ESTIMATING PROMOTIONAL EFFECTS
WITH RETAILER-LEVEL SCANNER DATA

Steven Tenn*

September 23, 2003

* Federal Trade Commission, Bureau of Economics. E-mail: stenn@ftc.gov. Telephone: (202) 326-3243. Fax: (202) 326-3443. Address: 600 Pennsylvania Ave NW, Washington, DC 20580.
Helpful comments from Hajime Hadeishi, Daniel Hosken, David Schmidt, Shawn Ulrick, and John Yun are greatly appreciated. The views expressed in this paper are those of the author, and do not necessarily represent the views of the Federal Trade Commission or any individual Commissioner.

ESTIMATING PROMOTIONAL EFFECTS

WITH RETAILER-LEVEL SCANNER DATA

Abstract

Estimating cross-brand promotional effects with aggregate data requires knowledge of the joint distribution of each brand's promotions. While such information is available in store-level scanner data, it is not included in more aggregated scanner datasets. This paper presents a technique for overcoming this difficulty and develops a retailer-level model that incorporates both own- and cross-brand promotions. Promotional activity is integrated into the specification in a manner consistent with the way store-level models control for promotions, thereby avoiding the problem of aggregation bias. The proposed methodology extends the usefulness of retailer-level scanner data by allowing it to answer important questions regarding how the promotions of competing products interact.

I. Introduction

The use of scanner data for analyzing the impact of promotional activity is widely seen as a "success story" by both academics and industry participants (Bucklin and Gupta 1999). A growing literature, however, questions whether aggregate-level scanner data can be reliably used for this purpose, or if store-level data is required instead (Allenby and Rossi 1991, Christen et al. 1997, Chung and Kaiser 2000, Link 1995). A drawback of aggregate data is that stores with differing levels of promotional activity are combined together. This promotional heterogeneity raises two econometric issues. First, estimation of cross-brand promotional effects requires knowledge of the joint distribution of each brand's promotions across stores, yet aggregate-level scanner data does not contain such information. Second, previous research demonstrates that demand estimation using aggregate data often leads to model misspecification known as "aggregation bias." We develop a demand model that overcomes these two issues, allowing for consistent estimation of both own- and cross-brand promotional effects with retailer-level scanner data (e.g., the Jewel supermarket chain in Chicago).[1]

The challenge of estimating cross-brand effects with aggregate data is easily demonstrated through a simple example. Suppose the manager of Brand A wants to estimate how a promotion by Brand B impacts his sales when Brand A is also on promotion. Aggregate scanner data only reports information regarding each brand's own promotional activity (Link 1995, Christen et al. 1997). While such data might indicate that Brands A and B are each on promotion in 50% of stores, it does not report the fraction of stores where Brands A and B are both on promotion. It could be the case that the two brands are never promoted at the same time in any given store. Alternatively, both brands may be on promotion in 50% of stores, with neither being promoted in the other half. Aggregate-level scanner data does not distinguish

[1] A confidentiality agreement with AC Nielsen prohibits retailer names from being revealed. This example does not indicate whether the dataset employed includes the Jewel supermarket chain in Chicago.

between these (or other) possibilities. Without information on the joint distribution of the two brands' promotional activity, it is impossible to analyze how their promotions affect each other.

In this paper, we develop an econometric model that provides consistent estimates of both own- and cross-brand promotions when used in conjunction with retailer-level scanner data. First, a store-level model is constructed that controls for the promotional activity of each brand. Product demand at each store is then added together to obtain retailer-level demand in a given city. Aggregation to the retailer-level is made possible by the development of a technique for estimating the joint distribution of each brand's promotions across stores. This is accomplished using information regarding the univariate distribution of each brand's promotions that is typically included in scanner datasets. By explicitly modeling store-level promotional heterogeneity, and then incorporating such activity into the aggregate demand specification in an internally consistent manner, aggregation bias is avoided for both own- and cross-brand promotional effects.

This methodological approach requires several modeling assumptions. To facilitate aggregation, it is assumed that all stores within a retail chain that have the same promotional activity for a given product also charge the same price. While price homogeneity clearly does not hold at the market level, for many products it will (approximately) hold for stores within the same retail chain in a given city.[2] This motivates the paper's focus on retailer-level scanner data. Additionally, it is assumed that the model parameters are homogeneous across stores. This is required since retailer-level scanner datasets do not include any information on particular stores, which makes it difficult to meaningfully incorporate parameter heterogeneity into the model. When these two conditions are satisfied, we demonstrate that retailer-level scanner datasets contain sufficient information to consistently estimate the effect of own- and cross-brand promotional activity.

[2] For example, Chevalier et al. (2003) conclude that there is relatively little inter-store price heterogeneity despite the presence of multiple price zones in their data.

Many important questions relate to how the promotions of competing products interact (Blattberg et al. 1995, Bucklin and Gupta 1999): How do the promotions of other brands impact one's own sales? Do cross-brand promotions enhance or diminish the promotional activity of your brand? Does this effect depend on the particular type of promotion? The methodology developed in this paper greatly enhances the usefulness of retailer-level data by allowing these issues to be addressed. The paper concludes with an empirical analysis demonstrating how retailer-level data can provide significant insight into cross-brand promotional effects. For example, we show that for two brands of premium ice-cream, each brand's promotional activity greatly diminishes the impact of the other's promotions. This likely explains why supermarkets rarely promote the two brands at the same time.

The paper is organized as follows. Section two details the reasons why aggregation bias arises in demand estimation. A demand model that is consistently estimated with retailer-level data is presented in section three. Section four uses the model to estimate the impact of promotional activity for two brands of premium ice-cream. Section five concludes.

II. Aggregation Bias

A common problem when estimating demand using aggregate data is that model misspecification leads to "aggregation bias." This section details the primary reasons why aggregate-level data often produces biased estimates. For a detailed consideration of these issues, Theil (1954) and Krishnamurthi et al. (1990) analyze the linear model; Lewbel (1992), Christen et al. (1997), and Chung and Kaiser (2000) analyze the constant elasticity model; and Allenby and Rossi (1991) and Krishnamurthi et al. (2000) analyze the Logit model. Stoker (1993) provides a general review of empirical approaches to the data aggregation problem.

Aggregation bias arises when the wrong explanatory variables are used to estimate product demand. This is easily demonstrated through a simple example. Suppose there are two types of stores that differ solely by the price each charges for a given product. In each time period *t*, let a fraction $\varphi \in (0,1)$ of retailers charge price $p_{1,t}$, with the remaining stores charging price $p_{2,t}$. All stores have the same demand curve: $q = \mu + \alpha p$. Suppose that the only sales data available is aggregated across both types of stores. Adding up product demand implies that the average quantity sold per store is $\overline{q}_t = \mu + \alpha\overline{p}_t$, where $\overline{p}_t = \varphi p_{1,t} + (1-\varphi)p_{2,t}$. Note that aggregate demand is a function of the average price across stores. Commonly available aggregate datasets, such as market-level scanner data, do not report this measure of price. Rather, average price is typically calculated as total dollar sales divided by total quantity sold, i.e., the volume-weighted average price. The use of this price measure to estimate aggregate demand leads to model misspecification. One cannot reliably estimate the model using volume-weighted average price simply because product demand is not determined by this variable. For this particular example, the estimated $\hat{\alpha}$ obtained using volume-weighted average price can be biased in either direction depending on the model calibration.

