Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions*

Michaela Draganska
Graduate School of Business, Stanford University†

Michael Mazzeo
Kellogg School of Management, Northwestern University‡

Katja Seim
Wharton School of Business, University of Pennsylvania§

October 21, 2007

*The authors thank MSI for financial support and Noriko Kakihara for excellent research assistance. This paper has greatly benefited from comments from Ulrich Dorazelski, Brett Gordon, Ken Wilbur and seminar audiences at University of Chicago, Duke, UNC, USC, and the Summer Institute for Competitive Strategy at Berkeley.

†Stanford, CA 94305-5015, U.S.A., draganska_michaela@gsb.stanford.edu.

‡Evanston, IL 60208-2001, U.S.A., mazzeo@kellogg.northwestern.edu.

§Philadelphia, PA 19104-6372, U.S.A., kseim@wharton.upenn.edu.

Abstract

In this paper, we take a first step toward exploring empirically the product assortment strategies of oligopolistic firms. Our starting point is a discrete-choice demand model for differentiated products. We incorporate the demand model into an equilibrium supply model, in which firms compete by first choosing which products to offer and then by setting prices. We show how modeling joint product assortment and pricing decisions enriches standard product choice models by allowing insights into how demand characteristics affect firms' product offerings in a competitive environment. We furthermore demonstrate that incorporating endogenous product choice into demand models is essential for policy simulations (e.g., mergers) as it entails at times dramatically different welfare assessments than the common assumption that product assortments are exogenous.

Keywords: product assortment decisions, multi-product firms, discrete games

JEL Classification: L0, L1, L2, L8, M3

1 Introduction

Decisions about product assortments and prices are among the most fundamental choices firms have to make. When selecting which products to offer, a firm in a competitive environment has to weigh the benefits of a "popular" product space location against the potential downsides of fiercer price competition. Ever since Hotelling's (1929) seminal paper, this fundamental tradeoff has been central to the literature. Deciding how to weigh demand against competitive considerations also remains a primary concern in applied contexts, with managers grappling over pricing and product assortment decisions.

In determining equilibrium product assortments, assumptions about the behavior of rivals and consumer preferences over product characteristics are crucial, in particular in product categories with multidimensional product differentiation.[1] Detailed modeling of demand and price competition is therefore of key importance in empirically assessing the determinants of product choices. In this paper we develop an integrated empirical framework that specifies consumer demand for differentiated products while endogenizing the pricing and product-assortment decisions of competing firms. Our model allows us to separate demand, marginal cost, and fixed cost contributions to profitability from alternative product offerings.

We demonstrate in a series of counterfactual experiments how changes in demand or market structure affect equilibrium product assortments and prices. Considering product choices as strategic variables to the firm when conducting policy analyses yields different predictions from a simpler model that holds these fixed. We show, for example, that a reduction in the number of competitors due to a merger may be profitable for the merging firm, while at the same time benefiting consumers in the form of higher product variety. To the extent that consumer surplus gains from product variety outweigh losses from higher prices in the more concentrated market, we illustrate that a merger may be unambiguously welfare enhancing, a prediction which critically depends on the ability of firms to respond in their assortment choices to the new market structure. These results complement recent theoretical work by Gandhi, Froeb, Tschanz & Werden (2006) that finds the potential for substantial differences in consumer welfare and profitability effects of a merger when allowing

[1] See for example, Vandenbosch & Weinberg (1995), Economides (1986), and Neven & Thisse (1990), for models of product competition with multiple vertical, horizontal, or both dimensions, respectively, and Gabszewicz & Thisse (1992) for a survey of location models.

post-merger product repositioning relative to a fixed product assortment.

The existing literature has made considerable progress in characterizing competition among heterogeneous firms by focusing on component parts of the product assortment decisions with separate streams of research. Structural demand models generate consistent estimates of price elasticities given the products that firms have chosen to offer, but they assume that these products and their characteristics are exogenous and fixed (see e.g., Berry 1994, Berry, Levinsohn & Pakes 1995, Nevo 2000). However, firms frequently adjust their product portfolios in response to changes in the economic environment such as a merger. Similarly, a national manufacturer can easily adapt offerings in a given market to reflect changing local demographics, seasonal demand spikes, or changes in the local competitive environment. Berry & Waldfogel (2001) and Berry, Levinsohn & Pakes (2004) provide empirical evidence of instances of product repositioning after consolidation or expansion in an industry. The assumption of fixed product assortments may thus be problematic.

At the same time, there is a growing literature on the supply side that endogenizes product-choice decisions for heterogeneous competitors, emphasizing the strategic aspects of product choice (Mazzeo 2002, Einav 2003, Seim 2006). These models focus on explaining entry and location decisions in situations where prices are not a choice variable of the firm or use a reduced-form profit function that does not explicitly incorporate the prices and quantities of the products offered. Firms' product-space locations and those of their competitors are the sole arguments of the firms' objective function, thereby also limiting the scope of counterfactual exercises one can conduct using the estimated parameters. Without an explicit model of demand and post-entry product market competition, for example, we cannot make inferences about equilibrium prices after a product portfolio change, e.g., due to a merger. An early attempt to tackle this issue is Reiss & Spiller (1989), albeit in the context of symmetric firms offering one of two products. Thomadsen (2007) uses estimated demand systems to conduct counterfactual analyses of location competition between single-outlet retailers. His work does not attempt to directly exploit the information entailed in firms' location choices to infer fixed cost determinants of entry decisions, but instead highlights the role of travel costs in determining equilibrium choices in simulations.

In addition, the entry literature typically relies on information contained in discrete firm decisions to infer bounds on profitability that would be consistent with the observed behavior, whereby, for example, the fact that a firm operates in a particular market allows the inference that it is more profitable to operate in that location than to exit. The coarseness of these discrete data make it difficult to base the profit function on all but the simplest of demand structures, ones which generally do not represent product-market competition in oligopolistic industries with differentiated products well. As a result, the majority of the literature focuses on relatively homogenous competitors, such as single-outlet retail stores in well-delimited, small markets. For frequently purchased products that differ in attributes, quality, and brand value, the interplay between consumer preferences for product attributes and their price sensitivities is arguably more central to the product offering decision than similar considerations would be in the context of, say, store location choices. For this reason we start with a discrete-choice demand model for differentiated products and from it develop an equilibrium model of joint product assortment and pricing decisions. The availability of richer data, in particular data on prices and quantities, allows us to better separate the strategic considerations in product assortment decisions of interest from market heterogeneity that drives consumer demand and marginal costs.

We estimate our empirical model of price and product selection by multi-product firms using data on supermarket ice cream sales to illustrate the empirical implementation. Industry analysts and regulators frequently discuss the interaction between flavor selection and pricing in shaping the competitive environment of ice cream markets. The U.S. Federal Trade Commission (FTC) recently sought a preliminary injunction to block a proposed merger between two competing ice cream manufacturers on the grounds that it would, "lead to anticompetitive effects ... including less product variety and higher prices."[2] We focus on two national manufacturers - Breyers and Dreyers - that meet in 64 separate regional markets. Since our data is aggregated across stores in a market area, we consider the manufacturers' product-choice decisions which flavors to offer at the market level abstracting from the manufacturer-retailer interaction. We model the possible offerings in the "vanilla" subcategory, which is by far the most frequently purchased flavor accounting for more than one quarter of all sales. Interestingly, in recent years there has been a number of new product introductions in this space - Breyers and Dreyers now offer up to six vari-

[2] Information from the FTC website at www.ftc.gov/opa/2003/03/dreyers.htm. Note that the FTC's concerns related primarily to Dreyers' super-premium brands (Dreamery, Godiva and Starbucks).

eties of vanilla. The size and evolution of the product category suggests that choices among vanillas are important in their own right, while also being representative of flavor offering decisions across the entire product assortment for these brands.

We consider a two-stage setup where firms initially make their assortment decisions in a discrete game that draws on their variable profits derived in the subsequent stage of price competition. In our set-up firms have at their disposal a set of possible, previously developed flavors from which they choose a subset of offerings depending on local product market and competitive conditions. We assume that competing firms have incomplete information about each others' profitability of offering particular assortments. This assumption allows us to avoid comparing all possible product configurations for all firms to ensure that no profitable unilateral deviation exists, which is necessary to compute the equilibrium in a complete information setting (Seim 2006). Instead, we derive the Bayesian Nash equilibrium conjectures - a computationally much easier task (Rust 1994). As such, the observed product offerings are optimal *ex ante* - if others had been chosen, the resulting price and quantity outcomes would have yielded lower profits for the market participants. The sequential structure of the game where firms choose prices after observing their competitors' first-stage assortment choices allows us to separately identify demand and marginal cost parameters from other determinants of the assortment decisions.

In summary, this paper makes three contributions. We extend prior research (Kekre & Srinivasan 1990, Bayus & Putsis 1999, Draganska & Jain 2005) on product-line length by considering not only how many, but also which of the vanilla varieties to offer. We show how data on prices and quantities can enrich the insights obtained from traditional location choice or entry models. Last, we demonstrate how incorporating endogenous product choice is essential for policy simulations and may entail very different conclusions from settings where product assortment choices are held fixed.

The remainder of this paper is organized as follows. In Section 2 we develop the modeling framework. Section 3 describes the ice cream market and the data we use for the empirical analysis. We outline our estimation approach in Section 4 and then discuss the estimation results and a number of counterfactual analyses that the proposed modeling framework allows us to conduct in Section 5. Section 6 concludes with directions for future research.

2 Model

A total of $b = 1, \ldots, B$ firms (brands)[3] decide which flavors to offer in a given market and how to price them given their expectation of their competitors' offerings, demand, and a fixed cost of offering each subset of flavors.

In the first stage, the firms decide which flavors to offer. Each firm starts with a predetermined set of potential flavors to offer and selects the optimal subset of flavors among this potential set. In the second stage, firms observe each others' flavor choices. Conditional on the firm's choice of flavors and its competitors' choice of offerings, firms choose prices.

Clearly, firms do not decide in each period and market on all potential flavors. There are certain flavors that a brand always offers. We call them staples. The assortment decisions being made concern only what we refer to as the optional flavors. The flavor choice model can be thus thought of applying to optional flavors of a brand that are not offered in all of the markets, as opposed to the staple flavors of a brand.[4] While we abstract from the product offering decision for staple flavors, our model takes into account the demand for staples in determining the price for all flavors in the market.

More formally, brand b has flavors $f = 1, 2, \ldots, O_b, O_b + 1, O_b + 2, \ldots, F_b$ at its disposal. The optional flavors are $1, \ldots, O_b$ and flavors $O_b + 1, \ldots, F_b$ are the staples that the firm always offers. Note that the optional and staple flavors may differ from brand to brand. Define the vector $d_{bt} = (d_{b1t}, \ldots, d_{bO_bt}) \in \{0, 1\}^{O_b}$, where d_{bft} indicates whether optional flavor f is offered by competitor b in market t.

2.1 Stage 2

In the second stage, we solve for equilibrium prices for every possible combination of flavor choices. These prices then flow back into the first stage to determine profits for each of the flavors that a firm is considering.

[3] In the remainder of the paper we use firms and brands interchangeably.

[4] The loss of information should not be too severe because all we can learn from the fact that a brand always offers a particular flavor is that the fixed cost of offering that flavor is smaller than the lowest incremental variable profit across periods from offering it, which would only yield an upper bound on the fixed costs.

Consumer demand. We assume a discrete choice model of demand. Let U_{bfkt} denote consumer k's utility for brand b's flavor f in market/period t. We specify

$$U_{bfkt} = X_{bft}\beta - \alpha p_{bt} + \epsilon_{bfkt} = \overline{U}_{bft} + \epsilon_{bfkt}, \tag{1}$$

where $\overline{U}_{bft}$ is the mean utility across consumers. In the above specification of utility, X_{bft} denotes observed characteristics of the flavor, such as firm and/or flavor fixed effects, whether the flavor is featured in the store ads or on display in the store in a given market. p_{bt} denotes the price charged by firm b in market t. Note that prices for all flavors within a brand are the same as is typical in product categories such as ice cream (Shankar & Bolton 2004, Draganska & Jain 2006). We assume that the random component of utility, ϵ_{bfkt}, is distributed according to an extreme value distribution. It is known to the consumer, but observed by the firms or the researcher only in expectation.

Normalizing utility from the outside good to zero results in logit market shares for the flavors that the brands offer:

$$\begin{aligned} & s_{bft}(p_{1t}, \ldots, p_{Bt}; d_{1t}, \ldots, d_{Bt}) \\ &= \frac{\exp(\overline{U}_{bft})}{1 + \sum_{b'} \sum_{f'=1}^{O_{b'}} \exp(\overline{U}_{b'f't}) d_{b'f't} + \sum_{b'} \sum_{f'=O_{b'}+1}^{F_{b'}} \exp(\overline{U}_{b'f't})}. \end{aligned} \tag{2}$$

Market shares depend on prices $p_{1t}, \ldots, p_{Bt}$ as well as flavor offerings $d_{1t}, \ldots, d_{Bt}$.

Demand models of this type typically incorporate unobserved (to the researcher) product attributes in consumer utility that are a potential source of price endogeneity (Berry 1994, Berry et al. 1995). These unobserved product characteristics may be constant over time such as brand quality perceptions or they may vary over time like shelf-space allocation (Villas-Boas & Winer 1999). While we can infer market/time-specific unobservables associated with product assortment that have been chosen, inferring the value of the unobservables for non-offered combinations without imposing additional (strong) assumptions is infeasible. For example, if we assumed that firms only observe the common demand shocks when they are making their pricing decision and not when they decide on assortments, then firms would need to form expectations over them. However, as will become clearer when we present the supply model below, the expectation of the variable profits that enter the product-choice stage is a highly nonlinear function of the unobservables, so taking this expectation is a nontrivial exercise. In particular, we would need to make some distributional assumption for the unobservables, thus implying that we know the distribution of the equilibrium prices (see Berry (1994) for an explanation of why this type of assumption is inconsistent with the equilibrium model). Our solution to this problem is pragmatic: We assume that in our empirical setting the brand-flavor-specific constants in the demand system along with the market characteristics captures most of the unobserved variation in brand-flavor shares across markets.

