

Understanding Geographical Markets of Online Firms Using Spatial Models of Customer Choice

Wolfgang Jank, P. K. Kannan

Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742,
{wjank@rhsmith.umd.edu, pkannan@rhsmith.umd.edu}

As the online channel matures, many firms are finding that an understanding of how their online customers' preferences and choices vary across geographical markets can be very useful. In this paper, we propose a spatial multinomial model of customer choice and illustrate how spatial modeling of choices of online customers across geographical markets provides useful insights in the context of a product mix and pricing decision of an online book publisher. The spatial multinomial model specifically accounts for the spatial correlations among customer choices among different product forms—print and PDF. The estimation results obtained using data generated from an online experiment show that the spatial model accounts for the geographical variation in many of the unobserved effects possibly due to locational differences and price sensitivities. The resultant spatial decision maps provide useful predictions as to how purchase rates vary across geographical markets as a function of the price differential between product forms, with implications for targeting customers through local market advertising, direct marketing, and cross-channel promotion.

Key words: spatial model; mixed multinomial logit; pricing; digital products; Internet

History: This paper was received December 4, 2003, and was with the authors 8 months for 3 revisions; processed by Michel Wedel.

1. Introduction

It has been long recognized in retail marketing that geographical factors play an important role in impacting consumers' purchase behavior. Spatial data such as trade area demographics and competitive landscape are routinely used in retail setting to guide strategic and tactical decisions in pricing and promotion. Recently, several articles have examined how response variables/parameters of interest vary across geographical markets by incorporating spatial structure into models of aggregate demand/sales (Bronnenberg and Mahajan 2001, Bronnenberg and Sismeiro 2002, Bronnenberg and Mela 2004), customer evaluation (Ter Hofstede et al. 2002), and customer satisfaction (Mittal et al. 2004). Their findings seem to suggest that spatial data capture efficiently not only the geographical variations in supply side factors (store-related variations in locations, products, layout, service, order of entry, etc.) but also variations in demand side factors (geographical characteristics and customer characteristics impacting preferences, product/service adoption and consumption patterns) through variations in physical and psychological landscapes.

While supply side factors could be constant in the online context, the extant findings suggest the possibility that spatial data could still be useful in capturing important variations in the demand side

factors. For example, customers for online grocery retailers typically tend to be time-starved shoppers who value convenience and hail from mid/upscale suburbs. Thus, online retailers could target their promotion in specific geographical areas where customers are more likely to derive value from their services. As another example, online debt consolidators soliciting sales leads online have greater success in converting leads to actual sales when they incorporate spatial factors in their selection criteria—those geographical areas where property values have risen significantly and/or houses at the appropriate age provide good prospects for debt consolidation through refinancing of homes. In such instances, spatial data can act as a proxy for many geography-related variables—income, education, age, house size, property values, etc.—and capture some of the variances in online behavior due to these variables. As the examples illustrate, spatial data can be very useful for predicting customer behavior and for targeting customers (Carton 2003), specifically in multichannel marketing contexts, where channel integration efforts are becoming important (e.g., online coupon promotion for offline sales and direct marketing to online customers; see Lipke 2001).

Specific online businesses such as search engines are increasingly using spatial data to provide local content and advertisements when customers search

using keywords. Google, for example, is introducing a service that uses spatial data inferred from user's search terms and Internet protocol (IP) address to target local advertisements (Parker 2004). Many online newspapers regularly require users, when registering, to provide demographic and zip code information, which they use to determine the positive and negative impact on circulation, readership, and ad revenues (Gartner 2003). Recent research focused on online customer choice indicates that while accuracy of purchase rate predictions is generally low (Sismeiro and Bucklin 2004, Montgomery et al. 2004), accounting for geographical variance explicitly in a choice model can significantly improve the predictive ability of the model, again highlighting the importance of accounting for variance in demand side factors. Proprietary products such as Spectra (which provides details of demographic variation across geographic markets) and GeoVALS (which provides variations in psychographics—VALS segments—across zip codes) indicate that such variations can be quite significant. These suggest that spatial data can play an important role in explaining customer behavior in specific online contexts.

Online businesses do have the potential to collect significant information about their customers through transaction records, website visits, and cookie-based information. However, the use of such data is limited in the case of product/service categories with infrequent purchases or for businesses in the beginning stages of online operations. In addition, most data collected online tend to be behavioral data, as opposed to demographics and/or psychographics. Sometimes, consumers' privacy concerns may inhibit collection of such data online. In such instances spatial data of customers can be very useful to better understanding customers' choices. For example, Bronnenberg and Sismeiro (2002) point out that catalog and Internet companies with sparsely sampled cross-section of customer data can estimate a priori the hit rate of mailings to potential customers in certain zip codes using spatial data and methods (Bronnenberg and Sismeiro 2002, p. 3; Steenburgh et al. 2003). As the example suggests, using spatial data is particularly appealing in multichannel situations when traditional channels or direct channels are used along with the online channel to target customers. Even when a substantial amount of information is available about customers, spatial data can enhance the understanding of customers' behavior by accounting for, in a parsimonious manner, the unobserved variables or the variables that managers have failed to account for in a model.

In this paper, we propose a spatial multinomial model of choice and apply it to predict the online choice of customers in the context of a product line

decision of a book publisher. The publisher sells its books in print form both online and offline (mostly, direct) and is in the midst of introducing online a PDF version of all the titles it sells in print form. The key issue for the publisher is to understand the impact of PDF price relative to print form on the demand for both forms and on promoting the new form appropriately. The publisher has limited customer data given the infrequent nature of purchases of its products—mainly research books, reference material, and monographs—and we illustrate how spatial data is able to provide insights into customers' PDF purchases relative to print form. The PDF form provides the advantages of online content and immediate access to content after purchase, while a reliable download of such content may require a broadband connection at the customers' end. Broadband connection could be a function of broadband availability and income, both of which are known to vary across geographical regions (Federal Communications Commission 2004). Preference for the PDF form may also be dependent on how technology-ready a customer is, and that could be a function of education and income of customers, both geographically varying factors. Diffusion of PDF in local markets may vary significantly, depending on early adopters of technology.¹ Finally, price sensitivity of customers could vary, depending on their income. While many of these variables are generally specified at the individual customer level, we use spatial data as a proxy for such variables, because the publisher does not have any data on customer demographics or technology readiness.

The publisher conducted a series of pricing experiments online with its customers to understand their preferences and determine their willingness to pay for PDF. We apply our proposed spatial multinomial model of choice to analyze the customer data from one experiment (1) to understand how customers' preferences for PDF relative to print form vary across the publishers' various geographical markets, (2) to examine price sensitivities concerning the likelihood of PDF purchase as a function of customer location, and (3) to predict the demand of PDF versus print across all of its potential customers and derive strategic insights into the product mix and promotion. Our empirical analysis shows that spatial methods can increase the predictive power of the choice models significantly, thus highlighting their usefulness in online situations.

¹ Recent research suggests that innovation diffusion and contagion mechanisms are clearly impacted by geography (Van den Bulte and Stremersch 2004, Garber et al. 2004). The geographical impact is not only limited to customers but also includes competitive firms (Debruyne and Reibstein 2005) and channel members (see Kalnins 2004, Bronnenberg and Mela 2004).