Linear demand is a special case where store and aggregate-level demand have the same functional form. For non-linear models, this will generally not be true. Consider the previous example, but with a constant elasticity store-level demand curve: $\ln q = \mu + \alpha \ln p$. Average store sales takes a different form: $\ln \overline{q}_t = \ln(\varphi e^{\mu+\alpha \ln p_{1,t}} + (1-\varphi)e^{\mu+\alpha \ln p_{2,t}})$. Further, aggregate demand is not a function of either store-weighted or volume-weighted prices. To properly estimate this demand specification, one must separately obtain the price in each type of store, and one must observe the fraction of stores that are of each type (φ and $1-\varphi$). Further, the model must be estimated using the aggregate-level demand specification implied by the store-level model, rather than with a constant elasticity demand model. Otherwise, the aggregate demand model will be incorrectly specified, and aggregation bias will likely occur.

An additional cause of aggregation bias is parameter heterogeneity across stores. Suppose that each type of store faces a distinct demand curve. For any given price *p*, let demand in the first type of store be determined by $q_1 = \mu_1 + \alpha_1 p$. Product demand in the second type of store is $q_2 = \mu_2 + \alpha_2 p$. Unit sales per store is equal to $\overline{q}_t = \overline{\mu} + \overline{\alpha}\,\overline{p}_t + \varphi(\alpha_1 - \overline{\alpha})p_{1,t} + (1-\varphi)(\alpha_2 - \overline{\alpha})p_{2,t}$, where $\overline{\mu} = \varphi\mu_1 + (1-\varphi)\mu_2$ and $\overline{\alpha} = \varphi\alpha_1 + (1-\varphi)\alpha_2$. This demand function takes a similar form to store-level demand, but contains two additional terms. If one estimates the linear model $\bar{q}_t = \bar{\mu} + \bar{\alpha}\,\bar{p}_t$, the obtained estimate $\hat{\bar{\alpha}}$ will be a biased relative to $\bar{\alpha}$ unless the last two terms of the aggregate-level model are jointly uncorrelated with $\bar{p}_t$. A special case where this occurs is when $\varphi = 1/2$ and prices are independently and identically distributed across each type of store.[3] In general, however, parameter heterogeneity will result in aggregation bias. This finding is not an artifact of linear demand; aggregation bias similarly arises when the store-level model is non-linear.

To summarize, one can obtain aggregate-level demand estimates that are untainted by aggregation bias under the following conditions. One must estimate aggregate demand using a functional form that is consistent with store-level demand. One must estimate the model using the correct variables, where the variables of the aggregate model may not be simple averages of their store-level counterparts. Lastly, one must take into account that the aggregate-level model may be a complex function of the store-level model's parameters.

Researchers have concluded that aggregation bias is less problematic when data is aggregated across stores with homogeneous marketing activity (Christen et al. 1997). The above discussion provides the intuition for this conclusion. First, homogeneity often implies that aggregate demand has the same functional form as store-level demand. For instance, if store-level demand is of the linear or constant elasticity form, then aggregate demand has the same specification when stores are identical (and face the same demand curve). Second, aggregate demand is often a function of the same variables (price, etc.) as store-level demand, simply because there is little or no variation in these variables across stores. This leaves store-level parameter heterogeneity as the leading determinant of aggregation bias. Store homogeneity in marketing activity does not necessary imply homogeneity in the demand-side response to

[3] These requirements may seem overly strict to readers familiar with Christen et al. (1997), which cites Krishnamurthi et al. (1990) as showing that a "bias exists only if there is (1) heterogeneity in both the parameters and the independent variables, and (2) a non-zero covariance between these two model elements." The non-zero covariance requirement must hold in finite sample for each observation across the units of aggregation. This condition is violated when the control variables and parameters are drawn from distributions that are independent of each other.

promotions (Hoch et al. 1995). Therefore, parameter heterogeneity can potentially lead to bias (relative to the average parameter values) even when data is aggregated across stores that undertake identical promotion and pricing decisions. The degree of bias depends on the model employed. Allenby and Rossi (1991) find that in the Logit model, parameter heterogeneity does not lead to significant aggregation bias under certain conditions, one of which is homogeneity in marketing activity. Krishnamurthi et al. (2000) conclude, however, that this finding is not robust to time-series variation in parameter heterogeneity.

Link (1995) suggests that aggregation bias be avoided by employing data that has been aggregated across stores with homogeneous marketing activity. Doing so does not eliminate aggregation bias due to parameter heterogeneity. Link argues, however, that data aggregation across stores with heterogeneous marketing activity is the most significant source of bias in practical applications. However, even if one obtains data that is aggregated across stores where a brand's own promotions are homogenous, heterogeneity in the promotions of competing products may still remain. Thus, Link's approach does not account for aggregation bias in cross-brand effects.

Christen et al. (1997) propose a methodology to "de-bias" demand estimates based on aggregate data. First, demand is estimated using simulated store-level data that has been aggregated across stores. The average difference between the true and estimated parameters from the simulation is then added to the estimates from an empirical application, so as to de-bias the results. A shortcoming of this approach is that it can be difficult to reliably estimate the magnitude of aggregation bias, as one may have insufficient information to calibrate the simulated data to the actual data. This is particularly problematic for cross-brand effects. Calibration of the simulated data requires knowledge of the joint distribution of each brand's promotions across stores, information which is not contained in aggregate-level scanner datasets. The de-biasing procedure may not entirely eliminate aggregation bias, and if done poorly, could exacerbate the problem.

A number of authors have analyzed the impact of aggregation bias in the following manner (Allenby and Rossi 1991, Christen et al. 1997, Chung and Kaiser 2000, Krishnamurthi et al. 2000). First, a store- or consumer-level model of demand is posited to be the true demand specification. The aggregate-level model is then obtained by adding up demand across each store (or individual). This specification is then compared to an alternative aggregate-level model to assess whether the latter is biased relative to actual aggregate demand. A commonly employed choice for the alternative model is the store-level model evaluated at the average value of each store-level variable. While this approach has been broadly employed to demonstrate the presence of aggregation bias, it can also be used to directly formulate an aggregate-level model of demand. Rather than addressing whether an alternative model is biased, one can simply estimate the aggregate-level model implied by the store-level model (Stoker 1993). While conceptually straightforward, data limitations complicate the estimation of the aggregate model. The following section presents an aggregate demand model that is based on this approach, and which is estimable with retailer-level scanner data.

III. Econometric Demand Model

This section develops a general demand specification that is consistently estimated with city specific, retailer-level data. First, a store-level model of demand is presented. This framework is then aggregated to the retailer-level, so that it can be feasibly estimated.

Store-Level Model

Each store *s* within a retail chain sells a set of brands *B*, with the total measure of stores normalized to one.[4] Denote the marketing activity of brand *i*, in store *s*, at time *t,* by $m_{ist} \in M$. Typical examples of marketing activity are “No Promotion,” “Feature,” “Display,” and “Feature

[4] The model assumes brands are composed of a single product. This is relaxed in the empirical example presented in section four, which controls for the number of different products contained within each brand (e.g., distinct flavors).

& Display."[5] The set consisting of each brand's marketing activity is denoted by $g_{st} = \{m_{ist}\}_{i \in B}$, with $g_{st} \in G$. For example, if M contains four types of marketing activity and B contains two brands, then G consists of 16 pairs of marketing activity. Similarly, the set of prices across all brands is defined as $p_{st} = \{p_{ist}\}_{i \in B}$. Store-level unit sales q_{ist} is determined by demand specification d, which is a function of price p_{st}, promotional activity g_{st}, and a vector δ_i of demand parameters:[6]

(3.1) $$q_{ist} = d(p_{st}, g_{st}; \delta_i).$$

Retailer-Level Model

Without store-level data, one cannot directly estimate model (3.1). Rather, one must aggregate the model to the level of the available data. For each $g \in G$, let π_{gt} denote the fraction of stores with said marketing activity. To facilitate aggregation, assume that an identical price is charged at all stores with the same promotional activity for a given brand:

(3.2) $$p_{ist} = p_{it}^{m}, \forall s : m_{ist} = m.$$

Note that price homogeneity is only required across stores *within* a given retail chain. Price homogeneity across retail chains is not a model requirement.

