bar Crunch		X				
Heavenly Hash					X	X
Honey Vanilla	X					
Macadamia Brittle	X					
Mint Chocolate Chip	X			X		X
Mint Chocolate Cookie		X				
Mocha Almond Fudge			X	X		
Neapolitan			X	X	X	X
New York Cherry						X
NY Super Fudge Crunch		X				
Peanut Butter Cup			X			
Rain Forest Crunch		X				
Rocky Road			X	X		X
Rum Raisin	X					
Strawberry	X		X	X	X	X
Toffee Coffee	X					
Vanilla	X	X	X	X	X	X
Vanilla Fudge	X			X	X	
Vanilla Swiss Almond	X					
Vanilla/Chocolate				X		
Total Flavors	16	12	14	15	9	13

Table 4: Number of brands offering a given flavor (among the data set's 520 store/week observations)

	# of brands	0	1	2	3	4	5	6
Butter Almond	1	416	104	—	—	—	—	—
Butter Pecan	5	0	1	57	181	161	120	—
Cherry Garcia	1	16	504	—	—	—	—	—
Chocolate	6	0	0	59	244	204	13	0
Chocolate Chip	5	25	142	271	82	0	0	—
Cookie Dough	3	192	279	37	12	—	—	—
Chocolate Chocolate Chip	1	6	514	—	—	—	—	—
Chocolate Fudge	2	151	322	47	—	—	—	—
Chocolate Fudge Brownie	2	46	256	218	—	—	—	—
Chocolate Peanut Butter	1	340	180	—	—	—	—	—
Chunky Monkey	1	50	470	—	—	—	—	—
Coffee	2	2	167	351	—	—	—	—
Coffee Heath Crunch	1	135	385	—	—	—	—	—
Cookies & Cream	4	12	204	116	172	16	—	—
French Vanilla	4	38	253	131	51	47	—	—
Heath Bar Crunch	1	15	505	—	—	—	—	—
Heavenly Hash	2	257	230	33	—	—	—	—
Honey Vanilla	1	261	259	—	—	—	—	—
Macadamia Brittle	1	131	389	—	—	—	—	—
Mint Chocolate Chip	3	11	421	88	0	—	—	—
Mint Chocolate Cookie	1	187	333	—	—	—	—	—
Mocha Almond Fudge	2	306	214	0	—	—	—	—
Neapolitan	4	3	94	300	123	0	—	—
New York Cherry	1	320	200	—	—	—	—	—
NY Super Fudge Crunch	1	16	504	—	—	—	—	—
Peanut Butter Cup	1	356	164	—	—	—	—	—
Rain Forest Crunch	1	123	397	—	—	—	—	—
Rocky Road	3	167	191	132	30	—	—	—
Rum Raisin	1	233	287	—	—	—	—	—
Strawberry	5	0	25	276	176	43	0	—
Toffee Coffee	1	175	345	—	—	—	—	—
Vanilla	6	0	0	2	13	147	236	122
Vanilla Fudge	3	0	1	258	255	—	—	—
Vanilla Swiss Almond	1	7	513	—	—	—	—	—
Vanilla/Chocolate	1	420	100	—	—	—	—	—

Table 5: Differences in offerings across brands: How many flavors are offered by how many other brands?

	Number of Flavors Offered by N Other Firms						
Firm	0	1	2	3	4	5	Total
Häagen Dazs	7	2	2	1	2	2	16
Ben & Jerry's	7	1	1	0	1	2	12
Dreyers	1	3	2	3	3	2	14
Breyers	2	2	3	3	3	2	15
Sealtest	0	1	1	2	3	2	9
Schoeps	1	1	3	3	3	2	13
Number of Flavors	18	5	4	3	3	2	35

Table 6: Flavor assortment by brand and store.

	Average Number of Brand's Flavors Sold per Week in Store					
Firm	A	B	C	D	E	Average
Ben & Jerry's	7.8	10.5	6.1	9.7	11.2	9.1
Häagen Dazs	11.7	11.1	10.4	13.2	14.1	12.1
Breyers	10.7	8.8	9.8	8.7	8.5	9.3
Dreyers	6.7	11.7	5.1	—	—	7.8
Sealtest	3	4	6.8	7.3	7.6	5.7
Schoeps	2.4	6.1	—	2.3	2.6	3.4

Note: Averages are calculated over the retailers that ever offer that brand.

