SPATIAL MODELS IN MARKETING

Eric T. Bradlow
University of Pennsylvania

Bart Bronnenberg
UCLA

Gary J. Russell
University of Iowa

Neeraj Arora
University of Wisconsin

David R. Bell
University of Pennsylvania

Sri Devi Deepak
University of Iowa

Frankel ter Hofstede
University of Texas, Austin

Catarina Sismeiro
Imperial College, London

Raphael Thomadsen
Columbia University

Sha Yang
New York University

September 2004

This paper is based upon the discussions of the Spatial Models in Marketing seminar at the Sixth Invitational Choice Symposium, June 2004. Eric T. Bradlow, Bart Bronnenberg and Gary J. Russell served as co-chairs of the session.

ABSTRACT

SPATIAL MODELS IN MARKETING

Marketing science models typically assume that responses of one entity (firm or consumer) are unrelated to responses of other entities. In contrast, models constructed using tools from spatial statistics allow for cross-sectional and longitudinal correlations among responses to be explicitly modeled by locating entities on some type of map. By generalizing the notion of a map to include demographic and psychometric representations, spatial models can capture a variety of effects (spatial lags, spatial autocorrelation, and spatial drift) that impact firm or consumer decision behavior. Marketing science applications of spatial models and important research opportunities are discussed.

INTRODUCTION

A consumer's decision to adopt a new Internet service is affected by interactions with other consumers who live in the same postal code area (Bell and Song 2004). The utility weights used by consumers to determine satisfaction ratings vary geographically due to the impact of demographics and lifestyle on choice behavior (Mittal, Kamakura and Govind 2004). Retailers develop promotional policies based on the policies of other retailers in the same trading area (Bronnenberg and Mahajan (2001). Each of these is an example of a marketing context in which the spatial location of a decision-maker plays a key role in the choice process. In each instance, the spatial component creates a process in which the choice outcomes of one individual are related to the choice outcomes of other individuals.

The basic tool for constructing models of choice interdependence is the stochastic theory of spatial statistics (Anselin 1988, Ripley 1988, Cressie 1993, Haining 1997). Simply put, spatial models assume that individuals (or, more generally, units of analysis, such as postal codes) can be located in a space. Typically, responses by individuals are assumed to be correlated in such a manner that individuals near one another in the space generate similar outcomes. (In a competitive context, individuals might generate dissimilar (negatively correlated) outcomes.) The methodology can integrate complex spatial correlations between entities into a model in a parsimonious and flexible manner. Because spatial statistics was originally developed as a modeling tool in the physical and biological sciences, much of the older literature in spatial statistics emphasizes the use of a geographical map. In marketing, however, it is more appropriate to regard the space as any type of map – geographic, demographic or psychometric -- that describes the relationship among individuals (or units). By generalizing the notion of a map, we can define a spatial model as a stochastic model which uses known or unknown (latent) relationships among individuals (consumers, managers, retailers etc.) to predict the outcome of a decision process.

The goal of this paper is to present a brief overview of spatial models in marketing science. We begin by defining the elements of a spatial model: a map, a distance metric, and model of spatial effects. We emphasize that the researcher need not use geography in developing an interesting and useful model. We then consider issues of model specification and calibration. We conclude with suggestions for new research in spatial models.

CONSTRUCTING SPATIAL MODELS

The key assumption in the traditional marketing science literature is that the behavior of one individual is conditionally independent of the behavior of another individual. Although researchers in marketing science now routinely pool information across individuals to allow heterogeneity in parameters (Allenby and Rossi 1998), the underlying model is still constructed by assuming that each individual acts in isolation while making a decision. In contrast, spatial models posit that a richer understanding of behavior can be obtained by assuming the actions of different individuals are correlated. The key questions are why these relationships exist and how they may be modeled.

Typology of Spatial Models

All spatial models are constructed using a number a key components. In addition to an outcome variable y, we assume that the researcher has available a set of covariates X and a set of spatial relationships Z. Examples of X include product attributes, demographics and marketing mix elements. In some cases, X can include lagged values of y, both over time and across individuals. The identity of the variables in Z largely depends on the application. However, Z can be viewed as the location of each individual on some type of map. Unlike X, the location information in Z is typically assumed to be exogenous. (However, in some applications, map positions Z are treated as parameters and estimated in the course of the analysis (see, e.g., DeSarbo and Wu 2001).) Finally, a spatial model includes a vector of parameters $\mathbf{\Omega}$ that determines the relationships among y, X and Z. Formally, the task of the researcher is to study the decision process by computing a reasonable estimate of $\mathbf{\Omega}$ from the available information.

