A New Model of Asymmetric Competitive Structure Using Store-level Scanner Data

Fang Wu, Paul R. Messinger*, and Terry Elrod

University of Alberta

*Corresponding Author. The authors thank participants of the Marketing Seminar Series at the University of Alberta for constructive comments. The analysis of the second data set comes from the Ph.D. dissertation proposal of the third author. For constructive suggestions, we thank Miguel Villas-Boas and participants of the University of Alberta Marketing Retreat.

A New Model of Asymmetric Competitive Structure Using Store-level Scanner Data

ABSTRACT

This article develops a model of competitive structure that can be applied to widely available store-level brand sales and price data. We estimate maps of the competitive brand relationships that are assumed to jointly underlie cross-price elasticities, own-price elasticities, and brand-specific intercepts. Our methodology uses an adaptive Bayesian approach that shares information across different brands and different terms in a set of demand equations. Drawing upon recent psychometric research, we express the asymmetries present in cross-price elasticities as the difference between what we refer to as brand power parameters, and we identify relationships between a focal brand's power parameter, clout, vulnerability, own-price elasticity, and spatial density. We apply the model separately for two datasets that consist of weekly sales and prices for beer and soft drinks. Our estimated maps of competitive brand relationships have implications for both brand management and antitrust policy.

Keywords: market structure analysis, information sharing, competitive maps, brand power, clout, vulnerability, spatial density

A New Model of Asymmetric Competitive Structure Using Store-level Scanner Data

1. Introduction

The problem of identifying competitive structure among brands in a market (i.e., estimating a map of product differentiation) is important for manufacturers, service providers, and retailers managing many aspects of marketing strategy (pricing, brand repositioning, etc.). In addition, government agencies need information about competitive structure in order to apply antitrust laws concerning mergers and monopolization.

Extant methods in marketing for calibrating competitive structure maps rely mainly on disaggregated data about consumers' perceptions, preferences, and choices (e.g., see Elrod et al. 2002) or on panel data (e.g., Hendry model; see Butler and Butler 1970, 1971). These forms of data, however, may be unavailable, costly, and unrepresentative of a market of interest. Scanner data at the store or chain level, by contrast, are readily available from retailers and provide timely information about the structure of competition specific to given localities, without requiring the modeling of household-specific behavior. In this article, we consider issues that arise when using store-level scanner data to estimate competitive structure for a product category.

With market-level or store-level scanner data, a common approach to calibrating market structure involves, first, estimating the parameters of a system of demand equations that relate quantity demanded for each product to the prices of the product and its rivals, and, then, post-processing the cross-price elasticity estimates to arrive at a structural map using such tools as factor analysis (e.g., Cooper 1988) or multidimensional scaling (e.g., Allenby 1989).

A problem with past applications of this approach is that market structure has been linked only to cross-price elasticities, and not to own-price elasticities or demand intercepts. By contrast, we think that competitive structure should underlie all of these demand parameters. Standard marketing intuition, for example, suggests that desirable brands (i.e., brands with high brand equity) might be associated with large brand intercepts, low own-price elasticities, and large cross-price effects on competitors' quantities.

A second problem with the above approach is that it uses a sequence of estimation steps (i.e., estimation of demand parameters in a first stage, and analysis of market structure in a later stage), which gives rise to three shortcomings: (a) conceptual inconsistency; (b) inefficient estimation; and even (c) infeasibility of estimation. The first shortcoming pertains to the latter-stage analysis making assumptions about competitive structure inconsistent with the first-stage analysis (such as assuming that cross-price effects are symmetric). The second shortcoming pertains to the estimation error not being passed from the first model to the second. This estimation error can be substantial when the number of brands is large which may lead to estimates that are unstable (i.e., high variance) or unreliable (i.e., incorrect signs or unreasonable magnitudes) (Montgomery and Rossi 1999). The third shortcoming pertains to the model in the first stage being inestimable. This can occur when manufacturers change the prices of all their brands in a product category in parallel, resulting in perfect price colinearity. The second data set analyzed in this article illustrates this problem.

The current article addresses these problems by (a) recognizing explicitly the connections between the underlying competitive structure and various brand-specific demand parameters, (b) estimating the demand parameters and the underlying structural map in one step, and (c) using adaptive Bayesian shrinkage to share information among brand-specific parameters. Our analysis provides contributions at three levels: conceptual, methodological, and managerial.

Conceptually, the proposed model obtains a market structure map that accounts for most brand-related components in the demand model. Thus, the model reveals the influence of a brand's location on the brand's demand-equation components (i.e., cross-price elasticities, own-price elasticities, and brand intercepts). A further conceptual contribution is that we propose a novel way to formulate the asymmetric features of competition contained in cross-price elasticities and we explore the relationships between these asymmetric features and several related concepts in marketing, economics, and psychometrics. Specifically, we model the asymmetric part of cross-price elasticities in terms of a single parameter (which we will call the brand power parameter). We then develop a proposition that shows the relationship between our brand power parameter and the marketing concepts of *competitive clout* and *vulnerability* (Cooper 1988; Kamakura and Russell 1989). We also develop a second proposition that shows the relationship between own-price elasticity, our brand power parameter, and the psychometric concept of *spatial density*.

Methodologically, we derive a market structure map directly from demand functions by combining two techniques – dimension reduction and adaptive Bayesian shrinkage. This approach enables information-sharing across brands (through soft constraints imposed by Bayesian shrinkage) and across other components of the demand function (through hard constraints imposed by dimension-reduction). This permits derivation of a market structure map even when price histories are perfectly collinear. A further methodological contribution is that we engage in model selection from among common specifications in marketing well-suited to the modeling purpose, including the vector model, a dominant point model, or a fully saturated formulation (i.e., a formulation that uses separate parameters for each brand or each pair of brands).

Managerially, analysts can use the proposed model to examine strategy scenarios involving changing competitive positioning in order to predict the impact on demand intercepts, own-price elasticities and cross-elasticities. And managers with access to scanner data at the store or chain level can also readily use the proposed model to track changes over time in asymmetric competitive structure.

Antecedents. Our conceptual contribution builds on the work of Pinkse, Slade, and Brett (2002) and Pinkse and Slade (2004), who are among the few researchers that explicitly relate several different parameters of a demand function to common underlying parameters. That work, however, requires collecting additional information describing product-specific characteristics or attributes, and, therefore, is constrained by the type of information that happens to be available. Our approach, by contrast, estimates latent product attributes and allows prominent features of the data to emerge, without the influence of data-availability constraints.

Our methodological approach of using adaptive Bayesian shrinkage to obtain and stabilize brand estimates utilizes insights gained from prior research in which Bayesian techniques are used to enable the estimation of parameters in the presence of severe multicolinearity, such as with ridge regression (Hoerl and Kennard 1970; Kubokawa and Srivastava 2004; Srivastava and Kubokawa 2005). We also build on the use of Bayesian techniques in marketing to improve the quality of demand parameters estimates (Blattberg and George 1991, Montgomery 1997, Wedel and Zhang 2004). Our application of Bayesian analysis is somewhat different from this past research, however. Unlike the current paper, the bulk of this past research assumes that the distribution of demand parameters in the model is known *ex ante* – using either weakly informative priors (which can result in highly unstable estimates when price information is collinear) or informative priors for all brand-related parameters. The *adaptive* Bayesian shrinkage technique used here, instead, (1) assumes a common prior distribution for a set of brand-related parameters (such as own-price elasticities), and (2) the parameters for this common distribution (known as hyperparameters), themselves, are estimated using weakly informative priors. This is analogous to random coefficient methods in classical (non-Bayesian) statistical modeling.

Our methodological approach of modeling asymmetries in cross-price elasticities adapts ideas from the psychometric asymmetric similarity literature (e.g., Holman 1979, Weeks and Bentler 1982). Our modeling of asymmetries in this way departs from Gonzalez-Benito et al. (2009), which constitutes the only paper in marketing of which we are aware that also estimates a latent market structure model directly from demand functions. That paper builds upon Russell's (1992) Latent Symmetric Elasticity Structure pattern (LSES) model, but, unlike our approach, the estimated competitive structure map only underlies the cross-price elasticities (there is no information-sharing across brands or across other parameters of the demand model).

The rest of this paper is organized as follows. Section 2 introduces our modeling framework and the various equation specifications that implement our framework. Section 3 describes our approach to model selection, identification, and estimation. Section 4 applies our modeling approach separately to two different datasets both describing weekly grocery sales and prices: one dataset describes beer sales in a test market in the U.S.; and the second dataset describes soft drink sales in a supermarket in St. Louis. Section 5 concludes with a summary of limitations of our approach and directions for future research.

2. Modeling Framework

Our approach uses several demand model components to situate brands in a multidimensional space. In particular, we assume that latent brand locations underlie (a) the symmetric structure of cross-price elasticities (CPEs), (b) asymmetric dominance relationships embedded within CPEs, (c) own-price elasticities (OPEs), and (d) brand intercepts.

We consider a set of constant-elasticity demand equations

$$\log(E(q_{it})) = \alpha_0 + \alpha_i + \beta_{ii}\log(p_{it}) + \sum_{j \neq i}\beta_{ij}\log(p_{jt}) + f(CV_{it}), \tag{1}$$

where $i, j = 1,\ldots,n$; $t = 1,\ldots,T$; n is the number of brands; T is the number of time periods covered in the data; $E(q_{it})$ is the expected unit sales for brand *i* at time *t*; β_{ij}, $i \neq j$, are cross-price elasticities; β_{ii} are own-price elasticities; α_i are brand intercepts (mean-centered around α_0); and CV_{it} are covariates, possibly time-dependent. We observe that in many standard econometric models, Equation (1) takes a form where q_{it} appears on the left-hand side in place of $E(q_{it})$ and an additive error term is included on the right hand side. We mean-centered log prices across brands and time in our model; and, as a result, α_i describes the log sales differences across brands when all prices are set equal to the mean price in the dataset (so α_i can be interpreted as a measure of the attractiveness of brand *i*).

Our basic idea is as follows. We suppose that the parameters β_{ij}, β_{ii}, and α_i are functions of latent coordinates of the associated brands, where the coordinates are described by $M \times 1$ vectors $\boldsymbol{\theta}_i$, for each brand $i = 1,\ldots,n$. In particular, we assume that the key parameters in (1) are described by the general relationships

$$\beta_{ij} = g(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j),\ \beta_{ii} = \beta(\boldsymbol{\theta}_i), \text{ and } \alpha_i = \alpha(\boldsymbol{\theta}_i), \tag{2}$$

where $i \neq j$. That is, the cross-price elasticity between brands *i* and *j* is described by the relationship between the product locations $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$; and the own-price elasticity and brand-specific intercept for brand *i* are also influenced by the product location $\boldsymbol{\theta}_i$. In this way, underlying structural information present in the data may inform the estimation of multiple model components. As discussed in the introduction, this is in the spirit of past work in marketing, which either post-processed elasticity estimates (e.g., Allenby 1989) or which involved one-step analysis of cross-price elasticities only (Gonzalez-Benito et al. 2009). This is also in the spirit of work in economics by Pinske and Slade (2004), who assume that relationships similar to g, β, and α are functions of measured product characteristics. Our point of departure is that g, β, and α are functions of *latent structural parameters*, $\boldsymbol{\theta}_i$, $i = 1,\ldots,n$. Arriving at such latent structure is more challenging, from an estimation perspective, than assuming that the structure is a function of measured product characteristics. In principle, one could think of the relationships $\beta_{ii} = \beta(\boldsymbol{\theta}_i)$ and $\alpha_i = \alpha(\boldsymbol{\theta}_i)$ as reflective of brand equity (Aaker 1991, 1996 and Keller 1993) and $\beta_{ij} = g(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$ as reflective of brand substitutability relationships.

In practice, the functions in (2) must be determined. This involves selecting suitable specifications for g, β and α and, then, estimating these specifications with the data. The following development outlines alternative specifications suggested by the literature. (As a backdrop to this development, Appendix 1 summarizes the notation of this paper.)

2.1. Decomposing Cross-Price Elasticities

In economics, the higher the cross-price elasticity between two brands, the more substitutable they are for each other. This constitutes our rationale for modeling β_{ij} as a function of the proximity of the two brand locations $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$ when we select a specification for $\beta_{ij} = g(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$.

Because β_{ij} is not necessarily equal to β_{ji}, it is desirable for an elasticity-based market structure analysis to be able to explicitly capture asymmetries. Indeed, competitive asymmetries in CPEs are well documented (DeSarbo, Grewal and Wind 2006), and asymmetric patterns are shown to exist between high-share vs. low-share brands (Sethuraman and Srinivasan 2002), high-quality vs. low-quality brands (Blattberg and Wisniewski 1989), national vs. store brands (Kamakura and Russell 1989), and high-priced brands vs. low-priced brands. To model such asymmetries, Cooper (1988) describes how an asymmetric elasticity matrix can be structured by three-way factor analysis. Similarly, Russell's (1992) LSES model and its applications (e.g., Gonzalez-Benito et al. 2009) capture competitive asymmetry by including brand specific coefficients that reveal the overall impact of a brand on its competitors. Innovative as they are, however, these two approaches appear somewhat cumbersome to use directly as the basis for our specification for $\beta_{ij} = g(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$. Russell's (1992) LSES model, for example, requires CPEs to be market-share based, and, our starting point is instead the standard econometric specification (1) for CPEs. And rather than build our model around factor analysis (as Cooper 1988 does), we find it more convenient to work in a multidimensional scaling context following the psychometric asymmetric similarity literature.

We follow a long history of consideration of systematic asymmetries in proximity data (e.g., Tversky 1977), all of which departs from traditional distance-based models of similarity (e.g., Shepard 1962a, 1962b). Several researchers in psychometrics have proposed what is called the *skew-symmetry* model, which decomposes matrices of (dis)similarity judgments among a set of objects into symmetric and asymmetric components, and represents the latter parsimoniously by as few as one dimension (e.g., see Weeks and Bentler 1982; Saito 1986; Saito and Takeda 1990; Satio 1991, Okada and Imaizumi 1987). Some approaches go a step further in parsimony by representing symmetric and asymmetric components in the same spatial configuration. Examples include the slide-vector model (Zielman and Heiser 1993), multidimensional unfolding models (DeSarbo and Grewal 2007), the hill-climbing model (Borg and Groenen 2005), and multidimensional scaling with a dominance point (Okada and Imaizumi 2007). We will use this last approach below. Yet another approach includes the *additive similarity-bias model* (see Holman 1979, Nosofsky 1991, Carroll 1976, and also Krumhansl 1978).