Denote the vector of cross-brand prices that corresponds to each $g \in G$ by p_t^g. Further, let $m_i(g)$ be the element of g corresponding to brand i. Model (3.1) implies that total unit sales for brand i, across stores with marketing activity m, can be written as follows:

(3.3) $$\bar{q}_{it}^{m} = \sum_{g \in G : m_i(g) = m} \pi_{gt} \; d(p_t^g, g; \delta_i).$$

[5] A "Feature" is typically defined as an advertisement in a promotional circular. A "Display" is a secondary sales location within a store which is used to draw special attention to a given product.

[6] The model can accommodate additional control variables. The empirical application of section four provides such an example.

Model (3.3) defines average quantity sold $\overline{q}_{it}^{m}$ as a function of each brand's price and marketing activity. Econometric models of demand account for deviations from expected sales through the use of an error term. A straightforward approach is to assume that the model can be written as follows, for a given function *f*, where q_{it}^{m} denotes actual quantity sold:

(3.4) $f(q_{it}^{m}) = f(\overline{q}_{it}^{m}) + \varepsilon_{it}^{m}$.

Although a standard assumption for the error term is $\varepsilon_{it}^{m} \sim N(0, \sigma_i^2)$, one can estimate the model using any distribution. If store-level demand is linear, then an obvious choice for *f* is the identity function $(f(x) = x)$. When store-level demand takes the constant elasticity form, one would likely choose *f* to be the natural log function $(f(x) = \ln x)$. Under these specifications for *f*, the error term enters model (3.4) in an analogous manner to the error term in the linear and constant elasticity specifications.

Homogeneity Assumptions

A key model requirement is pricing assumption (3.2). The empirical validity of price homogeneity, conditional on a given level of marketing activity, will depend on the application at hand. A potential violation is when a retailer employs more than one price zone within a city. For example, Hoch et al. (1995) report that the Dominick's supermarket chain in Chicago uses multiple price zones.

The model also assumes that stores have the same demand specification. Since a store can be defined in terms of "store-equivalent units," stores are allowed to differ in size so long as demand for all products uniformly scales up or down. For instance, if a store sells twice as much of each brand as the others, it can simply be viewed as two separate stores without violating the model assumptions.

Whether one can reasonably assume parameter homogeneity across stores is potentially problematic (Boatwright et al. 2001). However, research concerning the use of pooled datasets concludes that a violation of parameter homogeneity is not overly troublesome so long as the condition is close to being satisfied (Wallace 1972, Bass and Wittink 1975).

Data Requirements

The data used to estimate model (3.3) must meet several key requirements. Importantly, the dataset must separately provide price and unit sales for each brand by each type of marketing activity. Retailer-level scanner datasets typically report unit and dollar sales by promotion, where the set of marketing activity *M* contains four elements: "No Promotion," "Feature," "Display," and "Feature & Display." One can calculate price p_{it}^{m} as dollar sales for promotion *m* divided by unit sales. Note that equation (3.2) implies that the use of volume-weighted average price will not result in specification error, since prices are homogeneous across stores with a given level of marketing activity.

Through a variable known as "All Commodity Volume," or ACV, scanner datasets include an empirical proxy for the fraction of stores in which a brand has a given level of promotional activity. A product's ACV is the percentage of total sales, across all product categories, which are accounted for by the stores that carry that product. Similarly, the ACV for a given level of promotion is the fraction of all category sales that are accounted for by those stores where the brand has said promotional activity. The ACV for each level of promotion can be used as an empirical counterpart to π_{it}^{m}, the fraction of stores in which a brand has marketing activity *m*. However, scanner datasets do not directly report $\{\pi_{gt}\}_{g \in G}$, which is the joint distribution of each brand's marketing activity across stores. This data deficiency can be overcome by postulating that the joint distribution of marketing activity is a function *h* of the distribution of each brand's promotions:

$$(3.5) \qquad \{\pi_{gt}\}_{g \in G} = h(\{\pi_{it}^{m}\}_{\substack{i \in B \\ m \in M}} ; \theta).$$

Parameter θ is included in the model specification, and allows the joint distribution of marketing activity to be flexibly estimated. Alternatively, one can assume that each brand's promotions are independently distributed. Doing so eliminates the need for θ, since $\pi_{gt} = \prod_{i \in B} \pi_{it}^{m_i(g)}$ does not require any parameterization.

<u>Model Estimation</u>

The specification detailed above provides sufficient structure to estimate the model with retailer-level scanner data. Given a store-level demand function *d*, a function *h* that calculates the joint distribution of each brand's promotions across stores, and a function *f* that details how the econometric error term enters the demand equation, one can estimate the model parameters $\{\{\delta_i, \sigma_i^2\}_{i \in B}, \theta\}$ via maximum likelihood. An important property of maximum likelihood is that under standard conditions, it provides a consistent estimate of the model parameters (Greene 1997). An example of such estimation is provided in the following section.

IV. Empirical Application

The econometric model developed in the previous section is used to estimate the demand for premium ice-cream. The dataset employed is described below. Details concerning functional form assumptions and estimation technique are then provided. This is followed by a presentation of the results.

<u>Data Description</u>

The analysis utilizes retailer-level supermarket scanner data corresponding to AC Nielsen's "premium ice-cream" category. The dataset includes two of the five largest brands of premium ice-cream. Sales data for each brand is separately reported for four types of promotional activity: "No Promotion," "Feature," "Display," and "Feature & Display." A confidentiality agreement with AC Nielsen prevents brand or retailer names from being revealed. Henceforth, the two brands are referred to as "Brand A" and "Brand B." The dataset includes weekly scanner data for ten retailer-city combinations, covering the period December 1998 through June 2001 (132 weeks). Since premium ice-cream is predominantly sold in a half-gallon container, the dataset is restricted to items of that size. This limits potential bias when calculating the average price across SKUs, since each brand's ice-cream flavors for a given container size are line-priced.

Model Specification

A widely employed specification for store-level demand is the model presented in Wittink et al. (1988), where q_{ist} is the unit sales, p_{ist} is the price, and m_{ist} is the marketing activity of brand *i*, and where Z_{it}^{m} is a set of additional control variables:[7]

(4.1) $$\ln q_{ist} = \mu_i + \sum_{j \in B} \alpha_{ij} \ln p_{jst} + \sum_{j \in B} \sum_{m \in M} \gamma_{ij} 1_{(m_{jst} = m)} + Z_{it}^{m_{ist}} \beta_i .$$

This specification is relatively inflexible in how promotional activity affects unit sales. Promotions impact the intercept of the demand curve, but do not alter own- and cross-price elasticities.

An alternative approach is to let the model parameters vary by the promotional activity of each brand. As before, $g_{st} \in G$ denotes the set consisting of each brand's marketing activity. Since there are two brands and four distinct types of marketing activity *m*, the set *G* contains 16 pairs of marketing activity. A generalization of (4.1) is the following specification:

(4.2) $$\ln q_{ist} = \mu_i^{g_{st}} + \sum_{j \in B} \alpha_{ij}^{g_{st}} \ln p_{jst} + Z_{it}^{m_{ist}} \beta_i .$$

Each $g \in G$ has its own set of model parameters $\{\mu_i^g, \alpha_{ij}^g\}$. This allows promotions to impact product demand in a flexible manner. The disadvantage of this specification is that a large number of parameters must be estimated. As is typical, there is a trade-off between model flexibility and model precision (Greene 1997). Too restrictive a model can lead to

[7] In order for the model to aggregate to the retailer-level, the set of controls Z_{it}^{m} cannot vary by individual store *s*.

misspecification and bias, while excessive flexibility leads to parameter estimates that are too imprecise to be informative.