Using this notation, we can define two general types of spatial models of interest to researchers in marketing. Type I models, denoted by the notation f(y | X, Z, $\mathbf{\Omega}$), predict the choice outcome y, conditional on the X variables and the map locations. Type I models constitute the vast majority of models considered in regional sciences (Cressie 1993, Haining 1997) and in spatial econometrics (LeSage 1999; Anselin 2001, 2002). The simplest models in this area are spatial regression models with the general specification

$$y = X\beta + e, \qquad e \sim N(\mathbf{0}, \Sigma(Z,\theta)) \tag{1}$$

where $\Sigma(Z,\theta)$ denotes a properly specified covariance matrix in which the correlations between the responses of two individuals is monotonically decreasing in the distance between the individuals on the map. Because the errors in (1) have a spatial correlation pattern, the model can be used to predict the outcome variable of one individual at a specified location by using the known responses and locations of all other individuals. This approach, known as *kriging*, has been used in a marketing context to develop more accurate market-level estimates of brand sales (Bronnenberg and Sismeiro 2002). We consider more complex Type I models later in this article.

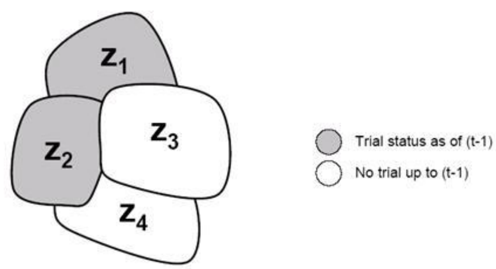
Type II models, denoted by the notation f(Z | X, y, Ω), reverse the logic of the modeling process. Instead of predicting outcome variables, we predict the locations at which certain outcomes occurred (Bradlow 2004). Models in this form are not generally discussed in the spatial statistics literature. However, marketing applications of Type II are clearly of interest. For example, consider the Path Tracker system developed by Sorensen Associates for category management applications (Murphy 2004, Sorensen Associates 2004). Using LPS (local positioning system) technology, the locations of a consumer's grocery cart in the store are recorded over time, providing information on the relationship between store layout and purchasing activity. Applied to these data, a Type II model would predict the consumer's path through the store, given information on purchases y and consumer characteristics X (Larson, Bradlow, and Fader 2004). Type II models are also potentially useful in the prediction of the sequence in which information is used in consumer behavior experiments (Wedel and Pieters 2000) and in market basket analysis (Manchanda, Ansari and Gupta 1999). In contrast to Type I models, Type II models do not generally make use of the tools of spatial statistics. For this reason, we restrict attention to Type I models in the remainder of this article.

Maps and Distance Metrics

Clearly, the most distinctive element of spatial models is the existence of a map. In regional sciences and spatial econometrics, the map is typically geographical in nature, indicating where the entities of interest (firms, consumers etc.) are located. The role of the map is similar to the role of time in time series models: spatial models typically assume that proximity on the map implies high correlation in the response variables. However, in contrast to time series models, the map is multidimensional – two or more dimensions – and can imply a rich variety of spatial relationships. For example, in ecological studies, spatial correlations are stronger in some directions (east-west) than others (north-south) due to prevailing wind patterns (Cressie 1993). Similarly, spatial models in marketing often seek to differentially weight information in modeling the correlation structure of response variables (Bronnenberg and Mahajan 2001, Yang and Allenby 2003).

The selection of an appropriate map is of singular importance in spatial modeling. In particular, it should be emphasized that a geographical map is not necessarily the best choice for marketing applications. The space in a spatial model represents additional variables (such as social networks, lifestyles or trading areas) that are not directly observed by the researcher, but which are likely to determine the response variable. For example, Yang and Allenby (2003) use both postal codes and demographics in defining the social network of consumers. Moon and Russell (2004) base their analysis solely on a pick-any map of consumer ideal points (a latent map), ignoring geographical location entirely. Clearly, the selection of a map implies an assumption about the variables that determine the relative similarities of individuals.