The spatial model results help the publisher in three key interrelated decisions: (1) they allow a precise estimation of the revenue impact of a specific pricing policy of the PDF form relative to the print form (e.g., revenue impact of pricing PDF form at 50% the price of print) and, thus, aid in setting an overall PDF pricing policy; (2) they aid in forecasting the demand for titles in specific subject areas in the alternative forms—PDF and print—which forms the input for deciding on lot sizes for print runs for new titles; and (3) they help in identifying specific geographic markets where the PDF form should be promoted—through direct mail and specific price discounts to customers in the customer base.

In the next section, we provide an overview of our spatial model in relation to the extant literature. We then provide a description of the model and pointers on estimation and prediction methods. In the empirical application, we briefly describe the experiment and data and illustrate the application of the model and discuss the substantive insights derived from the model. Finally, we conclude with a discussion of the limitation and future research.

2. Model Preamble

The spatial model of multinomial choice that we propose is related to the generalized linear mixed model (Zeger and Karim 1991, Breslow and Clayton 1993). The model has a structure similar to that of the random parameter (mixed) logit model (Train 2003), except that the unobserved errors in customer choices are correlated spatially. Spatial correlation between customers' responses is accounted for through the use of customers' zip code information and associated longitude and latitude data to derive distance measures (which is similar to the approach of Mittal et al. 2004). As the response variable is *choice*, our model forms a natural extension to the spatial linear models used in extant marketing literature, all of which consider Gaussian response variables (e.g., Bronnenberg and Sismeiro 2002, Mittal et al. 2004). While the normal distribution may be a good approximation for aggregate choice data in many instances and, thus, one could still use extant models, such approximations can lead to poor results, especially when sample sizes and/or success probabilities are small (Brown et al. 2002). Thus, our approach of modeling choice data directly could be a better option.

Our model provides an alternative approach to incorporate spatial heterogeneity using zip codes in a choice model context, as compared to the approach of Steenburgh et al. (2003), who use zip code information directly as categorical explanatory variables. Our spatial formulation, however, allows the flexibility of predicting customer behavior for regions where

we have no data ("krieging"), which is similar to other spatial models (e.g., Bronnenberg and Sismeiro 2002, Mittal et al. 2004, Ter Hofstede et al. 2002). As with the other spatial models, our model also allows for spatially varying parameters, in addition to spatially varying intercept terms.

Our model also shares similarities with the approach of Rust and Donthu (1995), who estimate, in a choice model setting, geographically localized misspecification errors that could arise due to omitted variables being correlated with geographic location. They combine, in a two-stage approach, conditional logit model with a nonparametric density estimation (Donthu and Rust 1989, Donthu 1991), which is in contrast to our single-stage parametric method, which allows for predictions of unobserved locations. Their empirical results suggest that estimating geographically localized errors can appreciably reduce prediction errors, even when the predictor model is well specified, which augurs well for our approach too, as it can potentially eliminate many of the errors associated in predicting customer choice in the online context.

We use the Monte Carlo EM algorithm for estimation, which provides an elegant frequentist solution to the problem. This point also forms a key difference between some of the extant Bayesian approaches (Bronnenberg and Sismeiro 2002, Steenburgh et al. 2003). While a Bayesian approach would also be feasible in our context, our model requires estimation of only a small set of parameters, which makes it attractively parsimonious in a frequentist setting. The Monte Carlo EM algorithm has also been established as a reliable tool for this class of spatial models (see, e.g., Zhang 2002, 2003). Our Monte Carlo EM implementation, drawing ideas from Caffo et al. (2005), results in a more accurate, yet more efficient, use of simulation resources and at the same time accommodates Markov chain Monte Carlo techniques under the same setting.

3. Spatial Logit Model of Multinomial Choice

In this section we describe our model that allows for spatial dependence of consumer choices. We model choices directly by postulating a generalized multinomial logit model for a consumer's choice decision. To allow for possible spatial dependence of the individual choice decisions, we include a stochastic term with suitably chosen geographically varying correlation structure. One of the appeals of this model is its similarity with the well-known random coefficient (mixed) logit models (Train 2003).

Let $\mathbf{z}_i = (z_{i1}, z_{i2})$, $i = 1, \dots, N$ denote the spatial coordinate or location of the observed response y_i . In our application, z_{i1} and z_{i2} denote the longitude

and latitude, respectively, of the location associated with y_i . Longitudinal and latitudinal information can be extracted from a transaction's zip code, as we describe later in our application.

Let $\mathbf{u}_i = \mathbf{u}(\mathbf{z}_i)$ denote a vector of stochastic terms associated with location \mathbf{z}_i . Since \mathbf{u}_i is unobserved we refer to it as a random effect. Our model assumes that the response variable takes on only one of J values, $y_i \in \{1, 2, \dots, J\}$, and that, conditional on \mathbf{u}_i , the y_i s are independent realizations of a multinomial random variable; that is,

$$y_i | \mathbf{u}_i \sim \text{Multinomial}(\pi_{i1}, \pi_{i2}, \dots, \pi_{iJ}), \quad (1)$$

where $\pi_{ij} = \text{Prob}(y_i = j | \mathbf{u}_i)$ is the probability of choosing category j , ($j = 1, \dots, J$), conditional on the random effect, \mathbf{u}_i . Since these probabilities are specified conditionally on \mathbf{u}_i , they are associated with location \mathbf{z}_i . Without loss of generality, let J denote the *baseline* category. We model the logit of π_{ij} as

$$\log\left(\frac{\pi_{ij}}{\pi_{iJ}}\right) = \mathbf{x}_i^T \boldsymbol{\beta}_j + \mathbf{x}_i^T \mathbf{u}_{ij}, \quad j = 1, \dots, J-1, \quad (2)$$

where $\boldsymbol{\beta}_j$ is a $p \times 1$ vector of unknown logit parameters associated with category j and \mathbf{x}_i is $p \times 1$ a vector of known covariates. Let $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \dots, \boldsymbol{\beta}_{J-1}^T)^T$ denote the complete logit parameter vector, and let $\mathbf{u}_i = (\mathbf{u}_{i1}^T, \dots, \mathbf{u}_{iJ-1}^T)^T$, where $\mathbf{u}_{ij} = (u_{ij1}, \dots, u_{ijp})^T$. Thus, if we assume zero mean for the \mathbf{u}_{ij} s, then model (2) implies that the logit parameters vary across geographical regions according to a distribution that is centered at $\boldsymbol{\beta}$. Note that in the logit model of Equation (2) the covariates \mathbf{x}_i are the same across all categories, whereas the effect $\boldsymbol{\beta}_j$ varies across categories. Thus, it is a generalized multinomial logit formulation, rather than the more familiar conditional logit formulation (McFadden 1981, Guadagni and Little 1983).

Our modeling objective is to allow for the (possible) spatial correlation between two choice observations. We note that the model in (2) implies that for two distinct locations \mathbf{z}_i and $\mathbf{z}_{i'}$, if their associated random effects \mathbf{u}_i and $\mathbf{u}_{i'}$ are correlated, so are the corresponding choices y_i and $y_{i'}$. We incorporate spatial correlation into (2) by modeling the distribution of the random effects suitably. We use the multivariate normal distribution, as it allows, among all parametric candidates, for greatest flexibility in choosing different correlation structures. We assume that the vector $\mathbf{u} = (\mathbf{u}_1^T, \dots, \mathbf{u}_N^T)^T$ follows a multivariate normal distribution with mean zero and a covariance structure $\boldsymbol{\Sigma}$ such that the correlation between two elements of \mathbf{u} decays with the geographical distance between the associated two observations.