As a balance between these two factors, the empirical analysis employs a model that is less restrictive than (4.1), but has far fewer parameters than (4.2). This is accomplished by employing a set of parameter restrictions:

$$
\begin{aligned}
&\alpha_{ii}^{g} = \alpha_{ii}^{m_i(g)}, \forall g \in G, \forall i \in B \\
(4.3) \qquad &\alpha_{ij}^{g} = \alpha_{ij}\alpha_{jj}^{g}, \forall g \in G, \forall i, j \in B \text{ s.t. } i \neq j .
\end{aligned}
$$

The first restriction requires that own-price elasticities only vary by a brand's promotional activity. This restriction assumes that a brand's own promotions dominate consumer response to changes in its price, rather than the promotions of competing products. The second restriction implies that any two price changes by brand *j* that have the same effect on brand *j*'s sales, also have the same impact on brand *i*'s sales. Suppose a 10% price increase leads to a 20% sales decrease for brand *j* when it is not on promotion, but only a 10% decrease when brand *j* is on "Feature." Since the fraction of consumers that discontinue their purchases in response to the price increase is twice as large when brand *j* is not on promotion, the second restriction implies that the percentage sales increase for brand *i* is also twice as large. Through these restrictions, the model produces relatively precise estimates while still remaining quite flexible.[8]

One must make additional assumptions regarding the specification of the econometric error term. It is assumed that ε_{it}^{m} is i.i.d Normally distributed. The variance of the error term is allowed to differ by a brand's own marketing activity: $\varepsilon_{it}^{m} \sim N(0, \sigma_{im}^{2})$. Additionally, the

[8] An attempt was made to test whether restrictions (4.3) hold in this empirical application. The Likelihood-Ratio test did not reject these restrictions. However, this test is statistically valid only under the assumption of an i.i.d error term. As discussed below, the estimation procedure allows for a general serially correlated error process. One can test the validity of restrictions (4.3) under this error structure by estimating the model without imposing the restrictions, and then employing a Wald test. Doing so requires an estimate of the variance matrix for the parameter estimates. For this empirical application, the estimated standard errors for the unrestricted parameters are implausibly small. This commonly occurs when one employs variance estimates that are only asymptotically valid in situations where certain parameters are identified by a small number of observations, as is the case for the unrestricted model. Since restrictions (4.3) cannot be formally tested, little can be said regarding their validity.

function f that determines how the error term enters the demand specification is chosen to be the natural log function: $f(x) = \ln(x)$. This specification allows the error term to enter the demand specification in a manner analogous to the error term in the constant elasticity model.

The parameterization of the error term assumes that there is no serial correlation. This assumption is a result of the dataset being an unbalanced panel of observations. The number of observations, for a given retailer in a given week, depends on how many different types of promotions have positive sales. This number ranges from between one observation ("No Promotion" only), to up to four observations ("No Promotion," "Feature," "Display," and "Feature & Display"). An unbalanced panel makes it difficult to formulate a sensible characterization of a serially correlated error process. However, as detailed in the following subsection, the standard errors of the parameter estimates are adjusted so as to correct for a general form of serial correlation.

These assumptions lead to the following retailer-level demand specification, subject to the constraints embodied in equation (4.3):

$$\begin{aligned} \ln(q_{it}^{m}) &= \ln\Big(\sum_{g \in G: m_i(g)=m} \Big(\frac{\pi_{gt}}{\pi_{it}^{m}}\Big) e^{(\mu_i^g + \sum_{j \in B} \alpha_{ij}^g \ln p_{jt}^g)}\Big) + \ln(\pi_{it}^{m}) + Z_{it}^{m} \beta_i + \varepsilon_{it}^{m} \\ \text{where } \pi_{it}^{m} &= \sum_{g \in G: m_i(g)=m} \pi_{gt} \text{, and } \varepsilon_{it}^{m} \sim N(0, \sigma_{im}^{2}). \end{aligned} \tag{4.4}$$

The first term is the log of the weighted average of unit sales across each set of promotional activity g. The subsequent term proportionally scales up unit sales by each brand's distribution π_{it}^{m}. The homogeneity assumptions of section three imply that this variable exhibits constant returns to scale, as seen by its unit coefficient. This is relaxed by including $\ln(\pi_{it}^{m})$ as one of the variables in Z_{it}^{m} (see below). One can thereby test whether distribution exhibits constant returns, rather than simply asserting unit elasticity of distribution.

The variable π_{it}^{m} is empirically measured as the fraction of each brand's "Total Distribution Points," or TDP. That is, $\pi_{it}^{m} = \frac{TDP_{it}^{m}}{\sum_{\widetilde{m} \in M} TDP_{it}^{\widetilde{m}}}$. The variable TDP_{it}^{m} is the sum, across all SKUs, of brand *i*'s ACV at time *t* that has promotional activity equal to *m*. TDP is an overall measure of product distribution, incorporating both the breadth and depth of a brand's availability. For example, a TDP of 300% is the distribution equivalent of a brand having three SKUs distributed at every store.

To fully specify the model, one must choose a particular function *h* to use in equation (3.5). Recall that this function specifies the cross-brand distribution of promotions, $\{\pi_{gt}\}_{g \in G}$, as a function of the univariate distribution of promotions for each brand, $\{\pi_{it}^{m}\}_{\substack{i \in B \\ m \in M}}$. One possibility in defining *h* is to assume that promotions are independently distributed across the two brands. To allow for greater flexibility, the following approach is taken instead.[9] Denote the fraction of stores in which brand *i* is being promoted ("Feature," "Display," or "Feature & Display") by $\widetilde{\pi}_{it}$. Since probabilities are restricted to $[0,1]$, the probability that Brand A and Brand B are both on promotion in a given store must be between $[\widetilde{\pi}_{t}^{\min}, \widetilde{\pi}_{t}^{\max}]$, where $\widetilde{\pi}_{t}^{\min} = \max(0, \widetilde{\pi}_{\text{Brand A,t}} + \widetilde{\pi}_{\text{Brand B,t}} - 1)$ and $\widetilde{\pi}_{t}^{\max} = \min(\widetilde{\pi}_{\text{Brand A},t}, \widetilde{\pi}_{\text{Brand B},t})$. One can then specify the probability that both brands are simultaneously on promotion in a given store by $\pi_{t}^{p,p} = \theta \widetilde{\pi}_{t}^{\min} + (1-\theta)\widetilde{\pi}_{t}^{\max}$, where θ is a model parameter that takes a value between zero and one. The probability that Brand A is on promotion, but Brand B is not, is equal to $\pi_{t}^{p,np} = \pi_{\text{Brand A},t}^{p} - \pi_{t}^{p,p}$. Similarly, the probability that only Brand B is on promotion is $\pi_{t}^{np,p} = \pi_{\text{Brand B},t}^{p} - \pi_{t}^{p,p}$. The probability that neither brand is on promotion is given by $\pi_{t}^{np,np} = 1 - \pi_{t}^{p,p} - \pi_{t}^{p,np} - \pi_{t}^{np,p}$. To complete the specification, it is assumed that conditional on whether Brand A and Brand B are each on promotion the particular type of promotion is independently distributed.

[9] Strictly speaking, this approach is not "more general" since the former is not nested within it.

The estimation procedure employs several additional control variables. For each brand, a set of time and retailer-city fixed effects is included in the specification.[10] As mentioned above, $\ln(TDP_{it}^{m})$ is also included as a control variable. This variable is a measure of each brand's retail distribution at those stores where it has promotional activity *m*.