In particular, following the skew-symmetric approach, we write the matrix of cross-price elasticities β_{ij}, $i \neq j$, together with zero diagonal entries, as $\mathbf{B}$ (i.e., the Greek letter capital beta). This constitutes an asymmetric proximity matrix. We decompose this asymmetric matrix $\mathbf{B}$ into symmetric and skew-symmetric components, $\mathbf{B} = (\mathbf{B} + \mathbf{B}')/2 + (\mathbf{B} - \mathbf{B}')/2 \equiv \mathbf{S} + \mathbf{A}$, which can be rewritten in scalar notation as

$$\beta_{ij} = s_{ij} + a_{ij}, \tag{3}$$

where $s_{ij} = (\beta_{ij} + \beta_{ji})/2$ has the property of symmetry ($s_{ij} = s_{ji}$) and $a_{ij} = (\beta_{ij} - \beta_{ji})/2$ has the *skew-symmetric property*, $a_{ij} = -a_{ji}$ for all $i \neq j$. Note that s_{ij} describes the average level of competition between brands *i* and *j*, and a_{ij} describes the dominance relationship between brands *i* and *j*. That is, when $a_{ij} > 0$, we have $\beta_{ij} > \beta_{ji}$, and we can interpret brand *j* as being more dominant than brand *i*.

We assume that the symmetric similarity measure s_{ij} is a decreasing linear function of an underlying distance metric,

$$s_{ij} = \phi_{CPE} - d_{ij}, \quad i \neq j, \tag{4a}$$

where ϕ_{CPE} is a constant and d_{ij} is a traditional (symmetric) inter-point distance between brands *i* and *j*.

Next, consistent with the common practice for the skew-symmetry model, we model a_{ij} with a one-dimensional linear form

$$a_{ij} = -x_i + x_j, \tag{4b}$$

for all *i* and *j,* where x_i and x_j are coordinates of objects *i* and *j* on this one dimension.

We combine (4a) and (4b) with (3) to yield

$$\beta_{ij} = \phi_{CPE} - d_{ij} - x_i + x_j. \tag{4}$$

Equation (4) constitutes our basic formulation for CPEs.

Prior to considering ways of modeling d_{ij} and x_i, we interpret this formulation. We observe that $x_i > x_j$ implies $\beta_{ij} < \beta_{ji}$, so the price of brand *j* has less of an influence on quantity demanded of brand *i* than the price of brand *i* has on the quantity demanded of brand *j*. This fact allows us to interpret x_i as the relative dominance, or power, of brand *i*. We refer to x_i as the *brand power parameter* of product *i*.

Our brand power parameter constitutes a one-dimensional measure that accounts for the asymmetry in cross-price elasticities. If the impact of price changes for brand *i* on the quantities demanded of other brands is consistently higher than the impact of price changes of these other brands on the quantity demanded of brand *i*, then the brand power parameter, x_i, will be large. Our definition of brand power is in contrast to two other definitions of brand power. (1) Na et al. (1999) measure "brand power image" as the weighted average of measured brand attributes, benefits, or values. Our brand power parameter is not a multi-item scale, but instead a measured latent parameter underlying cross-price effects.[1] (We will also later show that x_i is also related to own-price elasticity, β_{ii}.) (2) Steenkamp and Dekimpe (1997) operationalize power (of store vs. national brands) along two dimensions: "intrinsic loyalty" (to a brand) is defined to be a brand's ability to keep its current customers; and "conquesting power" is "the proportion of the market's non-loyal customers that one is able to attract in a given time period." We, instead, operationalize brand power in terms of one central dimension, not two, and we do not define brand power in terms of a smaller number of formative dimensions.

We do, nevertheless, recognize the implied relationship of our measure of brand power and the marketing concepts of *competitive vulnerability* and *clout* (see Cooper 1988 and Kamakura and Russell 1989).[2] Specifically, in the following proposition, clout describes the extent to which a focal brand *i* exerts influence on all other brands; and we define $Clout_i = \sum_{j \neq i} \beta_{ji}$. Vulnerability describes the extent to which all other brands collectively exert an influence on a focal brand; and, in particular, we define $Vul_i = \sum_{j \neq i} \beta_{ij}$.[3] We can now show the following relationships between x_i (power of brand *i*), β_{ij} (cross-price elasticity between brands *i* and *j*), $Clout_i$, and Vul_i:

Proposition 1. Equation (4) implies the following relationships:

(a) $x_i > x_j$ implies $\beta_{ij} < \beta_{ji}$ for all $i \neq j$.

(b) $x_i = \dfrac{Clout_i - Vul_i}{2J}$, if $\sum_i x_i = 0$.[4]

---

[1] We note that it might be valuable, in future research, to explore the relationship between our brand power parameter and other measured brand attributes, benefits, or values in order to understand the sources of brand power in particular circumstances.

[2] The concepts of competitive vulnerability and clout are closely related to Steenkamp and Dekimpe's (1997) dimensions "intrinsic loyalty" and "conquesting power," respectively. Indeed, these pairs of concepts are conceptually almost the same, although they are measured slightly differently. For a general discussion of brand power, see Stobart 1994.

[3] These definitions differ slightly from Kamakura and Russell (1989), who define $Clout_i = \sum_{j \neq i} \beta_{ji}^2$ and $Vul_i = \sum_{j \neq i} \beta_{ij}^2$, and from Cooper (1988), who defines $Clout_i = \sum_{j \neq i} \beta_{ji}^2 + \beta_{ii}^2$, $Vul_i = \sum_{j \neq i} \beta_{ij}^2 + \beta_{ii}^2$.

[4] We can also relate the brand power parameter, x_i, to an early model called the additive similarity-bias model (Holman 1979; Nosofsky 1991), also known as a hybrid model (Carroll 1976). This model assumes that a proximity measure δ_{ij} between two objects *i* and *j* is some increasing function (often assumed linear, for simplicity) of a

Proof. See Appendix 2.

Discussion. Statement (a) allows us to interpret x_i as a measure of *brand power* of product *i*, as we noted earlier. Statement (b) indicates that higher clout and lower vulnerability imply higher brand power, x_i, as might be expected. Thus, our proposed power parameter captures asymmetric pricing effects and links these with the concepts of clout and vulnerability.

2.2. Modeling the Symmetric and Asymmetric Components of Cross-Price Elasticities

We now consider modeling specifications for the symmetric and asymmetric components, d_{ij} and x_i. Since the asymmetric component is newer to marketing and economics, we devote more attention to this.

Symmetric Component. Note that the smaller the inter-brand distance, d_{ij}, the greater the cross-price elasticity, which constitutes greater substitutability between these brands. In particular, following the traditional multidimensional scaling (MDS) model (Shepard 1962a, 1962b), we assume that d_{ij} is the Euclidean distance (in an M-dimensional metric space) between brands *i* and *j*:

$$d_{ij} = \sqrt{\sum_{m=1}^{M} (\theta_{im} - \theta_{jm})^2} \ , \tag{5}$$

where θ_{im} describes the coordinate of brand *i* on dimension *m*.

Skew-Symmetric Component. We identify three ways of modeling the *brand power* term, x_i.

1. Estimate separate parameters, x_i, for each *i*.

2. Express x_i in terms of the brand locations θ_{im} using a form of *dominance (ideal) point model*:

$$x_i = -\omega_1 \sqrt{\sum_{m=1}^{M} (\theta_{im} - y_m)^2} \ . \tag{6}$$

3. Express x_i in terms of the brand locations θ_{im} using a *vector model*:

---

symmetric similarity measure, $\bar{s}_{ij}$; a row bias function, r_i; and a column bias function, c_j: $\delta_{ij} = \bar{s}_{ij} + r_i + c_j$. The justification for this form is that the row and column biases reflect distinguishing properties of individual items. For such a model, we can show a third relationship: (c) $x_i = (c_i - r_i)/2$, if $\sum_i x_i = 0$.

$$x_i = \sum_{m=1}^{M} \upsilon_{1m} \cdot \theta_{im} \,. \qquad (6')$$

In specifications 2 and 3, $\omega_1 \geq 0$, and y_m and υ_{1m}, $m = 1,...,M$, are all constants to be estimated.

Operationally, when the number of brands in the market is small, it may be feasible to estimate separate parameters, x_i, for each *i*. Since the same basic relationships among the brands are invariant with respect to an additive constant added to x_i; for all *i*, without loss of generality, we set $\sum_i x_i = 0$ as an identification restriction for the first formulation. Note that if the number of brands is large, this approach of estimating separate parameters, x_i, for each *i*, may be undesirable.

The formulation in (6) above is a modified adaptation of Okada and Imaizumi's (2007) model of MDS with a dominance point. Here, the hypothetical point of greatest dominance $Y = (y_1,...y_M)$ is situated in the same space as the symmetric components, and the dominance parameters are structured in relation to each brand's distance to this point, Y. Thus, the closer a brand is to Y, the stronger the brand's power. Furthermore, we can interpret ω_1 as the asymmetry weight because this represents the salience of the asymmetric model component in describing the cross-price elasticity relationships. Note that ω_1 is restricted to be nonnegative (so a brand closer to *Y* is more influential). When ω_1 equals zero in the dominance point model, there is no asymmetry present in the structure of cross-price elasticities. We graph the assumed structure for the dominance point model, when *M* = 2, in Figure 1. Here, the smaller the distance $-x_i / \omega_1$, the larger x_i, and the more dominant is brand *i*.

[Insert Figure 1 about here.]

Lastly, the vector model formulation (6') describes an increasing progression of magnitude of the dominance parameters x_i, along the direction of the estimated gradient vector $(\upsilon_{11},\ldots,\upsilon_{1M})$. Note that when $(\upsilon_{11},\ldots,\upsilon_{1M}) = 0$ in the vector model, there is no asymmetry present in the structure of the cross-price elasticities. Also note that the vector model can be shown to be a limiting case of the dominance-point model as the distance of the dominance point *Y* from the origin approaches infinity.

2.3. Modeling Own Price Elasticities and Brand Intercepts

We now specify formulations for the other model components $\beta_{ii} = \beta(\boldsymbol{\theta}_i)$ and $\alpha_i = \alpha(\boldsymbol{\theta}_i)$.

Own-Price Elasticities. We note three ways to model own price elasticities:

1. Estimate separate parameters, β_{ii}, for each $i = 1, \ldots, n$.

2. Express β_{ii} in terms of the brand locations θ_{im} using a form of *ideal (dominance) point model*:

$$-\beta_{ii} = \phi_{OPE} + \omega_2 \sqrt{\sum_{m=1}^{M} (\theta_{im} - z_m)^2}, \omega_2 > 0. \tag{7}$$

3. Express β_{ii} in terms of the brand locations θ_{im} using a form of *vector model*:

$$-\beta_{ii} = \phi_{OPE} + \sum_{m=1}^{M} \upsilon_{2m} \cdot \theta_{im}. \tag{7'}$$

In specifications 2 and 3 above, ω_2 is a constant constrained to be positive, and ϕ_{OPE} and υ_{2m}, $m = 1, \ldots M$ are constants without sign constraints. All these constants are to be estimated.

When the number of brands in the market is small, it may be simple enough to estimate separate parameters, β_{ii}, for each i=1,…,n. Nevertheless, this may be conceptually undesirable. These parameters reflect competitive structure of the market. When freely estimated, they neither contribute to, nor are explained by, the competitive structure of the market. This undermines the usefulness of the model for making predictions for repositioned, new, or deleted brands because the effects of such changes on own price elasticities and therefore on demand cannot be predicted.

Embedding the OPEs in competitive map solves this problem. Equations (7) and (7') achieve this by expressing own-price elasticities in terms of the brand locations θ_{im}. In particular, the formulation in (7) describes the magnitude (i.e., absolute value[5]) of brand i's own-price elasticity as a linearly increasing function of the squared distance between the brand and the hypothetical brand on the map located at $Z = (z_1, \ldots z_M)$. The closer the brand to the hypothetical brand $Z = (z_1, \ldots z_M)$, the smaller the OPE is in absolute value. By contrast, the vector model describes an increasing progression of magnitude of the OPEs, β_{ii}, along the direction of the estimated gradient vector $(\upsilon_{21}, \ldots, \upsilon_{2M})$.

[5] Since own-price elasticity should be negative, the negative of OPE equals its absolute value.

In order to interpret these formulations, we derive the following relationships.

Proposition 2. Assuming utility-maximizing consumers under a linear budget constraint and writing the income elasticity of brand *i* as β_{iI}, we have

(a) $-\beta_{ii} = (\sum_{j \neq i} \beta_{ij}) + \beta_{iI} = Vul_i + \beta_{iI}$; and (8)

(b) $-\beta_{ii} = (n-1)Density_i - nx_i + \beta_{iI}$, if $\sum_i x_i = 0$; (9)

where $Density_i \equiv (\phi_{CPE} - \sum_{j \neq i} d_{ij} / (n-1))$.

Proof. See Appendix 2. [Note (a) arises from utility maximizing behavior under a linear budget constraint; and (b) arises from (4).]

Discussion. We can make several observations from Proposition 2.

First, when the income effect β_{iI} is small, the OPE equals minus the vulnerability of brand *i*, as an approximation.

Second, when OPEs are represented according to the ideal (dominance) point formulation of (7) and income effects for the product category under study are small (or similar across brands)[6], the hypothetical brand $Z = (z_1, \ldots z_M)$ can be interpreted as the least vulnerable brand on the map.

Third, according to Proposition 2 (b), when the income effect β_{iI} is small, a brand's own-price elasticity may be viewed as depending on at least two factors: (i) its distinctiveness from all the other brands in the market, as indicated by a spatial density measure, $Density_i \equiv (\phi_{CPE} - \sum_{j \neq i} d_{ij} / (n-1))$; and (ii) its dominance parameter, x_i. Thus, Proposition 2 (b) indicates that the lower the spatial density around a product *i*, *ceteris paribus*, the less price sensitive the product will be (-β_{ii} low in magnitude). Intuitively, a product that is relatively unique will be less price elastic. Furthermore, the higher a product *i*'s power relative to all other products, *ceteris paribus*, the less price sensitive the product will be (i.e., the product will face inelastic demand.). Thus, a powerful product (high brand equity) will be less price elastic.

[Insert Figure 2 about here.]

---

[6] When the income effects are zero, the Slutsky equation together with the symmetry of substitution terms for Hicksian demand (Varian 1992, pp. 120 and 123) implies symmetric cross-price effects, i.e., $\partial q_i / \partial p_j = \partial q_j / \partial p_i$. This suggests further constraints that could be included in the model, which we do not impose in order to avoid loss of generality.

We illustrate these ideas in Figure 2, in which seven brands are shown, denoted by *A, B, C…, G*. Holding all the other factors constant, brand *G* should have the smallest absolute own-price elasticity and the hypothetical brand *Z* (interpreted as having the "least vulnerable" possible brand location) should be located in a sparse space, likely near brand *G* (or further out southwest in the third quadrant). Furthermore, any two brands that are very similar to each other should have similar own-price elasticities because brands similar to each other should also have similar distinctiveness relative to other brands.

Overall, our concept of brand power, x_i, enables us to identify the relationships in Propositions 1 and 2 between cross-price asymmetries, clout, vulnerability, own-price elasticity, and spatial density.

Note that our measure of spatial density (or distinctiveness) is similar to that of Krumhansl's (1978) distance-density model, in which an object's distinctiveness is related to the spatial density in a region surrounding it in the multidimensional configuration, and this spatial density is measured by the sum of a monotonically decreasing function of the distances of all other objects to the focal object. In this context, the denser this region, the more difficult it is for an object to distinguish itself.