We make several standard assumptions on $\boldsymbol{\Sigma}$ to facilitate the modeling of the spatial correlation (see

Cressie 1993). First, we assume that the variance of the random effects is homogeneous, which allows decomposition of the covariance into

$$\text{Cov}(u_{ijk}, u_{i'j'k'}) = \tau_j^2 + \sigma^2 r(\mathbf{z}_i, \mathbf{z}_{i'}; \boldsymbol{\alpha}). \quad (3)$$

The first term on the right hand side, τ_j^2 , captures pure modeling error and is often referred to as the *nugget effect*. Notice that we allow τ_j^2 to be alternative-specific. Spatial association is captured by the function $r(\cdot)$, scaled by σ^2 and dependent on a vector of spatial parameters $\boldsymbol{\alpha}$. We assume second-order stationarity and isotropy for $r(\cdot)$, which ensures translation and rotation invariance of the correlation structure. Under these assumptions, the correlation between any two points y_i and $y_{i'}$ is a function of the (Euclidean) distance $d_{i,i'} = \|\mathbf{z}_i - \mathbf{z}_{i'}\|$ separating them. Thus, we have:

$$\text{Cov}(u_{ijk}, u_{i'j'k'}) = \tau_j^2 + \sigma^2 r(d_{i,i'}; \boldsymbol{\alpha}). \quad (4)$$

A variety of different correlation functions can be used. We adopt a very flexible class of correlation functions called the *Matérn* family, which defines

$$r(d; \boldsymbol{\alpha}) = \frac{(d/\alpha_1)^{\alpha_2}}{2^{\alpha_2-1} \Gamma(\alpha_2)} K_{\alpha_2}(d/\alpha_1), \quad (5)$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^T$, $\alpha_1, \alpha_2 > 0$, and K_α is the modified *Bessel* function of (possibly fractional) order α . Note that α_1 is often referred to as the spatial scale parameter while α_2 is the shape parameter. Setting $\alpha_2 = 0.5$ recovers, as a special case, the *exponential* correlation function

$$r(d; \alpha_1) = \exp(-\alpha_1 d), \quad (6)$$

which implies that the correlation between two observations decays exponentially with the geographical distance between the associated two observations. We discuss estimation and prediction with the spatial choice model in the appendix.

4. Empirical Application

4.1. The Problem Context

The main objective of the publisher is to determine the specific pricing policy for the PDF form relative to print as it is introducing the form for the first time. Although each title could be priced differently relative to print price, to make the implementation of the PDF pricing policy straightforward, the publisher has decided that the price of the PDF form is to be set at a specific percentage of the print price for all titles. The relative price of PDF form has important implications for overall sales. A low relative price for PDF can increase the market size as visitors who had never bought any item from the publisher before could purchase titles attracted by the low price. At the same

time, a low relative price for PDF can increase the cannibalization of the print form, thereby impacting overall sales negatively. Since the publisher knows the distribution of its sales geographically from its customer database, estimates of customer preferences for print versus PDF forms and their price sensitivities across geographic markets can help the publisher estimate accurately the aggregate sales impact of the different relative pricing levels of PDF and, thus, determine the optimal price level that maximizes the overall sales.

A related issue is to determine the relative demand in units for print and PDF forms at the optimal price level, since this would impact the lot size for print runs, especially for the new titles. For example, if the estimated demand for a title in print form before the introduction of the PDF form is 50,000, its estimated demand with the PDF form priced at a lower level could be lower at 60% to 80%, depending on the relative price levels, the geographic distribution of demand for the title (or subject category), and the relative preferences and price sensitivities of customers in the different geographical markets. Estimating the demand for the different forms accurately can help the publisher in planning appropriate print runs and managing print inventory better.

Finally, the publisher has committed itself to encourage and promote the electronic form as the product of the future and improve its quality. Consequently, the publisher wants to promote the PDF form specifically through price discounts for the PDF form in those markets where their purchase probabilities are low. In this context, estimates of customers' relative preferences and price sensitivities in different geographical markets can help identify specific customers in the customer database for targeted promotion, using discount coupons mailed to their addresses. Such price discrimination using direct mail coupons can help promote the PDF form appropriately without having to extend this offer to all customers who visit the publisher's website.

Given the importance of determining the distribution of relative purchase probabilities across markets, the publisher conducted an online pricing experiment to gauge customers' interest in the PDF form and their purchase of the PDF form at different price levels relative to print price. The experiment was conducted online with the visitors to the publisher's website, rather than with the existing customer base, because the existing customer base does not contain visitors who have never bought at the website. Based on the relatively sparse experiment data on visitors, who hail from geographically diverse markets, the publisher is able to extrapolate customers' preferences and price sensitivities across geography using our spatial model of choice.

The experiment targeted visitors who made at least one click upon arriving at the publisher's website, thereby showing an interest in a specific title. This was necessary because only one-fifth of the 2,500 titles carried by the publisher were included in the pricing experiment. When a user clicked at the title information page upon arriving at website or in the free browse section for the title for which the publisher had a PDF version, he/she was intercepted with a chance of 1 in 4. These respondents had showed enough interest in the title to browse through the book information (summary) or content and could have moved on to purchase a print copy or could have left the website without purchasing the title. However, they could be classified as potential customers based on their website behavior showing some involvement with the title. Since the publisher wanted to focus on the population of potential online customers for its titles, sampling these customers at random and analyzing their choice behavior could provide insights into the potential customers' relative preference for the print and PDF forms and price sensitivities.

When a respondent was intercepted, he/she was provided the information on the PDF form along with its price. The respondent was then provided the option to buy the PDF form or decline the offer. They were also offered the chance to purchase the print form at the regular print price. Thus, the respondent either bought the print form, the PDF form, or none. After the experiment treatments, all of the respondents were debriefed, irrespective of whether or not they had purchased a product. Most of them also participated in a survey regarding their perceptions of the quality of PDF form vis-à-vis print form in return for an incentive (free shipping or a purchase discount or a free book, depending on whether they purchase a print, PDF, or none). All respondents were tracked by their IP addresses and their click-stream data collected for significant actions.

In the choice experiment, PDF prices were set at 110%, 100%, 75%, 50%, 25%, and 0% relative to the print book price. The PDF prices were displayed in absolute dollar terms. For each title, the price of the print form was held constant, while the PDF price was set at one of the above six levels relative to print form. The PDF prices of the 500 titles were set so that, within a subject category, each of the print book price levels was associated with each of the six different PDF price levels. For example, the PDF versions of titles in the education category, the print forms of which were priced at \$30, were priced at levels ranging from 110% to 0% of the print book price, with generally equal shares for all the price levels. The PDF prices were also assigned so as to achieve similar share balance across the number of visits for

the specific titles (a measure of their popularity and potential sales). Thus, among the \$40 “most popular” books, the PDF prices were assigned so as to cover all the price ranges (110% to 0%). This ensured that the experiment results were valid for all categories of subjects, price levels of print books, and for all levels of potential sales, in addition to minimizing the correlations between prices. The experiment was conducted for a total of three weeks. Given that the experiment involved only one-fifth of the total titles carried by the publisher, was conducted for only three weeks and, thus, not all geographical regions would have been represented in the data, the predictive ability of the spatial model to estimate customer preferences and choices in zip codes where no data was available would be very useful.