Table 7: Differences in products offered by retailers: How do differences in quality levels affect the flavor portfolio?

Flavor	None	One	Two			Three
			Sealtest & Schoeps	Schoeps & Breyers	Sealtest & Breyers	
Butter Pecan	7	87	1	30	230	165
Chocolate	1	194	12	45	243	0
Chocolate Chip	97	359	53	1	10	0
French Vanilla	68	279	40	6	60	67
Neapolitan	3	118	0	6	376	17
Strawberry	49	266	2	80	98	25

Table 8: Variety measures of retailers' assortments.

Store	Measure # 1	Measure # 2	Measure # 3
A	42.3	0.3	1162.12
B	52.2	0.42	1784.04
C	39.3	0.21	991.07
D	41.2	0.25	1090.71
E	43.9	0.29	1239.01
Total	43.8	0.29	1253.39

Table 9: Estimation results. All brand-flavor fixed effects included. Standard errors in parentheses.

Parameter	Standard Logit		Nested Logit BF		Nested Logit FB		Random Coefficients	
sigma			0.725	(0.036)	0.413	(0.042)		
price	-25.694	(2.858)	-24.074	(2.355)	-29.131	(2.724)	-21.253	(2.521)
SD price							5.767	(4.914)
display	1.056	(0.504)	2.436	(0.421)	0.856	(0.477)	1.552	(0.457)
SD display							0.084	(21.685)
feature	1.270	(0.203)	1.082	(0.167)	0.953	(0.194)	0.999	(0.175)
SD feature							0.040	(25.793)
a1001	-0.503	(0.450)	0.034	(0.371)	0.645	(0.441)	-2.279	(0.394)
a1003	-2.220	(0.450)	-0.437	(0.380)	-0.847	(0.447)	-3.994	(0.399)
a1011	-2.510	(0.450)	-0.516	(0.383)	-1.951	(0.429)	-4.281	(0.397)
a1012	-2.817	(0.456)	-0.570	(0.391)	-2.161	(0.436)	-4.640	(0.405)
a1013	-2.611	(0.482)	-0.453	(0.411)	-2.024	(0.459)	-4.564	(0.425)
a1014	-1.705	(0.450)	-0.297	(0.377)	-0.624	(0.439)	-3.475	(0.395)
a1017	-2.030	(0.450)	-0.381	(0.379)	-1.194	(0.434)	-3.800	(0.397)
a1019	-2.600	(0.454)	-0.525	(0.388)	-1.491	(0.444)	-4.379	(0.400)
a1023	-3.227	(0.445)	-0.781	(0.386)	-2.679	(0.424)	-5.030	(0.393)
a1024	-2.885	(0.453)	-0.606	(0.390)	-2.323	(0.432)	-4.673	(0.401)
a1027	-2.809	(0.528)	-0.402	(0.451)	-1.306	(0.522)	-4.916	(0.441)
a1038	-3.206	(0.457)	-0.764	(0.395)	-2.645	(0.436)	-5.080	(0.402)
a1040	-2.378	(0.449)	-0.481	(0.382)	-1.199	(0.441)	-4.155	(0.396)
a1041	-2.802	(0.457)	-0.607	(0.392)	-2.234	(0.436)	-4.573	(0.404)
a1044	-2.306	(0.448)	-0.470	(0.380)	-1.070	(0.441)	-4.055	(0.396)
a1045	-2.304	(0.450)	-0.459	(0.381)	-1.745	(0.429)	-4.055	(0.397)
a2001	-2.669	(0.460)	-0.953	(0.388)	-0.542	(0.485)	-4.495	(0.405)
a2007	-2.053	(0.465)	-0.730	(0.388)	-1.487	(0.443)	-3.904	(0.411)
a2009	-2.577	(0.467)	-0.861	(0.394)	-1.474	(0.455)	-4.467	(0.413)
a2010	-1.536	(0.468)	-0.536	(0.389)	-0.920	(0.447)	-3.363	(0.413)
a2014	-3.043	(0.467)	-0.949	(0.398)	-1.337	(0.474)	-4.915	(0.412)
a2015	-2.020	(0.464)	-0.716	(0.387)	-1.244	(0.445)	-3.834	(0.410)
a2016	-2.269	(0.466)	-0.755	(0.391)	-1.701	(0.444)	-4.115	(0.411)
a2018	-2.320	(0.466)	-0.741	(0.391)	-1.753	(0.444)	-4.156	(0.411)
a2021	-2.003	(0.464)	-0.730	(0.387)	-1.437	(0.442)	-3.838	(0.408)
