

Spatial Models in Marketing Research and Practice

Bart J. Bronnenberg*

19 February 2004

Forthcoming:
Applied Stochastic Models in Business and Industry

* Bart Bronnenberg is an Associate Professor at the John E. Anderson Graduate School of Management at UCLA. Correspondence to bart.bronnenberg@anderson.ucla.edu

Spatial Models in Marketing Research and Practice

Abstract

Spatial models formalize how cross-sectional observations relate to each other as a function of their spatial location. This paper discusses the generality of such models and how they are helpful to marketing practitioners in the description of marketing data, segmentation of markets, prediction of market behavior, and the pooling of data.

Keywords: Spatial statistics, autoregressive models, forecasting.

The analysis of data through spatial statistics is relatively new in marketing research. Spatial statistics are useful in describing cross-sections of markets, consumers, brands, or other units of analysis. These units of analysis (e.g., markets, consumers, or brands) can often be represented as a point or address in an underlying space (e.g., geographic, social, or attribute-based, respectively). Spatial analysis uses such addresses to directly place statistical structure on the cross-sections of data.

Spatial modeling contains similarities as well as differences with time-series analysis. For instance, a statistical formalization of local smoothness, i.e., that nearby observations are informative about the behavior of a process at a given point, is as useful in time series as it is in spatial analysis. Indeed, with equal justification insert the words “in space” or “in time” after the word “point” in the previous sentence. Further, spatial and temporal analysis both use similar fundamental blocks to build models, e.g., autoregressive or moving average specifications. A difference between spatial and temporal analysis is that space, in contrast to time, is not defined on a single dimension, doesn’t run in a single direction, and may not have constant units (e.g., geographical market segments of varying size). The purpose of this short essay is to present a –necessarily– selective synopsis of spatial models, to overview the nascent research on spatial models in marketing, to suggest several practical uses of these models, and to outline a subjective list of issues for future research in this field.

1 Modeling the cross-sectional dimension in marketing data

1.1 Spatial processes

A spatial process can be denoted as (see [1])

$$\{Y(\mathbf{z}) : \mathbf{z} \in D\} \tag{1}$$

where Y is the process of interest, observed at locations \mathbf{z} which are in turn defined on an L dimensional space D that is a subset of \mathbb{R}^L . The spatial process is said to be second-order stationary when the stochastic variables have a constant mean and when their covariance depends only on differentials of location, i.e.,

$$E(Y(\mathbf{z})) = \mu, \forall \mathbf{z} \in D \quad \text{and} \quad \text{cov}(Y(\mathbf{z}_n), Y(\mathbf{z}_{n'})) = C(\mathbf{z}_n - \mathbf{z}_{n'}), \tag{2}$$

where C is called the covariance function or covariogram. Interest in this essay is with spatial processes that are second-order stationary.

1.2 Measures of spatial proximity

In spatial models, Euclidean distance between two observational units is often taken as a relevant proxy for similarity. In a marketing context, Euclidean proximity helps account for several –possibly unobserved– variables including the organization of the distribution channel (e.g., distributors or retailers), certain types of demographics, climate factors, etc.

However, Euclidean distance is by no means the only interpretation of distance relevant to marketing. For instance, [2] combine demographic and zip code information to describe the location of consumers in geo-demographic space. Using these “locations” of consumers with the possibility that “nearby” consumers make similar choices, greatly improves prediction of choice compared to models where such information is ignored (i.e., models of spatially independent consumers). In an other interpretation of space, [3, 4] use the “address” of products in an attribute or perceptual space and take inter-product distance as a measure of similarity of (preferences for) alternative products. These perceptual distances are for instance used to make product recommendations [3].

Distances can be continuous or discrete. Discrete distances are used frequently when modeling social networks. For instance, [5, 6] define a social network with retailers as nodes and the degree of trade area overlap as links. A discrete measure of distance on such a network is simply the minimum degree of separation on a network (see e.g., [7]).

There are also applications wherein continuous space is partitioned into discrete areas. Various ad hoc methods to do this are readily available (the most widely used method being [8]). In other instances, predefined spatial tessellations are used such as zip codes, counties (FIPS codes), designated market areas (DMA's or advertising markets), states, or nations (see e.g., [5, 6, 9, 10]).