Firm profits. For a given choice of flavors determined in the first stage, firm b chooses prices to maximize expected profit. Consistent with observed pricing practices in the industry, we constrain prices to be identical across flavors. Firms are assumed to compete in Bertrand-Nash fashion, given their cost structures.

Firm b incurs a marginal cost of c_{bt} for each unit offered in market t. The marginal costs of offering a flavor include costs for ingredients such as milk, cream, sugar, and flavorings and costs of packaging, labeling, and distributing the product. We specify them as $c_{bt} = \sum_k w_{bkt}\gamma + \eta_{bt}$, where w_{bt} are brand-specific cost shifters k and η_{bt} is a brand-specific component of marginal cost.[5] We assume that firms observe each other's marginal costs when they choose prices, i.e., marginal costs are public information.

We follow the literature in allowing part of the marginal costs to be unobservable to the researcher (Berry et al. 2004). Similar to the demand-side problem of accounting for unobserved product characteristics for absent flavors, we have to confront the problem that we do not observe the value of the unobservable marginal cost components for a brand-flavor combination that is not offered. We solve this problem by assuming that the unobservable component of marginal cost varies by time and brand but not by flavor. Assuming that firms set their prices optimally (conditional on the chosen assortment), we can then recover the value of the unobservable from the pricing first-order conditions and use it to estimate the firm's marginal costs of offering a flavor that it ultimately does not include in its assortment.

In addition, we assume firm b has a fixed cost to offer flavor f in each market t, ν_{bft}, distributed according to probability distribution function G_{bf} that differs across brands and flavors. The fixed costs of offering a flavor may potentially in-

[5]While our model readily accommodates cost shifters that are brand-flavor specific, our application to ice cream does not require this additional generality, see Section 4.1 for details.

clude the operating costs of producing the flavor (foregone economies of scale due to smaller batches, cost of cleaning machines, labeling, etc.), the distribution costs of getting the flavor to customers (such as additional inventory and stocking costs that likely increase in the number of flavors offered), and advertising costs associated with promoting the flavor (which may vary on a flavor-by-flavor basis depending on the offerings of the local competition).

We assume that this fixed cost varies by flavor and is only observed by the firm itself, but not by its competitors, i.e., it is private information. In contrast to marginal costs, which are primarily driven by observable costs for homogeneous inputs, fixed costs may depend on the efficiency of each firm's processes or a proprietary strategic decision they have made.

If a firm decides to offer more than one optional flavor, we assume that its total fixed costs are the sum of the individual fixed costs. This additive formulation allows us to handle multi-product firms without adding too much complexity. The drawback is that we rule out economies of scope, i.e., the fixed cost of adding a particular flavor does not change with the products that are already being offered.

Firm b's objective is to maximize the profit from the staples and the optional flavors that it offers (as indicated by $d_{bt} = (d_{b1t}, \ldots, d_{bO_bt})$):

$$\max_{p_{bt}}(p_{bt} - c_{bt})M\left(\sum_{f=1}^{O_b} s_{bft}(\cdot)d_{bft} + \sum_{f=O_b+1}^{F_b} s_{bft}(\cdot)\right) - \sum_{f=1}^{O_b} \nu_{bft}d_{bft}, \tag{3}$$

where M is the size of the market. To simplify the notation, we suppress $(p_{1t}, \ldots, p_{Bt};$ $d_{1t}, \ldots, d_{bt})$ as arguments of s_{bft}.

Differentiating yields the competitors' first-order conditions with respect to prices:

$$p_{bt}(d_{1t}, \ldots, d_{Bt}) \quad = \quad c_{bt} - \frac{\sum_{f=1}^{O_b} s_{bft}(\cdot)d_{bft} + \sum_{f=O_b+1}^{F_b} s_{bft}(\cdot)}{\sum_{f=1}^{O_b} \frac{\partial s_{bft}(\cdot)}{\partial p_{bt}}d_{bft} + \sum_{f=O_b+1}^{F_b} \frac{\partial s_{bft}(\cdot)}{\partial p_{bt}}}. \tag{4}$$

Solving the system of equations (4) yields equilibrium prices for the specific flavor offerings considered. We emphasize the dependency of prices on flavor offerings by writing $p_{bt}(d_{1t}, \ldots, d_{Bt})$ for equilibrium prices. We solve for equilibrium prices for the remaining possible flavor sets analogously. This gives us a vector of $2^{\sum_b O_b}$ different prices for firm b, one for each possible bundle of flavors that could be offered. We let $s_{bft}(d_{1t}, \ldots, d_{Bt})$ denote the corresponding market share of flavor f offered by brand b

in market t and s_{bt} denote brand b's aggregate market share as a function of its and its competitors' flavor offerings, $s_{bt} = \left(\sum_{f=1}^{O_b} s_{bft}(d_{bt}, d_{-bt})d_{bft} + \sum_{f=O_b+1}^{F_b} s_{bft}(d_{bt}, d_{-bt})\right)$, where $d_{-bt} = (d_{1t}, \ldots, d_{b-1t}, d_{b+1t}, \ldots, d_{Bt})$ are the flavor offerings of all brands but b. There is no asymmetric information in the price-setting stage. Conditional on having made a flavor choice, prices are determined in a symmetric Nash equilibrium.

2.2 Stage 1

Each firm chooses the optimal set of flavors given its expectation of the other firms' choices and prices under each configuration. Firm b chooses $d_{bt} = (d_{b1t}, \ldots, d_{bO_bt})$ to maximize expected profits given by:

$$\begin{aligned}
& \mathrm{E}\left[\Pi_{bt}(d_{bt}, d_{-bt})\right] \\
= \; & \mathrm{E}\left[(p_{bt}(d_{bt}, d_{-bt}) - c_{bt})Ms_{bt}(d_{bt}, d_{-bt}) - \sum_{f=1}^{O_b} \nu_{bft}d_{bft}\right] \\
= \; & \sum_{d_{-bt}} \Big((p_{bt}(d_{bt}, d_{-bt}) - c_{bt})Ms_{bt}(d_{bt}, d_{-bt})\Big) \Pr(d_{-bt}) - \sum_{f=1}^{O_b} \nu_{bft}d_{bft} \\
= \; & \overline{\Pi}_{bt}(d_{bt}) - \sum_{f=1}^{O_b} \nu_{bft}d_{bft}.
\end{aligned} \tag{5}$$

The first part of the expression is the expected variable profit and the second represents the fixed costs. Since firm b does not know the fixed costs of its rivals, it cannot predict their flavor offerings with certainty. Hence, firm b forms expectations over its rivals' flavor offerings. In particular, $\Pr(d_{-bt})$ is the joint probability that its rivals offer the particular subset of flavors in d_{-bt}.

The marginal probability that firm b offers bundle d_{bt} is:

$$\begin{aligned}
\Pr(d_{bt}) \;&=\; \Pr\Big(\mathrm{E}\left[\Pi_{bt}(d_{bt}, d_{-bt}) \geq \mathrm{E}\left[\Pi_{bt}(d'_{bt}, d_{-bt})\right] \quad \forall d'_{bt} \in \{0,1\}^{O_b}\Big) \\
&=\; \int_{A(d_{bt})} \prod_{f=1}^{O_b} dG_{bf}(\nu_{bft}),
\end{aligned} \tag{6}$$

where we let $A(d_{bt})$ denote the set of values for $\nu_{bt} = (\nu_{b1t}, \ldots, \nu_{bO_bt})$ that induce the choice of flavor bundle d_{bt}:

$$A(d_{bt}) = \left\{ \nu_{bt} \middle| \overline{\Pi}_{bt}(d_{bt}) - \overline{\Pi}_{bt}(d'_{bt}) \geq \sum_{f=1}^{O_b} \nu_{bft}(d_{bft} - d'_{bft}) \quad \forall d'_{bt} \in \{0,1\}^{O_b} \right\}. \qquad (7)$$

Assuming independence across firm cost shocks, ν_{bft}, entails that the joint probability of observing a particular set of product offerings in the market $(d_{1t}, \ldots, d_{Bt})$ is the product of the marginal probabilities for d_{bt} defined in equation (6). Substituting the flavor choice probabilities defined above into each firm's expected profit yields a measure of the attractiveness of each choice as a function of the competitors' probabilistic choice. The probability that firm b chooses flavor offering d_{bt} is then the probability that the expected profit of offering d_{bt} exceeds expected profits of any other flavor offering d'_{bt}, given its conjecture of its competitors' behavior.

The expressions defined in equations (5) and (6) characterize a system of $\sum_{b=1}^{B} 2^{O_b}$ equations in $\sum_{b=1}^{B} 2^{O_b}$ unknown flavor choice conjectures. We solve for each firm's probability of offering a given product assortment by numerically integrating over its unobserved fixed cost ν_{bt}, as a function of its competitors' assortment choice probabilities. The equilibrium probabilities of offering each flavor combination are found by searching for the fixed point of the system of equations for all competitors, the solution to which are the $\sum_{b=1}^{B} 2^{O_b}$ flavor offering probabilities. We solve the system of equations defined in equation (6) with a nonlinear equation solver. The resulting fixed point in flavor offering probabilities is the Bayesian Nash equilibrium for the system of best response functions.

Two-firm-two-flavor example. As an illustration of the expected profit function and flavor choice conjectures, consider a two-firm problem ($B = 2$) where each firm has a choice of two optional flavors to offer ($O_1 = O_2 = 2$). To focus on the flavor choice stage, we restrict our attention to optional flavors only ($F_1 = O_1; F_2 = O_2$). Each firm then chooses to offer that set of flavors that maximizes expected profit in a given market. With two flavors, there are four possible choices, offering either, both, or none of the flavors, i.e., we have $d_b = (d_{b1}, d_{b2}) \in \{(0,0), (0,1), (1,0), (1,1)\}$. The firms thus compare four expected profit levels and choose the flavor(s) that corresponds to the highest level of expected profit. Figure 1 illustrates the example.

Suppressing market subscripts for ease of readability, firm 1's expected profit if it

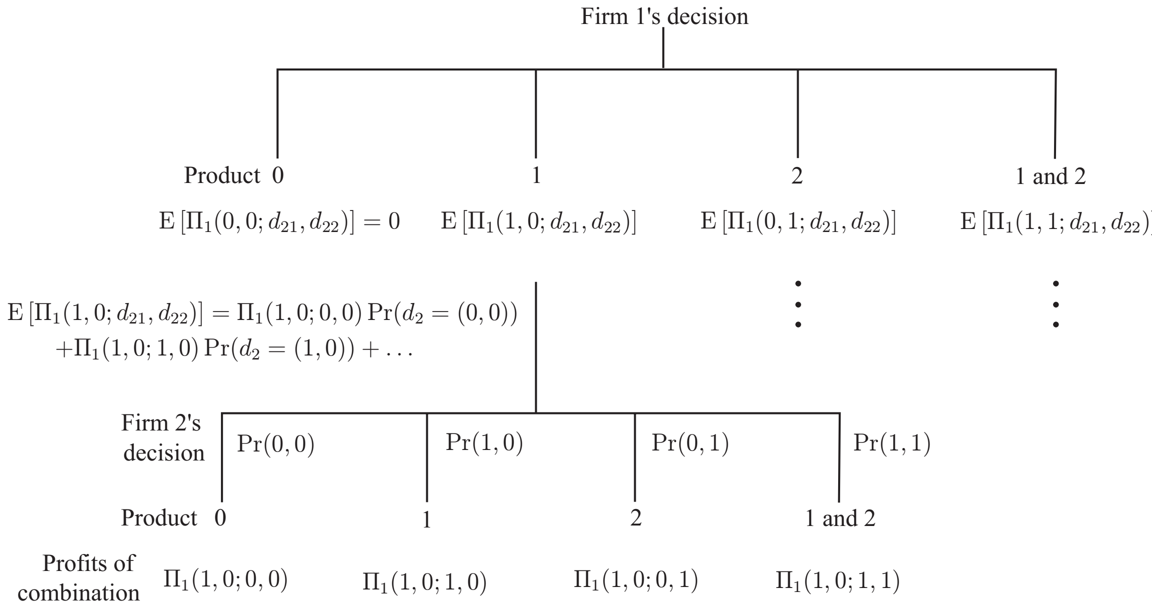

Figure 1: Expected profits.

chooses flavor 1, or $d_1 = (1,0)$, is given by:

$$\mathrm{E}[\Pi_1(1,0,d_{21},d_{22})] = \mathrm{E}[(p_1(1,0,d_{21},d_{22}) - c_1)Ms_{11}(1,0,d_{21},d_{22})] - \nu_{11}. \quad (8)$$

Since firm 1 does not observe firm 2's fixed cost, it has to form an expectation of firm 2's optimal flavor choice, that is, a probability assessment of how likely it is that firm 2 chooses any one of its four possible flavor sets. Integrating over firm 2's cost type yields expected profit of the form:

$$\begin{aligned} & \mathrm{E}[\Pi_1(1,0,d_{21},d_{22})] \\ = & \sum_{d_{21},d_{22}\in\{0,1\}} \Big(p_1(1,0,d_{21},d_{22}) - c_1\Big) Ms_{11}(1,0,d_{21},d_{22}) \Pr(d_{21},d_{22}) - \nu_{11} \\ = & \overline{\Pi}_1(1,0) - \nu_{11}, \end{aligned} \quad (9)$$

where $p_1(1,0,d_{21},d_{22})$ denotes firm 1's optimal price as determined in stage 2 if it offered flavor 1 and firm 2 offers the flavor set $d_2 = (d_{21},d_{22})$, while $\Pr(d_{21},d_{22})$ denotes the probability that firm 2 offers that flavor set. The flavor offering considered by firm 1 and the possible flavors offered by firm 2 are thus reflected in both the price firm 1 charges and its expected market share. Firm 1's expected profit for flavor 2 is computed similarly, while firm 1's expected profit if it does not offer any flavor is normalized to zero.