It is worth noting that the spatial density explanation of own-price elasticity has also been alluded to in the economics and marketing literatures. Pinkse and Slade (2004) allow a brand's OPE to depend on the number of neighbors that the brand has in an exogenously determined product-characteristic space. In a related development, Bronnenberg and Vanhonacker (1996) examine the implications of a consumer only responding to price variation of brands in the consumer's choice set (which they call *local price response*). They show that the fewer the number of other brands appearing with a focal brand in a consumer's choice set, the lower the focal brand's vulnerability (and OPE). They also argue that the more frequently a focal brand appears with other brands in consumers' choice sets, the higher the focal brand's clout.

Brand Intercepts. We next express the intercepts α_i in (1) in terms of the brand location parameters, θ_{im}, $m = 1,...,M$. In particular, we suggest two ways of modeling α_i:

1. Estimate separate parameters, α_i, for each $i = 1,\ldots,n$.

2. Express α_i in terms of the brand locations θ_{i1} on the first dimension in (4):

$$\alpha_i = \omega_0 \cdot \theta_{i1}, \omega_0 \neq 0\,, \tag{10}$$

where θ_{i1} is the coordinate of i^{th} brand on the first dimension, which we interpret as "perceived attractiveness" for brand *i*. Linking the intercept α_i to first dimension θ_{i1} provides a needed identification constraint and also aids in map interpretation, in the spirit of Bentler and Weeks (1978). As is well known, unconstrained MDS is otherwise invariant to rotations

Similar to our earlier treatment of OPEs, expressing intercepts in terms of the brand locations in the market structure map is conceptually appealing because a useful map of competition should have the possibility of accounting for all brand-specific components in a model. Here we recognize that the brand's perceived attractiveness is an important attribute that influences the observed switching patterns reflected in cross-price elasticities: two brands differing in perceived attractiveness cannot have identical locations in a map that purports to explain brand competition.

Covariates. Lastly, one must determine the function *f* according to which various covariates, CV_{it}, enter the model in (1). For the two data sets considered later in this article, we include seasonality and a time trend:

$$f(CV_{it}) = \gamma_1 t' + \gamma_2 \sin\frac{t'}{2\pi \times 52} + \gamma_3 \cos\frac{t'}{2\pi \times 52} \tag{11}$$

where $i = 1, \ldots, n$, $t = 1, \ldots, T$, and *t'* is the mean-centered *t* ($t' = t - \bar{t}$). In (11), *t'* is divided by 52, which represents the periodicity of the cycle when dealing with weekly data.

2.4. Modeling the Distribution of the Dependent Variable

For packaged-goods data sets at the store or chain level, the quantity demanded is integer-valued, with values ranging from 0 to several thousand. For such data, two suitable modeling alternatives consist of the Poisson distribution and the Negative Binomial distribution (NBD).

When the Poisson distribution is assumed to describe unit sales q_{it}, the probability mass function is

$$\Pr(q_{it} \mid \lambda_{it}) = (\lambda_{it})^{q_{it}} \exp(-\lambda_{it})/(q_{it})!, \;\; \lambda_{it} > 0\,, \tag{12}$$

where λ_{it} is the rate for this Poisson process. The expected sales in (1) is given by $E(q_{it}) = \lambda_{it}$.

The Poisson model is sometimes restrictive, however, because it assumes that the mean and the variance are both equal to λ_{it}. Often important determinants of the mean are unobserved, causing the variance to exceed the fitted mean. In such instances, the negative binomial distribution (NBD) may be used, which generalizes the Poisson distribution by allowing the rates λ_{it} to be distributed according to a gamma distribution across the population of customers. The gamma density is given by

$$\Pr(\lambda_{it} \mid \rho, \eta_{it}) = (\eta_{it})^{\rho} (\lambda_{it})^{\rho-1} \exp(-\eta_{it} \cdot \lambda_{it}) / \Gamma(\rho), \tag{13}$$

where η_{it} is the scale parameter for the gamma distribution, differing across brands and time, and ρ is the shape parameter. Combining Equations (12) and (13) and integrating out λ_{it} yields the negative-binomial density

$$\Pr_{NBD}(q_{it} \mid \rho, \eta_{it}) = \binom{q_{it} + \rho - 1}{\rho - 1} \left(\frac{\eta_{it}}{\eta_{it} + 1}\right)^{\rho} \left(\frac{1}{\eta_{it} + 1}\right)^{q_{it}}, \ \rho > 0, \eta_{it} > 0. \tag{14}$$

When the NBD model is adopted, the expected sales in (1) are given by $E(q_{it}) = \rho / \eta_{it}$.

[Insert Table 1 about here.]

2.5 Summary of Model Components

We conclude our discussion of model development by summarizing in Table 1 the various possible model specifications that are alternatives to freely estimating all the parameters of (1). Table 1 shows the decisions required for each model component. Component 1 pertains to the integer dependent variable, which is given either a Poisson or negative-binomial distribution (NBD). Component 2 involves the number of dimensions (*M*) for the symmetric CPE structure (the "map"). Component 3 concerns whether the brand power parameter assumed to underlie CPE asymmetries is freely estimated, dropped altogether, or embedded in the map using an dominance-point or vector formulation. Component 4 addresses whether the OPEs are freely estimated or included in the map using an ideal-point or vector formulation. Component 5 describes the brand intercepts, which can be freely estimated or tied to the horizontal axis of the map for the purpose of identification and interpretation. This paper restricts attention to formulations in Table 1 – which are all based on (1) and (11); but we acknowledge that future research may wish to consider other formulations as well.[7]

3. Model Selection, Identification, and Estimation

Table 1 describes many possible model specifications. Various approaches can be used to arrive at a final model specification. One approach begins by specifying values for the number of dimensions M (such as 1 or 2) and then doing one of the following: (a) Estimate all possible models. (b) Estimate most possible models, but without pursuing all variations of an empirically poor choice for one of the components. For example, if a model freely estimating a skew-symmetric structure (choice b for component 3) does not converge, then this same choice would not be repeated for all combinations of other component decisions. (c) Estimate the most parsimonious model and then explore one-component generalizations, in forward stepwise fashion. (d) Estimate a fractional factorial combination of all possible model decisions. (e) Use the most effective component decisions found for dimension M = 1 as an indicator of the key components to explore for dimensions M = 2 and higher. Of course, the modeling purpose and prior knowledge, and not just statistical criteria, must also be considered.

In terms of the order of estimation of models, we found it convenient to start with a saturated model (with as many parameters as possible freely estimated), and then to successively replace various model components with more parsimonious formulations. By proceeding in this sequence, we could then immediately identify problems with convergence and identification and

---

[7] The choice of model specification, of course, has an impact on the number of parameters that need to be estimated. We briefly provide an overview of the number of parameters that are implied by different model specifications. For Component 1, using the NBD formulation adds an additional parameter relative to the Poisson distribution. Components 2 and 3 include alternatives to freely estimating $n(n-1)$ CPEs. Component 2 imposes a symmetric CPE structure involving approximately nM location parameters (θ_{im}; $i=1,...n$; $m=1,...,M$) – actually after the identification restrictions described in Section 3.2 below, Component 2 involves adding *nM-(M-1)M/2* parameters. Component 3 (the skew-symmetric structure of CPEs) adds either (a) n – 1 dominance parameters, x_i, $i=1,\ldots,n$, (imposing $\bar{x}=0$), (b) $M+1$ parameters ($\omega_1, y_m, m=1,...,M$), or (c) M parameters ($\upsilon_{1m}, m=1,...,M$). Component 4 (the structure of OPEs) involves either (a) n OPE parameters (β_{ii}, $i=1,\ldots,n$), (b) $M+1$ parameters ($\omega_2, z_m, m=1,...,M$), or (c) M parameters ($\upsilon_{2m}, m=1,...,M$). Component 5 (the structure for the brand intercepts) entails either (a) n separate intercept parameters (α_i, $i=1,\ldots,n$) or (b) one parameter (ω_0) that links the intercepts with the horizontal axis. Adopting parsimonious forms can, therefore, reduce the parameter count of these five model components from as many as n^2+n to as few as $nM+2M+1-(M-1)M/2$. Thus, for $n=20$ brands and dimensionality $M=2$, there are 420 brand-related demand parameters in a fully saturated model and only 44 brand-related structural parameters in the most parsimonious form. These calculations are based on a simple count of the number of parameters in model components 2 through 5. These parameters constitute the focal parameters of interest.

trace them to their causes.

3.1 Model Selection

For model selection, we apply the deviance information criterion (DIC) to *focal* parameters (Spiegelhalter et al. 2002). In particular, let $\overline{D}$ denote the mean deviance for the model allowing all parameters to vary randomly according to their posterior distributions. Let φ denote a vector of model parameters deemed focal (see footnote 7), as distinct from other nuisance parameters. Finally, let $D(\overline{\varphi})$ denote the mean deviance for the model holding focal parameters at their posterior means while allowing all other parameters to vary randomly. Then the DIC is (Spiegelhalter et al. 2002)

$$DIC = 2\overline{D} - D(\overline{\varphi}) \,. \tag{15}$$

Generally, DIC is an adaptation of Akaike's information criterion (AIC) to a Bayesian context. (See Appendix 3 for a discussion of the differences between DIC, AIC, and BIC.)

Our purpose in modeling competitive market structure is to estimate all parameters characterizing the competitive map together with any brand-specific parameters in (1) that the researchers may have chosen to estimate independently of the competitive map (such as brand-specific intercepts). Thus, for our analysis these parameters are the focal ones. Calculating DIC using (15) requires first estimating all parameters, and then rerunning the model holding focal parameters at their posterior means (and letting nuisance parameters vary randomly). The DIC statistic produced by WinBUGS in the course of estimation of all parameters does not suffice. The program does not know the modeling purpose. Instead, it guesses that the purpose is to estimate the means of the observed responses, i.e., E(q_{it}) for (1). We accordingly had to calculate DIC by hand (with a routine we wrote) using (15).

3.2 Model Identification

As with all MDS procedures, our model requires identification conditions in order to obtain unique estimates. From (5), we see that $\Theta = [\theta_{.1}, \ldots \theta_{.m}, \ldots, \theta_{.M}]$ enters into the model mainly through the distance d_{ij}. However, inter-brand distance calculations are invariant to orthogonal, scale-preserving rotations of the brand locations given in $\mathbf{\Theta}$ and to mirror-image "flips" of the map. Therefore, some constraints must be imposed in order for the model to be identified. In particular, for the matrix of parameter locations **Θ**, we fix the elements above the diagonal to be zero and the diagonal elements to be positive. When M=2, this involves imposing the constraints $\theta_{12} = 0,\ \theta_{11} > 0,$ and $\theta_{22} > 0$.[8] Generally for any dimensionality *M*, this involves imposing *M(M+1)/2* constraints. In addition, the indeterminacy of the origin is resolved by requiring that $\sum_{i=1}^{M} \theta_{im} = 0, m = 1, 2.$

3.3 Model Estimation

We adopt a hierarchical Bayes modeling approach which we estimate using a Markov Chain Monte Carlo (MCMC) method. The hierarchical Bayes method is a powerful tool that recognizes that a family of parameters may be theoretically related and that can express such relationships with rich statistical formulations (Gelman and Hill, 2006). This section demonstrates the approach for the particular focal specification described in Table 1 by Components 1b, 2(*M*=2), 3d, 4c, 5b – analogous development applies for other specifications from Table 1.

The hierarchical Bayes version of the NBD model for our unit sales data is as follows. We write the posterior distribution as

$$\Pr(\rho, \{\eta_{it}\}, \{\lambda_{it}\} \mid \{q_{it}, p_{\cdot t}, t\}) \propto \prod_{t} \prod_{i} \Pr(q_{it} \mid \lambda_{it}) \Pr(\lambda_{it} \mid \eta_{it}, \rho) \Pr(\rho) \Pr(\eta_{it} \mid \{p_{\cdot t}, t\}), \quad (16)$$

with $\Pr(q_{it} \mid \lambda_{it})$ from (12); $\Pr(\lambda_{it} \mid \eta_{it}, \rho)$ from (13), and $\Pr(\rho)$ and $\Pr(\eta_{it} \mid \{p_{\cdot t}, t\})$ constituting the priors of the gamma distribution of λ_{it}.[9] We specify an uninformative gamma prior for $\rho \sim dgamma(0.01, 0.01)$. The heart of the model concerns our formulation of the prior distribution $\Pr(\eta_{it} \mid \{p_{\cdot t}, t\})$.

Hard constraints. We begin our characterization of the prior distribution of η_{it} under the focal specification by recognizing a set of conditional relationships arising from the applicable

---

[8] Occasionally, instead of imposing a constraint of the form $\theta_{mm} > 0$, it is more convenient to impose a constraint that $\theta_{im} > \theta_{jm}$ for two arbitrarily selected brands i and j. This equally well can avoid lack of model identification due to mirror-image "flips" of the map.

[9] Here, $i = 1,\ldots,n$, $t = 1,\ldots,T$; and $p_{\cdot t}$ represents the (vector of) prices of the various products at time *t*. The posterior can also be written as $\Pr(\rho, \{\eta_{it}\}) \mid \{q_{it}, p_{\cdot t}, t\}) \propto \prod_{t} \prod_{i} \Pr(q_{it} \mid \eta_{it}, \rho) \Pr(\rho) \Pr(\eta_{it} \mid \{p_{\cdot t}, t\})$ with the likelihood $\Pr(q_{it} \mid \eta_{it}, \rho)$ from (14). Regardless of which version used, note that $\rho/\eta_{it} = E(q_{it})$ under the NBD model, where we specify $E(q_{it})$ in (1). For the Possion model (Component 1a), we instead write the posterior probability as $\Pr(\{\lambda_{it}\} \mid \{q_{it}, p_{\cdot t}, t\}) \propto \prod_{t} \prod_{i} \Pr(q_{it} \mid \lambda_{it}) \Pr(\lambda_{it} \mid \{p_{\cdot t}, t\})$, with $\lambda_{it} = E(q_{it})$ and $E(q_{it})$ from (1).

model equations, as follows.

Conditional Relationship	**Applicable Equations**	
$\eta_{it} \mid \{p_{.t}, t\}, \alpha_i, \beta_{ii}, \{\beta_{i.}\}, \alpha_0, \gamma_1, \gamma_2, \gamma_3$;	(1) & (14)	
$\beta_{ij} \mid \theta_{i1}, \theta_{j1}, \theta_{i2}, \theta_{j2}, x_i, x_j, \phi_{CPE}$, $j \neq i$;	(4)	
$x_i \mid \theta_{i1}, \theta_{i2}, \upsilon_{11}, \upsilon_{12}$;	(6')	(17)
$\beta_{ii} \mid \theta_{i1}, \theta_{i2}, \upsilon_{21}, \upsilon_{22}, \phi_{OPE}$; and	(7')	
$\alpha_i \mid \theta_{i1}, \omega_0$.	(10)	

For example, Equation (1) establishes the relationship between the standard demand parameters (together with price data) and log($E(q_{it})$), which, in turn, is related to η_{it}, once we recognize that $E(q_{it}) = \rho/\eta_{it}$, under the NBD model (of Equation (14)). Similarly, Equations (4), (6'), (7') and (10) link the CPEs, power parameters, OPEs, and demand intercepts, respectively, to the underlying locations $(\theta_{i1}, \theta_{i2})$ and $(\theta_{j1}, \theta_{j2})$ of brands *i* and *j*. Recognizing these hard constraints, we complete the characterization of the prior distribution of η_{it} by specifying priors for the constant coefficients and the brand-specific coefficients in the model.