1.3 Incorporating the spatial organization of data in statistical models

Continuous measures of distance When space is continuous, the organization of observations is usually represented in a statistical model either through a polynomial function of the coordinates (i.e., by introducing a spatial trend over the coordinate system) or, more commonly, by introducing a spatial covariogram (see equation (2)). The simplest of such functions is a covariance function that depends on distance but not on direction (so called direction-invariant or isotropic covariograms). The covariance function can not be chosen freely but is subject to shape and range restrictions (see [1, 11]).

Discrete measures of distance To incorporate measures of discrete space in statistical models, the concept of a spatial lag (i.e., a spatial-shift function) is used. Spatial shifts can be coded in an $N \times N$ matrix \mathbf{W} whose row elements are zero if two locations are not contiguous and positive if they are (the diagonal of this matrix is zero, i.e., observation units are not a neighbor of themselves). Rows of this matrix are often standardized to add to 1 [12]. With a standardized matrix \mathbf{W} , the spatial lag of the $N \times 1$ vector of observations \mathbf{y} can be defined as the local averages $\mathbf{W}\mathbf{y}$. If \mathbf{W} is defined on the bases of direct neighbors, the spatial lag is of order 1. Alternatively, \mathbf{W} could express higher degrees of minimal separation (e.g., 2nd order spatial lags are formed by all observations that are two spatial units removed – excluding all circular paths). Graphically, spatial lags of increasing order can be visualized as widening concentric rings of observations around an observation of interest.

2 Models for cross-sections of data

2.1 Models with continuous distance measures

For brevity, discussion here is limited to Gaussian models. A simple model for an $N \times 1$ cross-section \mathbf{y} based on continuous distance can be stated as follows

$$\begin{aligned} \mathbf{y} &= \mathbf{x}\boldsymbol{\beta} + \mathbf{e}_t \\ \mathbf{e}_t &\sim \mathcal{N}_N(\mathbf{0}, \boldsymbol{\Sigma}), \\ \boldsymbol{\Sigma} &= f(\mathbf{Z}, \boldsymbol{\theta}) \end{aligned} \tag{3}$$

where \mathbf{y} contains the observations at locations $\mathbf{Z} \in \mathbb{R}^L$, \mathbf{x} is an $N \times P$ matrix of independent variables, $\boldsymbol{\beta}$ contains P exogenous effects, and the $N \times N$ covariance matrix $\boldsymbol{\Sigma}$ is a valid covariance function in \mathbb{R}^L with the coordinates \mathbf{Z} and parameters $\boldsymbol{\theta}$ as arguments. An example of a valid (isotropic) covariance function in \mathbb{R}^L is $\text{cov}(\mathbf{z}_1, \mathbf{z}_2) = \exp(-\|\mathbf{z}_1 - \mathbf{z}_2\|/\lambda)$, where $\|\mathbf{z}_1 - \mathbf{z}_2\|$ is the distance between locations, and the parameter $\lambda > 0$ is indicative of the strength of spatial dependence. Various methods for the estimation of the covariogram are discussed in [1, 13].

2.2 Models with discrete measures of distance

A descriptive model with discrete distance (or discrete proximity) measures can be constructed using a spatial autoregression

$$\begin{aligned} \mathbf{y} &= \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\xi} \\ \boldsymbol{\xi} &= \lambda \mathbf{W}\boldsymbol{\xi} + \mathbf{e} \end{aligned} \tag{4}$$

where \mathbf{y} is a $N \times 1$ vector of realizations of a spatial process, \mathbf{W} is a spatial shift operator as defined above, λ is a measure of spatial autoregression in ξ , $\mathbf{x}\beta$ are exogenous effects on \mathbf{y} , and ξ and \mathbf{e} are $N \times 1$ vectors of disturbances. Assuming that the e_n are distributed normally with mean 0 and variance σ_e^2 , the joint distribution of the observations is equal to

$$\mathbf{y} \sim \mathcal{N}_N \left(\mathbf{x}\beta, \sigma_e^2 (\mathbf{I} - \lambda\mathbf{W})^{-1} (\mathbf{I} - \lambda\mathbf{W}')^{-1} \right). \quad (5)$$

Estimation of this model through likelihood based methods is usually feasible (see [12]).