Estimation

All parameters from the eight equations of the model are jointly estimated via maximum likelihood (four types of promotion multiplied by two brands). As detailed above, the error structure assumes that ε_{it}^{m} is i.i.d Normally distributed. After obtaining the maximum likelihood estimates through the usual means, possible correlation between error terms is accounted for when the variance matrix of the parameter estimates is calculated. This adjustment is known as "Quasi Maximum Likelihood Estimation," or QMLE (White 1982, Hamilton 1994).[11] A Newey-West (1987) assumption on the error structure is employed, with a maximum lag difference of four weeks. That is, any two residuals that are from time periods within four weeks of each other are allowed to be arbitrarily correlated.[12] Robustness checks indicated that the estimated standard errors are not sensitive to the particular choice of how many time lags are allowed for. By constructing the variance matrix in this manner, robust standard errors are obtained without having to specify a particular process for the model's error structure.

[10] The inclusion of time fixed effects eliminates seasonal variation from the data. Due to a lack of instruments for price, it is important to do so to remove endogenous price variation related to seasonal demand changes.

[11] This is a special case of "Generalized Method of Moments," or GMM, estimation (Hansen 1982). The first order conditions from the log-likelihood function are used as moment conditions.

[12] The model "residual" is the gradient of the portion of the log-likelihood function that corresponds to a given observation.

Results

The remainder of this section details the results from estimating the model. To introduce the data, Table 1 presents the fraction of unit sales, revenue, and "Total Distribution Points" (TDP) that is accounted for by each type of promotion.[13] These percentages are calculated from variable totals for each retailer-city combination. The reported statistics are un-weighted averages across the 10 retailer-city combinations.

Promotions clearly play a significant role in the marketing of both brands. Approximately one half of unit sales are sold on promotion, with a "Feature" being the most common type of promotion. Unit sales are high, relative to distribution, for each level of promotional activity other than "No Promotion." This is potentially due to several distinct effects. First, promotions lead to an upward shift in the demand curve for a given brand. Promotional activity in the ice-cream category is also typically associated with a price reduction, which itself leads to increased sales (i.e., movement down the demand curve). In addition, promotions most often occur in the summertime, which is when demand for ice-cream is highest. See Blattberg et al. (1995) for a review of the empirical literature that considers the effect of promotions.

The price distribution for each type of promotional activity is reported in Table 2. Each retailer-city / week combination is given equal weight when constructing this distribution. For each retailer, all prices are re-scaled so that the modal price when not on promotion equals \$1.00. This transformation is used throughout this section to disguise the identity of each brand. On average, each brand's price is substantially lower when on promotion, particularly when on "Feature" or "Feature & Display." It is quite common for a promotional price cut to be 50% or more. Each brand is sometimes promoted with only a small price reduction, however. Thus, there is substantial price heterogeneity within each type of promotional activity. This allows the

[13] A confidentiality agreement with AC Nielsen prevents other summary statistics from being presented.

impact of promotions that are supported by a major price reduction to be separately identified from the impact of promotions that are not (see below).

Table 3 reports summary statistics from the estimation procedure. For both brands, the variance of log unit sales is highest when each brand is on “Display.” This potentially indicates that there is substantial heterogeneity in the quality of displays. While an end-of-aisle display may have a significant impact on sales, a display in a less attractive location likely has a much smaller effect. The low variance of log unit sales when on “Feature & Display” suggests that featured displays are uniformly high quality. This is consistent with the expectation that heavily promoted products receive a premier location within a store.

The model provides an estimate of the joint distribution of promotions across the two brands, which is presented in Table 4A. On average, neither brand is on promotion in 62% of stores. It is relatively common for one brand to be on “Feature” when the other brand is not being promoted. Rarely, however, are both brands simultaneously on promotion in a given store. For each month, Table 4B reports the fraction of stores that simultaneously promote both brands. As expected, retailers are relatively likely to do so in the summertime. In particular, simultaneous promotions are more likely to occur in August than in any other month.

Table 5 reports the impact of each brand’s distribution level (TDP). The estimates indicate that distribution has a highly significant effect on unit sales. For Brand A, a 1% increase in distribution leads to a .99% increase in sales (se=.03%). That is, distribution is (nearly) constant returns to scale for this brand. Although statistically distinct from one at conventional levels, the distribution elasticity of Brand B is close to constant returns to scale. A 1% distribution increase for Brand B leads to a .90% increase in unit sales (se=.04%). Note that each brand’s level of distribution is determined by the number of flavors that a supermarket carries for that brand. These results indicate that expanded distribution of an ice-cream flavor leads to a significant increase in brand sales, rather than simply a cannibalization of a brand’s other SKUs. This is consistent with previous research that documents the sales expansion effects of line-extensions (Reddy and Holak 1994, Lomax and McWilliam 2001).

Next, consider the impact of each brand's promotional activity on its own sales. To separate the direct effect of promotional activity from that of a promotional price decrease, Table 6 presents the impact of a promotion that is not accompanied by any price change. Parameter estimates from the demand model are used to predict the percentage change in unit sales from a promotion, relative to "No Promotion," while holding price constant at $.80. This is at the lower end of the price distribution when each brand is not on promotion, and is at the upper end of the distribution when each brand is being promoted (see Table 2).

Promotions generally have a significant impact on each brand's sales, even when not accompanied by a price reduction. The exception is when Brand B is on "Display," which has a statistically insignificant effect of only -1.2% (se=15.0%). This contrasts with a 52% sales increase when Brand A is on "Display" (se=10.9%). Brand A's price on "Display" is typically one half of its non-promoted price. Brand B is most commonly displayed without any price reduction (see Table 2). Since they are more often associated with significant price cuts, Brand A's displays may be of higher quality (e.g., end-of-aisle). This potentially explains why the impact of a display for Brand B is much smaller.

The parameter estimates from the demand model are used to predict the impact of a competing brand having a "Feature," relative to not having a promotion.[14] As before, each brand's price is held constant at $.80 when making these calculations. The results shown in Table 7 indicate that a "Feature" by one brand has little effect on the sales of the other when it is not on promotion. For example, when Brand A is not on promotion, its sales are reduced by only 2% (se=4.7%) when Brand B changes from "No Promotion" to being on "Feature." When both brands are simultaneously on promotion, however, each brand's promotions generally have a negative impact on the other. When Brand A is on "Feature", its sales fall by 36.9% (se=12.4%) when Brand B's promotional activity changes from "No Promotion" to "Feature." These results

[14] The cross-brand impact of other types of promotions is imprecisely estimated. This is a result of "Display" and "Feature & Display" occurring far less frequently than a "Feature."

indicate that although a brand's own promotional activity draws in additional consumers, a competing promotion takes away much of the increase.[15] This likely explains why retailers rarely promote Brand A and Brand B at the same time (see Table 4). Note, however, that not all of the cross-brand promotional effects are statistically significant at conventional levels.

We know from Table 6 that promotions have a major impact on sales even when not accompanied by a price reduction. Table 8 shows that price reductions have a highly significant impact as well. For both brands, the own-price elasticity when not on promotion is approximately -2. Each brand's own-price elasticity on "Display" is a less elastic -1.4. The results for "Feature" and "Feature & Display" differ by brand. Brand B has elastic demand during these promotions (-2.1 and -2.3, respectively), while Brand A's own-price elasticity is significantly smaller in magnitude (-1.5 and -1.3, respectively). Previous research offers mixed evidence concerning the relative magnitudes of promotional and non-promotional price elasticities (Guadagni and Little 1983, Lattin and Bucklin 1989, Mulhern and Leone 1991). Understanding their general relationship is an area that deserves additional attention (Blattberg et al. 1995, Bucklin and Gupta 1999).

Table 9 presents the effect of a price reduction by the competing brand. The results reveal substantial cross-price elasticities, with a price reduction in one brand leading to a substantial sales decline for the other. The cross-price elasticities vary between .17 and .43, depending on the competing brand's promotional activity. This corroborates the findings presented in Table 7, which indicate that Brand A and Brand B are substitutes for each other.

