Variety: Demand Functions and Price Adjustment

William C. Horrace [a]

Rui Huang [b]

Jeffrey M. Perloff [c]

December, 2010

Abstract

To optimally choose the number of varieties of a brand to carry, a retailer needs to know how changing the number affects the shape and position of the demand function. We use order statistics to analytically derive demand functions when consumers choose from among the varieties of two brands and an outside good. Given that consumers' valuations of the varieties are distributed independently uniform, demand functions are higher-order polynomials, where polynomial order is increasing in variety. We use our estimates of this model for soft drinks to illustrate how demand curves change with the number of varieties. Because these demand curves have convex and concave sections around an inflection point, firms are more likely to respond and make large price adjustments to increases in cost than to comparable decreases in costs.

key words: varieties, product line length, consumer surplus, welfare, demand, order statistics, oligopolistic

[a] Syracuse University
[b] University of Connecticut
[c] University of California, Berkeley

We are grateful to Brian Wright for extremely helpful advice on welfare issues and to George Judge and Derek Laing for insightful discussions on order statistics.

Variety: Demand Functions and Price Adjustment

1. Introduction

How do brands' product lengths—the number of varieties that each sells—affect consumers' brand choice? How does the shape of a brand's demand curve change when a retailer carries more varieties? Manufacturers and retailers need to know the answers to these questions to choose the optimal product lengths and prices. To address these questions, we use order statistics to develop a new theory of consumer choice across brands that have many varieties. After picking the best variety within each brand (an order statistic), the consumer selects the best choice across the brands (an order statistic over order statistics) and compares that best choice to an outside good. Given that consumers' tastes are distributed uniformly, we derive a complete set of analytic results. The demand functions are higher-order polynomials, where the order of the polynomial is increasing in the number of varieties. We analytically solve for the own- and cross-partial adjustments to consumers' demands from changes in the number of varieties. The demand curves have both convex and concave sections around an inflection point. This shape leads to asymmetric price response to cost shocks: Firms are more likely to respond to and to adjust price more to a cost increase than to a comparable cost decrease. We illustrate this theory for the soft-drink market, where we estimate a three-parameter version of the model, which does well in describing actual consumer choice. We use these estimates to determine the quantitative effects of changes in the number of varieties on demand curves and consumer welfare and discuss the implications for pricing.

According to many food and beverage manufacturing executives, brands are maintained through product differentiation (e.g., Nijssen and Van Trijp, 1998). Firms constantly innovate to keep up with changing consumer tastes.[1] Products that are not accepted by consumers are quickly dropped. One might think of this approach of constantly providing new products as a flagpole strategy: "Let's run it up the flagpole and see who salutes it." Firms differentiate by

---

[1] For example in 2000, Snapple introduced a new fruit drink, Diet Orange Carrot Fruit Drink ("Fruit Beverages Scope," *Beverage World*, February, 2000, p. 26), presumably reasoning that if they can sell that flavor, they can sell any flavor.

changing flavors or other aspects of the product as well as by altering the size or shape of the container.

We examine a market in which each of two firms (brands) produce many varieties of a good. Examples of such markets include sporting goods, yogurt, ice cream, and beverages. Sporting good firms that produce a variety of balls, gloves, shoes that differ only slightly in terms of which athlete endorses them or a variety of aesthetic bells and whistles. Yogurts vary by flavor, whether the fruit is on the bottom, and in other ways. Ice creams vary by flavor and fat content.

Beverage manufacturers offer many varieties. Our empirical example concerns retail demand curves for Coca-Cola and Pepsi in the U.S. soft-drink market. Coca-Cola sells many variations of its flagship product, including Coca-Cola, Cherry Coke, Diet Coke, Caffeine-Free Coke, Caffeine-Free Diet Coke, and Black Cherry Vanilla Coke. It also sells or has sold a wide variety of other soft drinks, including Tab, Sprite, Fresca, Fanta, Barq's Root Beer, Mello Yello, and Pibb Xtra, and, according to its website, 700 low-calorie or no-calorie drinks throughout the world. Pepsi has a similarly large number of varieties. Retailers differ as to how many of these varieties they carry, but no retailer carries all of them. For simplicity in the following, we refer to the Coke brand and the Pepsi brand, whereas each company actually has several brands with many varieties in each.

Most existing theoretical work on product differentiation or product length focuses on firms' behavior rather than on consumers' choices. Much of the theoretical literature abstracts from how consumers choose and assumes that there are only a small number of varieties (e.g., Brander and Eaton 1984, Gilbert and Matutes 1993, and Villas-Boas 2004).

Other theoretical papers model demand using specific functional forms that depend on the total number of varieties. By employing explicit utility functions, they can analyze the welfare effects of greater variety.

Four classic papers on product differentiation—Dixit and Stiglitz (1977), Spence (1976), Salop (1979), and Deneckere and Rothschild (1992)—assumed that each monopolistically competitive firm produces a single product and then asked if there are too many or too few products. The Chamberlin-representative-consumer competition papers of Dixit and Stiglitz (1977) and Spence (1976) used a constant elasticity of substitution (CES) utility function for a representative consumer that depends explicitly on the number of products. In the Hotelling-competition model of Salop (1979), consumers' tastes are uniformly distributed around a circle and products are evenly spaced around the circle (adjusting as new firms enter). Deneckere and Rothschild (1986) nested what they called a Chamberlin model (Perloff and Salop 1985) and the Hotelling-circle model of Salop (1979).

Some more recent papers modified these models to look at firms with product lines rather than firms that produce only one good each, but they maintain these functional assumptions. For example, Raubitschek (1987) extends the Spence-CES model to allow for brands to have varieties, and Klemperer (1992) modified the circle model to allow for an endogenous determination of spacing of varieties. Kim et al. (2002) chose a CES-like functional form for utility with a log-normal random utility component to estimate the compensating variation from removing an existing brand of yogurt. Draganska and Jain (2005) employed the Kim et al. model and assumed that all flavors yield the same utility, all flavors are offered with equal probability, and the cost of evaluating each flavor is convex.

The rest of the empirical literature falls loosely into three categories: flexible demand systems, nested or mixed logit, and demand systems that add measures of the length of product lines. Hausman (1996) used a flexible system of demand equations to estimate the welfare effects of adding one brand of cereal by assuming that the price went from infinity to a finite level. Kadiyali et al. (1999) employed a linear system of demands curves to study product extensions effects on prices. Israilevich (2004) estimated an AIDS system and concluded that a grocery store chain carries the optimal number or too many products. Nevo (2000) employed a mixed logit approach to address a similar question to Hausman's about the effects of adding a product. Bayus and Putsis (1999) estimated a three-equation system for share, price, and product-line length to examine the effect of product-line length on price and market share.

Our work is closest to the Perloff and Salop (1985) and Anderson et al. (1992) models that show the effects of greater product diversity on prices, market share, and welfare. In those models and in the current model, consumers place a value on each product and then choose the variety with the largest net surplus: the consumer's value minus the price. The values are drawn independently from a distribution. Each consumer buys one unit of one of these varieties if the net surplus from the consumer's favorite exceeds that of an outside good.

There are two main differences between our model and those of Perloff-Salop and Anderson et al. First, these earlier models assumed that each manufacturing firm produced a single product, whereas our model has two manufacturing firms (brands) each of which produces multiple varieties. Second, rather than focus on the decision of manufacturers, we examine the decision of the retailer. The reason for this later difference is that, at least in grocery markets, retailers determine the number of varieties to carry rather than manufacturers, who produce a much larger number of varieties than any one retailer carries.

The Perloff-Salop (1985) and Anderson et al. (1992) type random utility models provide the conceptual basis for the recent mixed-logit, market power empirical studies that follow the Berry et al. (1995) BLP model. We differ from that literature in that we emphasize developing a model that can be used for theoretical work—one in which we derive analytic solutions—whereas they are primarily interested in using a very flexible, general model to estimate. The very flexible mixed-logit models cannot be used to produce unambiguous analytical results. To get clear results, we make two assumptions. First, we assume that tastes are distributed uniformly; however, to actually estimate the mixed-logit models, the research must make an explicit distributional assumption (typically they assume a probit or logit distribution). Second, we place more structure on the relationship between brands and varieties than they do. Most mixed-logit models are similar to Perloff-Salop and Anderson et al. in that they treat each variety symmetrically (e.g., as would be case if each variety were manufactured by a distinct firm), whereas in our model, each brand produces many varieties.

We start by providing intuition for our model by using examples of a monopoly firm and a simplified duopoly problem. Next, we use order statistics to derive a general model of how consumer choice varies with brands' product lengths and derive a number of analytic, comparative statics properties. We use our model to estimate a demand system for Coke, Pepsi, and an outside good at supermarkets. Using our estimated demand system, we simulate various demand and consumer surplus comparative statics results. We then discuss the implications of the shapes of these demand curves for price adjustments.

2. Varieties and Consumer Choice

In our model, each consumer buys one unit of a good by choosing among all the available varieties offered by both brands. The value that a consumer places on each variety is drawn independently from uniform distributions that may differ across brands. Each consumer picks the variety with the largest net surplus if that net surplus is greater than the net surplus provided by the outside good.

We concentrate on soft drinks sold by Coke and Pepsi, where each variety of a brand sells for the same price, though the price may differ across brands. Before plunging into our order-statistics model with two brands and an outside good, we illustrate the basic idea with two simpler examples.

One Brand

Initially, suppose that there is a monopoly grocery store carries the varieties of only a single brand. Consumers choose their favorite variety, and the net surplus from the outside good is zero. A typical consumer places a value on each of the n varieties that is drawn independently from a uniform [0, 1] distribution. The price, p, lies within the range (0, 1) by appropriate scaling. The probability is $1 - p$ that a consumer will place a higher value on a given variety than its price. The probability that at least one variety is more valuable to the consumer than its price is 1 minus the probability that no variety has a value greater than the price: $1 - (1 - [1 - p])^n = 1 - p^n$.

The aggregate demand curve is the number of consumers, Z, multiplied by this probability: $[1 - p^n]Z$. (For simplicity, we henceforth normalize Z to equal one.) The slope of the demand curve with respect to price is $-np^{n-1} < 0$. If the number of varieties increases from n to $n + 1$, then the quantity purchased increases by $(1 - p)p^n$, which is positive because $p \in (0, 1)$. As n gets large, virtually everyone buys a variety from this brand.

Two Brands

Now, suppose that the grocery store carries two brands. The consumer might use a three-step procedure to pick which variety if any to buy. The consumer first picks the highest net-surplus variety within each brand, then the consumer selects between the two top choices for each brand, finally the consumer compares the best overall variety choice across the brands with the net surplus the consumer places on an outsider product. If the best variety across brands is more attractive than the outside good, the consumer buys that variety.

We can use a table to illustrate the effects from adding one more variety for one brand. The numbers in the table are the net surplus a consumer obtains from each variety. The net surplus from the outside good is .5. Initially, the store carries two varieties of Coke and two variety of Pepsi (the first two columns of the Pepsi section of the table). The first consumer (row one) receives a net surplus of .9 from the first variety of Coke and .6 from the second variety, so the consumer prefers the first variety. Similarly that consumer prefers the first Pepsi variety (.3) to the second one (.1). This consumer will buy one unit of Coke because the net surplus from the preferred variety of Coke (.9) exceeds the net surplus from the best Pepsi option (.3) and the net surplus from the outside good (.5). In the table, the consumer's overall choice is indicated by expressing the relevant surplus in italics (ignore the bolding for the moment).