An alternative to expressing spatial dependence through joint distributions is to condition on contiguous observations and to use a conditional independence assumption [1, 14, 15], i.e.,

$$y_n = x_n\beta + \gamma \sum_{\forall m \parallel n} (y_m - x_m\beta) + e_n, \quad n, m = 1, \dots, N \quad (6)$$

where the \parallel operator denotes contiguity, and γ measures a neighborhood effect. This process is akin to a Markov process, because observations are conditionally independent after conditioning on spatial lags. Therefore, the process above is called a Markov Random Field. Under the assumption of normality of e_n , the model is a Gaussian Markov Random Field (GMRF). The joint distribution of the GMRF of equation (6) is considered in [14, 15]. Let $\mathbf{\Gamma}$ be an $N \times N$ matrix with 0's for non-contiguous and γ 's for contiguous combinations of n and m . Then,

$$\mathbf{y} \sim \mathcal{N}_N \left(\mathbf{x}\beta, \sigma_e^2 (\mathbf{I} - \mathbf{\Gamma})^{-1} \right), \quad (7)$$

provided $(\mathbf{I} - \mathbf{\Gamma})^{-1}$ is positive definite. Because spatial processes usually contain many complex and looping relations among the observations, the joint and the conditional approach are generally not the same (in contrast to temporal processes).

The distribution of the GMRF has the useful property in estimation that the precision matrix (i.e., inverse of the covariance matrix) has a simple parametric structure. This is especially useful when combining spatial and temporal data, because one can directly specify how observations are putatively related to their neighbors in time and in space. For additional details, the reader is referred to the seminal and more general statement in [14, 15].

3 Applications of spatial data and spatial models in marketing

3.1 Determining the geographical extent of markets or segments

Early applications of spatial models in marketing focused on methods to determine the physical boundaries of geographical markets. For instance, so-called gravity models were proposed ([16, 17]) to determine the perimeter of store trade areas. Such models predict the extent of store trade-area because of a negative relation between store patronage and distance to consumers. The geographical extent of markets has also been studied using travel time of sales agents (as opposed to that of consumers), in the literature on sales territory alignment [18, 19]. Yet a different set of applications determines the extent of market segments by clustering contiguous regions on the basis of consumer responses to price and advertising [10]. Finally, in the literature on spatial price discrimination, the extent of the market is implicitly determined by the minimum distance that prevents consumer arbitrage [20]. Namely, two regions can be considered separate markets if consumers do not travel from one to the other to arbitrage away price differences, i.e., a practical definition of a geographical market is implied by the cost and the benefits of consumer travel. The determination of geographic territories continues to be an important area in marketing research.

3.2 Descriptions of marketing policies using multi-market data

Spatial data can be used to point out how local market conditions affect marketing policy. For instance, multi-market price data have been used to study competitive pricing of products when prices depend on manufacturer delivery cost and these manufacturers have different locations [21].

Another literature focuses on how U.S. brand managers set local advertising and promotion levels. It is often argued that such marketing variables are set on the basis of existing popularity of the brand or category. For instance, to defend an existing profitable market, a manager might decide to advertise more heavily in markets where a brand is already a large share player. Alternatively, a different manager might focus her advertising dollars in markets where the brand is small (for discussion see e.g., [22, 23]). The use of such heuristics is difficult to infer empirically from single market time-series data because base-line shares change very slowly. However, there is ample cross market variation in base-line shares across markets. This variation can be used to relate cross-sectional differences in advertising and promotion levels to unobserved existing brand popularity.

A simple example to express this relation is a system of equations in which local sales \mathbf{y}_t ($N \times 1$) depend on local promotion levels \mathbf{x}_t ($N \times 1$) which in turn are based on existing (market or retailer specific) base-line outputs $\boldsymbol{\mu}$ ($N \times 1$) with a spatial structure, e.g.,

$$\begin{aligned}\mathbf{y}_t &= \boldsymbol{\mu} + \mathbf{x}_t\boldsymbol{\beta} + \mathbf{e}_t \\ \boldsymbol{\mu} &= \lambda\mathbf{W}\boldsymbol{\mu} + \boldsymbol{\nu} \\ \mathbf{x}_t &= \alpha_0\boldsymbol{\nu} + \gamma\boldsymbol{\mu} + \boldsymbol{\xi}_t\end{aligned}\tag{8}$$

with the ($N \times 1$) vectors of shocks \mathbf{e}_t , $\boldsymbol{\nu}$, and $\boldsymbol{\xi}_t$ all spherical [24]. The dependence of promotion levels \mathbf{x}_t on base-line sales levels $\boldsymbol{\mu}$, is captured by γ . The matrix \mathbf{W} is the spatial lag operator as defined above (thus the second equation of the system contains a spatial autoregressive model of base-line demands).