Once a map has been selected, a distance metric must be defined. Again, the researcher faces a number of choices. One common approach is to assume that the correlation structure is *isotropic*: it depends only upon distances on the map, not upon the direction. For example, in the model of equation (1), it could be assumed that the covariance between two consumers (a and b) is proportional to $\exp(-d(a,b)/\theta)$ where d(a,b) is the Euclidian distance between the consumers on the map and $\theta > 0$ is a parameter to be estimated. Cressie (1993) and Haining (1997) provide extensive discussions on the specification of covariance structures using both Euclidian and non-Euclidian (spherical) geometry.



$$C = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \longrightarrow W = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 1/2 & 1/2 & 0 \end{pmatrix}$$

Figure 1 – Definition of Neighborhood

In many applications, however, a continuous measure of distance is not appropriate. This typically occurs when the unit of analysis is a collection of many individuals such as postal code area, county or state. For example, Bell and Song (2004) model the probability that at least one individual in a postal code area has purchased groceries from netgrocer.com at a given time point (during a three-year period in which the Internet firm was being established in the marketplace). In Figure 1 (adapted from Bell and Song (2004)), we illustrate the use of a *continguity matrix* C to determine a *neighborhood*. Because these authors are interested in modeling the impact of social interactions in the adoption process, they define the neighbors of a target postal code as all other postal codes that share boundaries with the target. For example, postal code 1 is a neighbor of postal codes 2 and 3 only. The contiguity matrix C is row conditional: the pattern of ones and zeros identifies all postal codes that are neighbors of the postal code in the given row. (For reasons of model identification, the main diagonal of C is set to zero.) Prior to model specification, the contiguity matrix is usually converted into a row standardized *spatial lag matrix* W by rescaling the rows of C to sum to unity.

There are two reasons to prefer a neighborhood structure over a continuous distance metric, when scientifically appropriate. First, as in Bell and Song (2004), the logic of the particular application may dictate the use of a contiguity matrix. For example, a neighborhood structure is the appropriate choice to represent a social network. A hybrid approach is provided by Anselin (2002). He suggests that distance between economic agents be determined by counting the number of nodes separating the agents on a graph representing the social network In this context, the elements of the contiguity matrix C are integers and the W matrix is a function of the inverse of these integers. Yang (2004) argues for a generalized contiguity matrix C that permits the possibility that influence between individuals is asymmetric (e.g., opinion leaders impact opinion followers more than followers impact leaders).

Second, in any spatial analysis, there will always be some individuals located near the edges of the map. When a continuous distance metric is adopted, these individuals have less surrounding points than those in the interior of the map. This can lead to biases in model parameters and poor forecasts at the edges of the map. A compromise is to define a neighborhood as the K nearest individuals using a Euclidian distance metric. This definition also has the useful property that individuals located in a sparse region of the space will have the same number of neighbors as those in a dense region of the space.

(However, given the larger distances, the impact of these individuals may be smaller in magnitude.) In effect, the Euclidian measure of unit distance is allowed to be larger when in sparse regions of the map. Haining (1997) provides an extensive discussion of the statistical problems of edge effects and possible solutions.

Modeling Spatial Effects

Spatial models can represent three different types of spatial patterns. Two of these effects have already been briefly discussed. First, using the W matrix noted above, spatial models can capture *spatial lags,* the idea that the individuals are directly affected by the known decisions of other individuals (Yang and Allenby 2003, Bell and Song 2004). Models of this sort are of particular interest in applications, such as spatial econometrics, in which economic agents are known to interact during the choice process. Second, as shown in equation (1), spatial models can capture *spatially correlated errors*, the idea that important latent variables that drive purchase behavior can be inferred from consumer proximity on the map (Russell and Petersen 2000, Bronnenberg and Sismeiro 2002, Yang and Allenby 2003). (Similar work by Chintagunta, Dube and Goh (2004) shows how omitted variables induce a form of time and brand spatial dependence.) Models of this sort can be regarded as a statistical adjustment for missing variables that determine the response variable, but are not available to the researcher.