a2028	-2.507	(0.457)	-0.885	(0.385)	-1.951	(0.436)	-4.334	(0.403)
a2032	-2.056	(0.465)	-0.731	(0.388)	-1.489	(0.443)	-3.894	(0.411)
a2036	-2.685	(0.462)	-0.888	(0.390)	-2.122	(0.440)	-4.515	(0.408)
a3001	-3.528	(0.186)	-2.651	(0.159)	-2.337	(0.213)	-4.509	(0.164)
a3003	-4.129	(0.185)	-2.780	(0.166)	-3.075	(0.205)	-5.106	(0.164)
a3009	-4.403	(0.189)	-2.471	(0.182)	-3.954	(0.184)	-5.338	(0.163)
a3010	-3.804	(0.244)	-2.422	(0.212)	-3.154	(0.239)	-4.939	(0.239)
a3012	-4.547	(0.188)	-2.550	(0.183)	-4.248	(0.180)	-5.426	(0.163)
a3014	-4.623	(0.189)	-2.631	(0.184)	-3.523	(0.210)	-5.571	(0.166)
a3015	-4.256	(0.188)	-2.758	(0.171)	-3.784	(0.184)	-5.216	(0.166)
a3019	-4.297	(0.187)	-2.680	(0.174)	-3.364	(0.201)	-5.257	(0.165)
a3020	-4.332	(0.184)	-2.563	(0.175)	-3.575	(0.190)	-5.270	(0.161)
a3029	-4.426	(0.191)	-2.674	(0.179)	-4.180	(0.183)	-5.347	(0.166)
a3030	-4.607	(0.189)	-2.608	(0.184)	-3.433	(0.214)	-5.538	(0.166)
a3034	-4.363	(0.188)	-2.644	(0.176)	-4.127	(0.179)	-5.295	(0.165)
a3037	-4.233	(0.187)	-2.794	(0.169)	-3.721	(0.184)	-5.200	(0.164)
a3040	-4.583	(0.207)	-2.495	(0.199)	-3.718	(0.214)	-5.565	(0.183)
a4001	-2.625	(0.179)	-1.717	(0.154)	-1.981	(0.182)	-3.621	(0.157)
a4002	-3.696	(0.174)	-1.819	(0.170)	-3.501	(0.166)	-4.700	(0.148)
a4003	-3.768	(0.180)	-2.030	(0.171)	-3.123	(0.182)	-4.754	(0.160)
a4009	-4.019	(0.296)	-1.922	(0.264)	-3.600	(0.283)	-5.109	(0.196)
a4014	-3.912	(0.180)	-2.069	(0.173)	-3.296	(0.181)	-4.895	(0.160)
a4017	-4.186	(0.193)	-2.337	(0.183)	-3.671	(0.190)	-5.186	(0.168)
a4019	-3.350	(0.171)	-1.645	(0.164)	-2.954	(0.167)	-4.283	(0.150)
a4020	-3.634	(0.192)	-1.924	(0.179)	-3.086	(0.189)	-4.581	(0.184)
a4027	-3.443	(0.179)	-1.937	(0.165)	-3.192	(0.171)	-4.427	(0.157)
a4029	-5.048	(0.237)	-2.863	(0.223)	-4.763	(0.226)	-6.154	(0.197)
a4030	-3.779	(0.179)	-2.028	(0.171)	-3.331	(0.175)	-4.759	(0.161)
a4037	-4.177	(0.184)	-2.072	(0.183)	-3.760	(0.179)	-5.135	(0.164)
a4040	-3.946	(0.182)	-2.094	(0.175)	-3.449	(0.179)	-4.947	(0.164)
a4042	-3.943	(0.175)	-1.888	(0.176)	-3.747	(0.166)	-5.012	(0.157)
a4044	-3.849	(0.179)	-2.054	(0.172)	-3.363	(0.176)	-4.834	(0.159)
a5001	-4.098	(0.153)	-3.139	(0.134)	-2.974	(0.184)	-5.006	(0.134)
a5003	-4.509	(0.153)	-3.192	(0.142)	-3.662	(0.168)	-5.408	(0.137)
a5009	-5.021	(0.160)	-3.536	(0.150)	-4.609	(0.157)	-5.961	(0.141)
a5014	-5.161	(0.164)	-3.398	(0.161)	-4.247	(0.181)	-6.096	(0.144)
a5020	-4.337	(0.152)	-3.220	(0.136)	-3.916	(0.150)	-5.246	(0.137)
a5022	-5.240	(0.165)	-3.622	(0.157)	-4.985	(0.158)	-6.218	(0.144)
a5030	-4.800	(0.152)	-3.332	(0.145)	-4.054	(0.163)	-5.695	(0.137)
a5040	-5.378	(0.177)	-3.507	(0.172)	-4.629	(0.184)	-6.385	(0.154)
a5044	-5.128	(0.165)	-3.409	(0.160)	-4.283	(0.178)	-6.057	(0.146)
a6001	-4.095	(0.164)	-3.690	(0.136)	-2.920	(0.195)	-5.143	(0.149)