The expected profit if firm 1 offers both flavors, i.e., chooses flavor set $d_1 = (1, 1)$, is given by:

$$
\begin{aligned}
& \mathrm{E}\left[\Pi_1(1,1,d_{21},d_{22})\right] \\
= & \sum_{d_{21},d_{22}\in\{0,1\}} \Big(p_1(1,1,d_{21},d_{22}) - c_1\Big) M \\
& \Big(s_{11}(1,1,d_{21},d_{22}) + s_{12}(1,1,d_{21},d_{22})\Big) \Pr(d_{21},d_{22}) - (\nu_{11}+\nu_{12}) \\
= & \overline{\Pi}(1,1) - (\nu_{11}+\nu_{12}).
\end{aligned} \tag{10}
$$

Firm 2's expected profits are derived analogously.

Each firm's expected profit depends on its assessment of how likely it is that its competitor offers each of its possible flavors and flavor combinations. Four flavor choice conjectures need to be formed: firm 1's assessment of firm's 2 probability of not offering any flavor, offering flavor 1, offering flavor 2, and offering both flavors. As in the entry literature (Bresnahan & Reiss (1991)), we normalize the profit from not offering any flavor to zero, yielding the traditional profit threshold crossing condition for offering a flavor.

Firm 1's assessment of firm 2's probability of offering flavor 1 is given by:

$$
\begin{aligned}
& \Pr(d_2 = (1,0)) \\
= & \Pr\Big(\mathrm{E}\left[\Pi_2(d_{11},d_{12},1,0)\right] > \mathrm{E}\left[\Pi_2(d_{11},d_{12},1,1)\right] \wedge \mathrm{E}\left[\Pi_2(d_{11},d_{12},1,0)\right] > 0 \\
& \wedge \mathrm{E}\left[\Pi_2(d_{11},d_{12},1,0)\right] > \mathrm{E}\left[\Pi_2(d_{11},d_{12},0,1)\right]\Big) \\
= & \Pr\Big(-\nu_{22} < \overline{\Pi}_2(1,0) - \overline{\Pi}_2(1,1) \wedge \nu_{21} < \overline{\Pi}_2(1,0) \\
& \wedge \nu_{21} - \nu_{22} < \overline{\Pi}_2(1,0) - \overline{\Pi}_2(0,1)\Big).
\end{aligned} \tag{11}
$$

Let the distributions of ν_{21} and ν_{22} be G_{21} and G_{22} with corresponding densities g_{21} and g_{22} and denote $\overline{\Pi}_2(1,0) - \overline{\Pi}_2(0,1)$ as a, $\overline{\Pi}_2(1,0)$ as b, and $\overline{\Pi}_2(1,0) - \overline{\Pi}_2(1,1)$ as c. The probability of offering flavor 1 is thus

$$\Pr(d_2 = (1,0)) = \Pr(-\nu_{22} < c, \nu_{21} < b, \nu_{21} - \nu_{22} < a), \tag{12}$$

which in $\nu_{21} \times \nu_{22}$ space in Figure 2 is the area left of b and above $-c$ minus the triangle spanned by $(b, -c)$, $(a-c, -c)$, and $(b, b-a)$. Hence,

$$\begin{aligned}
& \Pr(d_2 = (1,0)) \\
= \; & (1 - G_{21}(b))G_{22}(-c) - \int_{\nu_{21}=a-c}^{b} \int_{\nu_{22}=-c}^{\nu_{21}-a} g_{22}(\nu_{22}) d\nu_{22} g_{21}(\nu_{21}) d\nu_{21} \\
= \; & (1 - G_{21}(b))G_{22}(-c) - \int_{\nu_{21}=a-c}^{b} (G_{22}(\nu_{21} - a) - G_{22}(-c)) g_{21}(\nu_{21}) d\nu_{21} \\
= \; & (1 - G_{21}(b))G_{22}(-c) + G_{22}(-c)(G_{21}(b) - G_{21}(a-c)) \\
& + \int_{\nu_{21}=a-c}^{b} G_{22}(\nu_{21} - a) g_{21}(\nu_{21}) d\nu_{21}.
\end{aligned} \tag{13}$$

The above presumes $b \geq a - c$. If $b < a - c$, then the probability simplifies to:

$$\Pr(-\nu_{22} < c, \nu_{21} < b, \nu_{21} - \nu_{22} < a) = (1 - G_{21}(b))G_{22}(-c).$$

Depending on the distribution assumed for G_{21} and G_{22}, a closed-form solution for these probability expressions may not exist. However, one can easily find the probabilities using numerical integration techniques.

The probability that flavor 2 is chosen over no flavor, flavor 1, or flavors 1 and 2 together is obtained analogously as:

$$\begin{aligned}
\Pr(d_2 = (0,1)) \; = \; & \Pr(\mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 0, 1)\right] > \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 1)\right] \\
& \wedge \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 0, 1)\right] > 0 \\
& \wedge \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 0, 1)\right] > \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 0)\right]) \\
= \; & \Pr\left[-\nu_{21} < \overline{\Pi}_2(0,1) - \overline{\Pi}_2(1,1) \wedge \nu_{22} < \overline{\Pi}_2(0,1) \right. \\
& \left. \wedge \nu_{22} - \nu_{21} < \overline{\Pi}_2(0,1) - \overline{\Pi}_2(1,0) \right].
\end{aligned} \tag{14}$$

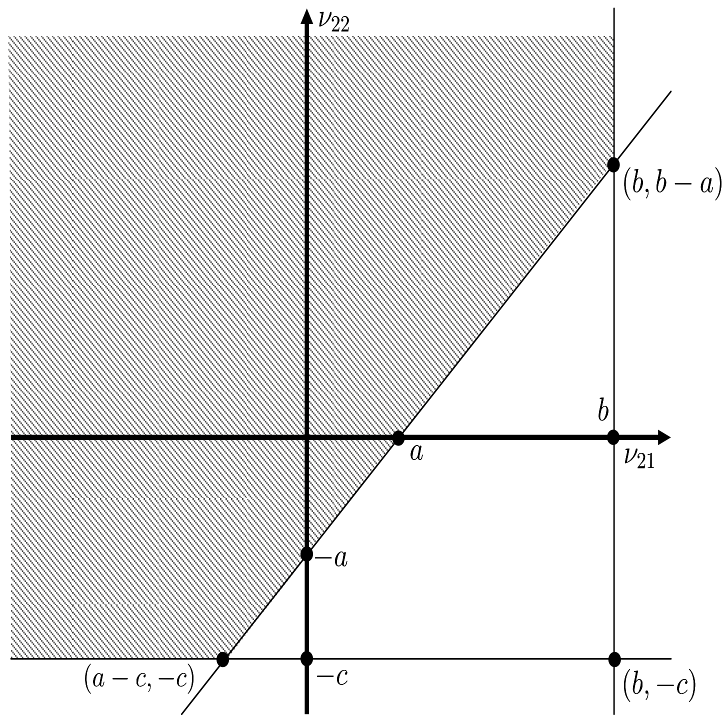

Figure 2: Region of integration.

The probability that firm 2 offers both flavors, flavors 1 and 2, is given by:

$$
\begin{aligned}
\Pr(d_2 = (1,1)) &= \Pr\Big(\mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 1)\right] > \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 0)\right] \\
&\quad \wedge \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 1)\right] > \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 0, 1)\right] \\
&\quad \wedge \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 1)\right] > 0\Big) \\
&= \Pr(\nu_{22} < \overline{\Pi}_2(1,1) - \overline{\Pi}_2(1,0) \wedge \nu_{21} < \overline{\Pi}_2(1,1) - \overline{\Pi}_2(0,1) \\
&\quad \wedge \nu_{21} + \nu_{22} < \overline{\Pi}_2(1,1)), \qquad (15)
\end{aligned}
$$

while the probability that firm 2 chooses not to offer any flavors equals

$$
\begin{aligned}
\Pr(d_2 = (0,0)) &= \Pr\Big(\mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 0)\right] < 0 \wedge \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 0, 1)\right] < 0 \\
&\quad \wedge \mathrm{E}\left[\Pi_2(d_{11}, d_{12}, 1, 1)\right] < 0\Big) \\
&= \Pr(\nu_{21} < \overline{\Pi}_2(1,0) \wedge \nu_{22} < \overline{\Pi}_2(0,1) \wedge \nu_{21} + \nu_{22} < \overline{\Pi}_2(1,1)), \qquad (16)
\end{aligned}
$$

which can be found similarly to the other probabilities.

Equations (11) – (16) together with their analogs for Firm 2's assessment of Firm 1's probabilities form a system of 8 equations in the 8 unknown equilibrium probabilities. One difficulty in estimating discrete games is the possibility of a multiplicity of equilibrium assortment choices. The literature has addressed this problem in a number of ways. Uniqueness generally ensues if one is willing to impose that the players make their assortment decisions sequentially in Stackelberg fashion. This assumption is difficult to justify in our environment both because of the frequent decision-making and the relative symmetry of the two companies in our context. Alternative two-step estimators that initially predict which equilibrium is chosen before computing profits (Bajari, Hong, Krainer & Nekipelov 2006) are difficult to implement for lack of exogenous shifters of each firm's equilibrium selection mechanism. Instead as in Orhun (2006), Seim (2006) and Zhu & Singh (2006), we investigate the prevalence of multiple equilibria in our context numerically, by computing the number of assortment equilibria that arise for each of a set of grid points that span a large part of the parameter space. At the estimated parameters, we find that there is always a unique equilibrium. We solve the system of equations (11) – (16) using a nonlinear equation solver. Relative to commonly used iterative fixed point algorithms which may not be able to reach certain solutions of the system of equations, this procedure is a more reliable, faster solution mechanism.

3 Data

The main data for our analysis were collected by Information Resources, Inc. (IRI) and cover 64 geographic markets across the U.S. for a period of 104 weeks from September 2003 to September 2005. We have weekly information on the units of ice cream sold, dollar sales, and percentage of sales sold on promotion for all UPCs in the markets. While retail prices and promotions may vary weekly, manufacturer decisions are made at a lower frequency. We are interested in the strategic decisions of manufacturers and therefore conduct the empirical analysis at the monthly level. Aggregating the data leaves us with 1600 observations (25 months, 64 markets) for each UPC.

Ice cream is one of the most popular categories in supermarkets: 92.9% of households in the United States purchase in the category (IRI Marketing Factbook, 1993). In the general category of ice cream, there is a distinction between ice cream, frozen

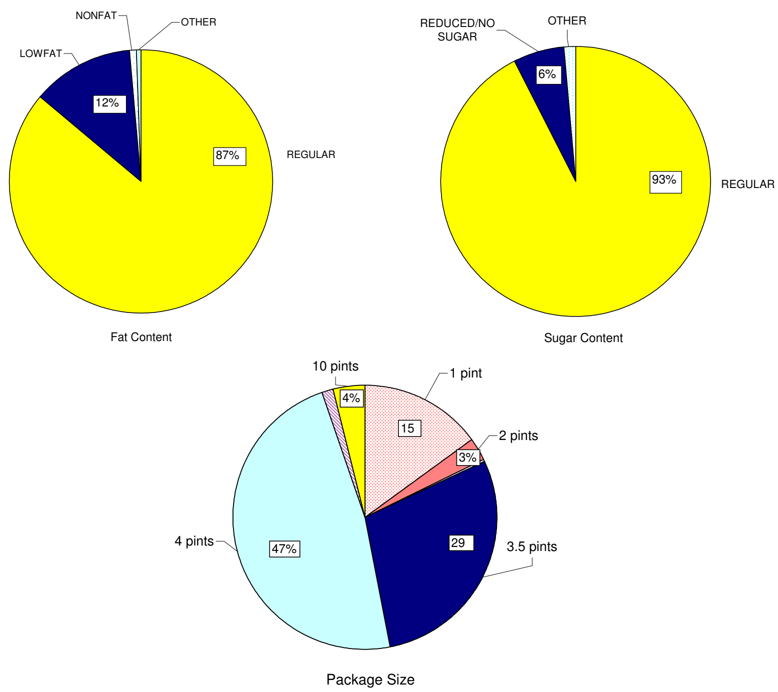

Figure 3: Dollar shares of ice creams by fat content, sugar content, and package size.

yogurt, sherbets and sorbets. Depending on butterfat content, ice cream is further disaggregated into superpremium, premium, and economy categories. So, while a half-cup serving of Häagen Dazs vanilla bean ice cream, a superpremium brand, has 18 grams of fat and 290 calories, the equivalent serving of Dreyers, a premium brand, has only 8 grams of fat and 140 calories. Furthermore, ice cream is offered in a multitude of package sizes, fat and sugar content levels. Figure 3 presents an overview.