4.2. Impact of PDF Price on Aggregate Sales and Unit Sales

To determine the impact of the introduction of PDF form at various price levels on the overall sales of the firm, we use the following procedure:

1. Run the spatial model of choice on the experimental data. Using these estimates, predict the relative probabilities of (1) purchasing print, (2) purchasing PDF, and (3) no purchase for each zip code (location) for each combination of different price levels of print and different relative prices of PDF (110%, 100%, 75%, etc.).

2. Determine from the customer data base for the firm:
 - (a) average unit sales in past two years for each zip code,
 - (b) the unit prices paid for each individual print title purchased in each zip code,
 - (c) the overall distribution of price paid for print books over the last two years, and
 - (d) using the above distribution as prior, and actual prices paid for print books in each zip code, obtain the posterior distribution of price paid for print books for each zip code.

3. Assume that the market remains the same size when the PDF form is priced at 100% of the print price (that is, some existing print customers may switch to PDF, but no new customers enter the market). This is a justifiable assumption, given that the quality of PDF is much lower than print and is generally valued lower than print. Therefore, we can equate the probability of making a purchase in a zip code (print or PDF, when PDF is priced 100% relative to print) to the current sale of print in that zip code. Given this phenomenon, the probability of making a purchase (print or PDF) at relative PDF prices lower than 100% provides the combined impact of cannibalization and market expansion at those specific PDF prices.

4. Using the above relative probabilities, the average unit sales in each zip code, and the posterior distribution of prices paid for print books in each zip code, determine the projected sales of print and PDF forms at various levels of print book price and relative prices of PDF combinations.

5. Aggregate the projected sales for all zip codes with positive sales in the past two years to arrive at the overall sales under different relative price levels of PDF and for different subject categories.²

4.3. Data Description, Estimation and Analysis

We focus on the U.S. market data from the experiment consisting of 186 print purchases, 129 PDF purchases, and 384 no purchases with print prices ranging from \$5 to \$94.50 (mean = \$36.41) and PDF price, relative to print, ranging from 0% to 110% (mean 60%). Approximately 17% of the data (120 observations) was set aside for validation. The explanatory variables included in the model consisted of price variables—price of print form, price of PDF form as a percentage of print, and the interaction between these two variables (which is a measure of absolute price of PDF), as well as eight subject category dummies to represent the categories of education, physical sciences, research issues, agriculture, food and nutrition, industry and economics, natural resources, public policy, and others.

Table 1 provides the details of the estimates from our spatial multinomial model of choice (Model 1) along with the estimates from a random parameters (mixed) logit model (Model 2) and a multinomial logit fixed effects model (Model 3). In Model 1 spatial effects are incorporated in the intercept as well as the price coefficients, as discussed in the model section. Model 2 incorporates random effects for the same parameters without the spatial effects. Model 3 is the pure fixed effects model with no random effect and no spatial effects. In all models, the base alternative was no purchase and, thus, the estimates of purchasing PDF and print forms are with respect to this base. Thus, the parameter β_1 (subscript 1 indicates PDF) represents the intercept in determining the relative purchase probability of PDF form, while parameter β_{pr1} indicates the impact of print price on PDF purchase probabilities, parameter β_{per1} the impact of PDF price as a percentage of print price on PDF purchase probabilities, and β_{pdf1} the impact of PDF absolute price on PDF purchase probabilities. The parameters with subscript 2 represent the corresponding impacts on the purchase probabilities of the print form. Since the experiment was set up to

² Since book sales in some zip codes tend to be sporadic over time, we chose a two-year horizon to capture zip codes with positive sales.

Table 1 Estimates of Spatial Model, Random Parameters Model, and Fixed Effects Models

Parameters	Multinomial logit spatial effects		Multinomial logit random parameters model		Multinomial logit fixed effects	
	Estimates (std. error)	p value	Estimates (std. error)	p value	Estimates (std. error)	p value
PDF						
β_1	1.8349	0.0014	1.7545	0.0097	1.7044	0.2110
Intercept	(0.5734)		(0.6783)		(1.3625)	
β_{pr1}	0.0126	0.0194	-0.0206	0.4945	0.1647	0.9282
Print price	(0.0054)		(0.0301)		(1.8281)	
β_{per1}	-0.0654	0.0000	-0.0728	0.0000	-0.0347	0.2462
% Price for PDF	(0.0134)		(0.0159)		(0.0299)	
β_{pdf1}	0.0384	0.0400	0.0292	0.0029	-0.0319	0.4767
Interaction	(0.0187)		(0.0098)		(0.0448)	
8 Subject dummies	0.0077 to 2.2963	0.0000 to 0.4531	-0.0019 to 2.2878	0.0001 to 0.2428	-0.4618 to 1.9560	0.0003 to 0.6791
Print						
β_2	0.3151	0.1787	0.2545	0.3991	2.1747	0.0076
Intercept	(0.2343)		(0.3018)		(0.8152)	
β_{pr2}	-0.0153	0.0000	-0.0221	0.0000	-0.2246	0.7614
Print price	(0.0011)		(0.0054)		(0.7399)	
β_{per2}	0.0085	0.0025	0.0003	0.0007	-0.0123	0.9899
% Price for PDF	(0.0028)		(0.0001)		(0.9696)	
β_{pdf2}	0.0969	0.0000	0.0587	0.0000	-0.3953	0.5792
Interaction	(0.0063)		(0.0097)		(0.7128)	
8 Subject dummies	-0.2010 to -1.3020	0.0011 to 0.6087	-0.2105 to -1.3119	0.0009 to 0.3651	-1.2767 to 2.2826	0.0091 to 0.6418
τ_1^2	1.2121	0.0002	1.9106	0.0024	*	*
	(0.3245)		(0.6294)			
τ_2^2	1.0131	0.0187	*	*	*	*
	(0.4309)					
σ^2	0.9291	0.0003	*	*	*	*
	(0.2589)					
α_1	0.0127	0.0001	*	*	*	*
	(0.0032)					
α_2	0.0153	0.0000	*	*	*	*
	(0.0021)					

minimize correlations, this formulation helps us to capture the impact of absolute prices, as well as the relative prices of forms (reference effects) on the purchase probabilities.

In comparing the three models estimated on the same data, the spatial model and the random parameters model cut down significantly on the unexplained

variance through parameter estimates τ^2 (random parameters model) and $\tau_1^2, \tau_2^2, \sigma^2$ and the α s (spatial model) and, thereby, reduce the standard errors of parameter estimates. The net impact is that more estimates are significant in the spatial and random parameters model (especially those associated with print probabilities). In addition, the model fit statistics (Table 2) suggest that those models also have higher adjusted ρ^2 values and lower AIC, BIC values, with the spatial model always dominating the random parameters model. This comparison suggests that while the fixed effects models are clearly inferior, the random parameter model could perform as well as the spatial model on the validation data. However, one of the downfalls of the random parameter model is that, unlike the spatial model, it has no inherent mechanism to predict at unobserved locations. We will revisit this issue later when we discuss the cross-validation performance.