Table 9: Estimation results. All brand-flavor fixed effects included. Standard errors in parentheses.

Parameter	Standard Logit		Nested Logit BF		Nested Logit FB		Random Coefficients	
a6003	-4.840	(0.159)	-3.662	(0.143)	-3.771	(0.185)	-5.762	(0.146)
a6009	-5.521	(0.178)	-3.940	(0.166)	-4.692	(0.188)	-6.404	(0.149)
a6010	-5.324	(0.224)	-3.758	(0.200)	-4.382	(0.233)	-6.219	(0.166)
a6014	-5.750	(0.215)	-3.589	(0.206)	-4.342	(0.248)	-6.415	(0.149)
a6019	-5.215	(0.201)	-3.749	(0.181)	-4.250	(0.214)	-6.315	(0.162)
a6020	-5.414	(0.176)	-3.900	(0.163)	-4.343	(0.199)	-6.291	(0.153)
a6022	-5.724	(0.191)	-4.094	(0.176)	-5.284	(0.186)	-6.643	(0.155)
a6027	-5.442	(0.180)	-3.733	(0.170)	-4.313	(0.205)	-6.326	(0.150)
a6030	-5.581	(0.254)	-4.120	(0.221)	-4.792	(0.253)	-6.606	(0.178)
a6031	-4.977	(0.160)	-3.698	(0.146)	-4.773	(0.152)	-5.889	(0.144)
a6037	-5.948	(0.218)	-3.644	(0.212)	-4.553	(0.250)	-6.803	(0.173)
a6040	-5.528	(0.178)	-3.856	(0.168)	-4.430	(0.202)	-6.395	(0.148)
SSE	490.524		270.060		422.445		535.502	

Table 10: Estimation results. Hedonic specification. Standard errors in parentheses.