Regular fat ice cream accounts for 86% of ice cream sales, and only 7.5% of all ice cream sold has reduced or no sugar content. The most popular size is 4 pints with about 48% of all sales, followed by the closely related 3.5 pint size with 29%,[6] and 1 pint with 15%. Most of the superpremium ice cream brands such as Ben & Jerry's and Häagen Dazs are sold almost exclusively in the smaller, 1 pint tubs, whereas the other brands are usually sold in larger sizes.

To illustrate the model developed in this paper, we focus our attention on non-diet ice cream (i.e., full fat and regular sugar) in the premium category, and in particular

[6]Some brands, like Breyers, replaced their 4 pint packages with 3.5 pint ones without changing the unit price. This strategy of increasing the per-ounce price is fairly common among manufacturers of frequently purchased consumer packaged goods because it is not as obvious to consumers as a change in the unit price.

on the decisions of the two leading national brands – Breyers and Dreyers – pertaining to their assortment of vanilla flavors in the most popular family size of 3.5/4 pints. Vanilla flavors represent up to one-third of total category sales. Our data reveal a total of 22 different varieties of vanilla ice cream, involving subtle differences in the ingredients. For example, Vanilla Bean flavors contain visible specks of vanilla, while French Vanillas have a higher egg content. The most popular vanilla varieties in the data are "French Vanilla," "Vanilla," "Vanilla Bean," "Natural Vanilla," and "Extra Creamy Vanilla." We do not include flavors with substantial additional ingredients or flavorings, such as Cherry Vanilla or Vanilla Fudge. Because manufacturers do not "specialize" in vanilla, but the number of vanilla flavors is highly correlated with the total number of flavors offered, an analysis of the vanilla market should shed considerable light on the firms' product assortment decisions in general.

Table 1 presents a market structure snapshot across the 64 geographic regions in our dataset. For the purposes of this analysis, we have classified brands with less than five percent market share in at least five percent of the markets (i.e., three markets) as "other." For each brand, the table presents the number of markets out of 64 for which the brand has each particular market share position. Note that the entries for "Private label" and "Other" in Table 1 are aggregates of all the private label (other brands) that are available in different regions and in different stores within a region. Hence, their competitive position is overstated.[7]

Breyers and Dreyers[8] are the only premium brands that are truly national and have a presence in all markets. However, given the production requirements and distribution economics associated with ice cream, many regional manufacturers established in the early and middle parts of the 20^{th} century have maintained their market position through the present. Brands such as Hood in the Northeast, Blue Bunny in the Midwest and the Southeast, and Tillamook in the Pacific Northwest have substantial sales; indeed, holding the top share in several markets. In addition, sales of private label brands vary in importance from one region to the next. The data in Table 1 suggest that Breyers and Dreyers face very different competitive conditions across the various geographic markets in which they compete.

Table 2 focuses on the vanilla flavors offered by the regional manufacturers, list-

[7] For this reason and because it is not clear what the individual vanilla flavors for the private label and other brands mean, we include them in the outside good.

[8] Dreyer's ice cream is sold under the brand name Edy's in the Midwestern and Eastern United States after Kraft (the makers of Breyers) raised objections in 1985.

Table 1: Market share rank of manufacturers. across the 64 regional ice cream markets.

	Number of Markets					
Market Share Rank:	1st	2nd	3rd	4th	5th-10th	Total
Breyers	14	21	23	5	1	64
Dreyers	5	11	14	20	14	64
Deans	0	0	0	1	10	11
Friendly	1	0	3	0	11	15
Hiland	0	2	0	0	5	7
Hood	1	2	0	2	3	8
Kemps	1	1	0	0	8	10
Mayfield	1	1	2	2	6	12
Pet	0	0	2	4	5	11
Prairie Farms	1	0	1	0	10	12
Tillamook	0	1	0	2	0	3
Turkey Hill	1	1	1	1	10	14
United Dairy	0	1	1	1	7	10
Wells Blue Bunny	3	0	4	6	15	28
Yarnells	1	0	0	2	2	5
Private Label	30	15	10	5	4	64
Other	5	8	3	13	32	61

ing the number of vanilla flavors offered by each across the geographic markets and over the 25 months in our sample period. The first column in Table 2 reports the maximum number of market-month observations, obtained by multiplying the number of geographic markets in which the regional brand has a presence by the number of months. Columns two and three indicate the maximum number of flavors that a brand ever offers in our sample period and the number of markets in which the brand is ever present, respectively. With a couple of exceptions (Kemps and Hiland), the regional players tend to offer fewer vanillas than Breyers and Dreyers. The remaining columns in the table report how frequently the brands carry a full assortment (or a subset) of their available flavors. Most of the regional brands exhibit relatively little variety in their product assortments across markets and over time - for ten of the thirteen brands, the modal number of flavors offered in the data occurs more than two-thirds of the time. We use this evidence to support our assumption that the regional brands do not act strategically with respect to product portfolio choice, leaving the national players to compete market-by-market taking the flavors offered by regional competitors to be exogenous. As such, this assumption provides an additional source of exogenous variation that can be helpful in identification of the model parameters.

Importantly, there is quite a bit of variation in the availability of some of the vanilla flavors for Breyers and Dreyers across geographic regions and months. Table 3 provides the details. Natural Vanilla, French Vanilla and Extra Creamy Vanilla for Breyers and Vanilla, French Vanilla and Vanilla Bean for Dreyers are (almost) always available and can thus be assumed to be staples. On the other hand, Breyers Homemade Vanilla and Dreyers Natural Vanilla, Double Vanilla and Vanilla Custard are the optional flavors, whose offering varies widely by markets and periods. Double Vanilla was introduced towards the end of our sample period, so it is a somewhat special case. Since we do not model the nationwide rollout of a new product, we drop it from the product-choice analysis. We also drop Breyers Vanilla because it only appears in two markets and a few months.

Table 4 presents a summary of the market shares and prices for the brands included in the demand analysis. Breyers is the clear market leader with an average market share of 21%, followed by Dreyers with a market share of almost 14%. Tillamook, Turkey Hill and Yarnells have also sizeable shares in their markets, reflecting their position as strong - albeit small - regional players. The brands vary in their pric-

Table 2: Distribution of flavor availability for regional manufacturers across markets/months in the data set.

	Market-month obs.	# of flavors	# markets	% of market-months in which # of flavors is offered						
				0	1	2	3	4	5	6
Wells Blue Bunny	700	4	28	0.1	-	-	27.7	72.1	-	-
Friendly	375	3	15	-	-	14.9	85.1	-	-	-
Turkey Hill	350	3	14	-	-	2	98	-	-	-
Prairie Farms	300	3	12	1	-	9.7	89.3	-	-	-
Mayfield	300	4	12	-	-	1.7	6	92.3	-	-
Deans	275	4	11	-	-	66.9	24	9.1	-	-
Pet	275	3	11	1.8	-	0.7	97.5	-	-	-
Kemps	250	6	10	3.2	4	22.8	10	11.6	20.8	27.6
United Dairy	250	4	10	-	1.6	16.4	80.4	1.6	-	-
Hood	200	3	8	-	-	24	76	-	-	-
Hiland	175	6	7	0.6	2.3	2.3	5.1	46.9	18.3	24.6
Yarnells	125	4	5	10.4	1.6	4.8	36.8	46.4	-	-
Tillamook	75	2	3	-	-	100	-	-	-	-

Table 3: Percentage of months in which a flavor is available in a geographic market.

Market	Breyers VANILLA	Breyers EXTRA CREAMY VANILLA	Breyers FRENCH VANILLA	Breyers HOMEMADE VANILLA	Breyers NATURAL VANILLA	Dreyers VANILLA	Dreyers FRENCH VANILLA	Dreyers NATURAL VANILLA	Dreyers DOUBLE VANILLA	Dreyers VANILLA BEAN	Dreyers VANILLA CUSTARD
Albany, NY	0	100	100	96	100	100	100	0	27	100	0
Atlanta, GA	0	100	100	100	100	100	100	65	27	100	65
Baltimore/Washington	4	100	100	100	100	100	100	88	27	100	42
Birmingham/Montgom	0	100	100	100	100	100	100	50	27	100	38
Boise, ID	0	100	100	54	100	100	100	50	19	100	31
Boston, MA	0	100	100	65	100	100	100	0	27	100	0
Buffalo/Rochester	0	100	100	96	100	100	100	50	27	100	0
Charlotte, NC	0	100	100	100	100	54	100	73	27	100	77
Chicago, IL	0	100	100	100	100	100	100	77	31	100	35
Cincinnati/Dayton	0	100	100	100	100	100	100	65	27	100	23
Cleveland, OH	0	100	100	100	100	100	100	73	23	100	42
Columbus, OH	0	100	100	100	100	100	100	50	23	100	88
Dallas/Ft Worth	0	100	100	100	100	100	100	100	27	100	88
Denver, CO	0	100	100	100	100	100	100	88	27	100	92
Des Moines, IA	0	100	100	100	100	100	100	54	27	100	23
Detroit, MI	15	100	100	100	100	100	100	42	23	100	38
Grand Rapids, MI	0	100	100	0	100	100	100	35	23	100	12
Green Bay, WI	0	100	100	100	100	100	100	81	27	100	50
Harrisburg/Scranton	0	100	100	100	100	100	100	54	27	100	0
Hartford/Springfield	0	100	100	96	100	100	100	0	27	100	0
Houston, TX	0	100	100	100	100	100	100	100	27	100	85
Indianapolis, IN	0	100	100	100	100	100	100	73	27	100	58
Jacksonville, FL	0	100	100	77	100	100	100	81	27	100	81
Kansas City, KS	0	100	100	100	100	100	100	62	27	100	35
Knoxville	0	100	100	100	100	81	100	58	27	100	46
Little Rock, AR	0	100	100	100	100	85	65	0	0	73	0
Los Angeles, CA	0	100	100	100	100	100	100	100	27	100	100
Louisville, KY	0	100	100	35	100	100	100	92	27	100	77
Memphis, TN	0	100	100	100	100	100	100	54	4	100	4
Miami/Ft Lauderdale	0	100	100	100	100	100	100	81	27	100	81
Milwaukee, WI	0	100	100	100	100	100	100	81	27	100	62
Minneapolis/St Paul	0	100	100	100	100	100	100	69	27	100	35
Mississippi	0	100	100	100	100	100	100	0	19	100	0
Nashville, TN	0	100	100	100	100	65	100	27	27	100	0
New England	0	100	100	100	100	100	100	0	27	100	0
New Orleans/Mobile	0	100	100	81	100	100	100	0	27	100	0
New York	4	100	100	100	100	100	100	100	27	100	92
Oklahoma City, OK	0	85	100	0	100	100	100	0	27	27	0
Omaha, NE	0	100	100	100	100	100	100	50	27	100	12
Orlando, FL	0	100	100	88	100	100	100	88	27	100	81
Peoria/Springfield	0	100	100	100	100	100	100	81	27	100	50
Philadelphia, PA	0	100	100	100	100	96	100	100	27	100	81
Phoenix/Tucson	0	100	100	100	100	100	100	100	27	100	58
Pittsburgh, PA	0	100	100	100	100	100	100	42	0	100	0
Portland, OR	0	100	100	46	100	100	100	81	27	100	31
Providence, RI	0	100	100	88	100	100	100	0	27	100	0
Raleigh/Greensboro	0	100	100	100	100	100	100	77	27	100	85
Richmond/Norfolk	0	100	100	100	100	100	100	81	27	100	0
Roanoke, VA	0	100	100	100	100	54	100	46	27	100	46
Sacramento, CA	0	100	100	100	100	100	100	100	27	100	88
Salt Lake City, UT	0	100	100	62	100	100	100	65	27	100	46
San Ant/Corpus Chr	0	85	100	100	100	100	100	100	27	100	92
San Diego, CA	0	100	100	100	100	100	100	100	27	100	85
San Fran/Oakland	0	100	100	100	100	100	100	100	27	100	77
Seattle/Tacoma	0	100	100	0	100	100	100	85	27	100	38
South Carolina	0	100	100	100	100	100	100	77	27	100	50
Spokane, WA	0	100	100	0	100	100	100	81	27	100	54
St. Louis, MO	0	100	100	100	100	100	100	88	27	100	42
Syracuse, NY	0	100	100	85	100	100	100	54	27	100	0
Tampa/St Petersburg	0	100	100	96	100	100	100	85	27	100	85
Toledo	0	100	100	100	100	100	100	85	27	100	65
Tulsa, OK	0	85	100	0	100	100	100	0	27	69	0
West Tex/New Mex	0	100	100	73	100	100	100	100	27	100	92
Wichita, KS	0	100	100	100	100	100	100	31	23	100	19

Table 4: Market shares and prices of brands included in the analysis.*

	Market Share		Price	
	average	std. dev.	average	std. dev.
Breyers	0.2118	0.0983	$3.78	$0.49
Dreyers	0.1379	0.0873	$3.43	$0.51
Deans	0.0236	0.0320	$3.64	$0.74
Friendly	0.0838	0.0724	$3.46	$0.62
Hiland	0.0563	0.0907	$3.53	$0.54
Hood	0.0898	0.1052	$2.80	$0.51
Kemps	0.0365	0.1054	$4.01	$1.01
Mayfield	0.0812	0.1080	$3.90	$0.66
Pet	0.0484	0.0562	$3.05	$0.54
Prairie Farms	0.0393	0.0739	$3.25	$0.54
Tillamook	0.1184	0.0491	$4.14	$0.48
Turkey Hill	0.1090	0.1049	$3.16	$0.54
United Dairy	0.0502	0.0513	$3.91	$0.87
Wells Blue Bunny	0.0710	0.1002	$3.69	$0.75
Yarnells	0.1201	0.1458	$3.80	$0.52

*Note: Market shares are with respect to the inside goods only and conditional on the brand being present in the market. Numbers do not add to 1 because private label and small brands are not reported.

ing strategies. Breyers and Dreyers occupy the middle ground, while many regional players have lower (Hood, Pet, Turkey Hill) or higher (Tillamook, Kemps) average prices.