In terms of future research, it is noted that elasticities or regression effects obtained from cross-sectional analysis are often very different from those obtained using the time-series of the same data. Although speculative, it is likely the case that multi-market time-series data reflect demand and supply side decisions to different degrees along the spatial vs. the temporal dimension (as implied by equation 8). This issue clearly deserves further study. Ultimately these differences need to be explained or reconciled.

3.3 Diffusion and growth

Spatial models are also useful for deeper understanding of new product growth. Essential in new product growth models are formalizations of how products diffuse. For many products, diffusion is not simply a process that takes place across time. Rather, diffusion of new products in marketing also takes place across geography, customers, or market segments. Spatial models that take into account such geographical, or social network effects, may lead to new insights of how new products diffuse across multiple markets [6, 25–27] or across multiple social actors [28–30].

As an example, a simple model of consumer adoption of a new brand can be constructed by letting adoption at t depend on whether other consumers have adopted at $t - 1$. In other words, if $y_{nt} = 1$ (0) when consumer n has (not) adopted by t , a simple model of first time adoption with cross-sectional or social dependence can be defined by

$$\Pr(y_{nt}) = \begin{cases} \Phi(\alpha_n + \lambda\mathbf{W}_n\mathbf{y}_{t-1}) & \text{if } y_{nt-1} = 0 \\ 1 & \text{if } y_{nt-1} = 1 \end{cases},\tag{9}$$

where \mathbf{W}_n denotes the n^{th} row of \mathbf{W} , α_n is a base-line “hazard rate,” λ is the diffusion effect, and Φ is the cumulative standard Normal distribution. Alternative (often nested) versions of \mathbf{W} can express anything from uniform contagion effects to contagion only among those consumers that are sufficiently “close” [26].

An important point for further study is that spatial diffusion models or spatiotemporal diffusion models are likely informative about which users or markets should be targeted for the purpose of “seeding” a new product. Currently, the implications of spatial models for this purpose are underexplored (but see [24, 26])

An other interesting phenomenon in modeling spatial diffusion is that different agents in the distribution and communication channel have different spatial footprints (e.g., trade areas of chains are of a different spatial scale than advertising markets or DMA’s). It is likely that spatial diffusion processes occur and interact at different spatial scales simultaneously. Currently, the integration of these various processes in a comprehensive new product diffusion model is lacking. Such integration is needed to study which of these processes dominates (if any) and who (consumer, retailers, manufacturers) is ultimately responsible for the “contagion” effect that is often found in reduced form models of new product growth.

3.4 Interpolation, and prediction

A further application of spatial models in marketing is to use similarities for the purpose of interpolation, and prediction. For instance, suppose that observations are generated from the model in equation (3). Suppose however, that we do not observe part of the $N \times 1$ vector \mathbf{y} . For instance, partition equation (3) as follows

$$\begin{aligned} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} \\ \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} &\sim MNV \left(\mathbf{0}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right) \\ \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{21} \\ \boldsymbol{\Sigma}_{12} & \boldsymbol{\Sigma}_{22} \end{bmatrix} &= f(\mathbf{Z}, \boldsymbol{\theta}) \end{aligned} \quad (10)$$

where \mathbf{y}_1 is a $N_1 \times 1$ vector of observed realizations of the process, and \mathbf{y}_2 is a $N_2 \times 1$ vector of the unobserved (to the analyst) realizations of the process. The observations \mathbf{y}_1 can be used to estimate the parameters $\boldsymbol{\theta}$ which in turn populate the entire $N \times N$ matrix $\boldsymbol{\Sigma}$. Then, using standard results

from Multivariate Normal distributions, the prediction for $[y_2|y_1]$ is

$$E[y_2|y_1]_{N_2 \times 1} = \underbrace{x_2\beta}_{\text{regression mean}} + \underbrace{\Sigma_{21}\Sigma_{11}^{-1}(y_1 - x_1\beta)}_{\text{adjustment for spatial dependence}} \quad (11)$$

and the variance-covariance matrix of the predictions is

$$V[y_2|y_1]_{N_2 \times N_2} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}. \quad (12)$$

The genesis of this and other interpolation predictors is discussed in [1, 106 ff]. Spatial prediction or interpolation is called “Kriging” (named after the mining engineer D.G. Krige by G. Matheron, both of whom were concerned early on with spatial prediction problems).