Table 2 Model Fits of Spatial Model, Random Parameters Model, and Fixed Effects Models

Model	Multinomial logit spatial effects	Multinomial logit random parameters model	Multinomial logit fixed effects
Sample size = 579			
Log-likelihood	-378.10	-401.19	-459.44
AIC	814.21	852.37	966.87
(Adj. ρ^2)	(0.36)	(0.33)	(0.24)
BIC	940.69	961.41	1,071.54

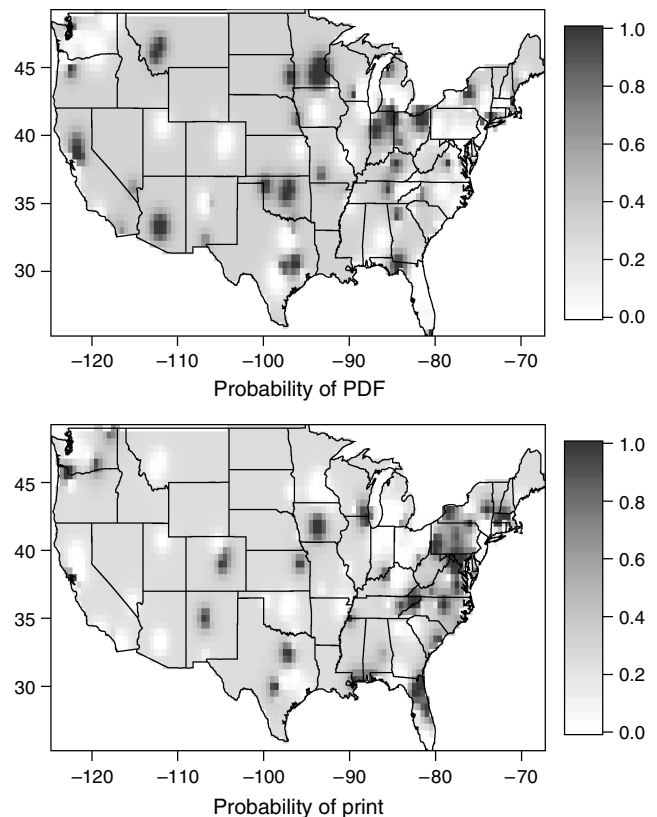
Based on the estimated coefficients of the spatial model, the subject category estimates (corresponding to the dummies) indicate that PDF purchase probabilities are higher for categories such as physical sciences, research issues, and industry and economics and relatively low for agriculture and education. With regard to the price variables, the probability of purchase of the print form (relative to no purchase) decreases as print price increases ($\hat{\beta}_{pr2}$ is negative), increases as PDF price as a percentage of print price increases ($\hat{\beta}_{per2}$ is positive) and increases as absolute price of PDF increases ($\hat{\beta}_{pdf2}$ is positive). These estimates have the expected signs. The probability of purchase of PDF (relative to no purchase) increases as PDF price as a percentage of print price decreases ($\hat{\beta}_{per1}$ is negative) and increases as absolute print price increases ($\hat{\beta}_{pr1}$ is positive). These are also expected results, all indicating the substitution effect between print and PDF, as well as the reference effect of print book price on the purchase probabilities.

The parameter estimate, $\hat{\beta}_{pdf1}$, corresponding to the interaction variable (a measure of absolute price of PDF form), is significantly positive. This indicates that the interaction terms moderates the effect of the relative price of PDF (PDF price as a percentage of print price) on PDF purchase probabilities. However, in comparing $\hat{\beta}_{pdf1}$ (0.0384) and $\hat{\beta}_{pdf2}$ (0.0969), we do find that increase in absolute price of PDF increases print purchase probabilities more than PDF purchase probabilities. The random effect parameter estimate $\hat{\tau}_1^2$ and $\hat{\tau}_2^2$ are both significant, indicating presence of alternative-specific significant correlations not attributable to spatial variables. We also estimate the between-parameter covariances and corresponding correlations for the six price parameters in the model (these are not reported in Table 1). The correlations tend to be small, ranging from 0.001 to 0.143, indicating that the spatial and alternative-specific correlations account for much of the correlations between u_{ijk} and $u_{i'j'k'}$.

The spatial parameter estimates, $\hat{\sigma}^2$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$ are also significant, indicating the significant impact of spatial correlations in the model. Note that $\hat{\alpha}_1$ and $\hat{\alpha}_2$ control the decay of the correlations between observations with distance, although they do not have much of a substantive interpretation in and of themselves. However, it is important to realize that the spatial correlations impact the model estimates in the price coefficients and, thus, can account for changes in price sensitivities across geographical regions, which we examine next.

Based on the estimates of the spatial model, the prediction of purchase probabilities for PDF and print for observed zip codes (from the data) and unobserved zip codes are derived for a grid of spatial coordinates

Figure 1 Print \$40 and Relative PDF 50%



outlining the geographical borders of the continental United States. For example, in Figure 1 (top) the purchase probabilities of PDF for a \$40 print price with a 50% relative PDF price are provided. The background shading (white/light gray) represents zero probability while the increasingly darker colors (darker gray and black) represent increasing probability values. Figure 1 (bottom) provides the corresponding purchase probabilities for the print form. The probabilities of purchase of PDF are higher in certain pockets in northern California and the southwest, Minnesota, the midwest, and the northeast. The map basically reflects the observed purchase patterns of PDF versus print in these geographical areas and captures customers' strong preference for PDF vis-à-vis print. The high PDF purchase pattern seems to be highly correlated with presence of centers of intellectually based institutions such as universities. The book publisher traditionally sells many of its titles to individuals in universities and institutions of higher learning, and this pattern suggests that these purchases show a strong preference for PDF, which could be due to customers' technology readiness and comfort with the technology and/or desire to obtain the content immediately rather than wait for it to be mailed. The pattern could also reflect the diffusion of PDF form and word-of-mouth effects in local geographical areas

Table 3 Cross-Validation on Hold-Out Sample

Predicted true	PDF	Print	No purchase	Total
A. Multinomial logit fixed effects—Overall hit rate 32.50%, RMSE, 26.89%				
PDF	9	0	13	22
Print	19	1	7	27
No purchase	40	2	29	71
Total	68	3	49	120
B. Multinomial logit random parameters model—Overall hit rate 28.33%, RMSE, 28.75%				
PDF	12	1	9	22
Print	21	1	5	27
No purchase	44	6	21	71
Total	77	8	35	120
C. Spatial multinomial logit model—Overall hit rate 52.50%, RMSE, 19.12%				
PDF	15	1	6	22
Print	15	7	5	27
No purchase	29	1	41	71
Total	59	9	52	120

(in certain universities but not all universities). The power of spatial modeling lies in uncovering such trends even when the underlying reasons may not be very obvious. As the validation exercise (discussed next) shows, accounting for such geographical patterns even when the reasons for such patterns cannot be clearly pinpointed, plays a crucial role in predicting customers’ purchases.

Table 3 provides the results of validation of the model on the holdout sample of respondents. Table 3A holds the validation results of the fixed-effects model, while Table 3B holds the results of using the random parameters model. Table 3C has the validation results of the spatial model. The overall hit rate for the fixed-effect model is 32.5%. The random parameter model performs poorly with a hit rate of 28.33%, while the overall hit rate for the spatial model is 52.5%. The root mean-squared-error (RMSE) values also favor the spatial model. This cross-validation exercise provides a clear indication of the importance

of accounting for spatial correlations in the purchases of PDF vis-à-vis print.