As mentioned above, the IRI data include measures of units sold and revenue (with which we calculate average prices) for each UPC in each market. To estimate the econometric model, we complement these data with information drawn from a variety of sources. Table 5 outlines the variables, their source, and the extent to which values differ across our observations. For example, the data that we have on individual demographics are from the 2000 Census - these data vary across geographic markets, but not over time. We have monthly information on several input cost measures; some (e.g., fuel prices) also vary across geographic markets while others (e.g., federal funds rate) do not. We have calculated the distance from each geographic market to the nearest production facility for Breyers and Dreyers. These are the only data that vary across the manufacturers (but are the same in each time period).

The panels of Table 5 are split based on the way we use these additional variables.

Table 5: Summary of Non-IRI Data.

Variable	Source	Level of Variation	Mean	Std. Dev.
Demographic and Demand Variables:				
Population	2000 U.S. Census	Market	3,164,796	3,044,238
% African American	2000 U.S. Census	Market	0.124	0.097
Avg. household size	2000 U.S. Census	Market	2.560	0.141
Per capita income	2000 U.S. Census	Market	21,831.210	2,917.420
% under 18	2000 U.S. Census	Market	0.257	0.019
% 18-24 years	2000 U.S. Census	Market	0.098	0.011
% 25-44 years	2000 U.S. Census	Market	0.306	0.018
% 45-64 years	2000 U.S. Census	Market	0.219	0.013
% over 65	2000 U.S. Census	Market	0.121	0.024
% Males	2000 U.S. Census	Market	0.489	0.006
Temperature	NOAA	Market & Month	67.454	17.245
Measures of Various Input Costs:				
Commercial paper rate	Datastream	Month	2.035	0.951
Cream II ($ per lb)	Dairy Market News	Month	2.247	0.405
Nonfat dry milk ($ per lb)	Dairy Market News	Month	0.926	0.092
Sugar (cents per lb)	Bloomberg	Month	9.039	1.560
Manufacturing wage (NAICS 3115)	Bureau of Labor Statistics	Month	688.407	17.316
Fuel Price ($ per gallon)	Energy Information Administration	Market & Month	147.471	31.746
Distance from closest production facility to market (Breyers)	Own calculations	Market & Firm	283.815	200.063
Distance from closest production facility to market (Dreyers)	Own calculations	Market & Firm	321.364	207.822
Market Structure - Complementary Industries:				
# of Walmart stores	Own calculations	Market	26.594	17.112
Local distributors (NAICS 424330) - population per establishment	County Business Patterns	Market	152,667	56,801
Local distributors (NAICS 424330) - share of employment in top-4 firms	County Business Patterns	Market	0.492	0.201

The top section of the table includes market demographics and temperature; we think that these may be associated with ice cream demand. There may be differences in input costs as well - the variables in the second panel possibly influence the costs of manufacturing and/or distributing the product. In the bottom panel, we have included some statistics on the market structure of complementary industries that may affect the ice cream market on either the supply or the demand side. Across categories including ice cream, prices and measured quantities sold in supermarkets may be affected if there are more Walmart stores in the local market. Since manufacturers rely on distributors that are specifically equipped to transport frozen dairy products, the market structure of these distributors may also be relevant.

4 Empirical Strategy

Below we first give details on the specification of our empirical model, which differs from the model presented in Section 2 by fully accounting for regional and private label brands in the demand estimation. We thus no longer assume that exactly the same brands appear in both stages of the game. We then discuss the estimation procedure in more detail.

4.1 Econometric Specification

We define the potential market size based on the total supermarket sales of regular, 4 pint ice cream in each market and calculate the shares of the competing brands relative to this size M.[9] While we consider only Breyers and Dreyers at the product-choice stage, our demand model also includes private labels and regional players. The utility of these alternatives is specified in the same way as for the branded flavors in equation (1). We assume that the prices for these alternatives are set in a non-strategic way, independent of the product offerings or prices of Breyers and Dreyers and therefore substitute their observed prices in the demand model. Because the identity of the smaller players changes from market to market, we write down a separate demand model for each market that includes the available flavors in that market.

[9]We tried several alternative definitions for M. In general, definitions based on ice cream consumption, which include non-supermarket ice cream sales (e.g., sales in ice cream parlors and specialty stores) were too broad to produce reasonable empirical results. Different definitions based on supermarket sales did, however, yield similar estimates to those reported here.

On the demand side, the observed characteristics of flavor f offered by brand b in market t, X_{bft}, include a brand-flavor dummy and the price. For the outside good we include in X_{00t} the market's monthly average temperature, monthly dummies and indictors for US regions (Northeast, Midwest, and South), the market population's breakdown by gender (%male), age (%18–24, %25–44, %45–64, and %65 and above), and race (% African American), as well as the average household size, per capita income, and lastly the number of Walmart stores operating in the market, capturing one of the primary alternatives to supermarket shopping. These variables affect demand for all inside goods relative to the outside good.

On the cost side, as evident from Figure 4, the flavor-specific "flavorings" component of total cost is relatively small; therefore, we assume that marginal costs are constant across flavors offered by a given firm. While our model can accommodate brand-flavor specific observable cost shifters, in our data the cost shifters are common to all flavors of a brand (formally, $w_{bft} = w_{bt}$). We further assume that these costs are common knowledge across players, which seems reasonable given that the primary cost components - dairy, packaging, and wages - are likely constant within regions and across manufacturers. In our empirical specification, we include as marginal cost shifters in w_{bt} a brand-specific constant, transportation costs (distance between the market and a brand's closest distribution center, average fuel cost), input prices (sugar, cream, dry milk, the local average weekly wage, and the commercial paper rate), and distribution costs (measures of market structure in local distribution: population per local distributor and share of employment in the top 4 local distributors).

We further assume that the flavor-specific fixed costs are drawn from a log-normal distribution with brand-flavor specific scale and shape parameters and a location parameter of zero, i.e., $G_{bf} = LN(\bar{\nu}_{bf}, \sigma^2_{bf})$, where $\bar{\nu}_{bf}$ and σ^2_{bf} denote the parameters of the normal distribution of the log of ν_{bf}. We use the log-normal distribution as a flexible distribution that ensures positive fixed costs and that allows us to compute in a tractable fashion the distribution of fixed costs when firms offer both flavors and the fixed costs equal to the sum of the two flavors' fixed costs. The mean of the distribution, $\exp\left(\bar{\nu}_{bf} + \frac{1}{2}\sigma^2_{bf}\right)$, captures all factors that determine product assortment choices that are not accounted for in the average estimate of variable profits, while its standard deviation captures deviations from the average decision across markets.

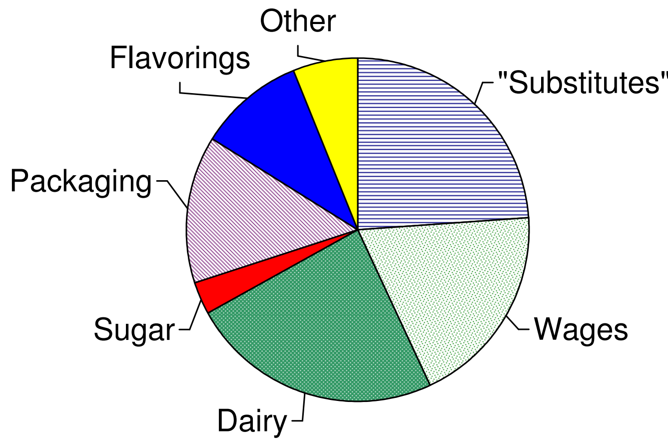

Figure 4: Breakdown of manufacturing cost in the ice cream industry. 1997 Economic Census.

4.2 Estimation

For a given set of parameters for the demand and pricing equations, the second stage of the model yields predicted market shares for the flavors offered in a given market. These market share values are then scaled by our estimates of market size M. In addition, the pricing stage generates estimates of marginal costs that are implied by the observed prices and an assumption on market conduct.[10] These marginal costs flow into the first-stage profit function to determine profits of all potential assortment choice combinations. The first stage then focuses on determining an equilibrium probability of each flavor in the firm's potential flavor set being offered in a given market.

For each brand, we observe its actual assortment decisions, denoted by $d^{\circ}_{bt} = (d^{\circ}_{b1t}, \ldots, d^{\circ}_{bO_Bt})$, the actual market share, s°_{bft}, for all flavors f that are part of the assortment chosen in the first stage (including both staples and optional flavors), and the price, p°_{bt}, charged by the brand for all flavors in the chosen assortment (recall that the price for a given brand is uniform across flavors). To estimate the parameters of the model, we match firms' behavior in terms of these three variables to the model predictions for these variables using a simulated method-of-moments estimator (Hajivassiliou & McFadden 1998).

[10]The data for one of the markets, Little Rock, AR, was suspect because Dreyers was not at all present for a couple of quarters. For this reason we could not back out marginal cost as described, and we drop this market from the analysis.

The first set of moment conditions matches the expected market shares as defined in equation (2) to the ones observed in the data. We define market share prediction errors, denoted by the F_b-dimensional row vector e^s_{bt} with elements

$$e^s_{bft} = \begin{cases} \{s^\circ_{bft} - s_{bft}(d^\circ_{1t}, \ldots, d^\circ_{Bt})\}d^\circ_{bft} & \text{if} \quad f = 1, \ldots, O_b, \\ \{s^\circ_{bft} - s_{bft}(d^\circ_{1t}, \ldots, d^\circ_{Bt})\} & \text{if} \quad f = O_b + 1, \ldots, F_b, \end{cases} \tag{17}$$

where predicted market shares are conditional on actual assortment decisions. The difference between observed and expected market shares is due to sampling error. Our first set of moment conditions is thus the sum of squared deviations of predicted from observed market shares:

$$Q_{1b}(\theta) = \sum_t e^s_{bt}(e^s_{bt})'.$$

Second, we exploit the assumption that observed and unobserved components in the pricing first-order condition, equation (4), are uncorrelated. We use equation (4) to back out the unobserved marginal cost contribution, η°_{bt}, that sets predicted prices equal to the observed prices for the chosen bundle. We then interact it with observed marginal cost shifters in a moment condition. Note that we cannot use a moment condition matching the predicted prices to the actual ones for the estimation because we already exploit the pricing first-order conditions to back out the cost shock. We use weights to combine the moment conditions pertaining to brand b into the least-squares objective:

$$Q_{2b}(\theta) = \eta_b' W_b (W_b' W_b)^{-1} W_b' \eta_b,$$

where η_b is a $T \times 1$ vector of marginal cost shocks for brand b and W_b is a $T \times K$ matrix of the exogenous marginal cost shifters w_{bt} (e.g., manufacturer transportation cost, price of milk and sugar for brand b). We obtain marginal cost estimates from minimizing this objective function.

Our third and last set of moment conditions results from matching the firms' actual assortment choices to the ones predicted by the model. Formally, we define assortment prediction errors (the difference between the predicted choice probability and the actual assortment choice), denoted by the 2^{O_b}-dimensional row vector e^a_{bt} with elements:

$$e^a_{b\cdot t} = 1(d^\circ_{bt} = d'_{bt}) - \Pr(d'_{bt}) \quad \forall d'_{bt} \in \{0,1\}^{O_b}, \tag{18}$$

where $1(\cdot)$ is the indicator function. We match observed to predicted choice probabilities:

$$Q_{3b}(\theta) = \sum_t e^a_{bt}(e^a_{bt})'.$$

We obtain fixed-cost estimates by minimizing this objective function. Reflecting the two-stage nature of the game, this last stage of the estimation takes the demand and marginal cost estimates as inputs.

To calculate the objective function we draw a large number of fixed costs ($S = 5000$) and obtain a nonparametric estimate of the frequency with which a firm offers a particular assortment given its beliefs about its rival's offerings. Because the frequency count can jump even for small changes in the parameter values, the objective function is discontinuous. Therefore we use a Nelder-Mead simplex algorithm for the minimization. In addition, we bootstrap standard errors. To this end, we create a large number (100) artificial data sets of the same size as our original data set by drawing observations with replacement from our original data set. We then apply our estimator to each of the artificial data sets. The empirical distribution of the estimates on the artificial data sets then approximates the distribution of our estimator.

5 Results

5.1 Monte Carlo Study

We first test the ability of our estimation procedure to recover the fixed costs using Monte Carlo simulations. We generate 100 replications of a simulated data set of 256 potential markets. We work with a very simple market structure scenario: there are two competitors and each has the option to offer zero, one, or two flavors. The firms are constrained to charge the same price for both products if they offer both varieties, similar to the current practice in the ice cream industry. We generate demand and cost shifters in the form of temperature and manufacturer-specific transportation costs by drawing from the empirical distribution of these variables in our data.