	Outside	**Coke**		**Pepsi**		
Consumer		*First*	*Second*	*First*	*Second*	*Third*
1	.5	***.9***	.6	.3	.1	.4
2	.5	.7	.6	.6	***.8***	.7
3	.5	.4	*.8*	.3	.6	**.9**
4	.5	.6	.3	***.7***	.4	.1
5	*.5*	.1	.2	.3	.2	**.6**

By similar reasoning, Consumer 2 buys the second variety of Pepsi, Consumer 3 buys the second variety of Coke, Consumer 4 chooses the first variety of Pepsi, and Consumer 5 opts for the outside good. In this market, 40% (2 out of 5) choose Coke, 40% choose Pepsi, and 20% consumes the outside good.

Now suppose that the store starts carrying a third variety of Pepsi. In the table, the net surplus corresponding to the consumer's choice is in bold type. The extra variety affects the decisions of only Consumers 3 and 5. Consumer 3 switches from buying the second variety of Coke to the third variety of Pepsi. Consumer 5 changes from the outside good to the third variety of Pepsi.

Pepsi's market share rises with the addition of another variety. The market shares are now 80% Pepsi and 20% Coke (though we wouldn't expect such an extreme outcome in general).

Consumers 1, 2, and 4 are unaffected by the additional variety, while Consumers 3 and 5 are better off, so total consumer surplus must rise. Initially, the total surplus was .9 + .8 + .8 +. 7 + .5 = 3.7. After the third variety of Pepsi is introduced, total surplus is .9 + .8 + .9 +. 7 + .6 = 3.9. This example illustrates that consumers benefit—have higher consumer surplus—from more choice where we hold price constant.

3. Order-Statistics Model

We now turn to a formal analysis of our model. We develop the model in four steps. First we discuss how a consumer would compare two sets of varieties if prices were zero and there is no outside good. Second, we introduce non-zero prices. Third, we allow the value distribution for each brand to have a different support, so that consumers might prefer one brand to another on average. Fourth, we introduce an outside good with a non-negative net surplus.

Distribution of the Difference of Independent Maxima

Each consumer's valuation of any variety of Brand 1 (Coca-Cola) or Brand 2 (Pepsi) is drawn independently from uniform distributions on [0, θ] with independent random sample of sizes n_1 and n_2 respectively, where n_1 is the number of varieties offered by Brand 1 (Coke, Diet Coke, …) and n_2 is the number of varieties offered by Brand 2 (Pepsi, Diet Pepsi,…). Let L_1 and L_2 be the maximal observations of a consumer's valuation of varieties for Brand 1 and Brand 2. That is, $L_1 = \max(L_{1,1},...,L_{1,n_1})$ and $L_2 = \max(L_{2,1},...,L_{2,n_2})$, where the valuations $L_{1,j}$ and $L_{2,j}$ are distributed independently uniform on [0, θ].

The distribution of the maximal valuation difference, $L_1 - L_2$, is the probability that a consumer selects Coke or Pepsi, or the market shares (relative demand): $s_1 \in [0,\ 1]$ and $s_2 \in [0,\ 1]$. For now, everyone buys one unit of either Coke or Pepsi—there is no outside good—so that $s_1 + s_2 = 1$. Also for now, we ignore prices. For example, suppose that a firm provides a free soft-drink at lunch, so that its employees simply have to decide which variety of which brand to choose independent of price. Using standard notation where ${}_aC_b = a!/(a-b)!/b!$, we derive

Proposition 1. The probabilities that the consumer chooses Brand 1 and Brand 2 are (respectively):

$$s_1(n_1,n_2) = \Pr(L_1 > L_2, n_1, n_2) = \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j},$$

$$s_2(n_1,n_2) = \Pr(L_1 \le L_2, n_1, n_2) = \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2}C_j}{{}_{n_1+j}C_j}. \quad \blacksquare$$

The proof is in the Appendix. The share functions are finite hypergeometric sums, which can be reformulated in terms of the hypergeometric function (Gauss 1813). In economic terms, cumulations in Proposition 1 are the probabilities of choosing Coke ($L_1 > L_2$) or Pepsi ($L_1 < L_2$), where the consumer's objective is to pick the Coke or Pepsi variety with the highest value (given that there is no outside good or price).

Using the Chu-Vandermonde identity for hypergeometric functions (Andrews et al., 2001, p.67), we can simplify the expressions for the shares:

Proposition 2. The shares in Proposition 1 can be reformulated as:[2]

$$s_1(n_1, n_2) = \frac{n_1}{n_1 + n_2},$$

$$s_2(n_1, n_2) = \frac{n_2}{n_1 + n_2}. \blacksquare$$

That is, the quantity share equals the share of varieties for each firm. If preferences are identical across consumers and the number of consumers is known, then multiplying the shares in Proposition 2 times the population produces demand curves for Coke and Pepsi, respectively. Therefore, normalizing the population to one, these probabilities are relative demands or market shares.

Relocation of the Distribution of the Difference by Prices

We now introduce prices, where $p_1 > 0$ is the price for all varieties of Coke and $p_2 > 0$ is the price for all varieties of Pepsi. This phenomenon of setting identical prices for all varieties of a specific brand occurs in soft drinks as well as a number of other industries including clothing, yogurt, movie theaters, books, and many teas and juices (see McMillan 2005, Draganska and Jain, 2005, Orbach and Einav 2007, and Heidhues and Koszegi 2008). The maximal net surplus of the choices are $\ell_1 = L_1 - p_1$ for Coke and $\ell_2 = L_2 - p_2$ for Pepsi. We can write the shares for each good i as $\tilde{s}_i(n_1, n_2, p_1, p_2)$:

Proposition 3. The brands' shares depend on the price differential $\pi = p_1 - p_2$ and the number of varieties:

$$\tilde{s}_1(n_1, n_2, p_1, p_2) = \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(1 - \frac{\pi}{\theta}\right)^{n_2+j} \text{ for } p_1 > p_2,$$

$$\tilde{s}_2(n_1, n_2, p_1, p_2) = \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2}C_j}{{}_{n_1+j}C_j} \left(1 + \frac{\pi}{\theta}\right)^{n_1+j} \text{ for } p_1 < p_2. \blacksquare$$

---

[2] Murty (1955) derived a similar result when he considered the distribution of L_1/L_2 and calculated the probabilities: $\Pr(L_1/L_2 > 1) = n_1/(n_1 + n_2)$ and $\Pr(L_1/L_2 \le 1) = n_2/(n_1 + n_2)$. In the following sections, we compare the differences between L_1 and L_2 rather than their ratio because relocation of a random variate by a constant (e.g., price) is generally simpler than rescaling.

The proof is in the Appendix. These shares are functions of the price difference, π, and are higher-order polynomials in the number of varieties, where the polynomial order is increasing in variety. If $p_1 > p_2$, then $\tilde{s}_1$ is the share of Coke purchased given the prices of Coke and Pepsi, and $1-\tilde{s}_1$ is the corresponding share of Pepsi purchased. If $p_1 < p_2$, then $1-\tilde{s}_2$ is the share of Coke, and $\tilde{s}_2$ is the share of Pepsi.

When $p_1 = p_2$, the shares in Proposition 3 are the same as those in Proposition 1: $\tilde{s}_1 = s_1$ = $\tilde{s}_2 = s_2$. If we constrain π/θ to the unit circle—which is equivalent to constraining prices such that $p_1, p_2 \in (0,\ \theta)$ so that $-1 < \pi/\theta < 1$—then $(1-\pi/\theta)^{n_2+j}$ and $(1+\pi/\theta)^{n_1+j}$ are on the unit interval,

$$\tilde{s}_1 < s_1 \text{ and } \tilde{s}_2 > s_2 \text{ for } p_1 > p_2,$$

$$\tilde{s}_2 < s_2 \text{ and } \tilde{s}_1 > s_1 \text{ for } p_1 < p_2.$$

If Coke's price rises above Pepsi's price, then the share of Coke falls ($\tilde{s}_1 < s_1$) and the share of Pepsi rises ($\tilde{s}_2 > s_2$). We can derive the comparative statics properties of shares with respect to prices and varieties:

Proposition 4. The brands' shares in Proposition 3 have the following properties:

(a) Share $\tilde{s}_i$ is decreasing in p_i and is first convex to the origin then concave in p_i with an inflection point at $p_i = p_j,\ i \neq j$.

(b) $\dfrac{\Delta\tilde{s}_1}{\Delta n_1} > 0,\ \dfrac{\Delta^2\tilde{s}_1}{\Delta n_1^2} < 0,\ \dfrac{\Delta\tilde{s}_1}{\Delta n_2} < 0,\ \dfrac{\Delta^2\tilde{s}_1}{\Delta n_2^2} < 0,$ and similarly for $\tilde{s}_2$.

(c) The "cross partial" with respect to the number of varieties is

$$\frac{\Delta\tilde{s}_1}{\Delta n_1 \Delta n_2} = \frac{n_2+1}{n_1+n_2+2}\left[\left(1-\frac{\pi}{\theta}\right)\tilde{s}_1(n_1+1) - \tilde{s}_1(n_1)\right] + \frac{n_1+1}{n_1+n_2+2}\frac{\Delta\tilde{s}_1}{\Delta n_2}$$ for $p_1 > p_2$. When prices are equal, $\dfrac{\Delta s_1}{\Delta n_1 \Delta n_2} = \dfrac{(n_2+1)}{(n_1+n_2+2)}\dfrac{\Delta s_1}{\Delta n_1} + \dfrac{(n_1+1)}{(n_1+n_2+2)}\dfrac{\Delta s_1}{\Delta n_2}$. If both prices are zero,

$$\frac{\Delta s_1}{\Delta n_1 \Delta n_2} = \frac{n_2 - n_1}{(n_1+n_2+2)(n_2+n_1)}$$, which is positive for $n_2 > n_1$ and negative for $n_1 > n_2$. ■

See the Appendix for the proof.

Different Supports

So far, we have treated the two brands symmetrically. However, consumers may prefer one brand to another in the sense that overall consumers will buy more of one brand than the other if prices are equal. One way to capture the difference in how much consumers like one brand relative to another is to allow the uniform distributions for each brand to have different supports: L_1 $L_1 \in [0, \theta_1]$ and $L_2 \in [0, \theta_2]$. The difference in the upper bounds of the supports represent the extent to which some consumers prefer one brand over the other given they can choose their most preferred variety from each. That is, θ_1 and θ_2 are preference parameters.

Proposition 5. Given that the brands' supports have different upper bounds, the shares of Proposition 3 become

$$\hat{s}_1(\theta_1, \theta_2, n_1, n_2, p_1, p_2) = \left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(1 - \frac{\pi}{\theta_1}\right)^{n_2+j}, \text{ for } p_1 - \theta_1 > p_2 - \theta_2,$$

$$\hat{s}_2(\theta_1, \theta_2, n_1, n_2, p_1, p_2) = \left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2}C_j}{{}_{n_1+j}C_j} \left(1 + \frac{\pi}{\theta_2}\right)^{n_1+j}, \text{ for } p_1 - \theta_1 \le p_2 - \theta_2. \blacksquare$$

The proof is similar to that in the symmetric case discussed in the Appendix. If $p_1 > p_2$, then $\hat{s}_1$ is the share of Coke purchased given the prices of Coke and Pepsi, and $1 - \hat{s}_1$ is the corresponding share of Pepsi purchased. If $p_1 < p_2$, then $1 - \hat{s}_2$ is the share of Coke, and $\hat{s}_2$ is the share of Pepsi. The inflection point occurring at $p_1 - \theta_1 = p_2 - \theta_2$.