A kriging predictor can be used to predict sales data in unsampled markets from a select sample of geographic markets [5]. Alternatively, the above predictor can also be used to make predictions of choices based on data from “similar consumers” [2]. A final application of spatial predictions can be found in Customer Relationship Management (CRM). Specifically, spatial methods have been used on sales data from a set of customers, to make predictions about the tastes of others in a product recommendation system [3]. An important advantage of the use of “similarity” in prediction is that it does not require one to be specific about the functional form of demographics that generates the behavior in consumers.

The success of spatial prediction depends on which similarity criterion (i.e., distance metric) is used. An opportunity for future research therefore lies in identifying which similarity criteria are useful for making predictions across different spatial units (e.g., city-blocks, zip-codes, counties, states, countries).

3.5 Combining cross-sectional information and pooling

A final application considered here is that one can use the data from “nearby” or similar observations to help make inferences about a particular observation if the data are sparse. For example, it is frequently the case that local data are relatively uninformative about, say, the effect of promotions (for the severity of this problem see [31]). When either the data or the promotion- and price-elasticities ([32]) are dependent across markets, a spatial model can help combining local estimators with their spatially related neighbors. This approach appears to work well in correcting for wrong signs (see [5]). “Local pooling” of information from “neighboring” observations is also used in

Geographically Weighted Regression (GWR). For a discussion see [33] and for an application to marketing see [34].

Cross-market pooling of data is done by several marketing research companies (see [35]). Cross-market pooling of data introduces the possibility that both promotion as well as sales data are generated by a common agent (the retailer). For example, retailers selectively accept promotion deals only from large share brands. Hence, across markets, promotion goes hand in hand with share or sales-per-capita. Obviously, this association combines both a demand effect, as well as the unobserved actions of the retailer. Filtering the spatial component from the data using a model like that in equation (8) helps to estimate promotion effects. A similar approach is used in [36] with a specific focus on demographic distance between stores to help estimating elasticities of shelf-space allocations, for which pooling across stores is necessary.

4 Conclusions

Spatial models offer a general statistical description of the cross-sectional dimension of data. Such models are currently being applied in marketing at both the aggregate level [5, 10, 19, 25, 26] as well as at the individual level [2, 6, 9]. Operationalizations of these models require different definitions of "space." Currently, no systematic attempt exists in marketing at studying spatial scaling issues or what drives spatial scale in a particular application.

A regularity that emerges from recent studies in marketing is that cross-sectional data are rarely independent realizations of a statistical process. However, many open questions about the origins of these cross-sectional or spatial dependencies remain. This is perhaps the chief opportunity for researchers interested in cross-sectional modeling. While the field of spatial statistics offers tools to account for spatial dependence, for the advancement of marketing practice, there is much value in studying the underlying spatial phenomena that give rise to the observed dependence.