Given the distribution of the unobservables, the exogenous characteristics, and a

Table 6: Monte Carlo analysis: fixed cost distribution estimates using simulated data.*

	True Value	Mean Est. Value	Bias	Std. dev.	RMSE
Mean					
brand 1, flavor 1	0.0100	0.0086	-1.36E-03	6.32E-03	6.47E-03
brand 1, flavor 2	0.0250	0.0220	-2.95E-03	1.65E-02	1.68E-02
brand 2, flavor 1	0.0100	0.0110	1.01E-03	7.14E-03	7.22E-03
brand 2, flavor 2	0.0200	0.0170	-3.05E-03	1.24E-02	1.27E-02
Standard deviation					
brand 1, flavor 1	0.1000	0.1061	6.13E-03	4.22E-02	4.26E-02
brand 1, flavor 2	0.2500	0.2758	2.58E-02	1.47E-01	1.49E-01
brand 2, flavor 1	0.1000	0.1052	5.19E-03	4.77E-02	4.80E-02
brand 2, flavor 2	0.2000	0.2133	1.33E-02	9.76E-02	9.85E-02

*Each estimation run is based on starting values of 0.0001 for all parameters.

reasonable, fixed set of parameters (listed in Table 6 under "True value"), we calculate the optimal choices of the operating firms with respect to the products they offer and the price they charge, as well as the corresponding market share for each offered product. Then we proceed to estimate the parameters of the model to see if we recover the true values that generated the predictions. We estimate the fixed-cost parameters taking demand and marginal cost parameters as given. As evident from Table 6, even when we start with values that are quite far from the truth (each estimation run is based on starting values of 0.0001 for all parameters), our procedure yields average estimates that are very close to the correct values. In unreported results, we find that our methods-of-moments estimator performs as well as an alternative maximum-likelihood procedure in recovering the fixed-cost parameters.

5.2 Merger Analysis

One compelling reason to model endogenous product choice together with demand is to generate more accurate merger simulations. As discussed previously, simulations based on demand models that do not allow for the possibility that a merged firm might change the composition or characteristics of its post-merger product portfolio do not necessarily reflect the firm's optimal behavior. The parameters of our model permit us to simulate more accurately, as both price and the set of offered products can be optimally adjusted. To illustrate the impact of this change, we computed a series of simple merger counterfactuals using the simulated 256 markets described above. The results of our counterfactual simulation demonstrate the potential pitfalls that can occur by ignoring endogenous product choice.

To obtain the effects of a merger and to demonstrate the impact of allowing for product choice in the model, we simulate optimal behavior in three different scenarios. First is the base duopoly case in which the two firms in question are competitors, choosing products to offer and then competing on price. We then allow the firms to merge, acting like a monopolist and potentially offering as many as four products. We distinguish between two alternatives, constraining the merged firm to either offer the same products that the duopolist did (the current standard in the literature) or allowing it to reoptimize in the product-choice stage. As a consequence, the monopolist potentially chooses a different set of products to offer than in the competitive environment. We simulate market outcomes under a low and high regime for the fixed costs of offering the individual flavors as presented in the left and right panels of Table 7.

To compute the statistics presented in Table 7, we use simulation techniques to integrate over the empirical distribution of flavor fixed costs. For a given draw from the cost distributions of each of the four flavors, we record the monopolist's optimal flavor choice given the realizations, together with the optimal price, variable profit, and total profit of the chosen assortment. We then solve the duopolist's assortment choice problem by computing each brand's expected profit of offering each assortment. As in the monopoly case, we record the realization of brand-flavor fixed costs, each firm's chosen assortment, and the associated optimal prices and profits. For the duopolists' chosen assortments, we recompute the monopoly prices and profits. We repeat this procedure to integrate over the distribution of fixed costs. This allows us to determine the expected profit and prices of offering each assortment under the three competitive scenarios and, for the monopolist, the empirical frequency with which each assortment is offered. For each of the 256 markets, we aggregate across assortments to obtain weighted average prices, consumer surplus, and variable and total profits, using as weights the empirical (in the case of the monopolist) or equilibrium (in the case of the duopolists) probability with which each assortment is offered.

Table 7 presents a summary of the key market-level outcomes under the scenarios described above, with all the figures representing the average outcomes across all the markets. Our "fixed products" merger simulation generates reasonable findings, in line with other studies using similar methodology. Comparing the first two columns of each panel, prices and profits are higher for the merged firm than for competing duopolists, while consumer surplus is lower. By construction, the number of flavors is the same in each of the first two columns. When no longer constrained, total industry profits are (necessarily) higher, as the newly merged firm chooses to offer a different assortment some of the time. In the case presented in Table 7, the resulting endogenous post-merger product assortment depends critically on the level of the fixed costs of offering additional flavors. In the low fixed cost regime the merged firm offers fewer flavors on average, while the merged firm occasionally offers more products in the high fixed cost scenario. Indeed, it appears that the reduction in price competition makes it worth spending the higher fixed cost to offer an additional flavor some of the time. As a consequence, in the high fixed cost simulation the merger results in both higher total profits and higher consumer surplus as compared with the duopoly case. Such a finding would not be possible without endogenizing the product assortment decision, as our methodology allows.

These simulated merger results also give some idea about magnitudes; in particular, whether ignoring product assortment endogeneity generates substantial changes between the results in the second and third columns (as compared with the differences between the first and second columns). As such, one could interpret the results in Table 7 as suggesting that ignoring product choice has minimal effect if the fixed-costs to offering each product are low. However, it is important to recognize that the example constrains the merged firm to optimize only among the previously offered flavors. In a case where the merged firm has the entire Hotelling line available to choose from (as in Gandhi et al. (2006)) or a larger flavor choice set at its disposal, the impact is likely to be more substantial. Additional market participants may also re-optimize portfolios post-merger, generating more changes to surplus and profits. Indeed, the results in any specific case will rely critically on the estimated parameters in the model. Nonetheless, this exercise clearly demonstrates the importance of endogenizing product choice in the context of a policy simulation.

Table 7: Merger Simulations.*

	Low Fixed Cost			High Fixed Cost		
		Merged Firm			Merged Firm	
	Duopoly	Fixed Products	Endog. Products	Duopoly	Fixed Products	Endog. Products
Price brand 1	4.1707	4.871	4.8317	2.3562	2.4076	3.6065
Price brand 2	3.9295	4.7381	4.6685	2.9565	2.9824	3.6111
Variable profits brand 1	0.409	0.6954	0.6152	0.2253	0.2255	0.2958
Variable profits brand 2	0.3076	0.4267	0.4496	0.3529	0.3530	0.3176
Industry variable profits	0.7166	1.1221	1.0648	0.5782	0.5785	0.6134
Total profits brand 1	0.2117	0.4981	0.4833	0.0487	0.0488	0.0646
Total profits brand 2	0.2075	0.3266	0.3822	0.0818	0.0819	0.0790
Industry total profits	0.4192	0.8247	0.8655	0.1305	0.1307	0.1436
Number of flavors	1.8585	1.8585	1.4361	0.4395	0.4395	0.4709
Consumer surplus	2.7593	1.2642	1.2261	0.6356	0.6348	0.6766

*Both scenarios assume the same demand parameters of $\beta_0 = [6.5; 6.0; 5.0; 5.5]$, $\beta_{price} = -2.5$, $\beta_{temp} = 0.1$, where $[\beta_0^1 ... \beta_0^4]$ denotes the four flavor-specific intercepts, and marginal cost parameters of $\gamma_0 = [0.45; 0.30]$, $\gamma_{distribution} = 0.001$, and $\gamma_{sugar} = 0.3$, where γ_0^1, β_0^2 denotes brand-specific intercepts. The low fixed cost scenario assumes the following parameter values for the four flavor fixed cost distributions: $\bar{\nu} = [0.034; 0.03; 0.01; 0.012]$ and $\sigma = [0.02; 0.02; 0.02; 0.02]$, while the high fixed scenario is based on $\bar{\nu} = [1.44; 1.20; 1.00; 1.12]$ and $\sigma = [0.16; 0.16; 0.16; 0.16]$.

5.3 Empirical Analysis

Demand and Marginal Cost. Table 8 presents the parameters of the demand and pricing equations for the ice cream data. We model the demand for all offered flavors as a function of a brand-flavor constant and the price. The demand for each flavor falls in the brand's price, with an implied elasticity ranging from -4.27 to -6.29, which is comparable to other frequently purchased consumer goods in mature categories.

In addition we control for variables that shift demand for all inside goods relative to the outside option such as market demographics and time dummies. Our estimates indicate that there is statistically significant seasonal and geographic variation in the demand for vanilla flavors in supermarkets. In addition, the demographic composition of a market has a pronounced impact on demand: Markets with a higher percentage of males and African Americans and higher per capita income tend to have higher demand for vanilla ice cream (lower demand for the outside good). Confirming our intuition, a large presence of Walmart in a given market takes away from supermarket sales for ice cream.

Most aggregate marginal cost shifters, such as the price of sugar and dry milk, are not statistically significant, possibly due to the lack of variation across markets and brands. As expected, marginal costs increase in brand-specific transportation (distance to the nearest distribution facility) and fuel costs, as well as the proxies for the size and density of the local distribution network.

Fixed Cost. Reasonable starting values for the flavor fixed cost distributions should reflect variation in actual fixed costs. To determine the likely magnitude for these costs, we use the following procedure. Beginning with initial estimates for demand and marginal cost, we calculate variable profits for each possible offering. We then loop through flavors and use data on whether the flavor is offered to infer bounds on fixed costs that would make the observed flavor offering decision optimal.

Take for example Breyers Homemade Vanilla. Assume first it is part of Breyers' actual flavor offering. We then consider the hypothetical offering that removes Homemade Vanilla, holding fixed the availability of all other flavors. Because of our assumption of cost additivity, the fixed costs of the actual offering equal those of the hypothetical offering plus the fixed cost of offering Homemade Vanilla. Since Breyers did not choose this hypothetical offering, the fixed cost draw for the first flavor must be smaller than the difference in variable profits between the actual and the hypothetical offering. This gives us an upper bound on the fixed cost draw for Homemade Vanilla. Conversely, if Homemade Vanilla is not offered, we consider adding it to the actually chosen offering, which allows us to derive a lower bound on the fixed cost draw in a similar fashion. Repeating this procedure for all flavors and all markets results in a number of bounds. In Figure 5 we graphically represent the obtained lower and upper bounds for the optional flavors offered. As evident from the box plots in Figure 5, there is large variation for both the lower and upper bound of the fixed costs obtained in this manner.

We use the bounds to generate starting values as follows: we take the average of the mean lower and upper bounds as a guess at the mean of that flavor's lognormal fixed cost distribution. Similarly, we take the average of the standard deviation of the lower and upper bounds as a guess at its standard deviation. Since we estimate the mean and standard deviation of the underlying normal distribution, we back out the $\bar{\nu}_{bf}$ and σ_{bf} associated with these two parameters of the lognormal distribution and use them as starting values in estimation.

Table 9 presents estimates of the distribution parameters of the underlying normal

Table 8: Demand and marginal cost estimates using ice cream data. Brand-flavor constants omitted for brevity.

Parameter	Estimate	Std. Error
Demand – Inside flavors		
Price	-0.5019	0.0209
Demand – Outside option		
Temperature	0.0009	0.0011
January dummy	-0.0080	0.0448
February dummy	0.0880	0.0384
March dummy	0.1193	0.0441
April dummy	0.0762	0.0448
May dummy	0.1198	0.0496
June dummy	0.1121	0.0560
July dummy	0.1134	0.0545
August dummy	0.1306	0.0641
September dummy	0.0745	0.0580
October dummy	0.0689	0.0479
November dummy	-0.0747	0.0453
Northeast dummy	0.6097	0.0449
Midwest dummy	0.3090	0.0365
South dummy	0.4451	0.0418
% African American	-1.1401	0.1566
% Male	-9.6801	1.7030
% 18-24 old	-4.4395	1.4749
% 25-44 old	-3.7634	1.5196
% 45-64 old	-2.9410	1.3352
% 65 and older	-8.0026	0.9295
Average household size	0.2340	0.1461
Per capita income	-0.0001	0.0000
Walmart	0.0015	0.0007
Marginal cost:		
Breyers constant	5.2320	0.9258
Dreyers constant	4.8952	0.9254
Transportation cost	0.0002	3.2E-05
Sugar price	-0.0027	0.0252
Wage	-0.0037	0.0014
Commercial paper	-0.0108	0.0600
Cream II price	-0.1180	0.0512
Dry milk price	-0.2712	0.2043
Distributor employment	0.4236	0.0584
Population per distributor	-2.0E-06	1.8E-07
Fuel cost	0.0029	0.0007

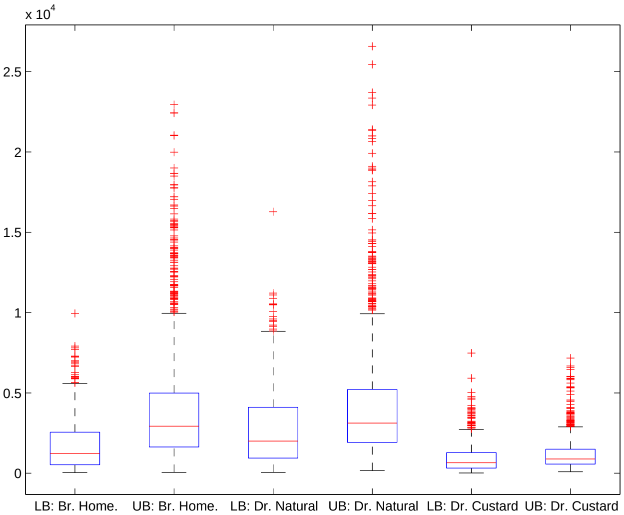

Figure 5: Fixed cost bounds obtained from demand and marginal cost estimates.

Table 9: Distribution parameters of log fixed cost estimated from ice cream data. Normal distribution.

Parameter	Estimate	Std. Error*	Confidence Interval	
Mean $\bar{\nu}_{bf}$				
Breyers Homemade Vanilla	5.4278	0.2249	4.9701	5.8783
Dreyers Natural Vanilla	7.2979	0.1128	7.0309	7.4837
Dreyers Vanilla Custard	6.6230	0.1180	6.4372	6.8551
Standard deviation σ_{bf}				
Breyers Homemade Vanilla	1.9129	0.2348	1.4932	2.4371
Dreyers Natural Vanilla	1.7195	0.1244	1.4626	1.9646
Dreyers Vanilla Custard	1.9988	0.1471	1.8115	2.3743

*Bootstrapped standard errors based on 100 bootstrap replications.