Tables 10 and 11 summarize each brand's demand curve as a function of its price and promotional activity. Each brand's demand curve is constructed under the assumption that the other brand is not on promotion. The demand curve for each level of promotional activity is shown for prices between the 10th and 90th percentile of the distribution for each type of

---

[15] See Van Heerde et al. (2003) for a decomposition of promotional effects according to cross-brand substitution, cross-period substitution, and category expansion effects.

promotion (see Table 2). For confidentially reasons, unit sales are normalized to one when each brand is not on promotion and has a price of \$1.00. These tables clearly demonstrate that promotions and price reductions both lead to significantly higher unit sales. For each brand, the combination of a 50% price cut and a “Feature & Display” yields sales that are more than six times typical sales when not on promotion.

V. Conclusion

Estimating cross-brand promotional effects requires knowledge of the joint distribution of each brand’s promotions, but this information is not reported in aggregate-level scanner data. We develop a model that overcomes this difficulty under certain conditions, namely parameter homogeneity and price homogeneity conditional on a given level of promotional activity. Two main factors make this possible. First, the model is consistent with adding up from store-level demand. This avoids misspecification due to estimating the “wrong” model. Additionally, the proposed framework is described in terms of store-level promotional activity. This is facilitated by the development of a technique for estimating the joint distribution of each brand’s promotions across stores. These factors allow an internally consistent demand model to be estimated with retailer-level scanner data. The model’s usefulness lies in the wide availability of such data. When store-level data is unavailable, or too costly to obtain, it provides a practical methodology for estimating both own- and cross-brand promotional effects with retailer-level scanner data.

A major benefit of disaggregate data is its ability to provide insight into how the promotional activity of one brand affects another. When store-level data is available, it is straightforward to estimate how different combinations of promotions for a given set of brands impact consumer demand. Aggregate-level scanner datasets do not contain information regarding the joint distribution of promotions, which has hereto precluded estimating cross-brand effects with such data. This paper illustrates how to use retailer-level scanner data to estimate not only the joint distribution of promotional activity, but also how the distribution of promotions impacts consumer demand for each product. The presented empirical application demonstrates how aggregate data can provide useful insights in this respect. For example, the results indicate that the impact of a brand's promotions on a competing product is far greater when that product is also on promotion.

Estimation of the model with market-level scanner data violates a key model requirement that price be homogenous conditional on a given level of promotional activity. Given the lack of alternative means for estimating both own- and cross-brand promotional effects with market-level data, however, application of the model may prove to be useful. Since price heterogeneity is the only assumption that is clearly violated by market-level data, the model should be able to provide (nearly) unbiased estimates of promotional effects when the use of volume-weighted average price does not significantly contaminate the estimates for the other variables. A promising area for future research is to analyze under what circumstances the model provides accurate results when estimated with market-level data.

Table 1

Summary Statistics

Brand A

	No Promotion	Feature	Display	Feature & Display
% of Unit Sales	57.1%	34.9%	1.9%	6.1%
% of Revenue	67.5%	26.5%	1.5%	4.4%
% of TDP	81.3%	16.6%	0.7%	1.5%

Brand B

	No Promotion	Feature	Display	Feature & Display
% of Unit Sales	46.6%	43.0%	0.9%	9.4%
% of Revenue	56.3%	35.3%	0.9%	7.5%
% of TDP	74.5%	22.1%	0.5%	2.8%

Notes: N=1,320, corresponding to a panel of 10 retailer-city combinations and 132 weeks. Percentages are calculated from variable totals for each retailer-city combination. The reported statistics are un-weighted averages across the 10 retailer-city combinations. Each row sums to 100%.

Table 2

Price Distribution by Promotional Activity

Brand A

	No Promotion	Feature	Display	Feature & Display
Average	**0.95**	**0.63**	**0.69**	**0.57**
Mode	**1.00**	**0.50**	**0.50**	**0.50**
Minimum	**0.48**	**0.30**	**0.22**	**0.30**
10th Percentile	**0.72**	**0.49**	**0.50**	**0.47**
25th Percentile	**0.95**	**0.50**	**0.50**	**0.50**
50th Percentile	**1.00**	**0.60**	**0.61**	**0.54**
75th Percentile	**1.00**	**0.72**	**0.83**	**0.64**
90th Percentile	**1.03**	**0.75**	**1.00**	**0.75**
Maximum	**1.19**	**1.02**	**1.13**	**0.83**
Number of Obs.	**1,279**	**319**	**294**	**211**

Brand B

	No Promotion	Feature	Display	Feature & Display
Average	**0.97**	**0.63**	**0.85**	**0.59**
Mode	**1.00**	**0.50**	**1.00**	**0.50**
Minimum	**0.47**	**0.25**	**0.35**	**0.25**
10th Percentile	**0.79**	**0.47**	**0.49**	**0.47**
25th Percentile	**0.98**	**0.50**	**0.69**	**0.49**
50th Percentile	**1.00**	**0.60**	**0.99**	**0.60**
75th Percentile	**1.01**	**0.75**	**1.01**	**0.67**
90th Percentile	**1.05**	**0.76**	**1.05**	**0.75**
Maximum	**1.13**	**1.03**	**1.12**	**1.04**
Number of Obs.	**1,090**	**338**	**244**	**254**

Notes: Brand A has 2,103 observations across all types of promotions. Brand B has 1,926 observations. For each retailer-city combination, the modal price of each brand when not on promotion is normalized to $1.00.

Table 3

Estimation Summary

	Brand A	Brand B
Number of Obs.	**2,103**	**1,926**
Number of Est. Parameters	**333**	
Root MSE by Promotion:		
No Promotion	**0.37** (0.02)	**0.35** (0.03)
Feature	**0.42** (0.03)	**0.51** (0.04)
Display	**0.62** (0.03)	**0.70** (0.04)
Feature & Display	**0.45** (0.03)	**0.46** (0.04)

Notes: Standard errors are reported in parentheses.