Third, spatial models can capture *spatial drift,* the idea that model parameters are a function of an individual's location on the map (Brunsdon, Fotheringham and Charlton 1998; Fotheringham, Brundson and Charlton 2002). Models of this sort can be regarded as a representation of unobserved heterogeneity in which the parameters (as opposed to the response variables per se) follow a spatial process. The theoretical justification for this type of model strongly depends upon the application. For example, Mittal, Kamakura and Govind (2004) argue that geography dictates the parameters of a satisfaction rating regression model due to differences in lifestyle and climate. Jank and Kannan (2003) argue that spatial patterning of utility model parameters can be expected because geographical location is a surrogate for the demographics that determine consumer tastes.

The book by Fotheringham, Brundson and Charlton (2002) provides an extensive discussion of a class of spatial drift models known as Geographically Weighted Regression (GWR). A Bayesian treatment of GWR models, designed to improve the statistical properties of GWR parameter estimates, is discussed by LeSage (2003). By exploiting a relationship between GWR and weighted maximum likelihood estimation, Deepak, Gruca and Russell (2004) develop a logit model with spatial drift. Intuitively, a GWR estimator can be regarded as an Empirical Bayes estimator where the prior for a particular individual is based upon the neighborhood structure of the contiguity matrix.

Formally, we can write a model specification which includes all three types of spatial effects by generalizing equation (1) as

$$y = \rho Wy + X\,\beta[Z] + e, \qquad e \sim N(\mathbf{0}, \Sigma(Z,\theta)) \tag{2}$$

where $\beta[Z]$ is a continuous function of the map coordinates Z and $\rho > 0$ is a scalar parameter. In this structure, ρWy represents spatial lag effects, $\Sigma(Z,\theta)$ represents spatial correlation effects, and $\beta[Z]$ represents spatial drift effects. The response variable y in equation (2) is typically assumed to be an observable outcome such as brand sales. Note that the outcome variable can also be ordinal in nature (such as preference scores) or constant sum data collected as part of allocation tasks; see Marshall and Bradlow (2002) for a unified computational approach to these types of models.

As noted by Anselin (2002), equation (2) can be adapted for choice modeling by replacing y with a continuous latent utility variable u

$$u = \rho W u + X \boldsymbol{\beta}[Z] + e, \qquad e \sim N(\mathbf{0}, \Sigma(Z,\theta)) \tag{3}$$

and by linking u to observed choice using a random utility theory argument. Although this generalization is simple in principle, it may not be appropriate for all marketing science applications. For example, the spatial lag term ρWu implies that the utility of one individual is influenced by the utilities of other consumers. This is clearly not the same as assuming that a given consumer's choice is influenced by the *observed* choices of other consumers (Anselin 2002). Although ρWu can be replaced by ρWy (thus, inducing a form of state-space dependence), model calibration must be approached carefully because the u values are correlated and u determines y.

Equation (2) can also be generalized to deal with cross-sectional time series data by allowing time to impact the model components as

$$y(t) = \rho W y(t) + X \boldsymbol{\beta}[Z,t] + e(t), \qquad e(t) \sim N(\mathbf{0}, \Sigma(Z,\theta)) \tag{4}$$

where the errors e(t) are correlated over time according to some stationary time series process. These models, known as *spatio-temporal* models, have been extensively studied in the biostatistics literature (see, e.g., Waller, Carlin and Xia 1997). Because spatio-temporal models provide considerable flexibility in capturing different types of dependence, they offer researchers in marketing science a promising direction for new work.

Statistical Issues

Collectively, the three types of spatial effects imply that the researcher must take into account interdependence in calibrating spatial models. Note that equation (2) can be rewritten as

$$y = [(I - \rho W)^{-1} X]\, \beta[Z] + v, \quad v \sim N(\mathbf{0}, (I - \rho W)^{-1} \Sigma(Z,\theta)\, (I - \rho W')^{-1}) \qquad (5)$$

where $v = \rho Wv + e$ can be interpreted as a spatially-lagged error structure. Even when the original errors e are not spatially correlated (i.e., $\Sigma(Z,\theta)$ is a diagonal matrix), all outcome variables y will be correlated due to the spatially-lagged error v. An analogous property holds for choice models constructed by replacing y by latent utilities u. In practical terms, interdependence means that calibration of a spatial model is considerably more complex than calibration of a traditional marketing science model because the standard assumptions of statistical conditional independence are inappropriate. Clearly, simultaneous equation estimation strategies must be used for model calibration. In this context, simulation technologies such as simulated maximum likelihood and Markov Chain Monte Carlo are apt to be attractive choices for the applied researcher (Tanner 1996, Train 2003).