Table 10: Implied means, standard deviations, and medians of estimated fixed costs.

Parameter	Estimate	Confidence Interval*	
Mean			
Breyers Homemade Vanilla	1,418.7	1,014.6	2,806.9
Dreyers Natural Vanilla	6,478.5	4,173.0	12,095.6
Dreyers Vanilla Custard	5,544.9	3,601.2	15,140.9
Standard deviation			
Breyers Homemade Vanilla	8,726.1	3,300.9	54,619.3
Dreyers Natural Vanilla	27,665.5	11,501.8	80,619.2
Dreyers Vanilla Custard	40,496.0	18,923.1	253,220.0
Median			
Breyers Homemade Vanilla	227.7	144.0	357.2
Dreyers Natural Vanilla	1,477.1	1,131.1	1,778.9
Dreyers Vanilla Custard	752.2	624.7	948.7

*Bootstrapped confidence intervals based on 100 bootstrap replications.

distribution of the log of fixed costs, while Table 10 contains the associated mean, standard deviation, and median for the level of fixed costs for each of the three optional flavors we consider in estimation, Breyers Homemade Vanilla and Dreyers Natural Vanilla and Vanilla Custard. Given the assumed log-normal distribution of fixed costs, the median level of fixed costs may be the most informative summary measure. As a check on their magnitudes, we compare the average fixed costs to the variable profits implied by the demand and marginal cost parameters presented in Table 8. The variable profits for each of the three optional flavors amount to \$4,175.58 (standard deviation of \$3,783.93) for Breyers Homemade Vanilla, \$4,505.43 (standard deviation of \$4,098.94) for Dreyers Natural Vanilla, \$6,537.92 (standard deviation of \$6,608.66) for Dreyers Vanilla Custard. They are comparable to the estimated fixed costs, suggesting that our fixed costs estimates are reasonable, as their value would translate into frequent, though not universal, offering of the three flavors in question.

Figure 6 shows how the flavor offerings change with the fixed costs. We plot changes in the optimal product portfolio offered by Breyers and Dreyers in response to uniform increases in the level of fixed costs across flavors. η is a scale factor that multiplies fixed costs across flavors, where the baseline fixed costs result from setting

η equal to one. In the case of Dreyers, the figure illustrates differential effects of higher flavor fixed costs on bundle offerings with the probabilities of both offering one of the optional flavors only as well as the option of not offering any optional flavor initially gaining steadily in fixed cost at the expense of the option of offering both flavors. For higher levels of fixed cost, however, the single-flavor options hold relatively steady assortment shares, while the option of offering neither of the two flavors continues to grow in likelihood. This finding suggests that the two flavors substitute for each other, such that with high fixed cost, demand is not sufficient to offer both, but more than outweighs the fixed cost of offering only one of the two flavors. We investigate the role of differentiation between optional flavors in greater detail in the next section.

With knowledge of the fixed cost estimates, one can conduct an analysis to compute the sort of endogenous product assortment merger effects that we show in the simulations to have important policy implications. Such a merger analysis is complicated in our case since the brands offer a number of overlapping staple flavors that we abstract from in the stylized merger analysis above. We instead use the estimated fixed cost parameters to investigate linkages between preferences and firms' pricing decisions on the one hand and product assortment decisions on the other to illustrate the benefits of incorporating a more fully specified demand side into a product assortment model.

5.4 Policy Experiments

We demonstrate the economic significance of the estimated structural parameters in several illustrative analyses. We consider how assortment depends on consumers' taste for variety and quality. To highlight the importance of product differentiation, we look at the effect of varying the degree of horizontal differentiation and the degree of vertical differentiation (or brand preferences) on assortment choices.

Horizontal differentiation. Given the logit specification for consumer demand in equation (1), we can investigate the role of horizontal preference heterogeneity by varying the logit scale parameter, σ (Anderson, de Palma & Thisse 1992). In estimation, we normalize σ to one. In a counterfactual, we compute how market shares, mark-ups, and ultimately assortment choices respond to changes in σ (or equivalently, to rescaling all demand estimates). Formally, we rewrite equation (1) as:

$$U_{bfkt} = X_{bft}\beta - \alpha p_{bt} + \sigma\epsilon_{bfkt} = \overline{U}_{bft} + \sigma\epsilon_{bfkt}. \tag{19}$$

Figure 7 shows how the likelihood that the two brands offer each of their optional flavors changes as we increase the scale parameter from zero to above two. We derive the predicted probabilities by using the estimated demand, marginal cost, and fixed cost parameters from Tables 8 and 9, adjusting the estimated demand-side parameters by σ, as in equation (19). The optional flavor assortment choice for Breyers is simply offering its optional flavor Homemade Vanilla versus not, and range for Dreyers from offering both of its optional flavors, offering Natural Vanilla or Vanilla Custard, to offering neither.

The figure illustrates that as the heterogeneity in consumer tastes increases, both Breyers (panel 1) and Dreyers (panel 2) are more likely to increase the number of flavors they offer. With increased horizontal differentiation, even small "pockets" of demand become more valuable, thus giving firms an incentive to crowd the product space. Dreyers, for example, is more aggressive in offering Natural Vanilla than Vanilla Custard alone. This reflects that Natural Vanilla, having a higher estimated brand-flavor preference, is on average more attractive to consumers than Vanilla Custard, thus yielding higher returns. The increasing attractiveness of the product as horizontal differentiation increases outweighs cost considerations: the estimated average fixed cost of offering Natural Vanilla exceeds that of Vanilla Custard. Eventually, of course, the degree of horizontal product differentiation is sufficiently large to warrant adding Vanilla Custard to Dreyers' portfolio.

Vertical differentiation. Next we turn to the role of vertical differentiation between the two brands in driving assortment choices. We consider the effect on each brand's assortment of increasing the dispersion in the brand-flavor constants for each brand's set of optional and staple vanilla flavors included in the demand system. We vary the degree of vertical differentiation between each brand's flavors by decomposing a brand-flavor constant into the mean brand effect $\bar{\beta}_b$ (-11.39 for Breyers and -11.88 for Dreyers) and deviations from the mean. Thus, $\beta'_{bf} = \lambda_b(\beta_{bf} - \bar{\beta}_b) + \bar{\beta}_b$. Our model estimates above are based on a specification where $\lambda_b = 1$. We vary the dispersion in brand-flavor constants by increasing λ_b from zero, equivalent to there being no vertical differentiation between the brand's flavors, to a value of five, which corresponds to significantly more vertical differentiation than in our estimates. In particular, if a given flavor dummy is estimated to be above (below) average for the brand, then it becomes more (less) attractive for $\lambda_b > 1$. By construction, we leave the average preference for the brand, and therefore the attractiveness of the brand's entire portfolio, unchanged.

As above, we use the estimated demand, marginal, and fixed cost parameters, together with varying values for λ_b, to trace out how the product assortment of each brand changes as the degree of vertical differentiation in its flavors changes. Figure 8 illustrates the changing assortment choices that increasing vertical differentiation in its own flavors has on Dreyers' own assortment choices, as well as the competitive effect that such a change has on Breyers' assortment choice.

In the case of Dreyers, the estimated flavor effects for the two optional flavors that we consider in the product choice stage (Natural Vanilla and Vanilla Custard) are below Dreyer's average (values of -12.48 and -13.75 for Natural Vanilla and Vanilla Custard, respectively). The vertical preferences for the two flavors thus falls as we increase the degree of vertical differentiation in the product line. Panel 2 in Figure 8 illustrates that in response Dreyers is increasingly likely not to offer the two flavors, an effect that is magnified by the fixed costs that Dreyers pays for offering the flavors (which is normalized to zero for all other flavors). Moreover, the probability that Vanilla Custard is offered as the single optional flavor decreases monotonically. In contrast, the probability that Natural Vanilla is offered on its own initially increases and then falls in $\lambda_{Dreyers}$. As Dreyers slowly removes Vanilla Custard from the market, it manages to redirect some of its demand to Natural Vanilla. Put differently, Natural Vanilla becomes a closer substitute for Vanilla Custard for at least some range of $\lambda_{Dreyers}$. Eventually, of course, Natural Vanilla becomes too unattractive relative to the rest of Dreyers' portfolio to be offered.

The top panel in Figure 8 shows that there is also a competitive effect of the varying degree of vertical product differentiation on Breyers' assortment choices. As the degree of vertical product differentiation rises, it puts downward pressure on the single price that Dreyers charges for all its flavors. Since in the Bertrand pricing game, prices are strategic complements, Breyers' price declines as well. The associated decline in variable profit implies that Breyers can no longer cover the fixed cost of offering its optional flavor, so that its likelihood of being offered declines in $\lambda_{Dreyers}$.

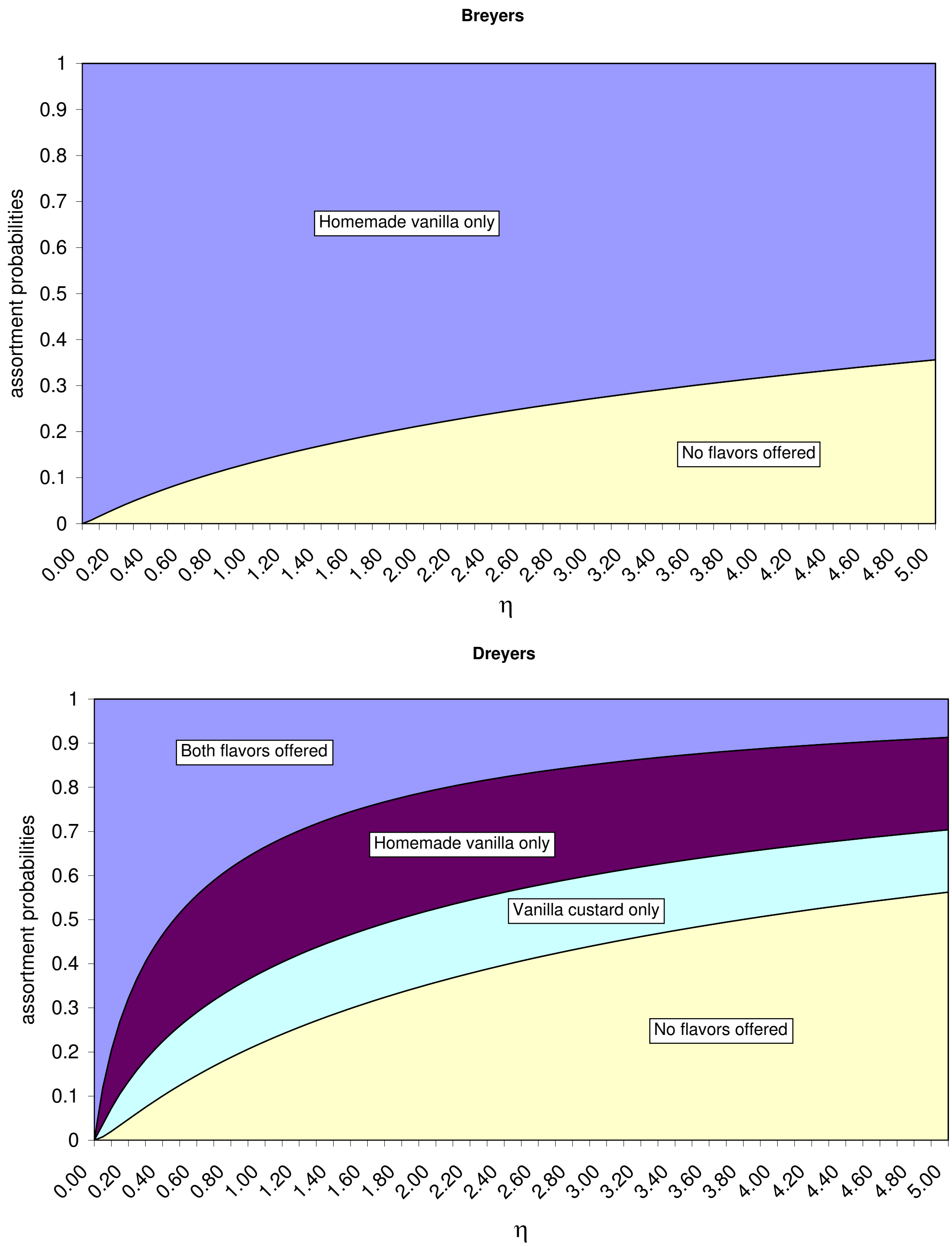

Figure 6: Assortment probabilities as a function of level of fixed costs.