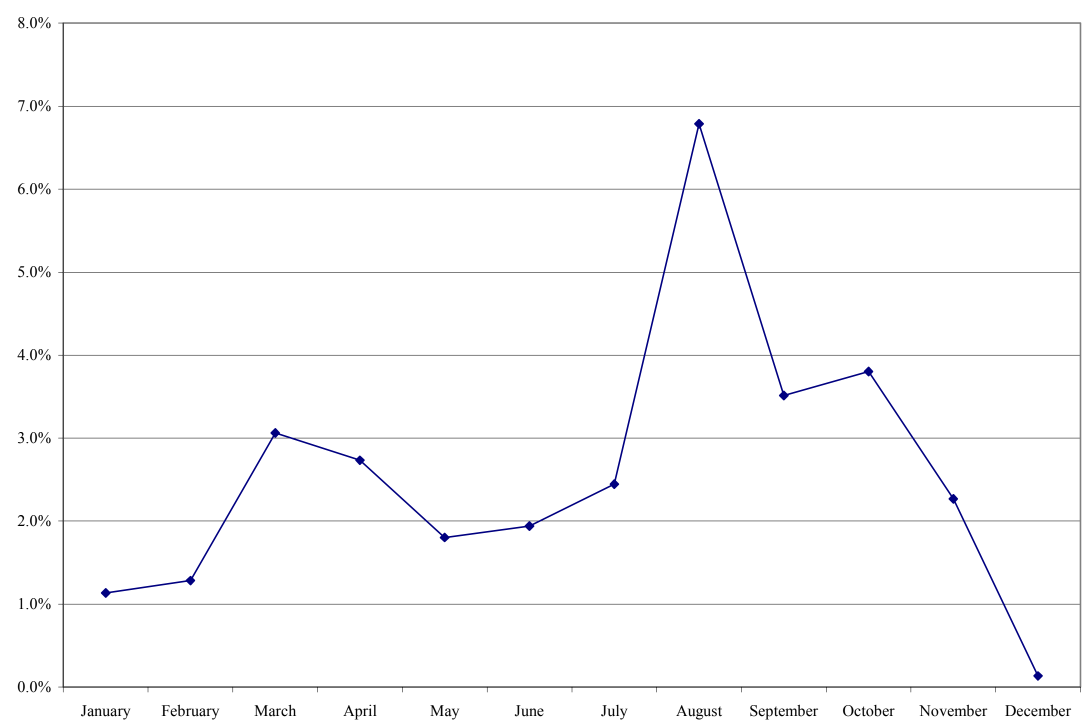
Table 4A

Joint Distribution of Promotional Activity

		Brand B Promotional Activity: No Promotion	Feature	Display	Feature & Display	Total
Brand A Promotional Activity	No Promotion	61.8%	18.4%	0.4%	2.4%	83.0%
	Feature	13.0%	1.9%	0.1%	0.2%	15.1%
	Display	0.4%	0.1%	0.0%	0.0%	0.6%
	Feature & Display	1.2%	0.1%	0.0%	0.0%	1.3%
	Total	76.4%	20.5%	0.5%	2.6%	100%

Table 4B

Average Percentage of Stores that Simultaneously Promote Both Brands



Notes: The reported statistics in Table 4A are un-weighted percentages across the 10 retailer-city combinations and 132 weeks. Table 4B reports un-weighted average percentages across retailer-city combinations for those weeks contained within each month.

Table 5

Distribution Elasticity

Brand A	Brand B
0.99	**0.90**
(0.03)	(0.04)

Notes: Standard errors are reported in parentheses.

Table 6

Impact of Own-Brand Promotional Activity,

Relative to "No Promotion"

Own-Promotion	Brand A	Brand B
Feature	**21.8%** (5.1%)	**22.0%** (5.2%)
Display	**52.3%** (10.9%)	**-1.2%** (15.0%)
Feature & Display	**80.1%** (9.6%)	**38.1%** (11.7%)

Notes: Each estimate is the predicted percentage change in unit sales from a brand's own promotional activity, relative to "No Promotion." The other brand's promotional activity is held constant at "No Promotion." All estimates assume a price of $.80 for each brand. Standard errors are reported in parentheses.

Table 7

Impact of a "Feature" Promotion by the Competing Brand, Relative to "No Promotion"

Own-Promotion	Brand A	Brand B
No Promotion	**-2.0%** (4.7%)	**5.8%** (6.1%)
Feature	**-36.9%** (12.4%)	**-41.5%** (14.8%)
Display	**-27.9%** (15.5%)	**7.6%** (13.1%)
Feature & Display	**-14.9%** (13.4%)	**-36.5%** (18.5%)

Notes: The first column reports the effect on Brand A's unit sales from Brand B having a "Feature," relative to "No Promotion," as a function of Brand A's own promotional activity. The second column reports counterpart results for the effect of a "Feature" by Brand A on Brand B's unit sales. All estimates assume a price of $.80 for each brand. Standard errors are reported in parentheses.

Table 8

Own-Price Elasticity by Own-Brand Promotional Activity

Own-Promotion	Brand A	Brand B
No Promotion	**-1.92**	**-2.01**
	(0.10)	(0.15)
Feature	**-1.53**	**-2.11**
	(0.16)	(0.19)
Display	**-1.40**	**-1.38**
	(0.19)	(0.25)
Feature & Display	**-1.31**	**-2.28**
	(0.22)	(0.22)

Notes: Standard errors are reported in parentheses.

Table 9

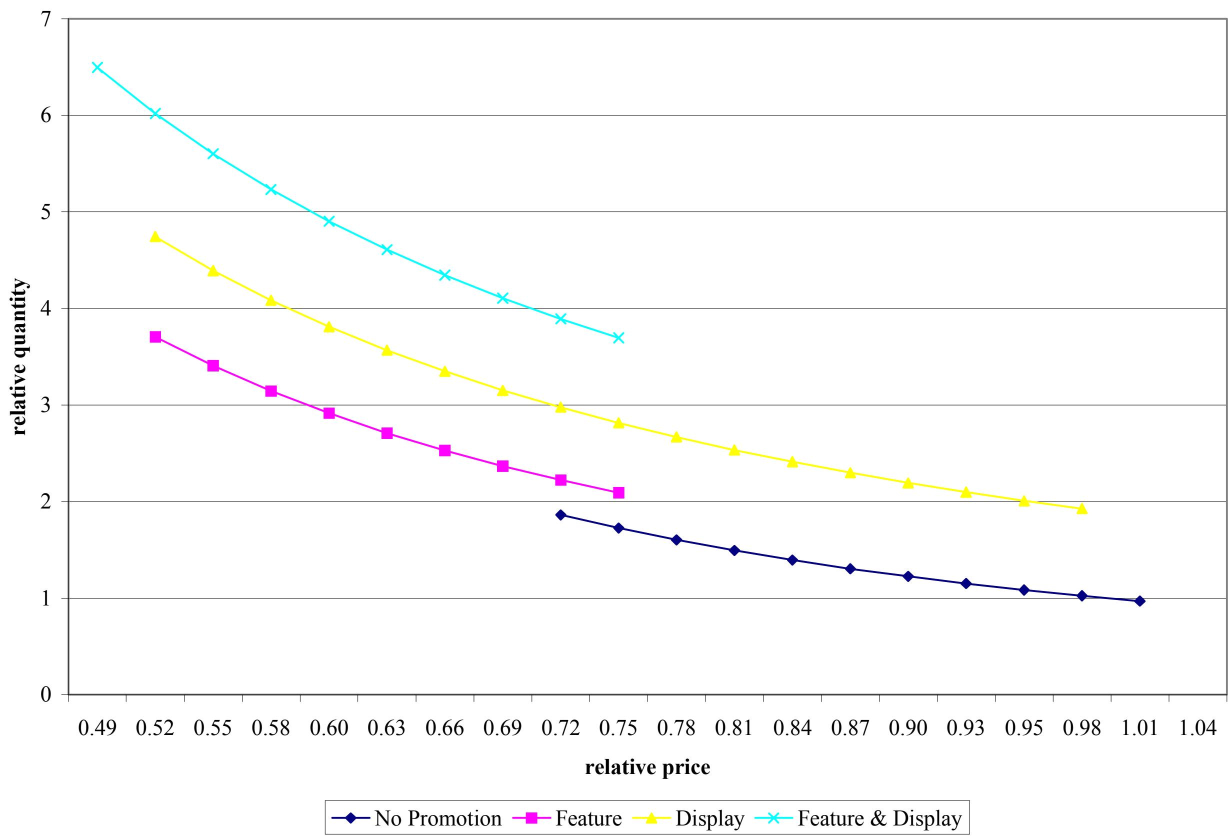
Cross-Price Elasticity by Competing Brand Promotional Activity

Cross-Promotion	Brand A	Brand B
No Promotion	**0.25** (0.09)	**0.43** (0.12)
Feature	**0.27** (0.09)	**0.34** (0.10)
Display	**0.17** (0.07)	**0.31** (0.09)
Feature & Display	**0.29** (0.10)	**0.29** (0.09)

Notes: The first column reports the cross-price elasticity of Brand A with respect to a price increase in Brand B, as a function of Brand B's promotional activity. Column two reports the cross-price elasticity for Brand B. Standard errors are reported in parentheses.