In calibrating a spatial model, the researcher needs to be aware that the general model in equation (2) cannot be aggregated analytically without changing the model structure. Aggregation in this context refers to grouping of individuals (e.g., analyzing segments instead of consumers) or aggregating geographical areas (e.g., analyzing counties instead of postal codes). This general issue, known in the spatial statistics literature as *ecological fallacy* or the *modifiable areal unit problem*, implies that spatial effects present at one unit of analysis may not be observed at another unit of analysis

(Anselein 2001, 2002). Sismerio (2004) addresses this problem by developing a generalized spatial model which simultaneously incorporates the different spatial effects for different levels of analysis. Using simulated data, she shows that both local and large-scale effects can be recovered if the model is appropriately specified. In general, the researcher should select a scale for the model specification which coincides with the intended use of the model (Anselin 2002).

RESEARCH OPPORTUNITIES

Spatial models are interesting new tools for analyzing interdependence in behavioral outcomes. Here, we briefly discuss topics that are of particular interest to researchers in marketing science. Our intent is to highlight aspects of spatial modeling that present opportunities for future research.

Dimensionality

Spatial data present a number of challenges for the researcher. The most obvious characteristic of spatial data is the sheer amount of information that must be stored. Specialized data storage formats, data retrieval tools and data presentation software now exist in the form of *geographical information systems* (Rigaux, Scholl, and Voisard 2002). GIS software tools emphasize the use of efficient strategies for representing spatial information. For example, consider again the contiguity matrix C in Figure 1. Although this typically is a large N by N matrix (where N is the number of individuals in the analysis), the number of ones (denoting neighbors) in C is often very small relative to the number of zeroes. By recording only the locations of the ones in C, the amount of data storage required is greatly reduced. Statistical software (such as the Matlab spatial statistics toolbox) which allows for sparse matrix representations can significantly reduce the time needed to calibrate a spatial model (LeSage 1999).

Several approaches exist that address the dimensionality problem of spatial models. An alternative to expressing the spatial dependence through joint distributions (presented earlier in equation (2)) is to assume conditional independence of neighboring locations and to create a model based upon a Markov Random Field (Besag 1974, 1975). The main difficulty of this approach is the need for "severe restrictions on the available functional forms of the conditional probability distributions in order to achieve a mathematically consistent joint probability structure" (Besag 1974, p. 196). For example, Moon and Russell (2004) estimate a conditional autologistic model and constrain the pairwise relationship between locations to be symmetric. In order to avoid the computation of a mathematically intractable joint likelihood function, parameters are estimated using a pseudo-likelihood algorithm based upon the conditional probability distributions.

Recent developments in the estimation and inference of joint autoregressive models with the large data sets also limit the number of direct relationships among locations. The goal is to simplify computations and reduce memory usage of likelihood-based approaches (including maximum-likelihood and Bayesian estimation). For example, Pace and Barry (1997, 1999) provide algorithms to quickly compute maximum-likelihood estimates when the dependent variables (or its errors) follow a general spatial autoregressive process with few direct relationships. LeSage and Pace (2000) introduce the matrix exponential spatial specification (MESS) that relies on a specific spatial transformation of the dependent variable. Pace and Zou (2000) provide closed-form maximum-likelihood estimates for the particular case of nearest-neighbor spatial dependence (allowing only one other location, the nearest-neighbor, to directly affect each sampled location). Similarly, LeSage (2000) derives expressions for the conditional distributions central in Bayesian estimation of nearest-neighbor models that avoid complex matrix computations.

<u>Analysis of Marketing Policies</u>

Spatial data can be used to understand the geographical patterns of marketing variables. In this context, the X variables may be endogenous to the model. For example, Anderson and de Palma (1988) study regional patterns of price discrimination, taking into account delivery costs and manufacturer locations. Using information on the location of consumer residences, Thomadsen (2004) calibrates a choice model for banking services and develops recommendations for optimal placement of automatic teller machines. Bronnenberg, Dhar and Dube (2004) show that current brand shares for a consumer packaged goods product have a spatial distribution dependent on order of entry into the region and the (endogenously determined) regional levels of advertising expenditure.