If prices are zero or equal, these shares are:

$$\hat{s}_1(\theta_1, \theta_2, n_1, n_2, p_1, p_2) = \left(\frac{\theta_1}{\theta_2}\right)^{n_2} \frac{n_1}{n_1 + n_2}, \text{ for } \theta_2 > \theta_1,\ p_1 = p_2$$

$$\hat{s}_2(\theta_1, \theta_2, n_1, n_2, p_1, p_2) = \left(\frac{\theta_2}{\theta_1}\right)^{n_1} \frac{n_2}{n_1 + n_2}, \text{ for } \theta_2 \le \theta_1,\ p_1 = p_2.$$

Thus with equal prices, each firm's share is its share in Proposition 2, $n_1/(n_1 + n_2)$, adjusted by a multiplicative term that depends on the ratio of its preference parameters raised to a power equal to the number of varieties of the "more popular" brand. Each brand's share is increasing in its preference parameter and decreasing in its rival's.

Non-negatively Valued Outside Good

In Proposition 3 (where we incorporated prices), we implicitly ignored a possible problem that a consumer would choose a Coke or Pepsi variety even though the net surplus for that good was negative. We could avoid that problem by starting the uniform distribution at a high enough level that a negative net surplus is impossible.[3] More realistically, we assume that consumers will not buy either Coke or Pepsi, if an outside good gives them greater net surplus. The outside good has non-negative, non-random net surplus, $\omega \in [0, \min(\theta_1, \theta_2)]$, which, for simplicity, we assume is the same for all consumers. We partition the domain of the joint distribution of $\ell_1 \in [-p_1,\ \theta_1 - p_1]$ and $\ell_2 \in [-p_2,\ \theta_2 - p_2]$ into four regions and outcomes:

Proposition 6. The general shares for Coke and Pepsi are s_1^* and s_2^*:

$$s_1^* = \hat{s}_1 + \left[1 - \left(\frac{p_1+\omega}{\theta_1}\right)^{n_1}\right]\left(\frac{p_2+\omega}{\theta_2}\right)^{n_2} - \left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(\frac{p_2+\omega}{\theta_1}\right)^{n_2+j}$$
$$- \frac{n_2}{n_1+1}\left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1+1}C_j}{{}_{n_2-1+j}C_j} \left(\frac{p_2+\omega}{\theta_1}\right)^{n_2-1+j} \left[1 - \left(\frac{p_1+\omega}{\theta_1}\right)^{n_1+1+j}\right], \quad p_1 - \theta_1 > p_2 - \theta_2,$$

$$s_1^* = \hat{s}_1 - \left(\frac{p_1+\omega}{\theta_1}\right)^{n_1} + \left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2}C_j}{{}_{n_1+j}C_j} \left(\frac{p_1+\omega}{\theta_2}\right)^{n_1+j}$$
$$+ \frac{n_1}{n_2+1}\left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2+1}C_j}{{}_{n_1-1+j}C_j} \left(\frac{p_1+\omega}{\theta_2}\right)^{n_1-1+j} \left[1 - \left(\frac{p_2+\omega}{\theta_2}\right)^{n_2+1+j}\right], \quad p_1 - \theta_1 \le p_2 - \theta_2,$$

$$s_2^* = \hat{s}_2 + \left[1 - \left(\frac{p_1+\omega}{\theta_1}\right)^{n_1}\right]\left(\frac{p_2+\omega}{\theta_2}\right)^{n_2} - \left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2}C_j}{{}_{n_1+j}C_j} \left(\frac{p_1+\omega}{\theta_2}\right)^{n_1+j}$$
$$- \frac{n_1}{n_2+1}\left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2+1}C_j}{{}_{n_1-1+j}C_j} \left(\frac{p_1+\omega}{\theta_2}\right)^{n_1-1+j} \left[1 - \left(\frac{p_2+\omega}{\theta_2}\right)^{n_2+1+j}\right], \quad p_1 - \theta_1 \le p_2 - \theta_2.$$

$$s_2^* = \hat{s}_2 - \left(\frac{p_2+\omega}{\theta_2}\right)^{n_2} + \left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(\frac{p_2+\omega}{\theta_1}\right)^{n_2+j}$$
$$+ \frac{n_2}{n_1+1}\left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1+1}C_j}{{}_{n_2-1+j}C_j} \left(\frac{p_2+\omega}{\theta_1}\right)^{n_2-1+j} \left[1 - \left(\frac{p_1+\omega}{\theta_1}\right)^{n_1+1+j}\right], \quad p_1 - \theta_1 > p_2 - \theta_2. \quad \blacksquare$$

---

[3] For example, if the support for Coke is $[a, \theta_1]$ and that for Pepsi is $[a, \theta_2]$ with $\theta_1 > \theta_2$ and $p_1 = p_2$, and a larger than either price, then Pepsi's share is

$$s_2 = \left(\frac{\theta_2 - a}{\theta_1 - a}\right)^{n_1} \frac{n_2}{n_1 + n_2}.$$

See the Appendix for a proof. These are Hicksian (income compensated) demand functions. The demand function for the outside good (not buying Coke or Pepsi) is the residual $s_\omega^* = 1 - s_1^* - s_2^*$.

Despite the relative complexity of the general share equations in Proposition 6, all of the qualitative comparative statics results of Proposition 4 still apply but with the inflection point occurring at $p_1 - \theta_1 = p_2 - \theta_2$:

Proposition 7. Share s_1^* is decreasing in p_1 and is first convex to the origin then concave in p_1 with an inflection point at $p_1 = p_2 - \theta_2 + \theta_1$. Share s_2^* is decreasing in p_2 and is first convex then concave in p_2 with an inflection point at $p_2 = p_1 - \theta_1 + \theta_2$.[4] ■

When the price of Coke is above its demand curve's inflection point, the price of Pepsi must be below its demand curve's inflection point, and vice versa. The inflection point is not a function of the numbers of varieties, n_1 and n_2, although increases in the number of varieties affect the curvature of the demand curve on either side of the inflection point.

4. Coke and Pepsi Estimation

Our analytical results show that increasing the number of varieties of one brand can have complex effects on the demand curves and consumer welfare measures for both brands. To illustrate the role of variety on demand and on welfare, we estimate the simplest possible version of our model for Coke and Pepsi and an outside good in U.S. grocery stores. These dominant oligopolistic firms the two companies accounted for three-quarters of the U.S. carbonated beverage market in 1999, the sample period for our empirical analysis (according to *Beverage Digest*). Previous estimates of the demand for Coke and Pepsi (Gasmi et al. 1992, Golan et al. 2000, Dhar et al. 2005, and Chan 2006) ignored or downplayed the role of variety.

Our order-statistic model has three parameters: The maximum value a consumer receives from a Coke variety, θ_1; the maximum value for a Pepsi variety, θ_2; and the surplus (value net of price) of the outside good, ω. Below, we discuss how various generalizations and variants of this model produce very similar results. We present this model as a plausible approximation of reality that allows us to simulate the demand and welfare effects of changes in variety.

---

[4] The proof is identical to the proof of Proposition 4a but with all results rescaled by $(\theta_1 / \theta_2)^{n_1}$ or $(\theta_1 / \theta_2)^{n_2}$ and with s_i^* replacing $\tilde{s}_i$, everywhere.

Data

We use Information Resources Incorporated's (IRI) InfoScan® store-level scanner data for 1998 and 1999 to obtain 5,114 weekly observations for prices and quantities at 50 randomly chosen traditional grocery stores for each soft-drink variety (as determined by Universal Product Codes, UPCs).[5] The number of varieties that stores carry and the prices they charge vary across stores and over time within a store. These traditional grocery stores belong to 32 grocery chains. Some of the grocery chains are national giants such as Kroger, Albertsons, and Safeway, while others are relatively small, regional chains such as City Markets and Piggly Wiggly.

We restrict our analysis to 12-packs of 12 ounce cans. This package is the best-selling one within our data set, accounting for 46% of the total observations in the canned soft drink category. The number of varieties that each manufacturer produces is determined by the number of unique UPCs for the relevant package. Varieties differ by flavor, whether diet or regular, whether caffeinated or not, as well as how the products are packaged. Across all the stores in our sample, Coca-Cola has 27 varieties and Pepsi has 36 varieties.[6] Across all stores, the average annual number of varieties within each store is 10.86 (with a standard deviation of 2.76) for Coke and 9.08 (2.52) for Pepsi, the maximum number of varieties is 16 for both Coke and Pepsi, while the minimum number is 5 for Coke and 3 for Pepsi.

Each brand's store/week price is a quantity-weighted average obtained by dividing the total revenue in cents from all the products of the two firms by total volume in ounces. The price across varieties for a given brand is identical, but because of sales, the prices fluctuate over time (including sometimes within a week). The average price for Coke is 2.347¢ per ounce and that for Pepsi is 2.363¢ per ounce, with standard deviations of 0.514¢ and 0.473¢, respectively. The average price ratio of Pepsi to Coke across the stores is 1.02 with a standard deviation of 0.09.

---

[5] We have less than two full years of data for a few of the grocery stores that dropped out of the sample shortly before the end of the period. We also experimented with a panel of 100 stores for a single year (5,523 observations) and the results are very similar.

[6] Coke varieties include Coke, Diet Coke, Coke Classic, Caffeine Free Coke Classic, Caffeine-Free Diet Coke, Citra, Diet Cherry Coke, Diet Sprite, Fresca, Mello Yellow, Minute Maid, Diet Minute Maid Orange, Minute Maid Strawberry, Minute Maid Grape, Minute Maid Fruit Punch, Mr. Pibb, Sprite, Surge, and Tab. Pepsi varieties include Pepsi, Diet Pepsi, Caffeine-Free Pepsi, Caffeine-Free Diet Pepsi, Diet Minute Maid, Diet Mountain Dew, Mountain Dew Caffeine Free, Mountain Dew Citrus, Mug Root Beer, Mug Cream Soda, Josta, Diet Wild Cherry Pepsi, Pepsi One, Diet Pepsi Lemon Lime, Slice Strawberry, Slice Grape, Slice Mandarin Orange, Slice Lemon Lime, Slice Red, and Wild Cherry Pepsi.

The correlation coefficient between Coke and Pepsi's price is 0.89. In other words, the prices of Coke and Pepsi are almost always equal.

There are a number of possibilities for the outside good. We use soda products of the same package and size products manufactured by firms other than Coca-Cola Co. and PepsiCo, including the stores' private label products. Data on the outside good other than share are not explicitly used in the estimation. We estimate a constant outside good net surplus ω, which is used to predict of shares of the outside good.