Priors for constant coefficients. We assume weakly informative priors for all constant coefficients (i.e., coefficients not subscripted by a particular brand) in the model (e.g., $\alpha_0, \phi_{CPE}, \phi_{OPE}, \upsilon_{11}, \upsilon_{12}, \upsilon_{21}, \upsilon_{22}, \omega_0, \gamma_1, \gamma_2, \gamma_3$). For example, we assume that $\alpha_0 \sim N(0, 10^4)$.[10]

Priors for brand-specific coefficients. We use a hierarchical Bayesian structure to describe the brand-specific parameters not predetermined by the assumed hard constraints (in (17)). We describe our approach for the estimation of brand locations θ_{im} ($i = 1, 2..., n$, $m = 1, ..., M$). The basic idea is to recognize that the brand locations θ_{im}, $i = 1, 2..., n$, in a given dimension *m*, are related; and these locations may be reasonably thought of as having been drawn from a common underlying distribution. This allows for information sharing across brands (because the hyper-parameters of the common underlying distribution are jointly estimated with the estimates of the locations θ_{im}, $i = 1, 2..., n$, themselves). This approach is referred to as *adaptive Bayesian shrinkage*.[11]

[10] Although not necessary for the focal model (Components 1b, 2(M=2),3d,4c,5b in Table 1), some other model formulations require imposing suitable sign constraints on the coefficients.

[11] For the focal specification of this section, the only brand-specific coefficients not predetermined by the hard constraints are the brand locations θ_{im} — all other brand-specific coefficients are conditional on the brand locations

Thus, we do not estimate the brand locations using informative priors, or using weakly informative priors, as in (18),

$$\theta_{im} \sim N(0,10^4),\ i=1,2...,n. \tag{18}$$

We, instead, estimate brand locations using a prior with hyper-parameters that are, themselves, given weakly informative priors on brand locations, as in (19):

$$\begin{aligned} &\theta_{im} \sim N(\mu_m, \sigma_m^2) \\ &\mu_m \sim N(0,10^4) \quad .i=1,2...,n \\ &\sigma_m \sim Unif(0,50) \end{aligned} \tag{19}$$

Following Gelman and Hill (2006), a broad uniform prior is used for the standard deviation of θ_{im}.

The distributions in (19) imply a prior distribution for θ_{im} that is virtually identical to (18). (The mean is the same and the variance is one percent larger.) However, (18) asserts the brand locations have nothing in common — that they are independently distributed. That is, it assumes that knowing the θ_{im} values for any n – 1 of the brands provides no information about the value for the remaining brand. Equation (19) makes no such assumption. Rather than assert, as in (18), that $\mu_m = 0$ and $\sigma_m = 100$, (19) estimates these values. Although the values of these two parameters assumed by (18) both lie well within the prior distributions given by these hyper-parameters in (19), comparison of these assumed values to their posterior densities invariably shows that the assumed value $\sigma_m = 100$ in (18) lies well outside its 95% credible intervals.

θ_{im} (or on constant coefficients) through the constraints in (17). Thus, for the focal specification of this section, adaptive Bayesian shrinkage is used only for the brand locations θ_{im}. For other specifications in Table 1, however, hard constraints do not necessarily predetermine OPEs, brand intercepts, or the power parameters, x_i. When this is the case, adaptive Bayesian shrinkage is applied for estimating such families of coefficients.

As background, a shrinkage estimator, in statistics, is an estimator that improves on an estimate by combining it with other information. Adaptive Bayesian shrinkage improves on a Bayesian estimate by combining it with information about other Bayesian estimates, by supposing that they share a common prior distribution, and the hyper-parameters of the common prior distribution are jointly determined with the family of related Bayesian estimates. Adaptive Bayesian shrinkage is in contrast with the use of informative priors, which express specific, definite information about a variable, and which, in this sense, impose a pre-specified level of shrinkage (i.e., non-adaptive information sharing). This is also in contrast with the use of weakly informative (uninformative) priors, which do not include other information, and which result in minimal shrinkage (information sharing).

Adaptive Bayesian shrinkage arises from assuming a hierarchical Bayes structure. Note that the use of separate informative or weakly informative priors for various parameters assumes independence of the estimates. In contrast, a hierarchical Bayes model removes the independence across the estimates. Generally, the posterior distributions arising from hierarchical Bayes models tend to move (or shrink) away from the maximum likelihood estimates towards their common mean.

The advantage of using (19) is that it allows for shrinkage of the brand parameters towards a common mean μ_m estimated by the data whenever $\sigma_m \ll 50$, and since the common mean μ_m is given a weakly informative prior, its estimate is determined by the data. However, if the θ_{im} values are greatly different from each other, then the posterior estimate of σ_m can be much larger than 100, which implies a prior for θ_{im} that is even less informative than (18).

Of course, many Bayesian models in marketing (and elsewhere) estimate distributions for model parameters (Rossi and Allenby 2003), but they generally do so when those model parameters are viewed as random effects. In such models, the parameters of the underlying distribution (μ_m and σ_m) are regarded as focal. When the purpose, instead, is to estimate those model parameters themselves (i.e., when the θ_{im} values for the $i = 1, 2..., n$ brands are focal), then a weakly informative prior such as (18) has been used, mirroring a fixed-effects specification. In such cases, we suggest that a prior for parameters should be estimated (as in (19)). This is the position taken by Gelman and Hill (2006, p. 246), who argue that the latter approach should be taken, regardless of the focus. However, two conditions should be met: (a) Weakly informative distributions should be used for the prior distribution's hyper-parameters. This ensures that the estimated prior subsumes both the fixed-effect and random-effect specifications. (b) There must be at least two (latent or observed) variables that share this prior distribution. Otherwise there is no opportunity for sharing information among parameters, and (19) is simply an inefficient means for implementing (18).

We note that the advantages of an "adaptive shrinkage" approach for the estimation of demand models such as (1) have been demonstrated by Montgomery and Rossi (1999). Their implementation differs from ours in several respects, however. First, they employ adaptive shrinkage for some parameters but not others, whereas we apply it consistently to all parameters that satisfy the two criteria (a) and (b) above. Second, they make use of informative priors at the highest level of the hierarchy, whereas we use pre-specified weakly informative priors at the highest level throughout. Third, they use strong additive utility as their theoretical model wherein all competitive effects are through the income effect only. This precludes differential competition among brands. The model here allows for some brands to compete more closely than others, and asymmetrically.

4. Applications

We now apply the model and estimation methods to two very different data sets: one is informative about brand-specific parameters, and the other has price histories that are perfectly collinear in thirteen of twenty-one dimensions.

4.1. Beer Market

We estimate our model using data from Information Resources consisting of aggregate weekly store-level scanner data for beer in a test market in the US from 1989 to 1996. The data set contains 365 weekly observations on unit sales, price, and other variables. We considered the top five brands: Budweiser, Miller, Busch, Old Milwaukee and Milwaukee's Best. Price is the inflation-adjusted average weekly price per ounce.

We used the same brand-level data that formed the basis for Srinivasan, et al.'s (2000) analysis. Those authors first created brand-level variables from SKU-level variables.[12] This aggregation from SKUs (stock keeping units) to brands may be subject to aggregation bias due to possible heterogeneity among promotional variables (Christen et al., 1997; Pesaran and Smith, 1995). Srinivasan et al. (2000) control for this bias by performing pooling tests to determine whether it was reasonable to pool the different varieties for a brand; over 95% of sales could be pooled.[13] They also found that this data set produced elasticities that matched prior estimates (e.g., Tellis, 1988), which also alleviated this concern to some extent. The benefit of aggregating from SKUs to brands is that it avoids problems associated with colinearity and a large state space, pointed out by Kopalle et al. (1999) and Bucklin and Gupta (1999). Indeed, our analysis of the second soft-drink data set will illustrates how severe colinearity can make estimation problematic, and we show how our modeling framework can assist with estimation. We begin with the less challenging beer data set.

Model Selection. Our model selection process (described earlier) identified the following

---

[12] We are indebted to Information Resources and the authors of this article for making these aggregated brand-level data available.

[13] Christen et al. (1997) point out that the aggregation bias is likely to be quite small in data characterized by three conditions: frequent promotions, frequent price cuts, and small own price elasticities. We note that the first two conditions are met for the beer data.

three models, which we will compare: [14]

Model 1. This comprises the NBD model of (14), together with the following demand equation (arrived at by combining (1) and (11)):

$$\log(1/\eta_{it}) = \alpha_0 + \alpha_i + \beta_{ii}\log(p_{it}) + \sum_{j\neq i}\beta_{ij}\log(p_{it}) + \gamma_1 t' + \gamma_2 \sin\frac{t'}{2\pi\times 52} + \gamma_3 \cos\frac{t'}{2\pi\times 52}, \tag{1'}$$

where $i=1,\dots,5$; $t=1,\dots,365$; $t'=t-\overline{t}$ (mean-centered); and prices are inflation-adjusted and also mean-centered across time and brands. This is a standard demand model in the literature. We estimate this model by imposing weakly informative priors for all demand parameters, which therefore leads to estimates that are equivalent to the ML estimators.

Model 2. This comprises Model 1 together with the following additional elements:

$$\beta_{ij} = \phi_{CPE} - d_{ij} - x_i + x_j,\ i \neq j \tag{4}$$

$$d_{ij} = \sqrt{\sum\nolimits_{m=1}^{M}(\theta_{im}-\theta_{jm})^2} \tag{5}$$

$$x_i = \sum_{m=1}^{M}\upsilon_{1m}\cdot\theta_{im} \tag{6'}$$

In this model, market structure is used to explain (only) the symmetric and asymmetric competitive patterns in CPEs. The adaptive Bayesian shrinkage approach is applied when estimating brands locations, OPEs, and brand intercepts.

Model 3. This comprises Model 2 together with the following additional elements:

$$-\beta_{ii} = \phi_{OPE} + \sum_{m=1}^{M}\upsilon_{2m}\cdot\theta_{im}, \tag{7'}$$

$$\alpha_i = \omega_0\theta_{i1},\ \omega_0 \neq 0 \tag{10}$$

For this model, the estimates of brand locations, θ_{im}, underlie all brand-related parameters in the demand function. The adaptive Bayesian shrinkage approach is applied when estimating the brand locations. Conceptually, we find this to be the most desirable model.

Tables 2 and 3 show estimates using WinBugs 1.4.3 of these three models. Model 1 yields DIC of 17407. For both models 2 and 3, the DIC suggest two dimensional models work better. Model 2 yields a DIC of 17403; while model 3 yields a DIC of 17405.

[Insert Tables 2 and 3 about here.]

[14] In this application, Model 1 is the fully saturated model, consisting of model components (1, 2, 3a, 4a, 5a) in Table 2. Model 2 consists of model components (1, 2, 3d, 4a, 5a) in Table 2. Model 3 consists of model components (1, 2, 3d, 4c, 5b).

It is worth mentioning that the beer data that we obtained for analysis already aggregated all SKUs for the same manufacturer with correlated price history. This was done by Srinivasan, et al.'s (2000) to permit estimation by standard models.

Overall, we observe that DIC slightly favors both structural models (Models 2 and 3) over Model 1. A close examination of Models 2 and 3 shows that most of the underlying structural parameters are estimated significantly (zero being outside their 95% credible intervals). Also, all own price elasticities are negative and in ranges similar to previous analyses (Tellis 1988). We further note evidence of seasonality, which is to be expected for beer consumption.

We emphasize that even when we impose underlying structure for the CPEs, OPEs, and intercepts, our methodology also recovers implied estimates for the CPEs, OPEs, and intercepts. That we get these implied estimates as a byproduct of the estimation procedure constitutes one of the strengths of MCMC Bayesian methodology. We observe that the estimated values of the intercepts, and OPEs are generally robust across Models 1, 2, and 3, even though the later two models impose substantial underlying structure. Furthermore, that the intercepts are so similar in Model 3 as with the other two models, suggests that interpreting axis one of the underlying Euclidean map as describing the intercept does not undermine our estimation. Lastly, many of the CPEs are robust across the three models, and where there are very large changes, the CPE in Model 1 was not significant to begin with. So imposing structure appears to be improving the accuracy of estimation of the cross-price elasticities. Moreover, Models 2 and 3 are very similar in estimates, and imposing structure on the intercepts and OPEs does not appear to change the underlying structure for the CPEs (both the symmetric structure and skew-symmetric brand power estimates). These are reassuring checks. Furthermore, notice that the hyperparameter estimates of brands location variances on each dimension, σ_1 and $\sigma_{\theta.2}$, are small, which indicates that shrinkage plays a role in estimating the brands' locations (Table 3).

Managerial Implications. Examining the underlying structure, we note that the OPEs are all significantly negative and reasonable, and the CPEs evidence various cross-price effects, some symmetric and others asymmetric. For example, Old Milwaukee and Milwaukee's Best are close symmetric substitutes with each other. On the other hand, Budweiser and Miller are more characterized by one-directional substitutability, where the price of Miller has significant effects on Budweiser's sales, but the price of Budweiser has less of an effect on Miller's sales.

As Model 3 is conceptually our most desirable model (despite its DIC being slightly higher than Model 2), we provide a biplot based on this model. Figure 3 displays the five brand locations and three variables (own-price elasticity, brand power, perceived attractiveness) on two latent attribute dimensions (Figure 3). Here, brands are displayed as points while the three variables are displayed as arrows. The horizontal dimension is interpreted as "perceived attractiveness" in accordance with specification component 5b (of Table 1). Several conclusions can be drawn from this map.

[Insert Figure 3 about here.]

First, this figure gives brand locations $(\theta_{i1}, \theta_{i2})$ for each brand i. Recall that the symmetric part of the cross-elasticity between any two products is modeled by the Euclidean distance in the map between the brands. Thus Old Milwaukee and Milwaukee's Best are quite similar, while Miller is distinct from all other brands.

Second, the cosine of the angle between arrows approximates the correlation between the variables. We can see from Figure 3 that "brand power" is highly correlated with "own-price elasticities," and both are positively correlated with "perceived attractiveness". This means that brands with more power have less elastic OPEs and are more attractive, which makes perfect sense. It is also worth noticing two other findings. (a) Since the cosine of the angle between a vector and an axis indicates the importance of the contribution of the corresponding variable to the axis dimension, we see that the second dimension is more strongly related to brand power and OPEs. (b) Since the length of an arrow is equal to the variance of the corresponding variable, it seems that the "perceived attractiveness" dimension is not that important for these data.