4.4. Managerial Implications

The spatial maps of predicted probabilities of PDF and print at various price levels can be useful for managers for several important decisions. First, as described in §4.2, the publisher can estimate the overall sales impact of introducing PDF forms at various price levels, keeping the price levels of print and PDF the same for all customers. Using the spatial model, the probabilities of PDF purchase, print purchase, and no purchase for all relevant combinations of print and PDF prices can be determined for all zip codes that exist in the publisher’s sales file. (Table 4 provides the mean purchase probabilities and confidence intervals as an illustration for print price \$20 and PDF price at 75% of print price for zip codes not observed in the experiment data. Note that probabilities vary significantly, as do the standard errors across these zip codes.) Zip code level forecasts are summed across all zip codes that exist in the sales file to arrive at the total sales forecast at different relative PDF price levels. The current sales of print are \$11.622 million, and a PDF price of 75% relative to print maximizes sales at \$12.608 million (the data are disguised, but the relative effects are real) while sales drop off to \$10.916 million at 50% relative price and to \$9.147 million at 25% relative price.

The spatial probabilities are also useful in estimating the demand of print form when the pricing for PDF form is set at 75% of print price (or any other relative price). While demand variations in PDF can be met without much cost, any overproduction of print form can result in excess inventory and waste. Given the distribution of demand of titles across zip codes, demand for print form for each subject category can be estimated for a specific relative price of PDF. These figures show significant variation in print demand across subject categories. For example,

Table 4 Mean Purchase Probabilities for Selected Zip Codes

Zip	City	State	Long.	Lat.	Mean*		Std. error		Lower 95% CI		Upper 95% CI	
					PDF	Print	PDF	Print	PDF	Print	PDF	Print
58102	Fargo	ND	96.8507	46.9259	0.2474	0.3948	0.0032	0.0012	0.2410	0.3924	0.2537	0.3973
82601	Casper	WY	106.3605	43.1744	0.1914	0.3657	0.0017	0.0043	0.1880	0.3571	0.1948	0.3742
97462	Oakland	OR	123.3526	43.4838	0.2263	0.4053	0.0054	0.0048	0.2155	0.3957	0.2370	0.4148
92037	La Jolla	CA	117.2497	32.8548	0.4545	0.2373	0.0325	0.0558	0.3895	0.1257	0.5195	0.3490
85364	Yuma	AZ	114.6726	32.7006	0.2194	0.4222	0.0060	0.0040	0.2075	0.4142	0.2314	0.4303
67201	Wichita	KS	97.3414	37.6897	0.2785	0.3312	0.0062	0.0073	0.2662	0.3165	0.2909	0.3458
70112	New Orleans	LA	90.0769	29.9574	0.2234	0.3693	0.0085	0.0621	0.2064	0.2451	0.2404	0.4936
31322	Savannah	GA	81.2597	32.1122	0.2858	0.3851	0.0268	0.0357	0.2322	0.3137	0.3395	0.4564
43601	Toledo	OH	83.5486	41.6525	0.1263	0.3904	0.0488	0.0284	0.0287	0.3336	0.2238	0.4473
40502	Lexington	KY	84.4842	38.0142	0.1612	0.3909	0.0506	0.0659	0.0600	0.2591	0.2623	0.5226

*Price of print \$20 and price of PDF 75% of the print price.

a 100,000-unit demand for a title in print form when only print is available reduces to around 65,000 when PDF is priced 75% relative to print in the “research issues” category, while it only decreases to 88,000 in the “agriculture” category. Across categories, the estimated mean print unit demand per 100 units of print demand under “print only” scenario varies significantly (for example, 87.2 for education, 68.6 for physical sciences, 78.1 for food and nutrition, 77.1 for natural resources, and 66.2 for public policy). Such estimation can help in appropriately planning the production of print form of titles in the different product categories.

The publisher is committed to encouraging the electronic form and moving entirely to producing and disseminating the electronic form, given the production and cost efficiencies. However, pricing PDF at 75% of the print form may not generate enough trial of the form among customers. In such a case, the publisher can use the spatial maps to selectively promote the PDF form through coupon discounts to customers in their database—those who may be more price sensitive, those who may have lower preference for PDF, etc. Figure 2 shows a map of statistically significant differences in PDF preference probabilities for different price levels—the white areas indicate a 95% level significant difference across price levels while gray areas indicate no difference. Such a price

sensitivity map can be helpful to managers in identifying areas where the preference for the PDF form varies significantly, even for a small change in PDF price levels. This can be useful for targeted cross-channel promotions.

5. Conclusion

Our contribution in this paper is twofold. From a substantive perspective, we have shown that in specific online contexts, spatial correlations can be useful in predicting customer online behavior. At the least, spatial modeling can be useful for confirming the hypotheses managers may have about the impact of geographical locations on online purchase. In such cases, it can more than complement a fixed-effects model with geographical dummy variables that form the basis for the hypotheses. At best, it can also provide significant reductions in error variances in customers’ choices, even if the reasons for such spatial variation are not clearly known, and improve predictions. While the significant potential of spatial modeling may not be realized until online technologies for customer location identification mature, the model can still be useful in multichannel contexts, as we have illustrated in the paper. Online visitors’ location can be identified through registrations and/or elicited through surveys, as in our experimental setting. Collecting online behavior and spatial data using such research studies, a firm can extrapolate the findings to the entire customer database and evaluate the impact of different strategies. From a methodological perspective, we extend spatial modeling techniques to model choice data directly. Our model incorporates spatial effects in all coefficients of the model, allows for the nugget effect in correlations, enables “kriegering,” and, thus, can be general in its application. In addition, the current implementation of Monte Carlo EM algorithm is capable of estimating a small number of covariances across parameters while a more efficient implementation is necessary if a larger number is desired.

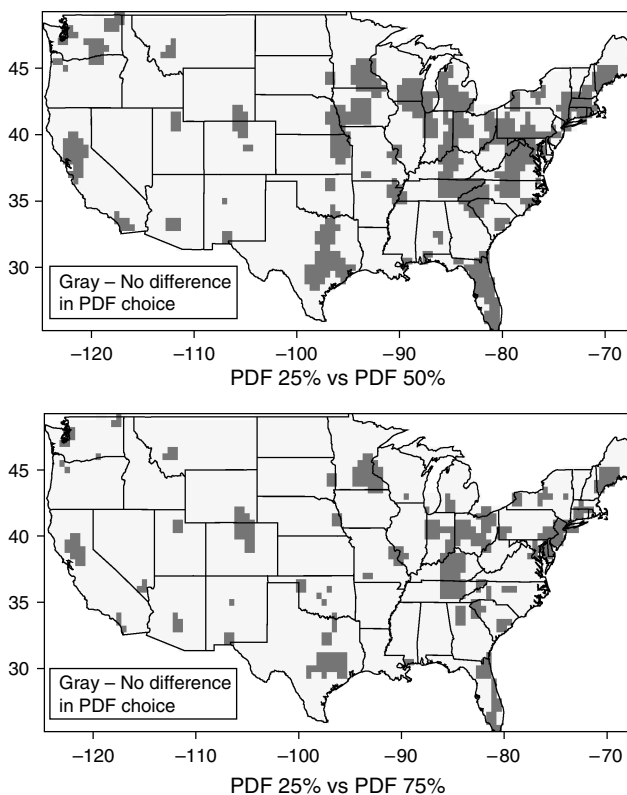
Acknowledgments

The authors thank the editor, associate editor, and four referees for invaluable comments and suggestions. P. K. Kannan acknowledges grant support from the Andrew Mellon Foundation for the research. The authors name are listed in alphabetical order, and both have contributed equally to the paper.