Share Estimates

We have observations over week *t* for store *i*. Our three-equation system of equations is

$$s_{1it}^* = s_1^*(n_{1it}, n_{2it}, p_{1it}, p_{2it}, \theta_1, \theta_2, \omega) + u_{1it},$$
$$s_{2it}^* = s_2^*(n_{1it}, n_{2it}, p_{1it}, p_{2it}, \theta_1, \theta_2, \omega) + u_{2it},$$
$$s_{1it}^* + s_{2it}^* = 1 - \left(\frac{p_{1it} + \omega}{\theta_1}\right)^{n_{1it}} \left(\frac{p_{2it} + \omega}{\theta_2}\right)^{n_{2it}},$$

where the shares are measured by volume. (The revenue shares are virtually identical as there is very little variation in the average absolute and relative prices of Coke and Pepsi.) The Coke and Pepsi shares are given by the shares in Proposition 7 with a random error term added to them. We estimate this three-equation system using a nonlinear, least squares, two-step method described by Davidson and MacKinnon (1993, p. 664).[7]

In our system of brand-share equations, we allow for the possibility that the prices and varieties are endogenously determined by using instrumental variables. Our instruments include cost shifters at the national level for the soft drink industry, the national share of each chain, and milk sales within each store (IRI). The cost shifters are the producer price index (PPI, from the Bureau of Labor Statistics) for high-fructose corn syrup, which is the main sweetener used in most soft drink; the PPI for aluminum, which is used to make cans; and the PPI for industrial electricity, and the national PPI for gasoline price interacted with city dummies, which is a proxy for variations in transportation costs across cities. We assume that national cost shifters are correlated with prices but are not correlated with underlying consumer preferences that vary

---

[7] The nonlinear portion of the estimation was performed using the built-in unconstrained nonlinear estimation routine in MatLab. At each step of the estimation algorithm, we checked whether the theoretical constraints, $\omega \in [0, \min(\theta_1 - p_1, \theta_2 - p_2)]$, were violated and observed no violations. Had we discovered violations, we would have used a constrained estimation algorithm.

from store to store or over time. The national shares of the chains are used as proxies for possible monopsony power. The within-store milk sales variable is a proxy for the size of the store. Presumably larger stores carry more varieties due to lower shelf-space costs; however, store size should not be explicitly correlated with the error terms of our share equations. We also include squared terms of these instruments as additional instruments because prices and varieties enter the share equations nonlinearly (Davidson and MacKinnon, 1993).

In our first stage, we estimate the prices and number of varieties as linear functions of these instruments. The first-stage R^2 is 0.30 for Coke price, 0.30 for Pepsi price, 0.81 for the number of Coke varieties, and 0.86 for the number of Pepsi varieties.

Our order-statistics model estimates are very precise: $\theta_1 = 25.6181$ (with an asymptotic standard error of 0.0110), $\theta_2 = 25.2096$ (0.0103), and $\omega = 21.0219$ (0.0091). That is, all else equal, consumers slightly prefer Coke soft drinks to Pepsi's: $\theta_1 > \theta_2$.

This extremely terse model fits the data very well. Our model predicts the mean shares of Coke, Pepsi, and the outside good well. For example, Coke's average share in the data is 0.49, our model's average predicted share is 0.49, and the correlation coefficient between the actual and predicted share is 0.54. Similarly for Pepsi, the share is 0.26, the model's average predicted share is 0.27, and the correlation is 0.31. Finally, for the outside good, the corresponding numbers are 0.25, 0.24, and 0.16.

Comparison Model

For comparison, we estimated the same model using a mixed logit model where the shares of the goods are a function of the prices, the number of varieties, a Coca-Cola dummy, and store dummies. The dependen

t variables are the logarithm share of Coca-Cola minus the log outside share and the log Pepsi share minus the log outside share. We use the same instruments as for our model. We randomly drew errors from a normal distribution with zero mean. The mixed logit fits well. However, this estimation model is unsatisfactory because the price coefficient is not statistically significantly from zero.[8]

---

[8] This bizarre result could be due to estimation problems with the mixed logit such as those described by Knittel and Metaxoglou (2008) and Dubé et al. (2009).

Robustness Checks

We estimated three variations of our basic model with an outside good to check for the robustness of our results. First, we eliminated the main variety for each brand (Coke and Pepsi), which have very large market shares, and re-estimated the model. The estimated coefficients and the fit are virtually unaffected.

Second, we modified our model so that the three coefficients were linear functions of the average household income and the average number of family members in a household at a store. Doing so reduced the correlation coefficients (the fit of the model), and the coefficients on the demographic variables were not statistically significantly different from zero. However, the estimated model was very close to our three-parameter version.

Third, using a sample of 106 stores over a single year (5,523 observations), we estimated a model that allowed the three parameters to vary over stores. That is, $\theta_{1i} = \theta_1 + \delta_{1i}$; $\theta_{2i} = \theta_2 + \delta_{2i}$; and $\omega_i = \omega + \delta_{\omega i}$, where the δ's are estimated coefficients for each store. Presumably, the store dummies capture differences in all (average) demographic and other variables that vary across stores. The results of this flexible model were similar to the three-parameter model. The basic parameter estimates at the means were similar: $\theta_1 = 25.5723$, $\theta_2 = 25.3866$, and $\omega = 21.4642$, which are close to those from our simpler, three-parameter specification: 25.6181, 25.2096, and 21.0219, respectively. Also, the estimated δ's were relatively small in magnitude (the average is 0.1287 with a standard deviation of 0.2590). For example, 98% of the estimated θ_{1i} and θ_{2i} were between 25 and 26, with the three exceptions being for a few stores with very high Pepsi prices and low Pepsi varieties. In all cases, Coke was preferred to Pepsi at the store level ($\theta_{1i} > \theta_{2i}$). Thus, we use our simple three-parameter model in the following simulations as it simplifies the analysis and does not differ substantially from the more complex estimation model.

5. Simulations

Using our estimated values for θ_1, θ_2, and ω, we can calculate demand curves and welfare measures. We can also simulate comparative statics results to illustrate our analytic results.

Unless otherwise stated, our results are calculated at sample average prices of p_1 = 2.35¢ per ounce for Coke and p_2 = 2.36¢ per ounce for Pepsi, and the sample average number of varieties of $n_1 = 11$ for Coke and $n_2 = 9$ for Pepsi.

Demand Curves

At the sample average prices, the own elasticity of share with respect to price for Pepsi is -0.5. That is, a 1% increase in the price of Pepsi holding the price of Coke and the number of varieties fixed lowers Pepsi's share by 1.5%. The corresponding Coke elasticity is –0.9. These elasticities have a different interpretation than traditional ones, because they are asking how share changes as the price of all the brand's varieties change at once.

The demand curves in Figure 1 panels a and b, illustrate the results in Proposition 7 that the demand curves are first convex then concave in own price. The Pepsi demand curves in panel have an inflection point at $p_2 = p_1 - \theta_1 + \theta_2$, and the Coke demand curves in panel b have an inflection point at $p_1 = p_2 - \theta_2 + \theta_1$.

By varying the number of varieties, we can show how the demand curves shift, thereby illustrating our analytic partial and cross-partial derivative results (e.g., in Proposition 4). As the effects of a change in the number of varieties is qualitatively the same for both Coke and Pepsi demand curves, in Figure 1 panels a and b, we show how the Pepsi demand curve and the Coke demand curve change as we increase the number of Pepsi varieties, holding the number of Coke varieties and the rival's price fixed.

In Figure 1 panel a, we fix the price and varieties of Coke at their sample averages, p_1 = 2.35¢ per ounce and $n_1 = 11$. This figure shows the effect of an own variety change. Moving from the demand curve on the left to the one on the right, we increase the number of Pepsi varieties in increments of two from $n_2 = 7$ to 9 (the average in the sample) and then to 11. The figure shows that, as the number of varieties of Pepsi increases, Pepsi's demand curve rotates around the price-axis intercept, becoming flatter.

We illustrate the effect of the increase in the number of Pepsi varieties on the Coke demand curve in panel b of Figure 1. As the number of Pepsi varieties increases from $n_2 = 7$ to 9 to 11, the Coke demand curve rotates in around the price-axis intercept. Thus, an increase in the number of Pepsi varieties causes the Coke demand curve to become steeper.

As expected, these graphs show that a change in the number of Pepsi varieties has a larger (own) effect on Pepsi's demand curve than its (cross) effect on Coke's demand curve. Similarly, changes in the number of Coke varieties have a larger effect on its demand curve than on Pepsi's. At the sample mean prices, increasing the number of Pepsi varieties from 9 to 10 increases Pepsi's share by 0.0233, reduces Coke's by 0.0099, and reduces the share of the outside good by 0.0134 (= 0.0233 – 0.0099).

Panel c of Figure 1 illustrates cross-partial effects for Pepsi's demand curve. The central demand curve is evaluated at the sample averages where $n_1 = 11$ and $n_2 = 9$. The one to its left has two fewer varieties of each brand, $n_1 = 9$ and $n_2 = 7$, while the one to its right has two additional varieties, $n_1 = 13$ and $n_2 = 11$. When the numbers of varieties of both brands rise, fewer consumers buy the outside good, so the demand curves for both brands shift to the right. The effect of increase the numbers of varieties for both brand equally is larger for Pepsi than for Coke because there are initially more Coke varieties.

As the outside good's net surplus rises, consumers who receive a relatively low net surplus from their favorite soft drink switch to the outside good. Figure 2 shows the degree to which Pepsi's demand curve shifts to the left as the net surplus on the outside good increases by 5 percent increments.

Consumer Surplus

The Appendix derives the formulas for calculating consumer surplus. Panel a of Figure 3 shows that consumer surplus is increasing at a decreasing rate in the number of varieties of both goods. The figure is slightly asymmetric because consumer prefer Coke to Pepsi ($\theta_1 > \theta_2$).

Panel b of Figure 3 shows the corresponding "iso-welfare" curves, where each curve holds consumer surplus constant and the number of varieties of each brand vary (treating the number of varieties as a continuous variable). These are curves are horizontal slices of the three-dimensional surface in panel a. These curves are virtually straight lines with slope less than –1. That is, consumers are willing to trade slightly more than one Pepsi variety for a single Coke variety to keep consumer surplus constant. This slight deviation from –1 is the result of a slight preference for Coke: $\theta_1 - p_1 > \theta_2 - p_2$. Indeed, the slope becomes slightly more negative as the number of varieties increase, so that the iso-welfare lines are not parallel. That is as varieties increase, it takes a greater increase in Pepsi varieties to offset the loss of one Coke variety. This same effect appears in panel a. Total surplus is increasing at a decreasing rate in variety, but the rate decreases more slowly for Coke than for Pepsi. The cross-partial analysis is consistent with these figures: If we increase by one the number of varieties of both goods, Coke's share rises by 0.0191, which is more than Pepsi's share increases, 0.0105.

Consumer surplus varies with the number of varieties and price, as Figure 4 demonstrates.

In panel a, as the number of varieties of Coke increases holding the number of varieties of Pepsi and prices fixed, the consumer surplus of Coke rises, while that of Pepsi and the outside good fall.[9] Consequently, total consumer surplus rises, but only slightly as the gain to Coke barely exceeds the combined losses from Pepsi and the outside good.

As the price of Coke increases, holding the price of Pepsi constant, the consumer surplus from Coke falls, while the consumer surplus from Pepsi and that from the outside good rise, as panel b of Figure 4 illustrates. As the price gets very large, total consumer surplus levels off and there is little decrease in total surplus because virtually all consumers have switched to the other goods.

6. Pricing Implications of the Shape of the Demand Curves

The alternating convex-concave shape of the demand curves has important implications for price setting behavior by retailers and manufacturers (whose derived demand curves must also have this shape).[10] For example, Cowan (2007) shows that the nature of the price discrimination and the welfare effects of third-degree price discrimination depend on the concavity or convexity of the demand curve.

Perhaps more importantly, the shape affects the rate of adjustment and the probability of a price adjustment. To illustrate the importance of the shape of the demand curve for the price response to a shift in marginal cost, we assume that the product length is arbitrarily fixed and consider the monopoly's pricing problem (the same type of argument holds for oligopolies). If the monopoly faces a constant marginal and average cost of m, the monopoly's profit is $[p(Q) - m]Q$, so its first-order condition is $Qp'(Q) + p(Q) - m = 0$. Totally differentiating the first-order condition, we learn that $\mathrm{d}Q/\mathrm{d}m = 1/[2p'(Q) + Qp''(Q)]$. By the chain rule, the change in the price in response to a change in the marginal cost is $\mathrm{d}p/\mathrm{d}m = p'(Q)\mathrm{d}Q/\mathrm{d}m = p'(Q)/[2p'(Q) + Qp''(Q)]$. The numerator is the slope of the inverse demand curve and the denominator is the slope of the marginal revenue curve. Dividing both the numerator and the denominator by $D'(p)$, we can rewrite this expression as $\mathrm{d}p/\mathrm{d}m = 1/[2 + x]$, where $x = Qp''(Q)/p'(Q)$. Thus, if the demand curve is linear, $p''(Q) = 0$ so $x = 0$ and $\mathrm{d}p/\mathrm{d}c = ½$. If $p''(Q) < 0$, then $x > 0$ and $\mathrm{d}p/\mathrm{d}c < ½$. Finally, if $p''(Q) > 0$, then $x < 0$ and $\mathrm{d}p/\mathrm{d}c > ½$.