Third, the points (brands) in the biplot can be projected perpendicular on the arrows, and the position of the points along the arrow gives information of the value of the brands on corresponding variables. As we expected, Miller has greater asymmetric dominance than Budweiser, which in turn has greater dominance than the other three brands. Notice the estimates of asymmetry weights on both dimensions in CPE are significantly different from zero (Table 2), which means that competitive asymmetry is an essential feature of this product category. Similarly, Miller has the greatest OPE in absolute value, followed by Budweiser, and then the other three brands. And finally, the brand intercepts (attractiveness), which are proportional to the horizontal axis, Budweiser is represented as the most popular (or mass marketed) brand, along with Busch.

Overall, Figure 3 presents a picture of quality tiers. Budweiser appears to be popular, attracting a broad market, but Miller is a beer with greater asymmetric dominance over Budweiser even if it has a smaller market share. There is, thus, a smaller quality market, of which Miller is one brand. Budweiser is at the center. Among the lower dominance segment, consisting of Busch, Milwaukee's Best, and Old Milwaukee, Busch has the highest share. This is intuitive, and perhaps sums up insights that resonate with beer drinkers.

4.2. Soft-drink Market and Historical Context

We now turn to analysis of store level aggregate data for carbonated soft drinks (CSD). Our selection of context was motivated in part by proposed soft-drink mergers in 1986 when the Federal Trade Commission deterred PepsiCo's proposed acquisition of Seven-Up Co. by threatening antitrust litigation. Unable to similarly dissuade Coca-Cola Co., the FTC went to court to block Coca-Cola Co.'s proposed acquisition of Dr Pepper (*FTC v. Coca-Cola Co.* 641 F. Supp. 1128; D.D.C. 1936). The FTC justified these actions by citing Section 7 of the Clayton Act, claiming in both antitrust issues, that the acquirer and acquiree were selling in the same line of commerce, defined in these cases as "the national carbonated soft drink market." We wish to answer a simple yet important question: To what extent do the brands carried by the four soft-drink manufacturers compete with each other?

Data Description. The data, recorded using the scanner UPC (Universal Product Code) system from a St. Louis supermarket, describes weekly quantity sales and price of the same two-liter soft drinks analyzed in the survey and covers 63 weeks commencing 2/22/88. There are in total 48 brands included in our dataset. We focus our analysis on the following 20 brands of four major manufacturers:

PepsiCo: Pepsi, Diet Pepsi, Pepsi Free, Diet Pepsi Free, Mountain Dew, Lemon-Lime Slice, Diet Lemon-Lime Slice

Coca-Cola Co.: Coke, Diet Coke, Caffeine-Free Coke, Caffeine-Free Diet Coke, Cherry Coke, Diet Cherry Coke, Sprite, Diet Sprite, MR. PiBB

Dr Pepper Co.: Dr Pepper, Diet Dr Pepper

Seven-Up Co.: 7-Up, Diet 7-Up

The remaining low sales brands (all shares less than 1.5%) were combined into a single "other" brand with price calculated as the average (quantity weighted) price of the constituent products. A preliminary examination of the data shows that the average weekly sales ranged from 838 for Pepsi to 5 for Mr.PiBB.

An important feature of our data is the presence of *perfect* multicolinearity among the products within a line of products sold by each of the four major companies. As shown in Table 4, the prices of the 21 brands vary in only 8 dimensions! From a managerial perspective, these product lines are being priced as groups—with identical prices over the course of the year. This parallel pricing of products within product lines or groups presents a challenge for estimation. A second challenge for this data set is that we estimate locations of 21 brands (including "other).[15] We show below how the information sharing feature of our modeling framework can enable estimation of the underlying structure among the brands – despite the presence of severe multicolinearity and the large number of brands.

[Insert Table 4 about here.]

Model Estimation Results. In general, we recommend using the NBD model above. However, when the information in the data is limited, the NBD model can suffer from an over-fitting problem, whereby the resulting map describes random error or noise rather than the underlying relationship. Therefore, for the soft drink data, we adopt only the Poisson distribution (12) for the dependent variable. So, instead of (1') above, we use the analogous form

[15] A managerially interesting feature is that the Coke and Pepsi lines are negatively correlated, apparently due to an implicit understanding between these companies that they offer alternative price promotions. In addition, the Sprite and Slice groups are closely correlated with the Pepsi and Coke groups, respectively, as we might expect.

$$\log(\lambda_{it}) = \alpha_0 + \alpha_i + \beta_{ii}\log(p_{it}) + \sum_{j \neq i} \beta_{ij}\log(p_{it}) + \gamma_1 t' + \gamma_2 \sin\frac{t'}{2\pi \times 52} + \gamma_3 \cos\frac{t'}{2\pi \times 52}, \qquad (1'')$$

where $\lambda_{it} \equiv E(q_{it})$; i, j = 1, …, 21; t = 1, …, 63; t' is mean-centered ($t' = t - \bar{t}$); and prices are mean-centered across time and brands.

Given the limitations of this data, the fully saturated model is not estimable, and we instead started by estimating a model that drops the cross-price elasticities ($\beta_{ij} = 0$ for all $i \neq j$) but retains the other terms in (1”). That is, this base model has components 1a, M = 0, 3a, 4a and 5a (see Table 1). The DIC for this model is 88059, with 42 parameters estimated. Many of the models for M = 1 or 2 and with partially saturated terms (Components 3b, 4a, or 5a) were not estimable, but the dominance-point formulations (3c and 4b) were estimable and worked better than the vector counterparts. The model with components 1a, 2 (with M = 1), 3c, 4b, and 5b has a DIC of 86584, with 25 focal parameters estimated,[16] while the M = 2 version of this model has DIC of 76340, with 47 focal parameters estimated, which we report on below.

Table 5 shows the estimated price elasticities from the two-dimension model. Own-price elasticities are all less than –1 and, consistent with previous literature on this industry (Dube 2004, 2005), own-price elasticities are very elastic, ranging from –2.5 to –5.5. In addition, the absolute values of cross-price elasticities are an order of magnitude smaller than the own-price elasticities, as expected and as found for the beer data. There are, however, many negative cross-price elasticities, which can arise when category expansion effects overwhelm substitution effects between brands (Russell et al. 2008). Both Diet 7Up and Diet Dr. Pepper have fairly large negative effects on the unit sales of many other brands, which seems to indicate that the price reduction of these two brands might benefit the whole category.

[Insert Tables 5 and 6 about here.]

Table 6 provides the estimates of three weights (w_1, w_2, w_3) associated with skew-symmetric component, own-price elasticities, and intercepts respectively—cf. (6), (7), and (10)—as well as their corresponding standard deviations. All of these estimates are statistically significant, and we

[16] The effective number of parameters estimated for both one-dimensional and two-dimensional models will be slightly smaller, as Bayesian shrinkage imposes information-sharing across parameters.

can conclude that asymmetry is an essential feature for our cross-price elasticities matrix.

Figure 4 gives the estimated two-dimensional map, which displays the 21 brands' locations and the two hypothetical brands locations (red circles) on two latent dimensions (describing the Dominance Brand Y and the Least Vulnerable Brand Z, respectively). The horizontal dimension is restricted to be proportional to the brand intercepts. From Figure 4, we observe some interesting findings.

[Insert Figure 4 about here.]

First, the horizontal dimension distinguishes the coke vs. non-coke products, with majority of the coke products considered superior to non-coke products. Here we interpret "*perceived attractiveness*" as increasing as we move to the right on the horizontal axis.

Second, the distance between points reflects the perceived similarity between the corresponding two brands and the competition between them. Figure 4 indicates that Pepsi and Coke are in head-to-head competition in both dimensions. On the other hand, Diet Pepsi, Diet Coke and Diet Pepsi Free are all diet products and they are similar, although they seem to share a different submarket from Pepsi and Coke. Dr. Pepper products are in direct competition with Sprite products and 7UP, with Diet Dr.Pepper competes with Diet Spirit, while Dr.Pepper competes with both Sprite and 7UP. Furthermore, there seems to be a close relationship between Cherry Coke products and Lemon-Lime Slices products. However, the structure of this competition seems somewhat counterintuitive, as Cherry Coke is in direct competition with Diet Lemon-Lime Slice, while Diet Cherry Coke with Lemon-Lime Slice. A probable explanation might be that there exists features other than Diet vs. Non-Diet that play a role in this competitive pattern.

Third, concerning brand power and own-price elasticity, the power of a brand should be interpreted as related to the distance between a brand and the hypothetical brand labeled as "Y" (the left circle), where a brand's power is larger, the closer the brand's location is to "Y". Similarly, the brand that is closest to the hypothetical brand labeled as "Z" (the right circle) in the map is considered the least sensitive to its own-price changes and also least vulnerable to other brands' price changes. As we can see from the map, most of the Coke and Pepsi products have relatively lower vulnerability, measured as smaller absolute value of own-price elasticity. This effect might be due to the high market share of these products, and also the relatively sparse spatial density of these products. The sparse spatial density of these products can be attributed to the fact that the Coke and Pepsi products have been differentiated (e.g., diet vs. non-diet, caffeine vs. caffeine-free, etc) so well that finer submarkets are formed among them than other companies' products. Besides, the low average prices of these products may also play a role. When prices are already low (average $1.1 per 2 liter), consumers' sensitivity to these products' price reductions would be very low. Now consider the power of all brands on the market. An interesting finding is that the brands with the highest power are 7UP, and Dr. Pepper, rather than Coke and Pepsi (because 7Up and Dr. Pepper are closest to the hypothetical Dominance Brand Y). Since both 7UP and Dr. Pepper are also quite near the least vulnerable brand Z, we can conclude that these two brands' high power is mainly due to their low vulnerability (rather than their high clout). One explanation for their low vulnerability can be that both 7UP and Dr. Pepper position their brands quite differently from the other brands in the market. In addition, these two brands are closer to non-coke products on the left panel of the Figure than those coke products on the right, but closer in "*perceived attractiveness*" to coke products. Pepsi and Coke, one the other hand, also have relatively low vulnerability, but their power (and their clout) appears less than that of 7Up and Dr. Pepper. On the other hand, both have very high brand intercepts (as shown by their position on the right on the horizontal axis).

Fourth, MrPiBB seems like a true outlier in the soft drink market. It does not exert influence on other brands, nor is it influenced by them. From the consumers' perspective, it has the lowest "*perceived attractiveness*" level but highest average price. This is a classic niche brand that enjoys a quasimonopoly, albeit with a very small segment.

Last, we derive each brand's average clout and vulnerability (Table 7). Recall the definition of clout and vulnerability are: $Clout_i = \sum_{j \neq i}^{J} \beta_{ji}$ and $Vul_i = \sum_{j \neq i}^{J} \beta_{ij}$. The entries in the table are actually $\frac{Clout_i}{J}$ and $\frac{Vul_i}{J}$.

[Insert Table 7 about here.]

Antitrust Implications. The findings from the model provide us an interesting feature of the competition in late 80's soft drink market: products that belong to 7Up and Dr. Pepper companies compete with each other, but in quite a different market than products of both the Coca-Cola and Pepsi companies. A direct implication of this finding is that the approval of both PepsiCo's proposed acquisition of Seven-Up Co and Coca-Cola Co.'s proposed acquisition of Dr Pepper may *not* necessarily "substantially lessen competition, or to tend to create a monopoly" in the soft drink market. Actually, with both the economies of scale and economies of scope, consumers may even benefit from these acquisitions. This would not be true if the mergers were between Pepsi and Coca Cola, or between 7Up and Dr.Pepper, since in either case, the competition happens within the same stragic submarket.

It is worth mentioning that the outcome in this industry was that a third party, Hicks and Haas, had, by October of 1986, acquired Dr Pepper, A&W Root Beer, and 7-Up, making Hicks and Haas the third largest U.S. soft drink maker. Our finding, however, provides contrary evidence that this outcome may not be better than permitting the previous two merger proposals to have become realized.

5. Conclusions

This article presents a unified framework for estimating a market-structure map that (a) represents the perceived substitutability between brands; (b) accounts for cross-price elasticity asymmetries; (c) reasonably integrates common marketing specifications (e.g., vector and dominant point formulations) to represent the various demand-model components; and (d) facilitates the estimation of price elasticities in the presence of severe colinearity in prices and other data limitations. From a microeconomic perspective, we identify underlying relationships between measures of cross-price elasticity, brand power, vulnerability, clout, own-price elasticity, and spatial density. From a methodological perspective, we demonstrate an adaptive Bayesian approach to estimation that shares information across different brands and different terms in a set of demand equations.

This framework can be applied to inform marketing managers who are selecting prices for the brands in their existing product lines and who are setting the positioning of new brands. The framework is also relevant to policy makers applying antitrust policy relating to mergers and to monopolization. We apply this approach to beer and soft-drink data sets and arrive at plausible estimates in light of both economic theory and past marketing literature.

We acknowledge limitations of this study and directions for extension of this research. In particular, it would be desirable do the following: (a) add covariates in the model such as promotions and displays (and the lack of such information may create imprecision or bias in the estimation of such elements as brand intercepts); (b) analyze the time dimension more explicitly, including lagged effects; (c) consider alternative functional forms (other than (1) and (5)); and (d) develop flexible, utility-maximization-based models, capable of estimating market structure that could also be useful for simulating consumer welfare effects for policy purposes (Bronnenberg et al. 2005). Concerning this last point, such utility-based models would serve as a complement to the price-elasticity, market-demand-based approach of the current paper.

Overall, price elasticities provide important information about competition in a market, and the availability of detailed scanner data in nearly all retail outlets serves as a potential source of this information. We look forward to continued work on elasticity-based market structure analysis that more fully exploits the potentialities of such data.

References

Aaker, David A. (1991), *Managing Brand Equity*, San Francisco: Free Press.

Aaker, David A. (1996), “Measuring Brand Equity Across Products and Markets,” *California Management Review*, 38 (3), 102–120.

Akaike, Hirotugu (1974), "A new look at the statistical model identification," IEEE Transactions on Automatic Control 19 (6), 716–723.

Allenby, Greg M. (1989), “A Unified Approach to Identifying, Estimating and Testing Demand Structures with Aggregate Scanner Data,” *Marketing Science*, 8(3), 265–280.

Bentler, P.M. and David G. Weeks (1978), “Restricted Multidimensional Scaling Models,” *Journal of Mathematical Psychology* 17, 138–151.

Berry, Steven T., (1994), “Estimating Discrete–Choice Models of Product Differentiation,” *RAND Journal of Economics*, 25 (2), 242–262.

Blattberg, Robert C. and George, Edward I. (1991), “Shrinkage Estimation of Price and Promotion Elasticities: Seemingly Unrelated Equations,” *Journal of the American Statistical Association*, 86(414), 304–215.

Blattberg, Robert C, and Ken J. Wisniewski (1989), “Price Induced Patterns of Competition,” *Marketing Science, 8 (4),* 291–309.

Borg, Ingwer and Patrick J.F.Groenen (2005), *Modern Multidimensional Scaling: Theory and Applications*, Second Edition, Springer.

Bronnenberg B.J, and VanHonacker W.R (1996), “Limited Choice Sets, Local Price Response, and Implied Measures of Price Competition,” *Journal of Marketing Research*, 33(2), 163–173.