Appendix A. Maximum Likelihood Estimation

Let $\mathbf{y} = (y_1, \dots, y_N)^T$ denote the vector of observed choice responses. Similarly, let $\mathbf{u} = (\mathbf{u}_1^T, \dots, \mathbf{u}_N^T)^T$ denote the associated vector of (unobserved) random effects. Let $f(\mathbf{y}, \mathbf{u}; \boldsymbol{\theta})$ denote the joint distribution of (\mathbf{y}, \mathbf{u}) , where $\boldsymbol{\theta}$ is the vector of parameters. Under maximum likelihood

Figure 2 Decision Maps of Price Sensitivities



estimation, the goal is to find the estimate $\hat{\theta}$ that maximizes the (marginal) likelihood, integrating out the random effects

$$L(\theta; \mathbf{y}) = \int f(\mathbf{y}, \mathbf{u}; \theta) d\mathbf{u}. \quad (1)$$

A.1. Parameter Estimation via the Monte Carlo EM Algorithm

We perform maximum likelihood estimation iteratively, using the Expectation-Maximization (EM) algorithm (Dempster et al. 1977). Let $\theta^{(t-1)}$ denote the current parameter value. Then, in the t th iteration of the algorithm, the E-step computes the conditional expectation $Q(\theta | \theta^{(t-1)}) = E[\log f(\mathbf{y}, \mathbf{u}; \theta) | \mathbf{y}; \theta^{(t-1)}]$. However, in our application this expectation involves an analytically intractable integral of high dimension, so we approximate it via Monte Carlo. Let

$$\tilde{Q}(\theta | \theta^{(t-1)}) = \frac{1}{m_t} \sum_{k=1}^{m_t} \log f(\mathbf{y}, \mathbf{u}^{(k)}; \theta), \quad (2)$$

where $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(m_t)}$ are simulated from the conditional distribution $f(\mathbf{u} | \mathbf{y}; \theta^{(t-1)})$. The Monte Carlo EM (MCEM) algorithm simply replaces Q by its Monte Carlo estimate \tilde{Q} . In the M-step, we maximize (2). This gives the t th update, $\theta^{(t)}$.

A.2. MCMC Sampling from $f(\mathbf{u} | \mathbf{y}; \theta)$

Notice that the Monte Carlo average in (2) requires a sample $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(m_t)}$ from $f(\mathbf{u} | \mathbf{y}; \theta^{(t-1)})$. To do so, we use a Metropolis algorithm similar to the one proposed by McCulloch (1997). Specifically, starting from an initial value $\mathbf{u}^{(1)}$, let $\mathbf{u} \equiv \mathbf{u}^{(k)}$ denote the current value of the Markov Chain. Let $h(\mathbf{u})$ denote the candidate distribution. We generate a new draw u_l^* , corresponding to the l th component of \mathbf{u} , using h . Let $\mathbf{u}^* = (u_1, u_2, \dots, u_{l-1}, u_l^*, u_{l+1}, \dots, u_L)^T$. We then accept \mathbf{u}^* as the next value of the chain with probability $A_l(\mathbf{u}, \mathbf{u}^*)$, where

$$A_l(\mathbf{u}, \mathbf{u}^*) = \min \left\{ 1, \frac{f(\mathbf{u}^* | \mathbf{y}; \theta) h(\mathbf{u})}{f(\mathbf{u} | \mathbf{y}; \theta) h(\mathbf{u}^*)} \right\}. \quad (3)$$

In particular, choosing the candidate $h(\mathbf{u}) = f(\mathbf{u}; \sigma^2, \alpha)$, $A_l(\mathbf{u}, \mathbf{u}^*)$ simplifies to

$$A_l(\mathbf{u}, \mathbf{u}^*) = \min \left\{ 1, \frac{f(\mathbf{y} | \mathbf{u}^*; \theta)}{f(\mathbf{y} | \mathbf{u}; \theta)} \right\}, \quad (4)$$

which involves only the conditional distribution of $\mathbf{y} | \mathbf{u}$.

A.3. Standard Errors of the Parameter Estimates

We obtain the standard errors of the parameter estimates directly through the MCEM algorithm. Suppressing the dependence on the observed data \mathbf{y} , we define the first and second derivatives of the complete data log-likelihood as:

$$\begin{aligned} l^1(\mathbf{u}; \theta) &= \frac{\partial}{\partial \theta} \log f(\mathbf{y}, \mathbf{u}; \theta) \quad \text{and} \\ l^2(\mathbf{u}; \theta) &= \frac{\partial^2}{\partial \theta \partial \theta^T} \log f(\mathbf{y}, \mathbf{u}; \theta). \end{aligned} \quad (5)$$

Using Louis' method (Louis 1982), we can approximate the observed information matrix (or the asymptotic variance-covariance matrix) via

$$\begin{aligned} & -\frac{\partial^2}{\partial \theta \partial \theta^T} \log L(\theta; \mathbf{y}) \\ & \approx -\frac{1}{m_t} \sum_{k=1}^{m_t} l^2(\mathbf{u}^{(k)}; \theta) + \frac{1}{m_t} \sum_{k=1}^{m_t} [l^1(\mathbf{u}^{(k)}; \theta) - \bar{l}^1(\theta)] \\ & \quad \cdot [l^1(\mathbf{u}^{(k)}; \theta) - \bar{l}^1(\theta)]^T, \end{aligned}$$

where $\bar{l}^1(\theta) = \sum_{k=1}^{m_t} l^1(\mathbf{u}^{(k)}; \theta) / m_t$.

B. Spatial Prediction

The key focus of our spatial model is on the estimation and prediction of the random effects. That is, we would like to predict a consumer's choice for a geographic location from which we do not have any previous data. To do so, we first estimate the random effects from the observed locations and then use these estimates to predict a consumer's behavior for unobserved locations.

B.1. Estimation of the Random Effects from Observed Locations

Let $\hat{\theta} = (\hat{\beta}, \hat{\sigma}^2, \hat{\alpha})^T$ denote the vector of estimated parameters from the MCEM algorithm. The minimum mean squared error estimate of \mathbf{u} is given by the conditional expectation $E[\mathbf{u} | \mathbf{y}; \hat{\theta}]$. While this expectation cannot be computed in closed form, we can approximate it via $\hat{\mathbf{u}} = (\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_N)^T = (1/m) \sum_{k=1}^m \mathbf{u}^{(k)}$, where $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(m)}$ are simulated from $f(\mathbf{u} | \mathbf{y}; \hat{\theta})$. In other words, $\hat{\mathbf{u}}_i$ is the estimated random effect corresponding to location \mathbf{z}_i .