---

[9] The curves are interpolated to smooth the discrete changes in the varieties of Coke.

[10] Similarly, Luan and Sudhir (2010) make an analogous argument that to advertise optimally, a firm has to know not just how advertising shifts up the demand curve but how it affects the slope.

Thus, given our estimated demand curves with alternating curvature around the inflection point at the initial equilibrium, the firm's price adjustment is asymmetric with respect to an increase or a decrease in cost. Holding the number of varieties constant, a cost shock has a larger price and smaller quantity effect for a positive cost shock than for a negative cost shock.

The shape also has implications for whether a firm adjusts its price if it incurs and adjustment or menu cost when it changes its price. It pays for a firm to adjust its price if the increase in gross profit from adjusting the price (ignoring adjustment costs) exceeds the menu cost. Thus, given that the shape of the demand curve causes asymmetric price adjustments, a firm is more likely to adjust its prices at all when faced with a positive cost shock rather than to a comparable size negative shock.

7. Summary and Conclusions

We examine markets in which duopoly brands sell many varieties. As the number of varieties of one brand increases holding the other brands' price and number of varieties fixed, consumers are more likely to buy that brand, as more consumers will find a variety that they prefer to those of the other brand and to the outside good.

Rather than imposing an explicit functional form on utility, we derive consumers' demand functions for each brand using order statistics where consumers' valuations of varieties are distributed independently uniform. Consumers rank choices based on net surplus—valuation minus price—of the varieties of two brands and an outside good.

The main advantage of this new model is that it allows us to examine many questions analytically (which cannot be done directly in mixed logit and other empirical studies). We derive analytic partial derivative results that show how changes in price and varieties affect demand.

We fit our model for Coke, Pepsi, and an outside good using U.S. grocery store data. Based on these estimates, we simulate the demand and welfare comparative statics results for changes in prices, varieties, and the value of the outside good.

We show both analytically and empirically that the brand demand curves have a convex and a concave section around an inflection point. Given that the prices of Coke and Pepsi are generally very close to each other, stores operate at or near the inflection point. As a result, a positive cost shock has a larger price effect and a smaller quantity effect than does a comparable negative cost shock. Given that grocery stores face adjustment or menu costs, stores are more likely to adjust price when faced with a positive cost shock than with a negative cost shock.

References

Abramowitz, M., and I. S. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1972.

Andrews, G.E., R. Askey, and R. Roy, *Special Functions*, Cambridge: Cambridge University Press, 2001.

Anderson, Simon P., André de Palma, Jacques-François Thisse, *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press, 1992.

Bayus, Barry L, and William P. Putsis, Jr., "Product Proliferation: An Empirical Analysis of Product Line Determinants and Market Outcome," *Marketing Science*, 18(2), 1999:137-153.

Brander, James A., and Jonathan Eaton, "Product Line Rivalry," *American Economic Review*, 74, June 1984:323-334.

Chan, Tat Y., "Estimating a Continuous Hedonic-Choice Model with an Application to Demand for Soft Drinks," *Rand Journal of Economics*, 37(2), Summer 2006:466-82.

Davidson, Russell, and James G. MacKinnon, *Estimation and Inference in Econometrics*, New York, Oxford University Press, 1993.

Deneckere, Raymond, and Michael Rothschild, "Monopolistic Competition and Preference Diversity," *Review of Economic Studies*, 59(2), April 1992:361-73.

Dhar, Tirtha, Jean-Paul Chavez, Ronald W. Cotterill, and Brian W. Gould, "An Econometric Analysis of Brand-Level Strategic Pricing Between Coca-Cola Company and PepsiCo," *Journal of Economics & Management Strategy*, 14(4), Winter 2005:905-31.

Draganska, Michaela, and Dipak C. Jain, "Product-Line Length as a Competitive Tool," *Journal of Economics and Management Strategy*, 14(1), Spring 2005:1-28.

Dixit, Avinash K., and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), June 1977:297-308.

Dubé, Jean-Pierre H., Jeremy T. Fox, Che-Lin Su, "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," NBER Working Paper, 2009.

Gasmi, Faid, Jean Jacques Laffont, and Quang H. Vuong, "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics & Management Strategy*, 1(2), Summer 1992:277-311.

Gauss, C. F., "Disquisitiones Generales Circa Seriem Infinitum $1+\frac{\alpha\beta}{1\cdot\gamma}x+\frac{\alpha(\alpha+1)\beta(\beta+1)}{1\cdot 2\cdot\gamma(\gamma+1)}x^2+\frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1\cdot 2\cdot 3\cdot\gamma(\gamma+1)(\gamma+1)}x^3$ etc. Pars Prior," *Commentationes Societiones Regiae Scientiarum Göttingensis Recentiores,* 2, 1813:123-162.

Gilbert, Richard J., and Carmen Matutes, "Product Line Rivalry with Brand Differentiation," *Journal of Industrial Economics*, 41(3), September 1993: 223-240.

Golan, Amos, Larry S. Karp, and Jeffrey M. Perloff, "Estimating Coke and Pepsi's Price and Advertising Strategies," *Journal of Business and Economic Statistics*, 18(4), October 2000:398-409.

Hausman, Jerry, "Valuation of New Goods Under Perfect and Imperfect Competition," in T. Bresnahan and R. Gordon, eds., *The Economics of New Goods*, Studies in Income and Wealth, 58, 1996, Chicago: National Bureau of Economic Research.

Heidhues, Paul, and Botond Koszegi, "Competition and Price Variation when Consumers are Loss Averse," *American Economic Review,* 98(4), May 2008:1245-1268.

Israilevich, Guillermo, "Assessing Supermarket Product-Line Decisions: The Impact of Slotting Fees," *Quantitative Marketing and Economics*, 2, 2004:141-167.

Kadiyali, Vrinda, Naufel Vilcassim, and Pradeep Chintagunta, "Product Line Extensions and Competitive Market Interactions: An Empirical Analysis," *Journal of Econometrics*, 89(1-2), March-April 1999:339-363.

Kim, Jaehway, Greg M. Allenby, and Peter E. Rossi, "Modeling Consumer Demand for Variety," *Marketing Science*, 21(3), Summer, 2002:229-250.

Klemperer, Paul, "Equilibrium Product Lines: Competing Head-to-Head May be Less Competitive," *American Economic Review*, 82(4), September 1992:740-755.

Knittel, Christopher R., and Konstantinos Metaxoglou, "Estimation of Random Coefficient Demand Models: Challenges, Difficulties and Warnings," NBER Working Paper, 2008.

Luan, Y. Jackie, and K. Sudhir, "Forecasting Marketing-Mix Responsiveness for New Products," *Journal of Marketing Research*, 47(3), June 2010:444–457.

McMillan, Rob, “Different Flavor, Same Price: The Puzzle of Uniform Pricing for Differentiated Products,” 2005, www.ftc.gov/be/seminardocs/050426diffflavorsmcmillan.pdf.

Murty, V. N., “The Distribution of the Quotient of Maximum Values in Samples from a Rectangular Distribution, *Journal of the American Statistical Association*, 50(272), December, 1955:1136-1141.

Nevo, Aviv, “New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems,” *Review of Economics and Statistics*, 85(2), May 2003:266-275.

Nijssen, Edwin J., and Hans C. M. Van Trijp, “Branding Fresh Food Products: Exploratory Empirical Evidence from the Netherlands,” *European Review of Agricultural Economics*, 25(2), 1998:228-42.

Orback, Barak Y., and Liran Einav, “Uniform Prices for Differentiated Goods: The Case of the Movie-Theater Industry,” *International Review of Law and Economics,* 27(2), 2007:129-153.

Perloff, Jeffrey M., and Steven C. Salop, “Equilibrium with Product Differentiation,” *Review of Economic Studies*, 52(1), January, 1985: 107-20.

Pochhammer, L., “Über hypergeometrische Funktion *n*ter Ordnung,” *Journal fur Die Reine und Angewandte Mathematik*, 71, 1870:316-352.

Raubitschek, Ruth S., “A Model of Product Proliferation with Multiproduct Firms,” *Journal of Industrial Economics*, 35(3), March, 1987:269-279.

Salop, Steven C., “Monopolistic Competition with Outside Goods,” *Bell Journal of Economics and Management Science*, 10(1), Spring, 1979:141-156.

Spence, A. Michael, “Product Selection, Fixed Costs, and Monopolistic Competition,” *Review of Economic Studies*, 43(2), June, 1976:217-36.

Villas-Boas, J. Miguel, “Communication Strategies and Product Line Design,” *Marketing Science*, 23(3), Summer, 2004:304-316.

Appendix: Proofs and Derivations

Proof of Proposition 1.

The marginal distributions of the maxima are $n_1\theta^{-n_1}L_1^{n_1-1}dL_1$ and $n_2\theta^{-n_2}L_2^{n_2-1}dL_2$, and the cumulations below c of each maxima are $\theta^{-n_1}c_1^{n_1}$ and $\theta^{-n_2}c_2^{n_2}$. The joint distribution of the maximal net surplus of Coke and Pepsi is:

$$f_L(L_1,L_2)=n_1n_2\theta^{-(n_1+n_2)}L_1^{n_1-1}L_2^{n_2-1},$$

with cumulations below $(c_1,\ c_2)$ as $\theta^{-(n_1+n_2)}c_1^{n_1}c_2^{n_2}$. The distribution of the difference of the maxima, $D_0=L_1-L_2\in[-\theta,\ \theta]$, has two parts:

$$f_{D_0}(D_0)=\begin{Bmatrix} \int_0^{\theta-D_0} f_L(D_0+L_2,L_2)dL_2 & D_0\in(0,\ \theta] \\ \int_0^{\theta+D_0} f_L(L_1,L_1-D_0)dL_1 & D_0\in[-\theta,\ 0] \end{Bmatrix}.$$

The upper-bound of the first integral maps $D_0\in[0,\ \theta]$ through $L_2=\theta-D_0$ to $L_2\in[0,\ \theta]$, and the upper-bound of the second maps $D_0\in[-\theta,\ 0]$ through $L_1=\theta+D_0$ to $L_1\in[0,\ \theta]$. After repeated integration by parts:

$$f_{D_0}(D_0)=\begin{Bmatrix} \dfrac{n_2}{\theta^{m+n}}\sum_{j=1}^{m}(-1)^{j-1}\dfrac{{}_{n_1}C_j}{{}_{n_2-1+j}C_j}(L_2+D_0)^{n_1-j}(L_2)^{n_2-1+j}\Big|_0^{\theta-D_0} & D_0\in(0,\ \theta] \\ \dfrac{n_1}{\theta^{m+n}}\sum_{j=1}^{n}(-1)^{j-1}\dfrac{{}_{n_2}C_j}{{}_{n_1-1+j}C_j}(L_1)^{n_1-j}(L_1-D_0)^{n_2-1+j}\Big|_0^{\theta+D_0} & D_0\in[-\theta,\ 0] \end{Bmatrix}$$

Evaluating the limits of integration:

$$f_{D_0}(D_0)=\begin{Bmatrix} \dfrac{n_2}{\theta}\sum_{j=1}^{m}(-1)^{j-1}\dfrac{{}_{n_1}C_j}{{}_{n_2-1+j}C_j}\left(1-\dfrac{D_0}{\theta}\right)^{n_2-1+j}, & D_0\in(0,\ \theta] \\ \dfrac{n_1}{\theta}\sum_{j=1}^{n}(-1)^{j-1}\dfrac{{}_{n_2}C_j}{{}_{n_1-1+j}C_j}\left(1+\dfrac{D_0}{\theta}\right)^{n_1-1+j}, & D_0\in[-\theta,\ 0] \end{Bmatrix}.$$

Evaluating cumulations above and below zero the results of Proposition 1 follow. ■

Proof of Proposition 3.