Bronnenberg, B.J., P.E.Rossi, N.J.Vilcassim (2005), “Structural Modeling and Policy Simulation,” *Journal of Marketing Research*, 42(1), 22–26.

Bucklin, R.E., Gupta, S. (1999), “Commercial use of UPC scanner data,” *Marketing Science,* 17(3), 247–273.

Butler, D. and B. Butler, Jr. (1970, 1971), Hendro–Dynamics: Fundamental Laws of Consumer Dynamics, Cambridge, Mass: The Hendry Corporation.

Carroll, J.D. (1976), “Spatial, Non–spatial and Hybrid Models for Scaling,” *Psychometrika*, 41, 439–463.

Christen, M., Gupta, S., Porter, J.C., Staelin, R., Wittink, D.R., (1997), “Using market–level data to understand promotion effects in a nonlinear model,” *Journal of Marketing Research* 34, 322–334.

Cooper, Lee G. (1988), "Competitive Maps: The Structure Underlying Asymmetric Cross Elasticities," *Management Science*, 34 (6), 707–23.

Cooper, Lee G, Daniel Klapper, and Akihiro Inoue (1996), “Competitive–Component Analysis: A New Approach to Calibrating,” *Journal of Marketing Research,* 33(2), 224–238.

DeSarbo, Wayne S. and Rajdeep Grewal (2007), "An Alternative Efficient Representation of Demand–Based Competitive Asymmetry," *Strategic Management Journal*, 28 (7), 755–66.

DeSarbo, Wayne S., Rajdeep Grewal, and Jerry Wind (2006), “Who Competes with Whom? A Demand–Based Perspective for Identifying and Representing Asymmetric Competition,” *Strategic Management Journal* (27), 101–129.

Dube, Jean–Pierre. (2004), “Multiple Discreteness and Product Differentiation: Demand for

Carbonated Soft Drinks," *Marketing Science,* 23, 66–81.

Dube, Jean–Pierre. (2005), "Product Differentiation and Mergers in the Carbonated Soft Drink Industry," *Journal of Economics and Management Strategy*, 14(Winter), 879–904.

Elrod, Terry, Gary J. Russell, Allan D. Shocker, Rick L. Andrews, Lynd Bacon, Barry L. Bayus, ..., and Venkatesh Shankar (2002), "Inferring Market Structure from Customer Response to Competing and Complementary Products," *Marketing Letters*, 13 (3), 221–32.

Gelman, Andrew and Jennifer Hill (2006), Data Analysis Using Regression and Multilevel/ Hierarchical Models, Cambridge.

Gonzalez–Benito, O., M. P. Martinez–Ruiz, and A. Molla–Descals (2009), "Spatial Mapping of Price Competition Using Logit-type Market share Models and Store-level Scanner data," *Journal of the Operational Research Society*, 60, 52-62.

Hoerl, A.E., Kennard, R.W. (1970), "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, 12, 55–67

Holman, E.W. (1979), "Monotonic Models for Asymmetric Proximities," *Journal of Mathematical Psychology*, 20, 1–15.

Kamakura, Wagner A., Russell, Gary J. (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 23 (November), 379–390.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer–Based Brand Equity," *Journal of Marketing*, 57, 1–22.

Kopalle, P., Mela, C. F., Marsh, L. (1999), "The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications," *Marketing Science* 18(3), 317–332.

Krumhansl, Carol (1978), "Concerning the Applicability of Geometric Models to Similarity Data: the Interrelationship Between Similarity and Spatial Density," *Psychological Review*, 85 (5), 445–463.

Kubokawa, Tatsuya and Muni S. Srivastava (2004), "Improved Empirical Bayes Ridge Regression Estimators Under Multicollinearity," *Communications in Statistics,* 33(9), 1943–1973.

Montgomery, Alan L. (1997), "Creating Micro–Marketing Pricing Strategies Using Supermarket Scanner Data," *Marketing Science*, 16(4), 315–337.

Montgomery, Alan L. and Rossi, Peter E.(1999), "Estimating Price Elasticities with Theory–based Priors", *Journal of Marketing Research*, 36 (4), 413–423.

Na, Woon Bong, Roger Marshall, and Kevin Lane Keller (1999), "Measuring brand power: validating a model for optimizing brand equity," *Journal of Product and Brand Management*, 8, 3, pp.170 – 184.

Nosofsky, R.M. (1991), "Stimulus Bias, Asymmetric Similarity, and Classification," *Cognitive Psychology*, 23, 94–140.

Okada, A. and Imaizumi, T. (1987), "Nonmetric multidimensional scaling of asymmetric proximities," *Behaviormetrika*, 21, 81–96.

Okada, Akinori and Imaizumi, Tadashi (2007), "Multidimensional Scaling of Asymmetric Proximities with a Dominance Point," *Advances in Data Analysis*, Springer.

Pesaran, M. H., Smith, R. (1995), "Estimating long–run relationships from dynamic heterogeneous panels," *Journal of Econometrics* 68, 79–113.

Pinkse, Joris, Margaret E. Slade, and Craig Brett (2002), "Spatial Price Competition: A

Semiparametric Approach," *Econometrica* 70 (3), 1111–1153.

Pinkse, Joris and Margaret E. Slade (2004), "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review 48*, 617–643.

Rossi, P. E., G. M. Allenby. 2003. Bayesian statistics and marketing. *Marketing Science*. 22(3) 304–328.

Russell, Gary J. (1992), "A Model of Latent Symmetry in Cross Price Elasticities," *Marketing Letters* 3(2), 157–169.

Russell, Gary J., Ann Petersen, and Suresh Divakar (2008), "Analysis of Brand Price Competition Using Measures of Brand Similarity," *working paper*.

Saito, T, (1986), "Multidimensional Scaling to Explore Complex Aspects in Dissimilarity Judgment," *Behaviormetrika*, 20, 35–62.

Saito, Takayuki (1991), "Analysis of Asymmetric Proximity Matrix By a Model of Distance and Additive Terms," *Behaviormetrika* (29), 45–60.

Saito, Takayuki and Shin–ichi Takeda (1990), "Multidimensional Scaling of Asymmetric Proximity: Model and Method," *Behaviormetrika* (1990), 28, 49–80.

Schwarz, Gideon (1978), "Estimating the Dimension of a Model," *Ann. Statist.* 6(2), 461–464.

Sethuraman, Raj and V.Srinivasan (2002), "The Asymmetric Share Effect: An Empirical Generalization on Cross–Price Effects," *Journal of Marketing Research,* 39 (3), 379–386.

Shepard, R.N. (1962), "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function," Part I. *Psychometrika*, 27, 125–140.

Shepard, R.N. (1962), "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function," Part II. *Psychometrika*, 27, 219–246

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde (2002), "Bayesian Measures of Model Complexity and Fit," *J. R. Statist. Sco. B,* 64(4), 583–639.

Srinivasan, Shuba, Popkowski Leszczyc, Peter T.L., and Bass, Frank M. (2000), "Market share response and competitive interaction: The impact of temporary, evolving and structural changes in prices," *International Journal of Research in Marketing*, 17, pp. 281–305

Srivastava, Muni S. and Tatsuya Kubokawa (2005), "Minimax Multivariate Empirical Bayes estimators Under Multicollinearity," *Journal of Multivariate Analysis,* 93, 394–416.

Steenkamp, Jan-Benedict E. M. and Marnik G. Dekimpe (1997), "The Increasing Power of Store Brands: Building Loyalty and Market Share," *Long Range Planning*, 30, 6, pp. 917 - 930.

Stobart, P. (1994), *Brand Power*, Macmillan, Basingstoke.

Tellis, G.J., (1988), "The price elasticity of selective demand: A meta–analysis of econometric models of sales," *Journal of Marketing Research* 25, 331–341.

Tversky, Amos (1977), "Features of Similarity," *Psychological Review*, 84 (4), 327–352.

Wedel, Michel and Jie Zhang (2004), "Analyzing Brand Competition Across Subcategories," *Journal of Marketing Research*, 41 (4), 448–456.

Weeks, D.C. and P. M. Bentler (1982), "Restricted multidimensional scaling models for asymmetric proximities," *Psychometrika*, 47, 201–208.

Zielman, Berrie and Willem J. Heiser (1993), "Analysis of Asymmetry by a Slide–Vector," *Psychometrika*, 58 (1), 101–114.

Appendix 1. Variable Definitions

Focal Variables

p_{it} price of brand i at time t (log prices are mean centered)

q_{it} unit sales of brand i at time t

Demand Equation Parameters

α_0 overall intercept of demand function

α_i brand-specific intercept (mean-centered)

β_{ii} own-price elasticity of brand i

β_{ij} cross-price elasticity between brands i and j

CV_{it} covariate in i^{th} brand equation at time t

f functional form of covariate term(s)

Latent Structural Parameters

θ_{im} location of brand i on *mth* dimension

$\boldsymbol{\theta}_i$ location (vector) of brand i

g specification linking β_{ij} and the relationship between $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$

β specification linking β_{ii} and $\boldsymbol{\theta}_i$

α specification linking α_i and $\boldsymbol{\theta}_i$

B (capital beta) asymmetric matrix of cross-price elasticities, with diagonal elements zero

s_{ij} $\equiv (\beta_{ij} + \beta_{ji})/2$

S symmetric matrix with typical element s_{ij}

a_{ij} $\equiv (\beta_{ij} - \beta_{ji})/2$

A skew-symmetric matrix with typical element a_{ij}

Indices

n number of brands in market

i index of brand, $i = 1,\ldots,n$

j index of brand, $i = 1,\ldots,n$

T number of weeks in data set

t index of week, $t = 1,\ldots,T$

t' is mean-centered version of index t

M number of dimensions in structural map

m index of map dimension, $m = 1,\ldots,M$

Covariates (time dependent)

γ_1 the coefficient for sine time trend

γ_2 the coefficient for cosine time trend

γ_3 the coefficient for linear time trend

Underlying Structural Model Parameters

ϕ_{CPE} constant in specification for β_{ij}

ϕ_{OPE} constant in specification for β_{ii}

d_{ij} inter-point distance between $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$

x_i brand power parameter for brand i

ω_0 weight in brand-specific intercept

ω_1 weight of dominance-point model for x_i

ω_2 weight of ideal-point model for β_{ii}

$Y = [y_m]$ most powerful location, $m = 1,\ldots,M$

$Z = [z_m]$ least vulnerable location, $m = 1,\ldots,M$

υ_{1m} m^{th} coefficient of vector model for x_i

υ_{2m} m^{th} coefficient of vector model for β_{ii}

Distribution Parameters

λ_{ij} the rate parameter of poison distribution

η_{ij} expected value for gamma distribution

ρ The shape parameter of gamma distribution

Hyper Parameters

μ_m hyper-parameter (prior mean) for brands' coordinates on *mth* dimension

σ_m hyper-parameter (prior standard deviation) for brands' coordinates on *mth* dimension

Brand-Related Measures[17]

$Clout_i \equiv \sum_{j \neq i} \beta_{ji}$

$Vul_i \equiv \sum_{j \neq i} \beta_{ij}$

$Density_i \equiv (\phi_{CPE} - \sum_{j \neq i} d_{ij} / (n-1))$

---

[17] From the formulation in footnote 4, we have

δ_{ij} proximity measure

c_j column bias function

r_i row bias function

$\bar{s}_{ij}$ the symmetric similarity measure of the additive similarity bias model

Appendix 2. Proofs of Propositions

Proof of Proposition 1. (a) This fact was established in the discussion following (4), which we restate here. Suppose $x_i > x_j$. Equation (4) states that $\beta_{ij} = \phi_1 - d_{ij} - x_i + x_j$. Then we have

$$\beta_{ij} = \phi_1 - d_{ij} - x_i + x_j < \phi_1 - d_{ij} - x_j + x_i = \phi_1 - d_{ji} - x_j + x_i = \beta_{ji}.$$

The first and last equalities are restatements of Equation (4); the inequality arises from $x_i > x_j$; and the next to last equality arises from the symmetry of $d_{ij} = d_{ji}$. So $\beta_{ij} < \beta_{ji}$, $i \neq j$.

(b) From the definition of clout and vulnerability, and after substitution (9), (10), and (12)

$$Clout_i = \sum_{j\neq i}^{J} \beta_{ji} = [\sum_{j\neq i}^{J} (\phi_1 - d_{ji} - x_j + x_i)]$$

$$= (J-1)\phi_1 - \sum_{j\neq i}^{J} d_{ji} - \sum_{j\neq i}^{J} x_j + (J-1)x_i$$

$$Vul_i = \sum_{j\neq i}^{J} \beta_{ij} = \sum_{j\neq i}^{J} (\phi_1 - d_{ij} - x_i + x_j)$$

$$= (J-1)\phi_1 - \sum_{j\neq i}^{J} d_{ij} - (J-1)x_i + \sum_{j\neq i}^{J} x_j$$

Subtracting the second equation from the first yields

$$Clout_i - Vul_i = \{(J-1)\phi_1 - \sum_{j\neq i}^{J} d_{ji} - \sum_{j\neq i}^{J} x_j + (J-1)x_i\} - \{(J-1)\phi_1 - \sum_{j\neq i}^{J} d_{ij} - (J-1)x_i + \sum_{j\neq i}^{J} x_j\}$$

$$= -2\sum_{j\neq i}^{J} x_j + 2(J-1)x_i = -2\sum_{j=1}^{J} x_j + 2Jx_i.$$

When $\sum_i x_i = 0$, $Clout_i - Vul_i = 2Jx_i$. So $x_i = \dfrac{Clout_i - Vul_i}{2J}$.

(c) [in Footnote 4] The additive similarity-bias model can be transformed into skew-symmetric model by properly decomposing its bias component.

$$\bar{s}_{ij} + r_i + c_j = \bar{s}_{ij} + \frac{(r_i + c_i)}{2} + \frac{(r_j + c_j)}{2} - \frac{(c_i - r_i)}{2} + \frac{(c_j - r_j)}{2} = s_{ij} - x_i + x_j, \tag{A4}$$

where $s_{ij} = \bar{s}_{ij} + \dfrac{(r_i + c_i)}{2} + \dfrac{(r_j + c_j)}{2}$, and $x_i = \dfrac{(c_i - r_i)}{2}$. Thus, we can interpret each brand's dominance effect (i.e., x_i) as the half of the difference between its column and row bias (i.e., $x_i = \dfrac{c_i - r_i}{2}$). Done.