5 References

1. Cressie N. Statistics for spatial data. Wiley and Sons: New York, 1993.
2. Yang S, Allenby GM. Modeling interdependent consumer preferences. *Journal of Marketing Research* 2003; 40 (August): 282-294
3. Moon S, Russell G. A spatial choice model for product recommendations. working paper, University of Iowa, 2003.
4. Carpenter G, Nakamoto K. Consumer preference formation and pioneering advantage. *Journal of Marketing Research* 1989; 26 (August): 285-98.
5. Bronnenberg BJ, Sismeiro C. Using multimarket data to predict brand performance in markets for which no or poor data exist. *Journal of Marketing Research* 2002; vol 39 (February): 1-17.
6. Bronnenberg BJ, Mela CF. Market roll-out and retail adoption for new brands of non-durable goods, working paper no. 372, Anderson Graduate School of Management, UCLA, 2003.
7. Wasserman S, Faust K. Social network analysis - methods and applications. Cambridge University Press, 1994.
8. Okabe A, Boots B, Sugihara K, and Chiu SN. Spatial tessellations. Wiley and Sons: New York, 2000.
9. Steenburgh TJ, Ainslie A, Engebretson PH. Massively categorical variables: Revealing the information in zip codes. *Marketing Science* 2003; 22 (1): 40-57.
10. ter Hofstede F, Wedel M, Steenkamp JB. Identifying spatial segments in international markets. *Marketing Science* 2002; 21 (2): 160:177.
11. Christakos G. On the problem of permissible covariance and variogram models. *Water Resources Research* 1984; 20: 251-265.
12. Anselin L. Spatial econometrics: Methods and models. Kluwer Academic Publishers: Boston, 1988.
13. Chen X, Conley TG. A new semiparametric spatial model for panel time series. *Journal of Econometrics* 2001; 1 (November): 59-83.
14. Besag JE. Spatial interaction and statistical analysis of lattice systems. *Journal of the Royal Statistical Society B* 1974; 36: 192-22
15. Besag JE. Statistical analysis of non-lattice data. *The Statistician* 1975; 24: 179-195
16. Huff D. Defining and Estimating a Trading Area. *Journal of Marketing* 1964; 28: 34-38.
17. Bell D, Ho TH, and Tang C. Determining Where to Shop: Fixed and Variable Costs of Shopping. *Journal of Marketing Research* 1998; 35: 352-369.
18. Lodish LM. Sales Territory Allignment to Maximize Profit. *Journal of Marketing Research* 1975; 12: 30:36

19. Skiera B, Albers S. COSTA: Contribution Optimizing Sales Territory Alignment. *Marketing Science* 1998; 17:3: 196-213
20. Anderson SP, de Palma A. Spatial Price Discrimination with Heterogenous Products. *the Review of Economic Studies*. 1988; 55: 573-592.
21. Greenhut ML. Spatial pricing in the United States, West Germany, and Japan. *Economica* 1981; 48: 79-86.
22. Kotler P. *Marketing management*. 11th edition. Prentice Hall: Upper Saddle River, NJ, 2003
23. Schultz DE, Martin D, Brown WP. *Strategic advertising campaigns*, Crain Books: Chicago, IL, 1984.
24. Bronnenberg BJ, Mahajan V. Unobserved retailer behavior in multimarket data: Joint spatial dependence in market shares and promotion variables. *Marketing Science* 2001; 20 (3): 284:299
25. Dekimpe MG, Parker PM, Sarvary M. Global diffusion of technological innovations: A coupled hazard approach. *Journal of Marketing Research* 2000; 37 (1): 47-59.
26. Putsis WP, Balasubramanian S, Kaplan EH, Sen SK. Mixing behavior in cross-country diffusion. *Marketing Science* 1997; 16 (4): 354-369.
27. Valente TW. *Network models of the diffusion of innovations*. Hampton Press: Cresskill NJ, 1995.
28. Siggelkow N, Van den Bulte C. Directions of growth: A test of positioning, resource, and social contagion arguments. working paper, University of Pennsylvania, Wharton School, 2002.
29. Strang D, Tuma NB. Spatial and temporal heterogeneity in diffusion. *American Journal of Sociology* 1993; 99 (3): 614-639.
30. Van den Bulte C, Lilien GL. Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology* 2001; 106 (5): 1409-1435.
31. Boatwright P, McCulloch R, Rossi PE. Account-level modeling for trade promotion: An application of a constrained parameter hierarchical model," *Journal of the American Statistical Association* 1999; 94: 1063-1073.
32. Hoch SJ, Kim BD, Montgomery AL, Rossi PE. Determinants of store-level price elasticity. *Journal of Marketing Research* 1995; 32 (February): 17-29.
33. Brunsdon C, Fotheringham AS, Charlton ME. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* 1996; 28: 281-298.
34. Mittal V, Kamakura W, Govind R. Geographic patterns in customer satisfaction: An empirical investigation. working paper, Duke University, 2003.
35. Bucklin RE, Gupta S. Commercial use of UPC scanner data: Industry and academic perspectives. *Marketing Science* 1999; 18:3: 247-273.
36. Van Dijk A, Leeflang PSH, Wittink DR. Similarity-based spatial methods to estimate shelf space elasticities, working paper, University of Groningen, 2003