The joint distribution of the maxima adjusted for price is:

$$f_\ell(\ell_1,\ell_2)=n_1n_2\theta^{-(n_1+n_2)}(\ell_1+p_1)^{n_1-1}(\ell_2+p_2)^{n_2-1},\ \ell_g\in[-p_g,\theta-p_g],\ g=1,2.$$

The distribution of the difference $D=\ell_1-\ell_2=D_0-\pi$ is:

$$f_D(D) = \left\{ \begin{array}{ll} \int\limits_{-p_2}^{\theta - p_1 - D} f_\ell(\ell_2 + D, \ell_2) d\ell_2, & D \in (-\pi,\ \theta - \pi],\ p_1 > p_2 \\ \int\limits_{-p_1}^{\theta - p_2 + D} f_\ell(\ell_1, \ell_1 - D) d\ell_1, & D \in [-\theta - \pi,\ -\pi],\ p_1 \le p_2 \end{array} \right\}.$$

The upper-bound of the first integral maps $D \in [-\pi,\ \theta - \pi]$ through $\ell_2 = \theta - p_1 - D$ to $\ell_2 \in [-p_2,\ \theta - p_2]$, and the upper-bound of the second maps $D \in [-\theta - \pi,\ -\pi]$ through $\ell_1 = \theta - p_2 + D$ to $\ell_1 \in [-p_1,\ \theta - p_1]$. After repeated integration by parts:

$$f_D(D) = \left\{ \begin{array}{ll} \dfrac{n_2}{\theta^{m+n}} \sum\limits_{j=1}^{m} (-1)^{j-1} \dfrac{{}_{n_1}C_j}{{}_{n_2-1+j}C_j} (\ell_2 + D + p_1)^{n_1 - j} (\ell_2 + p_2)^{n_2 - 1 + j} \Big|_0^{\theta - p_1 - D} & D \in (0,\ \theta - \pi] \\ \dfrac{n_1}{\theta^{m+n}} \sum\limits_{j=1}^{n} (-1)^{j-1} \dfrac{{}_{n_2}C_j}{{}_{n_1-1+j}C_j} (\ell_1 + p_1)^{n_1 - j} (\ell_1 - D + p_2)^{n_2 - 1 + j} \Big|_0^{\theta - p_2 + D} & D \in [-\theta - \pi,\ 0] \end{array} \right\}.$$

Evaluating the limits of integration:

$$f_D(D) = \left\{ \begin{array}{ll} \dfrac{n_2}{\theta} \sum\limits_{j=1}^{n_1} (-1)^{j-1} \dfrac{{}_{n_1}C_j}{{}_{n_2-1+j}C_j} \left(1 - \dfrac{\pi - D}{\theta}\right)^{n_2 - 1 + j} & D \in (-\pi,\ \theta - \pi],\ p_1 > p_2 \\ \dfrac{n_1}{\theta} \sum\limits_{j=1}^{n_2} (-1)^{j-1} \dfrac{{}_{n_2}C_j}{{}_{n_1-1+j}C_j} \left(1 + \dfrac{\pi + D}{\theta}\right)^{n_1 - 1 + j} & D \in [-\theta - \pi,\ -\pi],\ p_1 \le p_2 \end{array} \right\}.$$

As one might expect, this distribution is the same as the distribution of D_0 displaced by the price differential. Evaluating cumulations above and below zero the results of Proposition 3 follow. ■

Proof of Proposition 4.

(a) Applying Leibniz's rule to the probability integrals (evaluated at $D = 0$) in the proof of proposition 3, the derivatives of the shares with respect to prices are,

$$\frac{\partial \tilde{s}_1(n_1, n_2)}{\partial p_2} = -\frac{\partial \tilde{s}_2(n_1, n_2)}{\partial p_2} = \frac{n_2}{\theta} \tilde{s}_1(n_1, n_2 - 1) \ge 0, \quad p_1 > p_2,$$

$$\frac{\partial \tilde{s}_2(n_1, n_2)}{\partial p_1} = -\frac{\partial \tilde{s}_1(n_1, n_2)}{\partial p_1} = \frac{n_1}{\theta} \tilde{s}_2(n_1 - 1, n_2) \ge 0, \quad p_1 \le p_2.$$

Interestingly, derivates of the shares are functions of the shares evaluated at lower numbers of varieties. Second derivatives follow directly,

$$\frac{\partial^2 \tilde{s}_2(n_1,n_2)}{\partial p_2^2} = -\frac{\partial^2 \tilde{s}_1(n_1,n_2)}{\partial p_2^2} = -\frac{n_2(n_2-1)}{\theta}\tilde{s}_1(n_1,n_2-2) \le 0, \quad p_1 > p_2,$$
$$\frac{\partial^2 \tilde{s}_1(n_1,n_2)}{\partial p_1^2} = -\frac{\partial^2 \tilde{s}_2(n_1,n_2)}{\partial p_1^2} = -\frac{n_1(n_1-1)}{\theta}\tilde{s}_2(n_1-2,n_2) \le 0, \quad p_1 \le p_2.$$

Taking derivatives of the equations in Proposition 3, we have,

$$\frac{\partial \tilde{s}_1(n_1,n_2)}{\partial p_1} = -\frac{\partial \tilde{s}_1(n_1,n_2)}{\partial p_2} = -\frac{n_2}{\theta}\tilde{s}_1(n_1,n_2-1), \quad p_1 > p_2,$$
$$\frac{\partial \tilde{s}_2(n_1,n_2)}{\partial p_2} = -\frac{\partial \tilde{s}_2(n_1,n_2)}{\partial p_1} = -\frac{n_1}{\theta}\tilde{s}_2(n_1-1,n_2), \quad p_1 \le p_2.$$

Taking derivatives again,

$$\frac{\partial^2 \tilde{s}_1(n_1,n_2)}{\partial p_1^2} = -\frac{\partial^2 \tilde{s}_1(n_1,n_2)}{\partial p_1 \partial p_2} = -\frac{n_2}{\theta}\frac{\partial \tilde{s}_1(n_1,n_2-1)}{\partial p_1}, \quad p_1 > p_2,$$
$$\frac{\partial^2 \tilde{s}_2(n_1,n_2)}{\partial p_2^2} = -\frac{\partial^2 \tilde{s}_2(n_1,n_2)}{\partial p_2 \partial p_1} = -\frac{n_1}{\theta}\frac{\partial \tilde{s}_2(n_1-1,n_2)}{\partial p_2}, \quad p_1 \le p_2.$$

To sign these last two results, we must sign the two derivates of the right-hand sides by again applying Leibniz's rule to the probability integrals in the proof of proposition 3, yielding,

$$\frac{\partial \tilde{s}_1(n_1,n_2-1)}{\partial p_1} = -\frac{n_1}{\theta}\left[\left(1-\frac{\pi}{\theta}\right)^{n_2} - \tilde{s}_1(n_1-1,n_2-1)\right] \le 0, \quad p_1 > p_2.$$
$$\frac{\partial \tilde{s}_2(n_1-1,n_2)}{\partial p_2} = \frac{n_2}{\theta}\left[\left(1+\frac{\pi}{\theta}\right)^{n_1} - \tilde{s}_2(n_1-1,n_2-1)\right] \le 0, \quad p_1 \le p_2.$$

Hence, we have that $\partial \tilde{s}_i / \partial p_i \le 0$ everywhere, but $\partial^2 \tilde{s}_1 / \partial p_1^2 \ge 0$ and $\partial^2 \tilde{s}_2 / \partial p_2^2 \le 0$ for $p_1 > p_2$, and $\partial^2 \tilde{s}_1 / \partial p_1^2 \le 0$ and $\partial^2 \tilde{s}_2 / \partial p_2^2 \ge 0$ for $p_1 \le p_2$. So the main results hold.

(b) First consider the case where $p_1 > p_2$. If we increase n_1 by one in Proposition 3,

$$\tilde{s}_1(n_1+1) = \sum_{j=1}^{n_1+1}(-1)^{j-1}\frac{{}_{n_1+1}C_j}{{}_{n_2+j}C_j}\left(1-\frac{\pi}{\theta}\right)^{n_2+j} = \sum_{j=1}^{n_1+1}(-1)^{j-1}\frac{{}_{n_1}C_j}{{}_{n_2+j}C_j}\left(1-\frac{\pi}{\theta}\right)^{n_2+j}\left(\frac{n_1+1}{n_1+1-j}\right)$$
$$> \sum_{j=1}^{n_1+1}(-1)^{j-1}\frac{{}_{n_1}C_j}{{}_{n_2+j}C_j}\left(1-\frac{\pi}{\theta}\right)^{n_2+j} = \sum_{j=1}^{n_1}(-1)^{j-1}\frac{{}_{n_1}C_j}{{}_{n_2+j}C_j}\left(1-\frac{\pi}{\theta}\right)^{n_2+j} = \tilde{s}_1.$$

This result follows because $(n_1 + 1)/(n_1 + 1 - j) > 1$ and the $(n_1 + 1)^{th}$ term in the sum of the second to last line is zero because ${}_{n_1}C_{n_1+1} = 0$. Thus we have:

$$\frac{\Delta \tilde{s}_1}{\Delta n_1} = \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(1 - \frac{\pi}{\theta}\right)^{n_2+j} \left(\frac{j}{n_1 + 1 - j}\right) > 0.$$

Given this, it must be true that $\frac{\Delta^2 \tilde{s}_1}{\Delta n_1^2} < 0$, for if it were not, then as $n_1 \to \infty$, $\tilde{s}_1 > 1$, which violates the axioms of probability. Increasing n_2 in Proposition 3,

$$\begin{aligned}
\tilde{s}_1(n_2 + 1) &= \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+1+j}C_j} \left(1 - \frac{\pi}{\theta}\right)^{n_2+1+j} \\
&= \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(1 - \frac{\pi}{\theta}\right)^{n_2+j} \left(\frac{n_2 + 1}{n_2 + 1 + j} \cdot \left(1 - \frac{\pi}{\theta}\right)\right) \\
&< \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(1 - \frac{\pi}{\theta}\right)^{n_2+j} = \tilde{s}_1(n_2).
\end{aligned}$$

This result follows because $(n_2 + 1)/(n_2 + 1 + j) < 1$ and $(1 - \pi/\theta) \leq 1$. Thus we have for $p_1 > p_2$:

$$\frac{\Delta \tilde{s}_1(n_2)}{\Delta n_2} = -\sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j} \left(1 - \frac{\pi}{\theta}\right)^{n_2+j} \left(1 - \frac{n_2 + 1}{n_2 + 1 + j} \cdot \left(1 - \frac{\pi}{\theta}\right)\right) < 0$$

Using the difference of the differences, it is easy to show that $\frac{\Delta^2 \tilde{s}_1}{\Delta n_2^2} < 0$. By showing similar results for the case $p_1 < p_2$, the proof is complete.