Proof of Proposition 2. (a) Assuming utility-maximizing consumers under a linear budget constraint, we evoke the microeconomic property of demand homogeneity of degree zero in prices and income (because, for example, doubling all prices and incomes would have no real effect). This property arises directly from standard utility maximization under a linear budget constraint, and it can be shown to apply at the individual consumer level and at the market level. This property implies that, for any good, the sum of the own price elasticity and all of cross price elasticities equals minus the income elasticity. That is,

$$\beta_{ii} + \sum_{j \neq i} \beta_{ij} + \beta_{iI} = 0 .$$

Note that $\sum_{j \neq i} \beta_{ij}$ is our definition of vulnerability of brand i. Thus, we have (8), which we repeat below:

$$-\beta_{ii} = (\sum_{j \neq i} \beta_{ij}) + \beta_{iI} = Vul_i + \beta_{iI} . \tag{8}$$

Incidentally, for some categories it is approximately true that $\beta_{iI} = 0$,

(b) From the (4) and (8), we have

$$-\beta_{ii} = \sum_{j \neq i} \beta_{ij} + \beta_{iI} = \sum_{j \neq i} (\phi_{CPE} - d_{ij} - x_i + x_j) + \beta_{iI} = \sum_{j \neq i} \phi_{CPE} - \sum_{j \neq i} d_{ij} - \sum_{j \neq i} x_i + \sum_{j \neq i} x_j + \beta_{iI}$$

$$= (n-1)\phi_{CPE} - \sum_{j \neq i} d_{ij} - \sum_{j \neq i} x_i - x_i + \sum_{j \neq i} x_j + x_i + \beta_{iI}$$

$$= (n-1)\phi_{CPE} - \sum_{j \neq i} d_{ij} - nx_i + \sum_{j=1}^{n} x_j + \beta_{iI}$$

$$= (n-1)Density_i - nx_i + \beta_{iI} + \sum_{j} x_j \text{ , where } Density_i \equiv (\phi_{CPE} - \sum\nolimits_{j \neq i} d_{ij} / (n-1)) .$$

And, without loss of generality, if $\sum_{j=1}^{n} x_j = 0$, we have (9), which we repeat below:

$$-\beta_{ii} = (n-1)Density_i - nx_i + \beta_{iI} . \tag{9}$$

Appendix 3. Discussion of Alternative Specification Criteria

AIC (Akiake's Information Criteria) is defined as the deviance of a model plus two times the number of parameters estimated. It is designed to identify the "best" model, where best is taken to be the model that has the highest likelihood for additional data generated by the same process. Since AIC was developed in a frequentist and nonhierarchical modeling context, all parameters are estimated as fixed effects. The near equivalent in a Bayesian context is a model in which all parameters are given uninformative priors. For such a model, AIC and DIC agree.

However, most Bayesian models are hierarchical. That is, some parameters (such as the means of observed responses) are given distributions that are not uninformative and in fact are often functions of other, higher-level, parameters which also have distributions. In such cases, direct application of AIC is insufficient for two reasons. First, the number of parameters being estimated is unknown *ex ante* because the prior distributions for parameters introduce information sharing (and hence dependence) in their estimates. Thus the effective number of parameters being estimated is less than the number of parameters that appear in the model. Spiegelhalter et al. solve this problem using (15). Second, the criterion of best model is no longer clear, because a model's ability to predict depends upon which parameter estimates are held constant when making these predictions. In other words, models often contain nuisance parameters in addition to the parameters central to the purpose of the analysis (*focal parameters*). Spiegelhalter et al. recommend that the second term in (15) be calculated with φ including only focal parameters that are constant. If all focal parameters are given uninformative prior distributions, then they may be simply counted and AIC can be used instead of (15).

A popular alternative to AIC and DIC in a non-Bayesian context is BIC (Schwarz, 1978), and in a Bayesian context is the mean posterior of the likelihood of the data holding no parameters constant (the Bayes factors approach). Under certain conditions, these criteria will identify the true model from those estimated with probability one as the sample size grows to infinity, and no modeling purpose need be specified. However, DIC avoids assuming that one of the models considered is true, and instead chooses which model is best for a specific modeling purpose. Also, unlike BIC and Bayes factors, DIC is not subject to Lindley's paradox, so it can be used to compare models that make some use of weakly informative priors.

Table 1. Specification Decisions for the Model Components

Model Component	*Alternative Formulations*	*Equation*
1. Dependent variable distribution	(a) Poisson formulation	(12)
	(b) Negative-binomial formulation	(14)
2. Symmetric CPE structure	(a) Number of dimensions (*M*) for map	(4a), (5)[1]
3. Skew-symmetric CPE structure	(a) Leave x_i out of model	–
	(b) Freely estimate x_i for each brand *i*	(4)
	(c) Dominance-point formulation	(4), (6)[2]
	(d) Vector model formulation	(4), (6')
4. OPE structure	(a) Freely estimate β_{ii} for each brand *i*	(1)
	(b) Dominance -point formulation	(7)
	(c) Vector formulation	(7')
5. Intercept structure	(a) Freely estimated α_i for each brand *i*,	(1)
	(b) Make α_i proportional to horizontal axis	(10)

[1]The symmetric CPE structure is set in (3), (4a), and (5), assuming $a_{ij} = 0$ Equation (5) determines the structure underlying d_{ij}.

[2]The asymmetric CPE structure is set, in particular, by (4b) – the relationship $a_{ij} = -x_i + x_j$, $i \neq j$ — in Equation (4). Equation (6) and (6') specify the structure underlying x_i.

Table 2. Parameter Estimates for Beer Data

		Model 1	Model 2 (M=2)	Model 3: (M=2)
Intercepts	Budweiser	**1.28***	**1.36***	1.33*
	Miller	**-0.88***	**-0.96***	-0.93*
	Busch	**0.82***	**0.81***	0.93*
	Old Milwaukee	**-1.31***	**-1.35***	-1.35*
	Milwaukee's Best	**0.10**	**0.13***	0.02*
	Weight ω_3			**0.07***
Own-Price Elasticities (OPEs)**	Budweiser	**-4.33***	**-4.25***	-4.28*
	Miller	**-4.80***	**-5.14***	-5.21*
	Busch	**-3.19***	**-3.36***	-3.72*
	Old Milwaukee	**-3.43***	**-3.53***	-3.48*
	Milwaukee's Best	**-3.91***	**-3.89***	-3.67*
	Weight 1 υ_{21}			**1.08***
	Weight 2 υ_{22}			**2.16***
Cross-Price Elastities (CPEs)**	β_{ii} $[i = Budweiser, j = Miller]$	**0.63***	0.50*	0.40*
	β_{ii} $[i = Budweiser, j = Busch]$	**0.19**	0.08	0.16*
	β_{ii} $[i = Budweiser, j = Old\ Milwaukee]$	**0.03**	-0.04	-0.09
	β_{ii} $[i = Budweiser, j = Milwaukee's\ Best]$	**-0.21**	-0.02	0.12
	β_{ii} $[i = Miller, j = Budweiser]$	**0.27**	-0.21	-0.48*
	β_{ii} $[i = Miller, j = Busch]$	**-2.05***	-1.07*	-0.92*
	β_{ii} $[i = Miller, j = Old\ Milwaukee]$	**-0.77***	-0.89*	-0.91*
	β_{ii} $[i = Miller, j = Milwaukee's\ Best]$	**-0.27***	-0.90*	-0.86*
	β_{ii} $[i = Busch, j = Budweiser]$	**0.07**	0.60*	0.65*
	β_{ii} $[i = Busch, j = Miller]$	**0.48***	0.17	0.44*
	β_{ii} $[i = Busch, j = Old\ Milwaukee]$	**0.08**	0.13	0.25*
	β_{ii} $[i = Busch, j = Milwaukee's\ Best]$	**-0.07**	0.16	0.49*
	β_{ii} $[i = Old\ Milwaukee, j = Budweiser]$	**-0.06**	0.71*	0.57*
	β_{ii} $[i = Old\ Milwaukee, j = Miller]$	**0.36***	0.58*	0.62*
	β_{ii} $[i = Old\ Milwaukee, j = Busch]$	**0.90***	0.36*	0.42*
	β_{ii} $[i = Old\ Milwaukee, j = Milwaukee's\ Best]$	**0.93***	0.85*	0.53*
	β_{ii} $[i = Milwaukee's\ Best, j = Budweiser]$	**0.04**	0.74*	0.64*
	β_{ii} $[i = Milwaukee's\ Best, j = Miller]$	**0.27***	0.58*	0.53*
	β_{ii} $[i = Milwaukee's\ Best, j = Busch]$	**1.18***	0.39*	0.52*
	β_{ii} $[i = Milwaukee's\ Best, j = Old\ Milwaukee]$	**0.99***	0.85*	0.39*
Brand Power Parameters	Budweiser		0.13*	0.08*
	Miller		0.49*	0.52*
	Busch		-0.13*	-0.17*
	Old Milwaukee		-0.24*	-0.24*
	Milwaukee's Best		-0.25*	-0.18*
	Asymmetry Weight 1 [υ_{11}]		**0.56***	**0.41***
	Asymmetry Weight 2 [υ_{12}]		**0.21**	**0.96***
Coordinates for brands (θ_1)	Budweiser		**0.25***	**0.193***
	Miller		**0.31**	**-0.136***
	Busch		**0.09**	**0.135***
	Old Milwaukee		**-0.34***	**-0.196***
	Milwaukee's Best		**-0.32***	**0.003**
Coordinates for brands (θ_2)	Budweiser		0***	0***
	Miller		**0.68***	**0.61***
	Busch		**-0.52***	**-0.24***
	Old Milwaukee		**-0.07**	**-0.18***
	Milwaukee's Best		**-0.09**	**-0.19***
Number of Focal Params	**(in bold above)**	**30**	**21**	**14**

* The 95% credible interval excludes zero.

** The CPEs and Brand Power parameters in the last two columns and the OPEs and Intercepts in the last column are derived from the related underlying parameters (shown below them in bold in this table).

***Identification constraint.

Table 3 Other Parameter Estimates for the Beer Data

		Model 1	Model 2 (M=2)	Model 3 (M=2)
Covariates	γ_1	0.004	0.006	-0.002
	γ_2	0.006	0.003	0.008
	γ_3	-0.09*	-0.09*	-0.095*
Shrinkage Parameters*	σ_1		1.42 (1.18)	0.1834 (0.1275)
	σ_2		0.95 (2.04)	0.8191 (1.043)
NBD Model parameter	ρ	22.72*	20.96*	20.47*
DIC		17407	17403	17405

* σ_1 and σ_2 are the standard deviations of $\theta_{\cdot 1}$ and $\theta_{\cdot 2}$, respectively. These parameters are constrained to be nonnegative. We report (in parentheses) the posterior standard deviations of these estimates.

Table 4 Soft Drink Price Correlations[18]

Manufacturers	Products	Pepsi	Diet Pepsi	Mount Dew	Pepsi Free	Diet Pepsi Free	Lemon Lime Slice	Diet Lemon Lime Slice	Cherry Coke	Diet Cherry Coke	Coke	Diet Coke	Caff.- Free Coke	Caff.- Free Diet Coke	Sprite	Diet Sprite	Mr. PiBB	Dr. Pepper	Diet Dr. Pepper	7-Up	Diet 7-Up	Other
Pepsi Co.	Pepsi	1	1	1	1	1	0.7	0.7	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.2	-0.2	-0.1	0.1	0.1	0.1	0.1	0.2
	Diet Pepsi	1	1	1	1	1	0.7	0.7	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.2	-0.2	-0.1	0.1	0.1	0.1	0.1	0.2
	Mountain Dew	1	1	1	1	1	0.7	0.7	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.2	-0.2	-0.1	0.1	0.1	0.1	0.1	0.2
	Pepsi Free	1	1	1	1	1	0.7	0.7	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.2	-0.2	-0.1	0.1	0.1	0.1	0.1	0.2
	Diet Pepsi Free	1	1	1	1	1	0.7	0.7	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	-0.2	-0.2	-0.1	0.1	0.1	0.1	0.1	0.2
	Lemon-Lime Slice	0.7	0.7	0.7	0.7	0.7	1	1	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0	0	-0.3	-0.2	-0.2	-0.1	-0.1	0.3
	Diet Lemon-Lime Slice	0.7	0.7	0.7	0.7	0.7	1	1	-0.2	-0.3	-0.3	-0.3	-0.3	-0.3	0	0	-0.2	-0.1	-0.1	-0.1	-0.1	0.3
Coca-Cola Co.	Cherry Coke	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.2	1	1	1	1	1	1	0.7	0.7	0.1	-0.1	-0.1	-0.2	-0.2	-0.2
	Diet Cherry Coke	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	1	1	1	1	1	1	0.7	0.7	0.1	-0.1	-0.1	-0.2	-0.2	-0.2
	Coke	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	1	1	1	1	1	1	0.7	0.7	0.1	-0.1	-0.1	-0.2	-0.2	-0.2
	Diet Coke	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	1	1	1	1	1	1	0.7	0.7	0.1	-0.1	-0.1	-0.2	-0.2	-0.2
	Caffeine-Free Coke	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	1	1	1	1	1	1	0.7	0.7	0.1	-0.1	-0.1	-0.2	-0.2	-0.2
	Caff-Free Diet Coke	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	1	1	1	1	1	1	0.7	0.7	0.1	-0.1	-0.1	-0.2	-0.2	-0.2
	Sprite	-0.2	-0.2	-0.2	-0.2	-0.2	0	0	0.7	0.7	0.7	0.7	0.7	0.7	1	1	0	0	0	0	0	-0.3
	Diet Sprite	-0.2	-0.2	-0.2	-0.2	-0.2	0	0	0.7	0.7	0.7	0.7	0.7	0.7	1	1	0	0	0	0	0	-0.3
	Mr. PiBB	-0.1	-0.1	-0.1	-0.1	-0.1	-0.3	-0.2	0.1	0.1	0.1	0.1	0.1	0.1	0	0	1	0.2	0.2	0.1	0.1	0.2
Dr Pepper Co.	Dr. Pepper	0.1	0.1	0.1	0.1	0.1	-0.2	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0	0	0.2	1	1	0.9	0.9	-0.1
	Diet Dr. Pepper	0.1	0.1	0.1	0.1	0.1	-0.2	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0	0	0.2	1	1	0.9	0.9	-0.1
Seven-Up Co.	7-Up	0.1	0.1	0.1	0.1	0.1	-0.1	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	0	0	0.1	0.9	0.9	1	1	0.1
	Diet 7-Up	0.1	0.1	0.1	0.1	0.1	-0.1	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	0	0	0.1	0.9	0.9	1	1	0.1
	Other	0.2	0.2	0.2	0.2	0.2	0.3	0.3	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.3	-0.3	0.2	-0.1	-0.1	0.1	0.1	1

[18] "Other" refers to the composite of private labels, generics, and lesser known brands. Note the cases of correlation of 1 are identical prices to within one cent (except for one case of a 3-cent price difference in week 55 for Caffeine-free Diet Coke).