Table 10

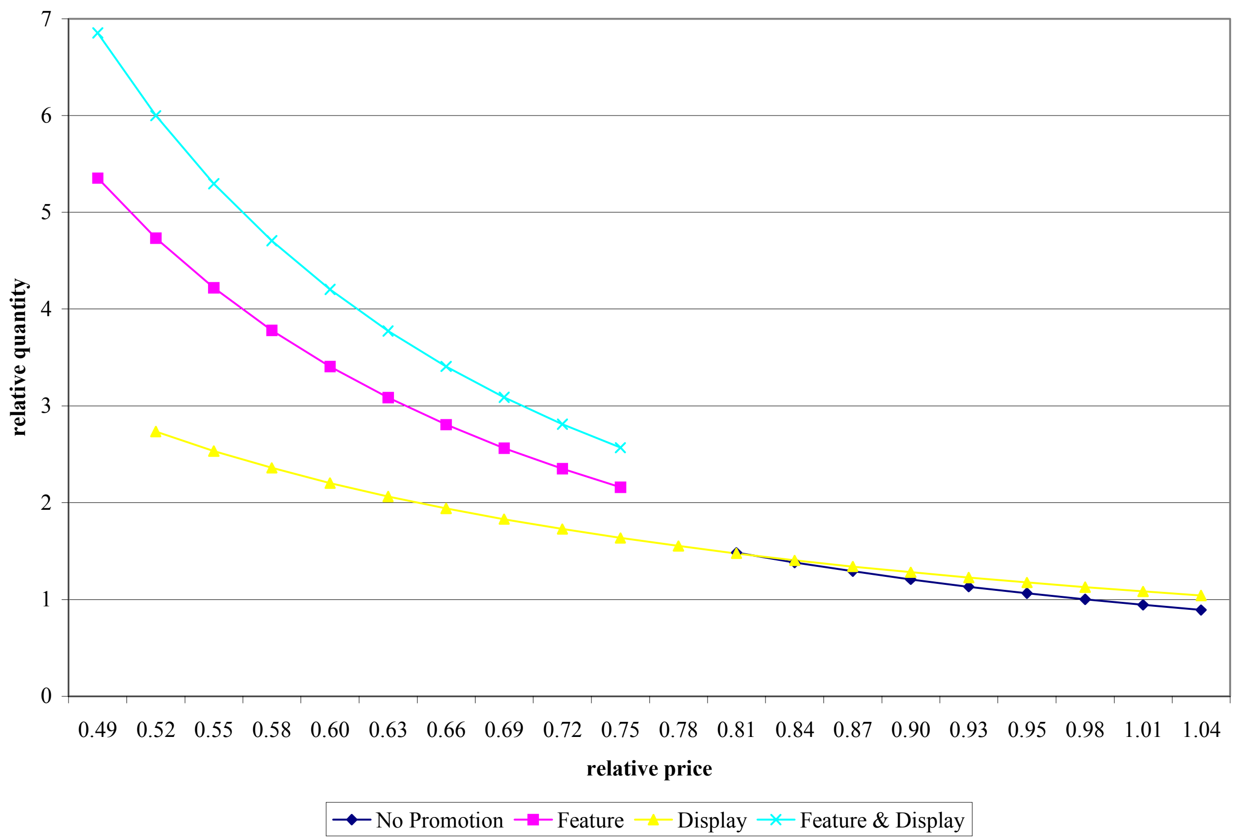
Brand A Demand Curves



Notes: All prices are rescaled so that the modal price when not on promotion equals \$1.00. Unit sales when not on promotion, and when the price is \$1.00, are normalized to one. The demand curve for each level of promotional activity is shown for prices between the 10th and 90th percentile for each type of promotion (see Table 2). The promotional activity of Brand B is held constant at "No Promotion."

Table 11

Brand B Demand Curves



Notes: All prices are rescaled so that the modal price when not on promotion equals $1.00. Unit sales when not on promotion, and when the price is $1.00, are normalized to one. The demand curve for each level of promotional activity is shown for prices between the 10th and 90th percentile for each type of promotion (see Table 2). The promotional activity of Brand A is held constant at "No Promotion."

References

Allenby, Greg and Peter Rossi (1991), "There is No Aggregation Bias: Why Macro Logit Models Work," *Journal of Business and Economic Statistics*, 1-14.

Bass, Frank and Dick Wittink (1975), "Pooling Issues and Methods in Regression Analysis with Examples in Marketing Research," *Journal of Marketing Research*, 414-425.

Blattberg, Robert, Richard Briesch, and Edward Fox (1995), "How Promotions Work," *Marketing Science*, G 122-132.

Boatwright, Peter, Sanjay Dhar, and Peter Rossi (2001), "The Role of Retail Competition and Account Retail Strategy as Drivers of Promotional Sensitivity," *Journal of Business*, Forthcoming.

Bucklin, Randolph and Sunil Gupta (1999), "Commercial Use of UPC Scanner Data: Industry and Academic Perspectives," *Marketing Science*, 247-273.

Chevalier, Judith, Anil Kashyap, and Peter Rossi (2003), "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 15-37.

Christen, Markus, Sachin Gupta, John Porter, Richard Staelin, and Dick Wittink (1997), "Using Market-Level Data to Understand Promotion Effects in a Nonlinear Model," *Journal of Marketing Research*, 322-334.

Chung, Chanjin and Harry Kaiser (2002), "Advertising Evaluation and Cross-Sectional Data Aggregation," *American Journal of Agricultural Economics*, 800-806.

Greene, William (1997), *Econometric Analysis*. Upper Saddle River, NJ: Prentice-Hall.

Guadagni, Peter and John Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 203-237.

Hamilton, James (1994), *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hansen, Lars (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1029-1054.

Hoch, Stephen, Byung-Do Kim, Alan Montgomery, and Peter Rossi (1995), "Determinants of Store-Level Price Elasticity," *Journal of Marketing Research*, 17-29.

Krishnamurthi, Lakshman, S.P. Raj, and Raja Selvam (1990), "Statistical and Managerial Issues in Cross-Sectional Aggregation," Working paper, Northwestern University.

Krishnamurthi, Lakshman, Raja Selvam, and Michaela Draganska (2000), "Inference Bias in Cross-Sectional Aggregation," Unpublished working paper.

Lattin, James and Randolph Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior," *Journal of Marketing Research*, 299-310.

Lewbel, Arthur (1992), "Aggregation with Log-Linear models," *Review of Economic Studies*, 633-642.

Link, Ross (1995), "Are Aggregate Scanner Data Models Biased?," *Journal of Advertising Research*, RC 8-12.

Lomax, Wendy and Gil McWilliam (2001), "Consumer Response to Line Extensions: Trial and Cannibalisation Effects," *Journal of Marketing Management*, 391-406.

Mulhern, Francis and Robert Leone (1991), "Implicit Price Bundling of Retail Products: A Multiproduct Approach to Maximizing Store Profitability," *Journal of Marketing*, 63-76.

Newey, Whitney and Kenneth West (1987), "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 703-708.

Reddy, Srivivas and Susan Holak (1994), “To Extend or Not to Extend: Success Determinants of Line Extensions,” *Journal of Marketing Research*, 243-262.

Stoker, Thomas (1993), “Empirical Approaches to the Problem of Aggregation Over Individuals,” *Journal of Economic Literature*, 1827-1874.

Theil, Henri (1954). *Linear Aggregation of Economic Relations*. Amsterdam: North-Holland.

Van Heerde, Harald, Peter Leeflang, and Dick Wittink (2003), “Flexible Unit-Sales Decompositions of Price Promotion Effects Based on Store Data,” Unpublished working paper.

Wallace, T. (1972), “Weaker Criteria and Tests for Linear Restrictions in Regression,” *Econometrica*, 689-698.

White, Halbert (1982), “Maximum Likelihood Estimation of Misspecified Models,” *Econometrica*, 1-25.

Wittink, Dick, Michael Addona, William Hawkes, and John Porter (1988), “SCAN*PRO: The Estimation, Validation, and Use of Promotional Effects Based on Scanner Data,” Working Paper, AC Nielsen.