A promising use of spatial models in marketing science is the correction of endogeneity in marketing mix response models (Bronnenberg and Mahajan 2001). Bronnenberg (2004) proposes that the spatial regression model of equation (1) be modified to

$$y = \tau + X(\tau)\,\beta + e, \qquad e \sim N(\mathbf{0}, \sigma^2 I) \tag{6}$$

where τ follows a spatial lag pattern and $X(\tau)$ is (partially) dependent on τ. In words, this model asserts that the base level of the response y exhibits a spatial pattern.

Moreover, this base level is used by marketing managers to set the observed marketing mix expenditures (such as advertising budgets) found among the $X(\tau)$ variables. A representation of this sort both corrects biases in the estimation of β and provides a structural view of managerial decision rules with respect to marketing mix variables. An analysis based upon equation (6) is particularly important if the goal of the research is to develop optimal policy recommendations for marketing managers.

Identifiability of Spatial Effects

Most researchers in marketing are interested in spatial models primarily as a means of understanding choice behavior. Yang and Allenby (2003) and Bell and Song (2004) use spatial models to measure the impact of social influence on choice behavior. Ter Hofstede, Wedel and Steenkamp (2002) use spatial priors (in a hierarchical Bayes analysis) to understand geographical dispersion of preference segments in Europe. Ter Hofstede (2004) extends this work by linking spatial segmentation to the *means-end chain* framework (in which abstract values lead to desired benefits which lead, in turn, to desired product features). In each of these examples, the structure of the spatial model suggests that specific behavioral mechanisms determine choice outcomes.

In this context, it is important to understand the spatial models – like most statistical models – may not be informative about the constructs underlying behavior. This can be seen by rewriting the spatial correlation model of equation (1) as

$$y = \rho W^* y + X^* \beta + e^*, \qquad e^* \sim N(\mathbf{0}, I) \qquad (7)$$

where $W^* = \rho^{-1}[I - T^{-1}]$, $X^* = T^{-1}X$ and T is the Cholesky factor of the spatially-correlated error covariance matrix $\Sigma(Z,\theta)$. That is, the spatial correlation process of equation (1) can be formally expressed in the form of a spatial lag process. Recall that we can interpret a spatial correlation process as the influence of missing variables that have a spatial distribution. In contrast, the usual interpretation of a spatial lag process is that individuals influence one another by some sort of social interaction. Although equation (1) and (7) are mathematically equivalent, they have entirely different behavioral interpretations.

The only way that a researcher can persuasively argue for one type of spatial mechanism versus another is to collect additional data that strongly favors one type of model specification. For example, Arora (2004) argues that spatial models could be an effective tools in studies of group decision-making (see, e.g., Arora and Allenby 1999; Aribarg, Arora and Bodur 2002). Because the data collection process would include *direct* observation of dyadic interactions (such as discussions between parent and child), the researcher would be justified in using a spatial lag model to model joint decision behavior. Again, the strong implication is that substantive knowledge of the application area must guide the use of spatial models. For this reason, joint work between behavioral and quantitative researchers should be encouraged. This work must include laboratory or field experiments that seek to confirm the hypothesized spatial structure.

CONCLUSION

Spatial models are a new research area in marketing science. Because spatial models allow for correlations in response variables across individuals, they represent an entirely new way of understanding decision processes. Major opportunities exist for researchers who understand the methods of spatial statistics and can craft specialized models to study substantive marketing issues.