Parameter	Standard Logit		Nested Logit BF		Nested Logit FB		Random Coefficients	
sigma			0.876	(0.019)	0.725	(0.020)		
price	-40.461	(2.686)	-36.147	(2.208)	-42.018	(2.528)	-49.831	(2.543)
SD price							11.239	(2.902)
display	2.734	(0.505)	3.567	(0.415)	1.971	(0.475)	3.817	(0.505)
SD display							0.000	(19.087)
feature	-0.195	(0.187)	0.336	(0.154)	0.111	(0.176)	-0.674	(0.166)
SD feature							0.021	(13.008)
ab1	4.845	(0.280)	4.793	(0.230)	4.390	(0.263)	4.702	(0.262)
ab2	3.911	(0.293)	4.342	(0.241)	4.378	(0.276)	3.674	(0.277)
ab3	1.023	(0.041)	1.273	(0.034)	0.752	(0.040)	1.109	(0.037)
ab4	1.750	(0.039)	2.000	(0.033)	0.650	(0.048)	1.799	(0.034)
ab5	0.496	(0.037)	0.509	(0.031)	-0.105	(0.039)	0.487	(0.033)
ab6	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
af1	-3.496	(0.147)	-2.826	(0.122)	-1.481	(0.149)	-3.824	(0.138)
af2	-4.647	(0.157)	-2.774	(0.135)	-3.470	(0.152)	-5.101	(0.135)
af3	-4.384	(0.147)	-2.874	(0.125)	-2.517	(0.148)	-4.753	(0.139)
af7	-3.547	(0.158)	-2.781	(0.131)	-3.769	(0.149)	-3.909	(0.150)
af9	-4.405	(0.152)	-2.886	(0.129)	-3.526	(0.145)	-4.790	(0.145)
af10	-3.195	(0.161)	-2.654	(0.132)	-3.132	(0.151)	-3.495	(0.156)
af11	-4.979	(0.153)	-2.936	(0.133)	-4.302	(0.145)	-5.378	(0.145)
af12	-5.050	(0.155)	-2.832	(0.136)	-4.337	(0.148)	-5.416	(0.152)
af13	-4.927	(0.187)	-2.707	(0.161)	-4.237	(0.177)	-5.503	(0.173)
af14	-4.513	(0.149)	-2.843	(0.127)	-2.755	(0.149)	-4.877	(0.143)
af15	-3.824	(0.155)	-2.843	(0.129)	-3.446	(0.146)	-4.179	(0.146)
af16	-3.759	(0.159)	-2.764	(0.132)	-3.982	(0.150)	-4.117	(0.150)
af17	-4.627	(0.154)	-2.968	(0.131)	-3.293	(0.150)	-5.030	(0.147)
af18	-3.817	(0.160)	-2.739	(0.134)	-4.037	(0.151)	-4.183	(0.153)
af19	-4.610	(0.148)	-2.788	(0.127)	-3.051	(0.146)	-4.949	(0.139)
af20	-4.207	(0.148)	-2.819	(0.125)	-2.952	(0.143)	-4.595	(0.140)
af21	-3.501	(0.157)	-2.795	(0.130)	-3.724	(0.148)	-3.861	(0.149)
af22	-4.800	(0.160)	-3.084	(0.136)	-4.189	(0.151)	-5.140	(0.150)
af23	-5.741	(0.153)	-3.138	(0.137)	-5.063	(0.145)	-6.190	(0.149)
af24	-5.329	(0.157)	-2.950	(0.139)	-4.655	(0.149)	-5.765	(0.149)
af27	-4.257	(0.150)	-2.824	(0.127)	-2.996	(0.146)	-4.582	(0.142)
af28	-4.050	(0.153)	-2.913	(0.128)	-4.275	(0.144)	-4.406	(0.147)
af29	-4.769	(0.168)	-2.988	(0.143)	-4.213	(0.159)	-4.997	(0.157)
af30	-4.508	(0.148)	-2.865	(0.126)	-2.917	(0.146)	-4.860	(0.140)
af31	-4.212	(0.154)	-2.812	(0.130)	-4.130	(0.145)	-4.501	(0.147)
af32	-3.550	(0.158)	-2.782	(0.131)	-3.772	(0.149)	-3.908	(0.150)
af34	-4.457	(0.164)	-2.808	(0.139)	-4.097	(0.155)	-4.781	(0.153)
af36	-4.193	(0.156)	-2.855	(0.131)	-4.419	(0.147)	-4.544	(0.149)
af37	-4.621	(0.153)	-2.907	(0.131)	-3.435	(0.148)	-4.993	(0.144)
af38	-5.649	(0.163)	-3.063	(0.144)	-4.967	(0.154)	-6.190	(0.153)
af40	-4.758	(0.150)	-2.892	(0.129)	-3.108	(0.148)	-5.119	(0.141)
af41	-5.239	(0.162)	-2.959	(0.141)	-4.557	(0.153)	-5.631	(0.157)
af42	-4.892	(0.158)	-2.806	(0.137)	-3.716	(0.153)	-5.423	(0.147)
af44	-4.677	(0.149)	-2.900	(0.128)	-2.958	(0.148)	-5.010	(0.141)
af45	-4.772	(0.153)	-2.910	(0.132)	-4.096	(0.145)	-5.145	(0.147)
SSE	1783.000		453.148		775.182		1735.210	

Table 11: Regression of brand-flavor constants estimated in 2SLS on product characteristics.

Parameter	Estimate	Std. Err.	t-statistic	$P > \|t\|$
calories	-0.0099	0.0078	-1.262	0.211
fat calories	0.0229	0.0092	2.488	0.015
sugar	0.0611	0.0413	1.480	0.143
chocolate	-0.0147	0.1578	-0.093	0.926
fruit	0.1288	0.2819	0.457	0.649
nuts	-0.0855	0.2126	-0.402	0.689
pure	-0.0807	0.2099	-0.384	0.702
vanilla	1.0251	0.3157	3.247	0.002
intercept	-6.5596	0.3777	-17.368	0.000
adj. R^2	0.589			

Table 12: Regression of average prices on product characteristics.