(c) To understand the cross-difference (cross-partial difference) with respect to n_1 and n_2 it is useful to restate the shares in terms of the Gaussian hypergeometric function (Gauss 1813 or Pachhammer 1870):

$$F(a,\ b;\ c;\ z) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j\, j!} z^j,$$

where $(a)_j = a(a+1)...(a+j-1)$ is the Pochhammer rising factorial. When the first argument is negative, the hypergeometric sum is finite. As a result, the probability in Proposition 3 can be restated as:

$$\tilde{s}_1(n_1, n_2, p_1, p_2) = \left(1 - \frac{\pi}{\theta}\right)^{n_2} \left[1 - F(-n_1,\ 1;\ n_2 + 1;\ 1 - \pi / \theta)\right].$$

Rearranging:

$$F(-n_1,\ 1;\ n_2 + 1;\ 1 - \pi / \theta) = 1 - \left(1 - \frac{\pi}{\theta}\right)^{-n_2} \tilde{s}_1.$$

Incrementing n_1 and n_2:

$$F(-n_1-1,\ 1;\ n_2+1;\ 1-\pi/\theta)=1-\left(1-\frac{\pi}{\theta}\right)^{-n_2}\tilde{s}_1(n_1+1),$$

$$F(-n_1,\ 1;\ n_2+2;\ 1-\pi/\theta)=1-\left(1-\frac{\pi}{\theta}\right)^{-n_2-1}\tilde{s}_1(n_2+1),$$

$$F(-n_1-1,\ 1;\ n_2+2;\ 1-\pi/\theta)=1-\left(1-\frac{\pi}{\theta}\right)^{-n_2-1}\tilde{s}_1(n_1+1,n_2+1).$$

Substituting these into Equation 15.2.17 of Abramowitz and Stegun (1972) yields:

$$\tilde{s}_1(n_1+1,n_2+1)=\frac{n_2+1}{n_1+n_2+2}\left(1-\frac{\pi}{\theta}\right)\tilde{s}_1(n_1+1)+\frac{n_1+1}{n_1+n_2+2}\tilde{s}_1(n_2+1).$$

Subtracting $\tilde{s}_1$ yields the result. ■

Proof of Proposition 6.

Consider the following regions in two-dimensional Cartesian space.

Region	*Market Share*
$R_{\omega\omega}=\{(\ell_1,\ \ell_2):\ \ell_1\in[-p_1,\ \omega),\ \ell_2\in[-p_2,\ \omega)\}$	$s_1^*=s_2^*=0$
$R_{1\omega}=\{(\ell_1,\ \ell_2):\ \ell_1\in[\omega,\ \theta_1-p_1],\ \ell_2\in[-p_2,\ \omega)\}$	$s_1^*\in[0,\ 1]$; $s_2^*=0$
$R_{\omega 2}=\{(\ell_1,\ \ell_2):\ \ell_1\in[-p_1,\ \omega),\ \ell_2\in[\omega,\ \theta_2-p_2]\}$	$s_1^*=0$; $s_2^*\in[0,\ 1]$
$R_{12}=\{(\ell_1,\ \ell_2):\ \ell_1\in[\omega,\ \theta_1-p_1],\ \ell_2\in[\omega,\ \theta_2-p_2]\}$	$s_1^*\in[0,\ 1]$; $s_2^*\in[0,\ 1]$

Given that some consumers purchase the outside good instead of a Coke or Pepsi variety, $s_1^*+s_2^*\le 1$. Coke and Pepsi shares equal the probabilities:

$$(*)\quad s_1^*(n_1,n_2,p_1,p_2,\theta_1,\theta_2)=\left\{\begin{matrix}\Pr(\ell_1>\ell_2\cap R_{12})+\Pr(R_{1\omega}) & p_1-\theta_1>p_2-\theta_2\\ 1-\Pr(\ell_1\le\ell_2\cap R_{12})-\Pr(R_{\omega 2})-\Pr(R_{\omega\omega}) & p_1-\theta_1\le p_2-\theta_2\end{matrix}\right\},$$

$$(**)\quad s_2^*(n_1,n_2,p_1,p_2,\theta_1,\theta_2)=\left\{\begin{matrix}\Pr(\ell_1\le\ell_2\cap R_{12})+\Pr(R_{\omega 2}) & p_1-\theta_1\le p_2-\theta_2\\ 1-\Pr(\ell_1>\ell_2\cap R_{12})-\Pr(R_{1\omega})-\Pr(R_{\omega\omega}) & p_1-\theta_1>p_2-\theta_2\end{matrix}\right\}.$$

The probability masses on $R_{1\omega}$, $R_{\omega 2}$, and $R_{\omega\omega}$ are cumulations:

$$\Pr(R_{1\omega})=\iint_{R_{1\omega}} f_\ell(\ell_1,\ell_2)d\ell_1 d\ell_2=\left[1-\left(\frac{p_1+\omega}{\theta_1}\right)^{n_1}\right]\left(\frac{p_2+\omega}{\theta_2}\right)^{n_2},$$

$$\Pr(R_{\omega 2}) = \iint_{R_{\omega 2}} f_\ell(\ell_1, \ell_2) d\ell_1 d\ell_2 = \left[1 - \left(\frac{p_2 + \omega}{\theta_2}\right)^{n_2}\right]\left(\frac{p_1 + \omega}{\theta_1}\right)^{n_1},$$

$$\Pr(R_{\omega\omega}) = \iint_{R_{\omega\omega}} f_\ell(\ell_1, \ell_2) d\ell_1 d\ell_2 = \left(\frac{p_1 + \omega}{\theta_1}\right)^{n_1}\left(\frac{p_2 + \omega}{\theta_2}\right)^{n_2}.$$

The last equation is the probability that the outside good is purchased, and is, hence, the share function for that good. Probabilities in R_{12} are:

$$\Pr(\ell_1 > \ell_2 \subset R_{12}) = \left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1}C_j}{{}_{n_2+j}C_j}\left[\left(1 - \frac{\pi}{\theta_1}\right)^{n_2+j} - \left(\frac{p_2 + \omega}{\theta_1}\right)^{n_2+j}\right]$$
$$- \frac{n_2}{n_1 + 1}\left(\frac{\theta_1}{\theta_2}\right)^{n_2} \sum_{j=1}^{n_1} (-1)^{j-1} \frac{{}_{n_1+1}C_j}{{}_{n_2-1+j}C_j}\left(\frac{p_2 + \omega}{\theta_1}\right)^{n_2-1+j}\left[1 - \left(\frac{p_1 + \omega}{\theta_1}\right)^{n_1+1+j}\right],$$

for $p_1 - \theta_1 > p_2 - \theta_2$, and

$$\Pr(\ell_1 \le \ell_2 \subset R_{12}) = \left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2}C_j}{{}_{n_1+j}C_j}\left[\left(1 + \frac{\pi}{\theta_2}\right)^{n_1+j} - \left(\frac{p_1 + \omega}{\theta_2}\right)^{n_1+j}\right]$$
$$- \frac{n_1}{n_2 + 1}\left(\frac{\theta_2}{\theta_1}\right)^{n_1} \sum_{j=1}^{n_2} (-1)^{j-1} \frac{{}_{n_2+1}C_j}{{}_{n_1-1+j}C_j}\left(\frac{p_1 + \omega}{\theta_2}\right)^{n_1-1+j}\left[1 - \left(\frac{p_2 + \omega}{\theta_2}\right)^{n_2+1+j}\right],$$

for $p_1 - \theta_1 \le p_2 - \theta_2$. Substituting the probabilities into equations (*) and (**) produces the main result of Proposition 6. ■

Calculating Consumer Surplus.

The area under the demand curves s_1^* and s_2^* can be used to calculate the effect of increasing varieties (Δn_1 or Δn_2) on total consumer surplus (*CS*). For instance, if we increase the varieties of Coke by one ($\Delta n_1 = 1$), demand for Coke increases so that demand for Pepsi and the outside good change, and the change (increase) in total CS can be represented as the difference of the areas under the compensated demand function for Coke (and above its price) before and after the change. That is, when $p_1 - \theta_1 > p_2 - \theta_2$,

$$\frac{\Delta CS(p_1 - \theta_1 > p_2 - \theta_2)}{\Delta n_1} = \int_{p_1}^{\theta_1 - \omega} s_1^*(n_1 + 1, p_1 - \theta_1 > p_2 - \theta_2) dp_1 - \int_{p_1}^{\theta_1 - \omega} s_1^*(n_1, p_1 - \theta_1 > p_2 - \theta_2) dp_1.$$ In this

equation, with a slight abuse our notation, the share $s_1^*(n_1, p_1 - \theta_1 > p_2 - \theta_2)$ corresponds to Coke demand when $p_1 - \theta_1 > p_2 - \theta_2$. Because the demand functions are compensated, we need calculate only areas under the Coke function. We integrate with respect to price from the given price, p_1, to the upper bound, $\theta_1 - \omega$, because demand for Coke is zero if the price is higher (the outside good dominates for all consumers).

When $p_1 - \theta_1 \le p_2 - \theta_2$ the calculation is slightly different:

$$\frac{\Delta CS(p_1-\theta_1 \le p_2-\theta_2)}{\Delta n_1} = \int_{p_1}^{p_2-\theta_2+\theta_1} s_1^*(n_1+1,\ p_1-\theta_1 \le p_2-\theta_2)dp_1 - \int_{p_1}^{p_2-\theta_2+\theta_1} s_1^*(n_1,\ p_1-\theta_1 \le p_2-\theta_2)dp_1$$
$$+\int_{p_2-\theta_2+\theta_1}^{\theta_1-\omega} s_1^*(n_1+1,\ p_1-\theta_1 > p_2-\theta_2)dp_1 - \int_{p_2-\theta_2+\theta_1}^{\theta_1-\omega} s_1^*(n_1,\ p_1-\theta_1 > p_2-\theta_2)dp_1.$$

To calculate the change in total consumer surplus, we must use both parts of the demand function, s_1^*. These CS calculations are complicated because the demand equation for a brand differs depending on relative prices.

Similarly, the equation for s_2^* can be used to calculate the increase in total consumer demand when the varieties of Pepsi are increased by one. For example,

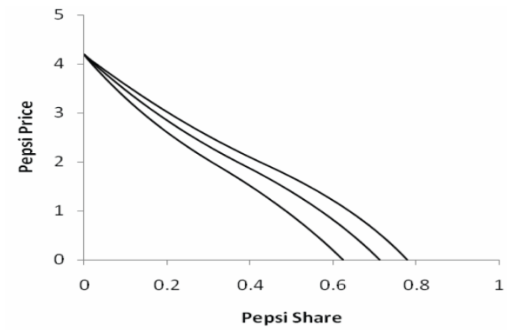
$$\frac{\Delta CS(p_1-\theta_1 \le p_2-\theta_2)}{\Delta n_2} = \int_{p_2}^{\theta_2-\omega} s_2^*(n_2+1,\ p_1-\theta_1 \le p_2-\theta_2)dp_2 - \int_{p_2}^{\theta_2-\omega} s_2^*(n_2,\ p_1-\theta_1 \le p_2-\theta_2)dp_2,$$

while the integrals are straight-forward to calculate, the resulting formulae are long and not presented here.