REFERENCES

1. Allenby, Greg M. and Peter E. Rossi (1998), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89, 57-78.
2. Anderson S. P. and A. de Palma (1988), "Spatial Price Discrimination with Heterogeneous Products," *The Review of Economic Studies*, 55, 573-592.
3. Anselin, Luc (1988), *Spatial Econometrics: Methods and Models*, Dorddrecht: Kluwer Academic Publishers.
4. Anselin, Luc (2001), "Spatial Econometrics," in B. Baltagi (ed.), *A Companion to Theoretical Econometrics*, Oxford: Basil Blackwell, 310-330.
5. Anselin, Luc (2002), "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models," *Agricultural Economics* ,17 (3), 247-267.
6. Aribarg, Anocha, Neeraj Arora and H. Onur Bodur (2002), "Understanding the Role of Preference Revision and Concession in Group Decisions," *Journal of Marketing Research*, 39 (August), 336-349.
7. Arora, Neeraj (2004), "Joint Decision Making and Spatial Models," Position Paper, Session on Spatial Models in Marketing, Invitational Choice Symposium.
8. Arora, Neeraj and Greg M. Allenby (1999), "Measuring the Influence of Individual Preference Structures in Group Decision Making", *Journal of Marketing Research*, 36 (November), 476-487.
9. Bell, David R. and Sangyoung Song (2004), "Social Contagion and Trial on the Internet: Evidence from Online Grocery Retailing," Working Paper, Wharton School of Management, University of Pennsylvania.
10. Besag, Julian (1974), "Spatial Interaction and the Statistical Analysis of Lattice Systems," *Journal of the Royal Statistical Society B*, 36, 192-236.
11. Besag, Julian (1975), "Statistical Analysis of Non-Latice Data," *The Statistician*, 24, 179-195.
12. Bradlow, Eric (2004), "New Types of Spatial Data in Marketing," Position Paper, Session on Spatial Models in Marketing, Invitational Choice Symposium.
13. Bronnenberg, Bart J. (2004), "Spatial Models in Marketing Research and Practice," *Applied Stochastic Models in Business and Industry* (forthcoming).
14. Bronnenberg, Bart J. and Vijay Mahajan (2001), "Unobserved Retailer Behavior in Multimarket Data: Joint Spatial Dependence in Market Shares and Promotion Variables," *Marketing Science*, 20 (Summer), 284-299
15. Bronnenberg, Bart J. and Catarina Sismeiro (2002), "Using Multimarket Data to Predict Brand Performance in Markets for Which No or Poor Data Exist," *Journal of Marketing Research*, 39 (February), 1-17.

16. Bronnenberg, Bart J., Sanjay Dhar and Jean-Pierre Dube (2004), "Order of Entry and the Geographical Distribution of CPG Brand Shares: An Application of Endogenous Sunk Cost Theory," Working Paper, Anderson School of Management, UCLA.

17. Brunsdon, Chris, Stewart Fotheringham and Martin Charlton (1998), "Geographically Weighted Regression: Modeling Spatial Non-Stationarity," *The Statistician*, 47 (3), 431-443.

18. Chintagunta, Pradeep, Dube, Jean-Pierre, and Goh, Kim Yong (2004), "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models", *Management Science,* forthcoming.

19. Cressie, Noel A. C. (1993), *Statistics for Spatial Data*, New York: John Wiley and Sons

20. Deepak, Sri Devi, Tom Gruca and Gary J. Russell (2004), "Modeling Household Response Heterogeneity Using Spatial Measures of Similarity," Presentation at the MRSIG Special Session, AMA Educators' Conference.

21. DeSarbo, Wayne S. and Jianan Wu (2001), "The Joint Spatial Representation of Multiple Variable Batteries Collected in Marketing Research", *Journal of Marketing Research*, 38 (May), 244-253.

22. Fotheringham, A. S., C. A. Brunsdon, and M. Charlton (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, New York: John Wiley and Sons.

23. Haining, Robert (1997), *Spatial Data Analysis in the Social and Environmental Sciences*, New York: Cambridge University Press.

24. Jank, Wolfgang and P. K. Kannan (2003), "Understanding Geographical Markets of Online Firms Using Spatial Models of Consumer Choice," Working Paper, Maryland Business School, University of Maryland.

25. Larson, Jeffrey L., Bradlow, Eric T., and Fader, Peter S. (2004), "An Exploratory Look at Supermarket Shopping Paths", Wharton School Marketing Department Working Paper.

26. LeSage, James P. (1999), *Spatial Econometrics*, Morgantown: Regional Research Institute, University of West Virginia.

27. LeSage, James P.(2000), "Bayesian Estimation of Limited Dependent-Variable Spatial Autoregressive Models," *Geographical Analysis*, 32 (1),19-35.

28. LeSage, James P. (2003), "A Family of Geographically Weighted Regression Models," Working Paper, Department of Economics, University of Toledo.