B.2. Prediction of the Random Effects for Unobserved Locations

The goal of spatial prediction (or "kriegering") is to make inference about unobserved locations, borrowing information from the neighboring areas for which data is available. Let \mathbf{z}_* denote a location for which no data has been observed. Let \mathbf{u}_* denote the corresponding random effect. The linear predictor $\hat{\mathbf{u}}_*$ of \mathbf{u}_* is given by $\hat{\mathbf{u}}_* = E[\mathbf{u}_* | \mathbf{y}] = \mathbf{C}_* \hat{\mathbf{u}}$, where the coefficient matrix \mathbf{C}_* is such that $E[\mathbf{u}_* | \mathbf{u}] = \mathbf{C}_* \mathbf{u}$ (see Zhang 2002). In other words, if $\hat{\Sigma}$ denotes the covariance matrix of \mathbf{u} evaluated at the MLE $\hat{\theta}$, then the joint distribution of \mathbf{u}_* and \mathbf{u} is given by

$$\begin{bmatrix} \mathbf{u}_* \\ \mathbf{u} \end{bmatrix} \sim \text{MVN} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \hat{\sigma}_*^2 & \hat{\rho}_*^T \\ \hat{\rho}_* & \hat{\Sigma} \end{bmatrix} \right), \quad (6)$$

where $\hat{\sigma}_*^2 = \text{diag}(\hat{\sigma}^2)$ denotes the diagonal matrix containing the variances of the components of \mathbf{u}_* , and $\hat{\rho}_*$ denotes the matrix of covariances between \mathbf{u}_* and \mathbf{u} , all evaluated at $\hat{\theta}$. It follows that for the matrix of coefficients, $\mathbf{C}_* = \hat{\rho}_*^T \hat{\Sigma}^{-1}$.

Given the predicted random effect $\hat{\mathbf{u}}_*$, we can calculate the predicted choice probability for an unobserved location \mathbf{z}_* via

$$\pi_{*j} = \exp(\mathbf{x}_*^T \hat{\beta}_j + \mathbf{x}_*^T \hat{\mathbf{u}}_{*j}) / \left[1 + \sum_{j=1}^{J-1} \exp(\mathbf{x}_*^T \hat{\beta}_j + \mathbf{x}_*^T \hat{\mathbf{u}}_{*j}) \right]. \quad (7)$$

We can assess the variability of the prediction in (7) by repeating the estimation, say, R independent times, obtaining R independent estimates for the predicted choice probabilities. The variance of these R estimates can then be used for inference, for instance, by computing desired level- α confidence intervals.

References

Breslow, N. E., D. G. Clayton. 1993. Approximate inference in generalized linear mixed models. *J. Amer. Statist. Association* 88 9–25.

- Bronnenberg, Bart J., Vijay Mahajan. 2001. Unobserved retailer behavior in multimarket data: Joint spatial dependence in market shares and promotion variables. *Marketing Sci.* **20**(3) 284–299.
- Bronnenberg, Bart J., Carl F. Mela. 2004. Market roll-out and retailer adoption for new brands. *Marketing Sci.* **23**(4) 500–518.
- Bronnenberg, Bart J., Catarina Sismeiro. 2002. Using multimarket data to predict brand performance in markets for which no or poor data exist. *J. Marketing Res.* **39**(February) 1–17.
- Brown, Lawrence D., T. Tony Cai, Anirban Dasgupta. 2002. Confidence intervals for a binomial proportion and asymptotic expansions. *Ann. Statist.* **30**(1) 160–201.
- Caffo, Brian S., Wolfgang S. Jank, Galin L. Jones. 2005. Ascent-based Monte Carlo EM. *J. Roy. Statist. Soc., Series B* **67**(2) 235–252.
- Carton, Sean. 2003. Geotargeting: Why it matters to marketers. *ClickZ*, 4-7-2003, available at www.clickz.com/tech/lead-edge/article.php/2176361.
- Cressie, Noel A. C. 1993. *Statistics for Spatial Data*, Revised ed. Wiley, New York.
- Debruyne, Marion, David J. Reibstein. 2005. Competitor see, Competitor do: Incumbent entry in new market niches. *Marketing Sci.* **24**(1) 55–66.
- Dempster, A. P., N. M. Laird, D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. B* **39** 1–22.
- Donthu, Naveen. 1991. Comparing market areas using kernel density estimation. *J. Acad. Marketing Sci.* **19**(4) 323–332.
- Donthu, Naveen, Roland T. Rust. 1989. Estimating geographic customer densities using kernel density estimation. *Marketing Sci.* **8**(Spring) 191–203.
- Federal Communications Commission. 2004. Section 706 report, Advanced services inquiry. (March) Available at <http://www.fcc.gov/broadband/706.html>.
- Garber, Tal, Jacob Goldenberg, Barak Libai, Eitan Muller. 2004. From density to destiny: Using spatial dimension of sales data for early prediction of new product success. *Marketing Sci.* **23**(3) 419–428.
- Gartner Report. 2003. Print publishers are uncertain about their digital future. (September) Available at <http://www.gartner2.com/qa/qa-0903-0051.asp>.
- Guadagni, P. M., J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2**(3) 203–238.
- Kalnins, Arturs. 2004. An empirical analysis of territorial encroachment within franchised and company-owned branded chains. *Marketing Sci.* **23**(4) 476–489.
- Lipke, David J. 2001. You are here. *Amer. Demographics* **23**(30) 65–68.
- Louis, Thomas A. 1982. Finding the observed information matrix when using the EM algorithm. *J. Roy. Statist. Soc. B* **44** 226–233.
- McFadden, D. 1981. Econometric models of probabilistic choice. C. Manski, D. McFadden, eds. *Structural Analysis of Discrete Data*. MIT Press, Cambridge, MA.
- Mittal, Vikas, Wagner Kamakura, R. Govind. 2004. Geographic patterns in customer service and satisfaction: An empirical investigation. *J. Marketing* **68**(July) 48–62.
- Montgomery, Alan L., Shibo Li, Kannan Srinivasan, John Liechty. 2004. Modeling online browsing and path analysis using clickstream data. *Marketing Sci.* **23**(4) 579–595.
- Parker, Pamela. 2004. Google focuses local ad targeting. *ClickZ News* (April 15). Available at <http://www.clickz.com/news/article.php/3340591>.
- Rust, Roland T., Naveen Donthu. 1995. Capturing geographically localized misspecification error in retail store choice models. *J. Marketing Res.* **32**(February) 103–110.
- Sismeiro, Catarina, Randolph E. Bucklin. 2004. Modeling purchase behavior at an E-commerce website: A task completion approach. *J. Marketing Res.* **41**(3) 306–323.
- Steenburgh, T., A. Ainslie, P. H. Engebretson. 2003. Revealing the information in zipcodes: Bayesian massively categorical variables and aggregated data in direct marketing. *Marketing Sci.* **22**(1) 40–57.
- Ter Hofstede, Frenkel, Michel Wedel, Jan-Benedict E. M. Steenkamp. 2002. Identifying spatial segments in international markets. *Marketing Sci.* **21**(2) 160–177.
- Train, Kenneth E. 2003. *Discrete Choice Models with Simulation*. Cambridge University Press, Cambridge, UK.
- Van den Bulte, Christophe, Stefan Stremersch. 2004. Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Sci.* **23**(4) 530–544.
- Zeger, S. L., M. R. Karim. 1991. Generalized linear models with random effects: A Gibbs sampling approach. *J. Amer. Statist. Association* **86** 79–86.
- Zhang, Hao. 2002. On estimation and prediction for spatial generalized linear mixed models. *Biometrics* **58** 129–136.
- Zhang, Hao. 2003. Optimal interpolation and the appropriateness of cross-validating variogram in spatial generalized linear mixed models. *J. Comput. Graphical Statist.* **12** 698–713.