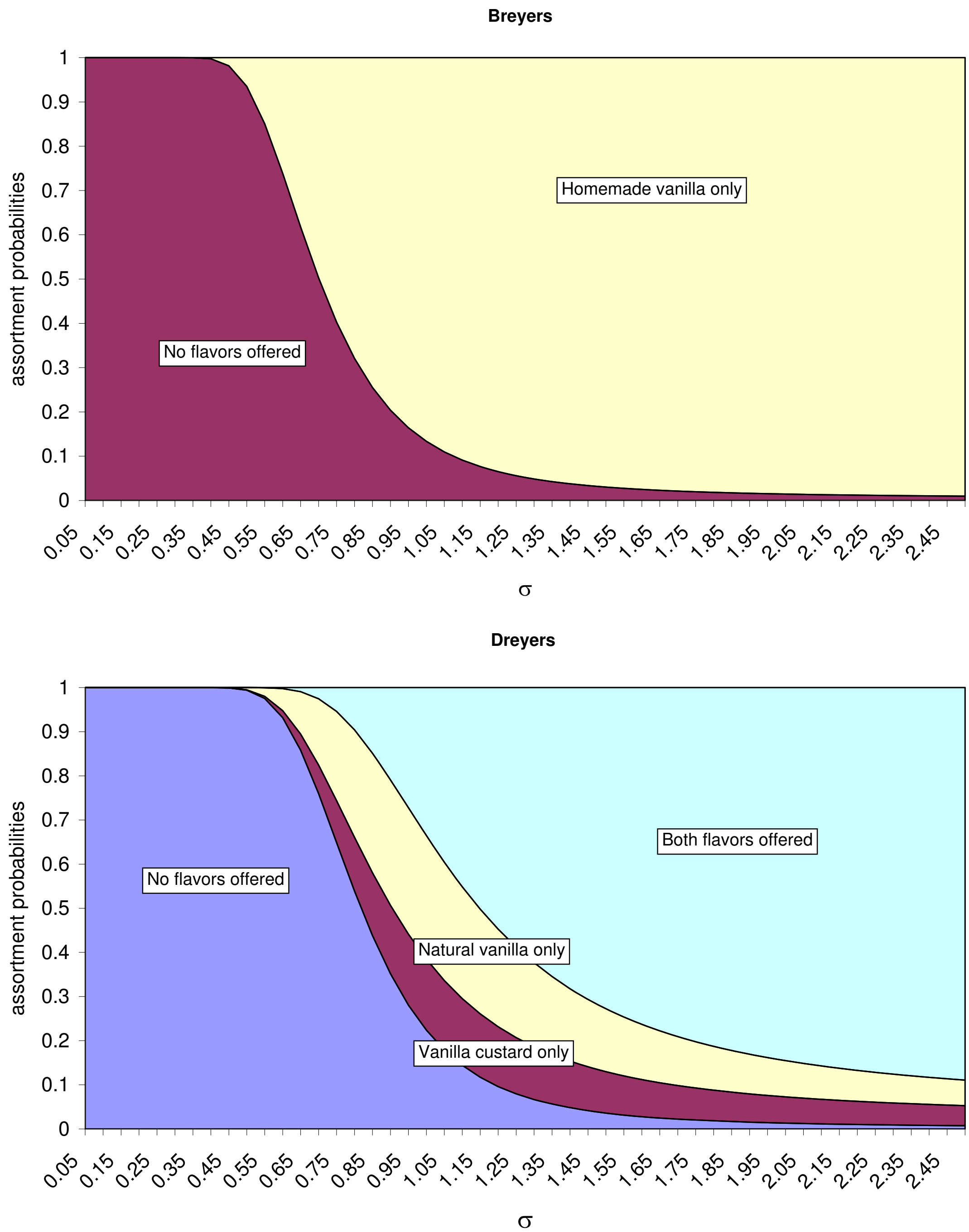

Figure 7: Assortment probabilities as a function of degree of horizontal differentiation.

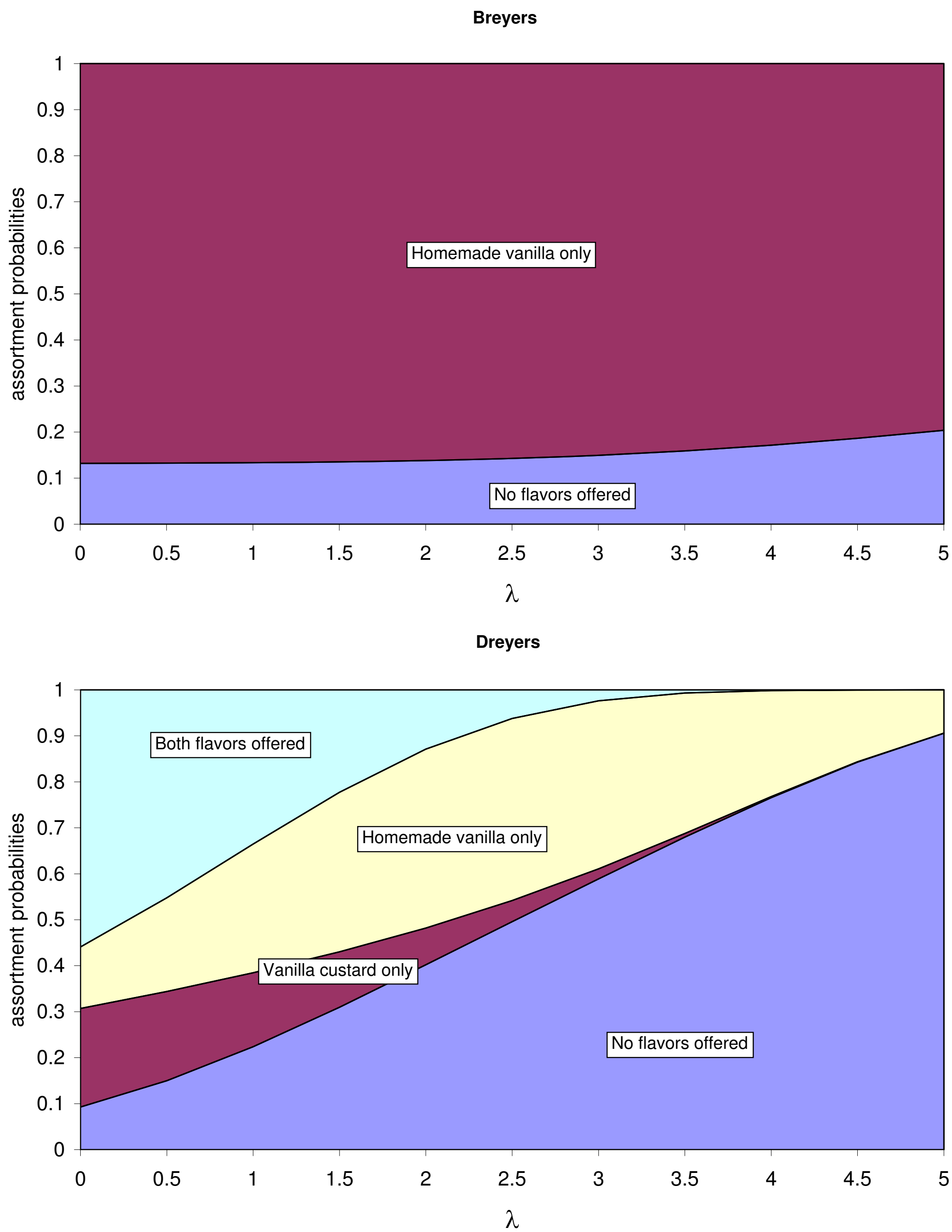

Figure 8: Assortment probabilities as a function of Dreyers' degree of vertical differentiation.

6 Conclusions

In this paper, we develop a framework for incorporating endogenous product choice in a supply-and-demand model of competition in a differentiated product market. The empirical model generates estimates of the fixed costs associated with offering particular products in addition to the typical demand and marginal cost parameters.

With these estimates in hand the researcher is better able to conduct counterfactual experiments by allowing competitors to change their product offerings optimally as part of the exercise. We demonstrate the impact of endogenizing product-assortment decisions in the context of a merger simulation, in which the merged firms often choose a different set of products than those previously offered, generating higher profits. The impact of abstracting from endogenous product choice may or may not be large, depending on the estimated cost and demand parameters. What is clear though, is that sometimes we reach fundamentally different conclusions by modeling joint product assortment and pricing decisions. For example, a reduction in the number of competitors due to a merger may benefit consumers by leading to increased product variety. The gain accruing to consumers due to the availability of more products may offset the higher prices due to reduced competition. Hence a merger may be unambiguously welfare enhancing contrary to the inferences based on the commonly used methodology.

Unlike reduced-form approaches used in the entry literature, by explicitly modeling price competition we can show how demand-side factors affect product-assortment decisions. In particular, we investigate the effect of both horizontal and vertical differentiation on equilibrium assortments and prices. With increased horizontal differentiation, even small consumer segments can become valuable enough to give firms an incentive to crowd the product space. The effect of a change in vertical product differentiation is more subtle and depends on how exactly consumers value the various products alternatives that a firm may consider offering. There is no doubt, however, that product assortment decisions are not made in a competitive vacuum: As our empirical findings indicate, when a rival's products become more differentiated, the price level in the market may fall and the firm may be inclined to cull the variety offered since variable profits no longer can cover fixed costs.

In sum, deriving the variable profits that enter the product-choice decision from a structural model of product-market competition is a big step forward from the reduced-form profit functions typically used in the entry and location choice literature. Given the importance of price in consumer purchase decisions, this is a critical element when attempting to model product assortment decisions. In addition, relative to the literature on structural demand models, our results show that incorporating endogenous product choice is essential for policy simulations and may entail very different conclusions from settings where product assortment choices are held fixed.

Our game-theoretic model abstracts from a number of complicating factors for the sake of empirical tractability. While our two-stage game partially captures the relative irreversibility of assortment decisions, ideally the model would reflect the different periodicity of the pricing and product choice decisions. One may also want to allow for serial correlation in firms' assortment decisions over time. Short of specifying and estimating a fully dynamic model, one could possibly introduce state-dependence into the model, thus allowing the distribution of fixed costs to differ systematically depending on whether the product has been offered in the previous period.

While our results indicate that deriving the variable profits from a structural model of product-market competition is critical to modeling product assortment decisions, it has a cost: We abstract from unobserved product characteristics that would introduce selection effects into the assortment and pricing decisions, which would significantly limit our ability to use information on demand and prices for offered products to infer the profitability of those products that the firms chose not to offer. Formulating a model that confronts this issue and developing an econometric method to deal with the ensuing endogeneity bias in the demand estimation is of critical importance for future work.

Another venue to pursue is to relax the restriction that firms select among a prescribed set of already developed alternatives. The initial product development decision would be very interesting to analyze, and allowing firms greater choice among product characteristics would certainly increase the value and importance of incorporating product selection. In addition, addressing dynamic new product development as part of the analysis is a promising area for future research.

References

Anderson, S., de Palma, A. & Thisse, J. (1992). *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge, MA.

Bajari, P., Hong, H., Krainer, J. & Nekipelov, D. (2006). Estimating static models of strategic interaction, *NBER Working Paper No. W12013* .

Bayus, B. & Putsis, W. (1999). Product proliferation: An empirical analysis of product line determinants and market outcomes, *Marketing Science* **18**: 137–153.

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.

Berry, S., Levinsohn, J. & Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new vehicle market, *Journal of Political Economy* **112**(1): 68–104.

Berry, S. & Waldfogel, J. (2001). Do mergers increase product variety? Evidence from radio broadcasting, *Quarterly Journal of Economics* **116**: 1009 – 1025.

Bresnahan, T. & Reiss, P. (1991). Entry and competition in concentrated markets, *Journal of Political Economy* **99**(51): 977–1009.

Draganska, M. & Jain, D. (2005). Product-line length as a competitive tool, *Journal of Economics and Management Strategy* **14**(1): 1–28.

Draganska, M. & Jain, D. (2006). Consumer preferences and product-line pricing strategies: An empirical analysis, *Marketing Science* **25**(2): 164–174.

Economides, N. (1986). Nash equilibrium in duopoly with products defined by two characteristics, *RAND Journal of Economics* **17**(3): 431–439.

Einav, L. (2003). Not all rivals look alike: Estimating an equilibrium model of the release date timing game. Working Paper, Stanford University.

Gabszewicz, J. & Thisse, J.-F. (1992). Location, *in* R. Aumann & S. Hart (eds), *Handbook of Game Theory*, North-Holland, Amsterdam.

Gandhi, A., Froeb, L., Tschanz, S. & Werden, G. (2006). Post-merger product repositioning, *Journal of Industrial Economics* **forthcoming**.

Hajivassiliou, V. & McFadden, D. (1998). The method of simulated scores for the estimation of ldv models, *Econometrica* **66**(4): 863–896.

Hotelling, H. (1929). Stability in competition, *Economic Journal* **39**: 41–57.

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

Mazzeo, M. (2002). Product choice and oligopoly market structure, *RAND Journal of Economics* **33**: 221–242.

Neven, D. & Thisse, J. (1990). On quality and variety competition, *in* J. Gabszewicz, J. Richard & L. Wolsey (eds), *Economic Decision-Making: Games, Econometrics and Optimisation*, Elsevier, Amsterdam, pp. 175–199.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry, *RAND Journal of Economics* **31**(3): 395–421.

Orhun, Y. (2006). Spatial differentiation in the supermarket industry. Working Paper, GSB Chicago.

Reiss, P. C. & Spiller, P. T. (1989). Competition and entry in small airline markets, *Journal of Law & Economics* **32**(2): S179–202.

Rust, J. (1994). Estimation of dynamic structural models, problems and prospects: Discrete decision processes, *in* C. Sims (ed.), *Advances in econometrics: Sixth World Congress*, Vol. 2, Cambridge University Press, Cambridge.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices, *RAND Journal of Economics* **37**(3): 619–642.

Shankar, V. & Bolton, R. (2004). An empirical analysis of determinants of retailer pricing strategy, *Marketing Science* **23**(1): 28–49.

Thomadsen, R. (2007). Product positioning and competition: The role of location in the fast food industry, *Marketing Science* **forthcoming**.

Vandenbosch, M. & Weinberg, C. (1995). Product and price competition in a two-dimensional vertical differentiation model, *Marketing Science* **14**(2): 224–249.

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

Zhu, T. & Singh, V. (2006). Spatial competition and endogenous location choices: An application to discount retailing. Working Paper, Carnegie Mellon University.