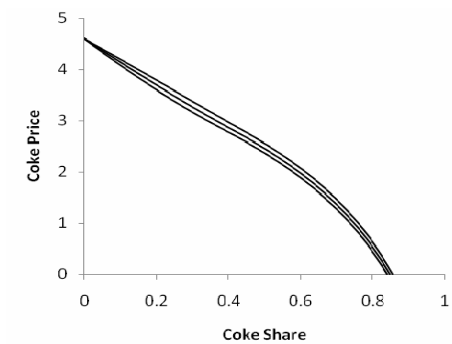
Figure 1

Effect on Demand Curves as the Number of Pepsi Varieties Change

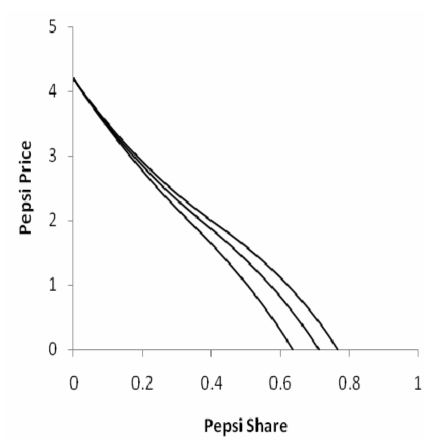
(a) Shift in Pepsi's demand curve as the number of Pepsi varieties increases from $n_2 = 7$ to 9 to 11 (left to right)



(b) Shift in Coke's demand curve as the number of Pepsi varieties increases from $n_2 = 7$ to 9 to 11 (right to left)

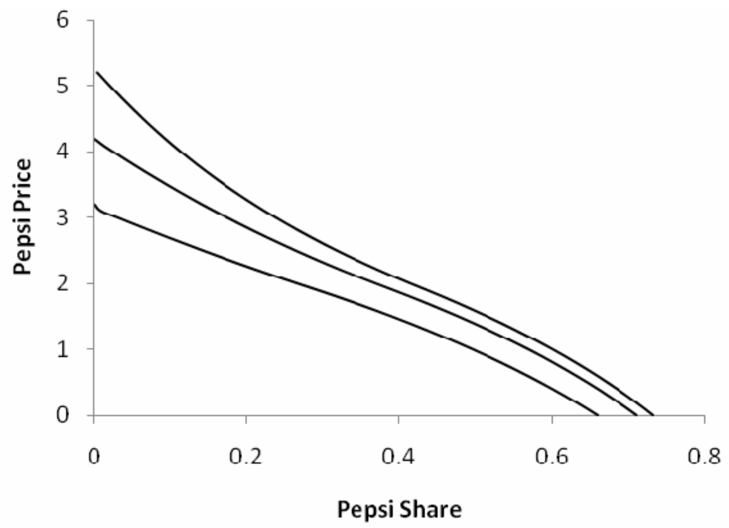


(c) Pepsi demand curves: Cross-partial effect (left: $n_1 = 13$, $n_2 = 11$; middle: $n_1 = 11$, $n_2 = 9$; right: $n_1 = 9$, $n_2 = 7$)



Note: Unless other stated, these simulations are based on $n_1 = 11$, $n_2 = 9$, $p_1 = 2.347$¢ per ounce, $p_2 = 2.363$¢ per ounce, $\theta_1 = 25.6181$, $\theta_2 = 25.2096$, and $\omega = 21.0219$.

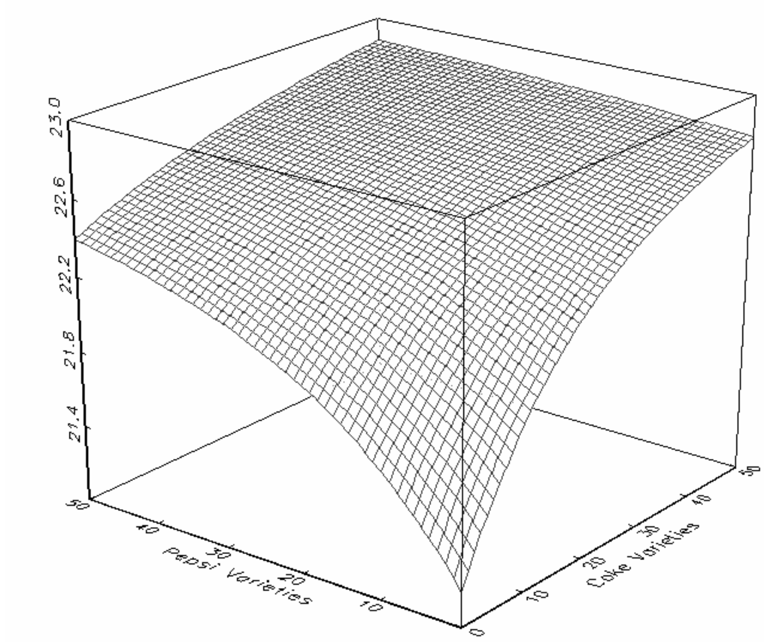
Figure 2
Effect on Pepsi's Demand Curve of a 5% Change in the Net Surplus of the Outside Good



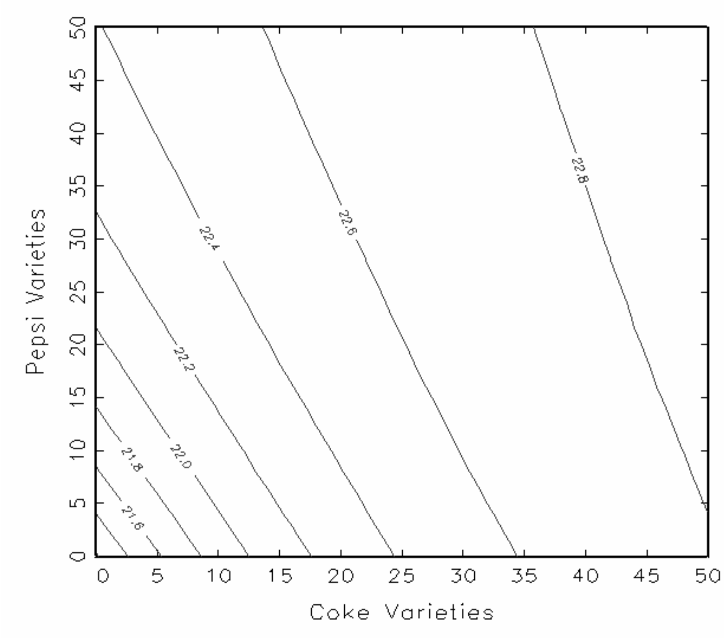
Note: In these simulations, n_1 = 11, n_2 = 9, Pepsi's price = 2.363¢ per ounce, θ_1 = 25.6181, and θ_2 = 25.2096. In the middle demand curve, we use the estimated ω = 21.0219; while ω is 5% larger in the demand curve on the left, and ω is 5% smaller in the demand curve on the right.

Figure 3
Total Consumer Surplus as a Function of the Number of Coke and Pepsi Varieties

(a) Consumer surplus as a function of the number of Coke and Pepsi varieties from 0 to 50



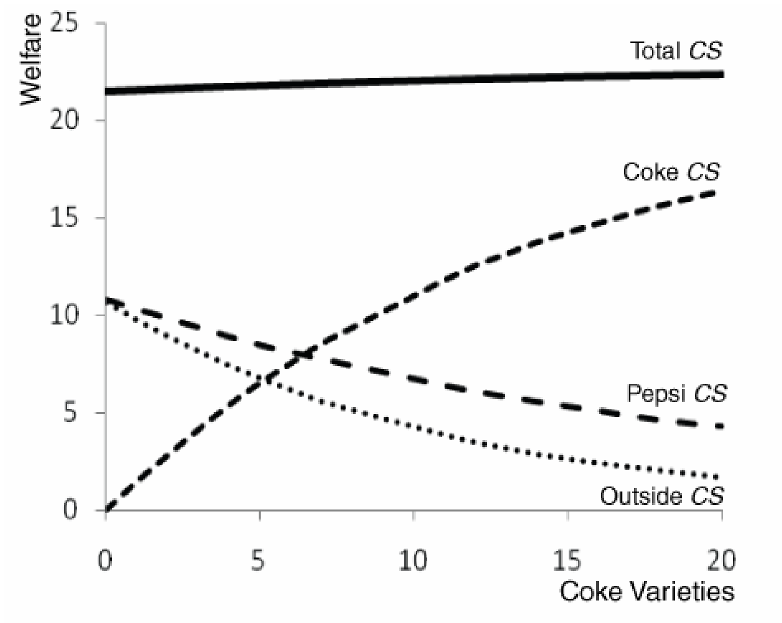
(b) Iso-welfare curves



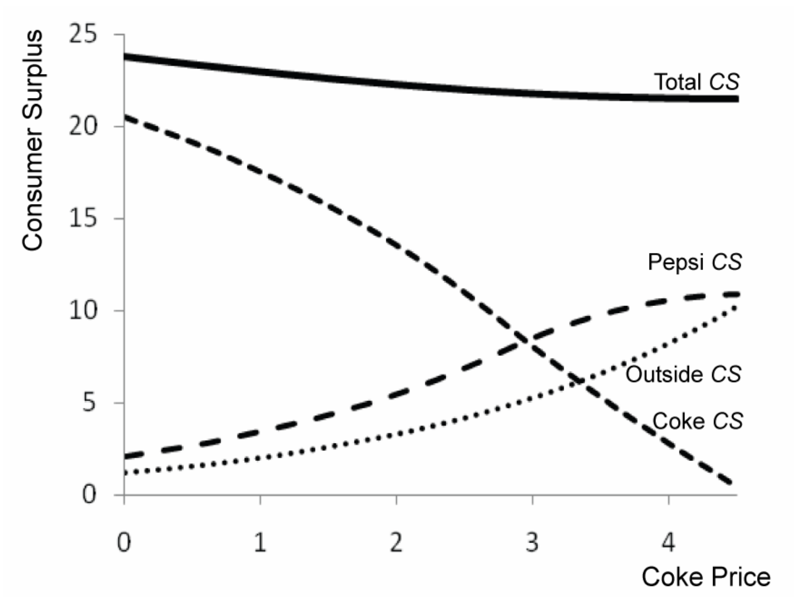
Note: These simulations assume that Coke's price = 2.347¢ per ounce, Pepsi's price = 2.363¢ per ounce, θ_1 = 25.6181, and θ_2 = 25.2096, ω = 21.0219. When Coke and Pepsi varieties are 0, the total consumer surplus equals ω = 21.0219.

Figure 4
Variation in Consumer Surplus with the Number of Coke Varieties or Price

(a) Consumer surplus and the number of Coke's varieties



(b) Consumer surplus and Coke's price



Note: These simulations assume that $n_1 = 11$ in (b), $n_2 = 9$, Coke's price = 2.347¢ per ounce in (a), Pepsi's price = 2.363¢ per ounce, $\theta_1 = 25.6181$, $\theta_2 = 25.2096$, and $\omega = 21.0219$.

Table 1
Profit Maximization vs. Social Optima

		Planner Sets		
	Profit Max	*Prices & Varieties*	*Varieties Only*	*Prices Only*
n_1	11	26	27	11
n_2	9	11	11	9
p_1	2.35¢	1¢	2.35¢	1¢
p_2	2.36¢	1¢	2.36¢	1¢
p_ω	2.21¢	1.5¢	2.21¢	1.5¢
Profit	$2,447	-$797	$2,086	-$273
Variety Cost	$273	$797	$840	$273
CS	$48,729	$52,662	$49,719	$51,510
Welfare	$51,176	$51,825	$51,805	$51,237

Notes: We set Z = 220581.7, m = 1¢, m_ω = 1.5¢, so that c = 666.89¢, k = 69.05¢