Parameter	Estimate	Std. Err.	t-statistic	$P > \|t\|$
calories	0.0006	0.0002	3.363	0.001
fat calories	0.0001	0.0002	0.364	0.717
sugar	0.0020	0.0009	2.284	0.025
chocolate	-0.0065	0.0034	-1.895	0.062
fruit	0.0049	0.0061	0.797	0.428
nuts	-0.0039	0.0046	-0.838	0.405
pure	0.0092	0.0045	2.015	0.048
vanilla	0.0015	0.0068	0.225	0.823
intercept	-0.0666	0.0082	-8.141	0.000
adj. R^2	0.938			

Table 13: Flavor substitution measure.

	Vanilla	Chocolate	Butter Pecan	Chocolate Chip	Strawberry	Average
Haagen Dazs	1.38	1.30	1.20	—	1.52	1.35
Ben & JerryÆs	1.48	1.44	—	1.84	—	1.58
Dreyers	1.14	1.27	1.29	1.31	1.42	1.33
Breyers	1.05	0.95	1.00	1.84	1.15	1.20
Sealtest	1.35	1.67	1.21	2.08	1.24	1.51
Schoeps	0.93	1.52	1.51	1.72	2.00	1.54
Average	1.26	1.36	1.24	1.76	1.47	

Table 14: Own price elasticities.

	Random Coefficients	Quality Tiers
Haagen Dazs	-2.54	-3.01
Ben & Jerry's	-2.57	-3.18
Dreyers	-1.24	-2.61
Breyers	-1.18	-2.47
Sealtest	-1.03	-4.64
Schoeps	-1.08	-4.87

Table 15: Number products offered in each utility ranking group.

	Ranking Group				
Store	1-20	21-40	41-60	61-79	Total*
A	13	10	8	11	42
B	16	16	13	11	56
C	14	11	10	7	42
D	10	5	12	15	42
E	11	8	13	16	48
Average	13	10	11	12	46

*Note: 21 brand/flavor combinations are offered by all 5 retailers.

References

Aydin, G. & Ryan, J. (2000). Product-line selection and pricing under the multinomial logit choice model, *working paper*, Stanford University.

Berry, S. (1994). Estimating discrete-choice models of product differentiation, *RAND Journal of Economics* **25**: 242–262.

Berry, S., Levinsohn, J. & Pakes, A. (1995). Automobile prices in market equilibrium, *Econometrica* **63**: 841–890.

Besanko, D., Gupta, S. & Jain, D. (1998). Logit demand estimation under competitive pricing behavior: An equilibrium framework, *Management Science* **44**: 1533–1547.

Bultez, A. & Naert, P. (1988). S.H.A.R.P.: Shelf allocation for retailers' profit, *Marketing Science* **7**: 211–231.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity, *Econometric Theory* **13**: 185–213.

Chong, J.-K., Ho, T.-H. & Tang, C. (2001). A modeling framework for category assortment planning, *Manufacturing and Service Operations Management* **3**(3): 191–210.

Costjens, M. & Doyle, P. (1981). A model for optimizing retail shelf space allocations, *Management Science* **27**: 822–833.

Dobson, G. & Kalish, S. (1988). Positioning and pricing of a product-line: Formulation and heuristics, *Marketing Science* **7**(2): 107–125.

Fader, P. & Hardie, B. (1996). Modeling consumer choice among SKUs, *Journal of Marketing Research* **33**: 442–452.

Hoch, S., Bradlow, E. & Wansink, B. (1999). The variety of an assortment, *Marketing Science* **18**: 527–546.

Marshall, R. & Arbuckle, W. S. (2000). *Ice Cream*, Aspen Publishers, Gaithersburg, MD.

McFadden, D. (1980). Econometric models for probabilistic choice among products, *Journal of Business* **53**(3): S13–S29.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry, *Econometrica* **69**(2): 307–342.

Shugan, S. (1989). Product assortment in a triopoly, *Management Science* **35**: 304–320.

Villas-Boas, M. & Winer, R. (1999). Endogeneity in brand choice models, *Management Science* **45**: 1324–1338.

Villas-Boas, S. (2001). Vertical contracts between manufacturers and retailers: An empirical analysis, *working paper*, UC Berkeley.