29. LeSage, James P. and R.K. Pace (2000), "Using Matrix Exponentials to Explore Spatial Structure in Regression Relationships," Working Paper, Department of Economics, University of Toledo.

30. Manchanda, P., A. Ansari and S. Gupta (1999), "The Shopping Basket: A Model for Multicategory Purchase Incidence Decisions," *Marketing Science*, 18 (2) 95-114.

31. Marshall, P. and Bradlow, E.T. (2002), "A Unified Approach to Conjoint Analysis Models," *Journal of the American Statistical Association*, 97 (459), 674-682.

32. Mittal, Vikas, Wagner A. Kamakura and Rahul Govind (2004), "Geographic Patterns in Customer Service and Satisfaction: An Empirical Investigation," *Journal of Marketing*, 68 (July), 48-62.

33. Moon, Sangkil and Gary J. Russell (2004), "A Spatial Choice Model for Product Recommendations," Working Paper, Tippie School of Business, University of Iowa.

34. Murphy, Edward D. (2004), "Tracking Grocery Hot Spots," *Portland Press Herald*, Tuesday, January 27, 2004 edition (http://www.MaineToday.com/).

35. Pace, R.K. and Barry (1997), "Quick Computation of Regressions with a Spatially Autoregressive Dependent Variable," *Geographical Analysis*, 29 (1).

36. Pace, R.K. and Barry (1999), "Monte Carlo Estimates of the Log Determinant of Large Spatial Matrices," *Linear Algebra and its Applications*, 289 (1-3), March, 41-54.

37. Pace, R.K. and Dongya Zou (2000), "Closed-Form Maximum Likelihood Estimates of Nearest Neighbor Spatial Dependence," *Geographical Analysis*, 32 (2), 154-172.

38. Ripley, B. D. (1988), *Statistical Inference for Spatial Processes,* New York: Cambridge University Press.

39. Rigaux, Philippe, Michel Scholl, and Agnes Voisard (2002), *Spatial Databases with Application to GIS*, New York: Academic Press.

40. Russell, Gary J. and Ann Petersen (2000), "Analysis of Cross-Category Dependence in Market Basket Selection," *Journal of Retailing*, 76 (Fall), 367-392.

41. Sismeiro, Catarina (2004), "Microlevel Spatial Data: Challenges and the Effects of Macrolevel Structure," Position Paper, Session on Spatial Models in Marketing, Invitational Choice Symposium.

42. Sorensen Associates (2004), *Introducing Path Tracker*, Product Brochure, http://www.sorensen-associates.com/.

43. Tanner, Martin A. (1996), *Tools for Statistical Inference: Methods for the Exploration of Poseterior Distributions and Likelihood Functions*, New York: Springer-Verlag.

44. Ter Hofstede, Frenkel, Michel Wedel and Jan-Benedict E.M. Steenkamp (2002), "Identifying Spatial Segments in International Markets," *Marketing Science*, 21 (Spring), 2002, 160–177.

45. Ter Hofstede, Frenkel (2004), "On the Spatial Organization of Consumer Needs," Position Paper, Session on Spatial Models in Marketing, Invitational Choice Symposium.

46. Thomadsen, Raphael (2004), "Optimal Location Choice: Which Products Should Firms Offer?", Presentation, Session on Spatial Models in Marketing, Invitational Choice Symposium.

47. Train, Kenneth E. (2003), *Discrete Choice Methods with Simulation*, New York: Cambridge University Press.

48. Waller, Lance A., Bradley P. Carlin and Hong Xia (1997), "Structuring Correlation Within Hierarchical Spatio-temporal Models for Disease Rates," in Timothy G. Gergoire et al. (eds.), *Modelling Longitudinal and Spatially Correlated Data*, New York: Springer-Verlag, 309-319.

49. Wedel, Michel and Rik Pieters (2000), "Eye Fixations on Advertisements and Memory for Brands: A Model and Findings," *Marketing Science*, 19 (Fall), 297-312.

50. Yang, Sha and Greg M. Allenby (2003) "Modeling Interdependent Consumer Preferences," *Journal of Marketing Research*, 40 (August), 282-294.

51. Yang, Sha (2004), "Understanding the Interdependence in Consumer Preferences via Spatial Modeling," Position Paper, Session on Spatial Models in Marketing, Invitational Choice Symposium.