Table 5 Implied Soft Drink Cross-Price Elasticities

	Pepsi	Diet Pepsi	Mountain Dew	Pepsi Free	Diet Pepsi Free	Lemon-Lime Slice	Diet Lemon-Lime Slice	Cherry Coke	Diet Cherry Coke	Coke	Diet Coke	Caffeine-Free Coke	Caff-Free Diet Coke	Sprite	Diet Sprite	Mr. PiBB	Dr. Pepper	Diet Dr. Pepper	7-Up	Diet 7-Up	Other
Pepsi	-3.91	-0.40	0.50	0.71	-0.57	0.17	-0.27	0.14	0.44	0.95	-0.43	-0.25	0.65	0.69	-0.89	0.02	0.67	-1.16	0.63	0.01	0.63
Diet Pepsi	-0.36	-4.09	0.03	-0.30	0.81	-0.64	0.55	0.50	-0.26	-0.38	0.91	0.58	-0.63	0.38	0.39	0.04	0.59	0.16	0.67	-1.20	0.56
Mountain Dew	0.12	-0.39	-3.77	0.46	-0.52	0.31	0.00	0.46	0.68	0.09	-0.44	0.01	0.05	0.87	-0.72	0.49	0.83	-0.99	0.67	-0.33	0.50
Pepsi Free	0.59	-0.46	0.71	-3.94	-0.62	0.43	-0.20	0.24	0.70	0.56	-0.51	-0.19	0.58	0.78	-0.88	0.22	0.73	-1.16	0.63	0.09	0.55
Diet Pepsi Free	-0.43	0.91	0.00	-0.35	-4.35	-0.67	0.60	0.51	-0.29	-0.45	0.87	0.63	-0.69	0.34	0.56	0.06	0.55	0.33	0.62	-1.24	0.50
Lemon-Lime Slice	0.28	-0.57	0.80	0.67	-0.70	-4.63	-0.16	0.30	0.88	0.26	-0.62	-0.16	0.37	0.71	-0.89	0.44	0.66	-1.16	0.51	0.24	0.36
Diet Lemon-Lime Slice	-0.43	0.35	0.22	-0.24	0.30	-0.43	-4.12	0.81	-0.08	-0.45	0.28	0.96	-0.62	0.49	0.25	0.40	0.68	-0.02	0.66	-1.09	0.45
Cherry Coke	-0.31	0.02	0.40	-0.08	-0.08	-0.25	0.52	-3.72	0.09	-0.33	-0.05	0.52	-0.47	0.63	-0.21	0.53	0.80	-0.48	0.72	-0.91	0.48
Diet Cherry Coke	0.27	-0.47	0.89	0.66	-0.61	0.60	-0.10	0.37	-4.09	0.24	-0.52	-0.08	0.28	0.81	-0.81	0.44	0.76	-1.09	0.61	-0.03	0.46
Coke	0.97	-0.40	0.49	0.71	-0.57	0.16	-0.28	0.13	0.43	-3.93	-0.43	-0.25	0.66	0.68	-0.89	0.01	0.66	-1.16	0.63	0.01	0.63
Diet Coke	-0.35	0.94	0.01	-0.31	0.80	-0.66	0.51	0.47	-0.28	-0.37	-4.13	0.54	-0.63	0.37	0.37	0.01	0.57	0.15	0.66	-1.20	0.56
Caffeine-Free Coke	-0.42	0.37	0.22	-0.23	0.32	-0.44	0.94	0.80	-0.09	-0.44	0.29	-4.08	-0.62	0.49	0.25	0.38	0.68	-0.02	0.67	-1.09	0.46
Caff-Free Diet Coke	0.80	-0.52	0.58	0.85	-0.68	0.40	-0.32	0.12	0.59	0.79	-0.56	-0.30	-4.33	0.67	-0.97	0.09	0.63	-1.25	0.55	0.34	0.51
Sprite	0.04	-0.30	0.60	0.26	-0.45	-0.05	-0.01	0.43	0.33	0.01	-0.35	0.01	-0.12	-3.21	-0.69	0.28	0.93	-0.96	0.78	-0.61	0.61
Diet Sprite	-0.56	0.68	-0.01	-0.43	0.74	-0.67	0.73	0.56	-0.31	-0.58	0.63	0.75	-0.79	0.29	-4.86	0.15	0.49	0.71	0.53	-1.30	0.36
Mr. PiBB	-0.22	-0.23	0.63	0.10	-0.33	0.09	0.31	0.73	0.37	-0.25	-0.30	0.31	-0.30	0.69	-0.42	-4.19	0.76	-0.68	0.61	-0.63	0.38
Dr. Pepper	-0.13	-0.25	0.42	0.06	-0.39	-0.25	0.04	0.45	0.13	-0.15	-0.30	0.06	-0.31	0.78	-0.63	0.20	-3.07	-0.91	0.81	-0.82	0.60
Diet Dr. Pepper	-0.63	0.65	-0.08	-0.50	0.72	-0.74	0.67	0.50	-0.39	-0.65	0.61	0.68	-0.86	0.22	0.91	0.09	0.42	-5.25	0.45	-1.38	0.30
7-Up	-0.11	-0.11	0.31	0.01	-0.26	-0.35	0.08	0.43	0.03	-0.13	-0.16	0.10	-0.34	0.69	-0.54	0.11	0.87	-0.82	-2.80	-0.88	0.74
Diet 7-Up	0.53	-0.71	0.58	0.74	-0.86	0.66	-0.40	0.06	0.66	0.52	-0.75	-0.39	0.72	0.56	-1.10	0.13	0.51	-1.38	0.39	-5.22	0.31
Other	0.06	-0.05	0.31	0.11	-0.22	-0.32	0.03	0.36	0.06	0.04	-0.09	0.06	-0.20	0.68	-0.54	0.05	0.82	-0.81	0.91	-0.79	-2.46

Table 6. Estimates of Weights for the Soft-drink Data

	w_1 Asymmetry Weight	w_2 Own-price Elasticity Weight	w_3 Intercept Weight
Posterior Mean	0.58	2.27	4.17
Posterior Standard Dev.	0.01	0.07	0.18

Table 7. Clout and Vulnerability

	Pepsi	Diet Pepsi	Mountain Dew	Pepsi Free	Diet Pepsi Free	Lemon-Lime Slice	Diet Lemon-Lime Slice	Cherry Coke	Diet Cherry Coke	Coke
Vul/21	0.29	0.31	0.28	0.32	0.32	0.33	0.31	0.25	0.32	0.29
Clout/21	0.17	0.15	0.54	0.33	0.06	0.09	0.35	0.58	0.37	0.15

	Diet Coke	Caffeine-Free Coke	Caff-Free Diet Coke	Sprite	Diet Sprite	Mr. PiBB	Dr. Pepper	Diet Dr. Pepper	7-Up	Diet 7-Up	Other
Vul/21	0.30	0.31	0.32	0.19	0.32	0.28	0.12	0.30	0.12	0.29	0.14
Clout/21	0.11	0.37	0.05	0.72	-0.12	0.40	0.79	-0.36	0.74	-0.36	0.59

Figure 1. Modeling Asymmetric CPEs with a Dominance Brand

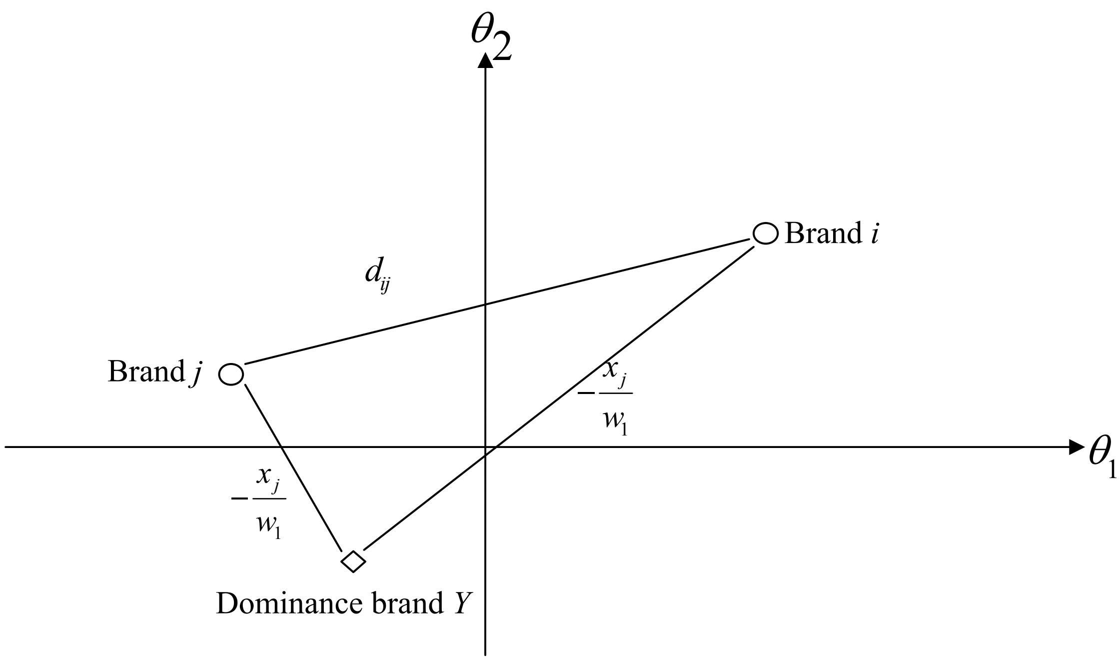


Figure 2. Modeling OPEs with a Least Vulnerable Brand

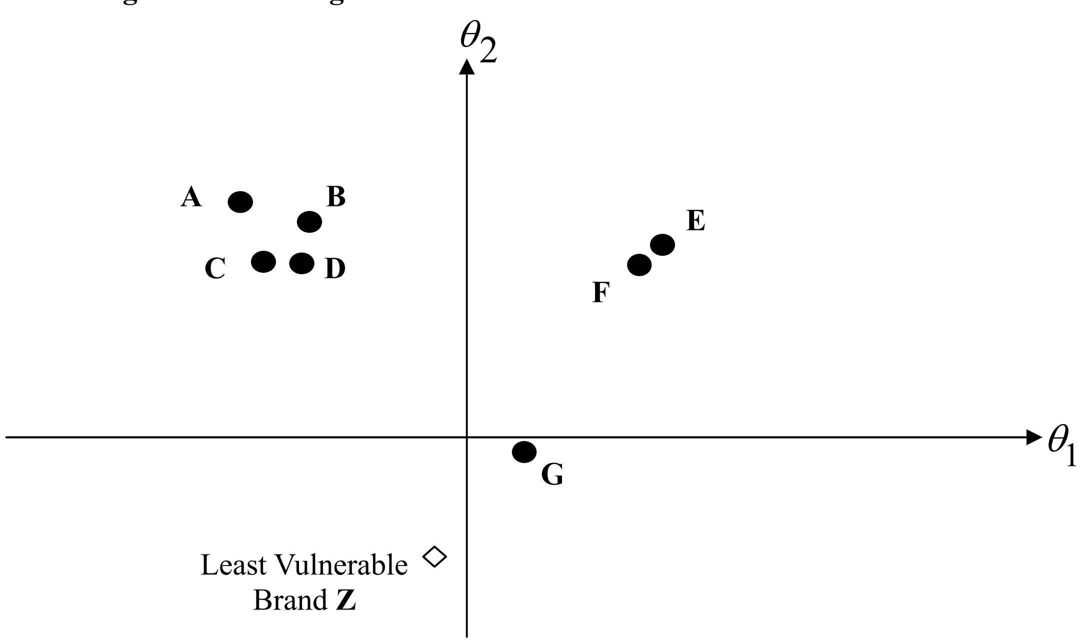


Figure 3. Plot of Beer Brand Locations in Model 3

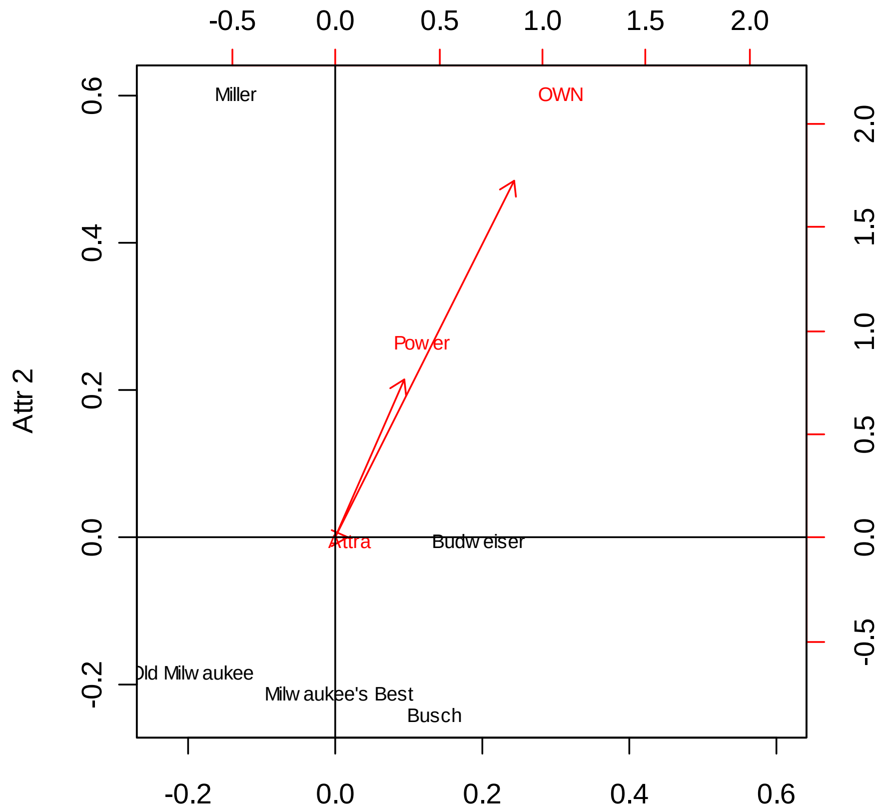
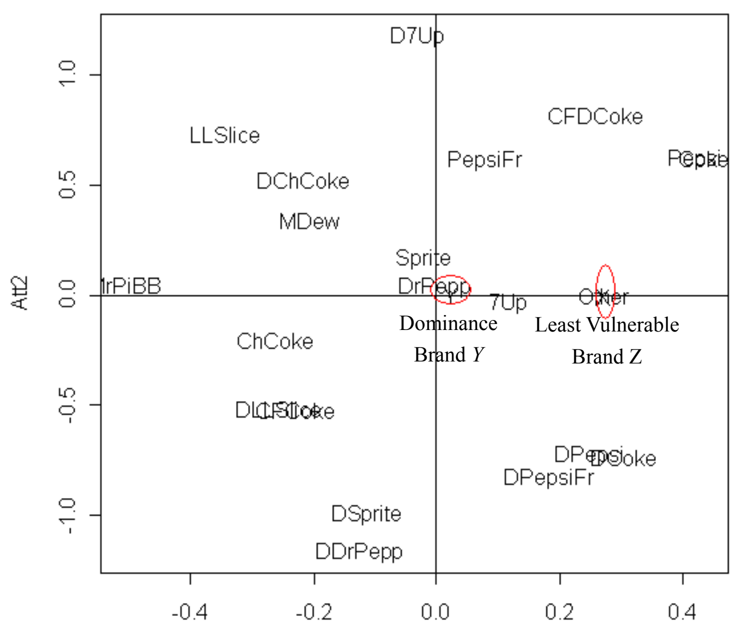


Figure 4 Two-Dimensional Map for Soft Drinks



Perceived Attractiveness