

GEOGRAPHY AND MARKETING STRATEGY IN CONSUMER PACKAGED GOODS

Bart J. Bronnenberg and Paulo Albuquerque

ABSTRACT

A significant portion of academic research on marketing strategy focuses on how national brands of repeat-purchase goods are managed or should be managed. Surprisingly little consideration is given in this tradition to the extended role of geography, i.e. distance and space. For instance, manufacturers of brands in non-durable product categories are well aware of the fact that their national brands perform very different across domestic U.S. markets. This holds even for product categories with limited product differentiation. In this chapter, we outline various processes through which the influence of geography on performance of national brands materializes. We discuss a number of alternative explanations for the emergence and sustenance of spatial concentration of market shares. Several of these explanations are modeled empirically using data from the United States packaged goods industry. This chapter closes with avenues for further academic research on spatial aspects of the growth of new products.

INTRODUCTION

Geography has become an important practical component of marketing strategy. This is driven to a large extent by organizational expansion goals that force managers to take into account increasingly more complex spatial delivery and advertising systems during the launch and management of new products. In step with this trend, researchers in marketing and economics have developed an interest in the spatial aspects of growth and market structure. The resulting research tradition has been called the "new economic geography." This research stream – which started in the 1970s in the field of industrial organization – is aimed at answering two questions (Fujita, Krugman & Venables, 1999)

- When is a symmetric equilibrium, without spatial concentration, unstable?
- When is a spatial concentration of economic activity sustainable?

The main goal of the "new economic geography" is thus to describe competitive processes driving the growth and subsequent stability of spatial concentration in economic activity (Bonanno, 1990; Fujita & Thisse, 2002). In spirit of these two central questions, this chapter is concerned with the empirical stylized fact that market shares of undifferentiated packaged goods (e.g. food or convenience items) are spatially concentrated. To this end, we outline empirical and analytical models of spatial concentration and growth in the context of packaged goods even when such goods are not meaningfully differentiated. Using these models, we speculate on the reasons why strong spatial concentration in market shares emerges for undifferentiated goods, and we offer several explanations for why such concentration, once established, tends to persist.

The rest of this chapter is organized as follows. In the next section, we commence by looking at some of the basic reasons for why market outcomes in packaged goods should be expected to be spatially dependent and outline some of the geographical aspects of the distribution and advertising infrastructure needed to connect manufacturers and consumers. Then we describe various methods to account for the spatial market-dependence that is caused by this infrastructure. In this section, we also offer a small empirical example of how spatial concentration in market shares can be accounted for. The fourth section, focuses on the first question above and outlines two path-dependent processes that create spatial concentration of outcomes. The fifth section focuses on the second question and discusses several strategic competitive processes that tend to enforce spatial concentration across time and explain why spatial concentration persists. We conclude with directions for future research.

GEOGRAPHICAL ASPECTS OF MARKETING STRATEGY

Two spatially relevant dimensions of new product strategy are distribution and advertising. These two factors are controlled by manufacturers at different levels of spatial aggregation and cause marketing strategies as well as their outcomes to be linked through space. Therefore, when investigating the spatial concentration of market shares, it is useful to commence by looking at how distribution and communication channels are structured geographically.

The Geographical Organization of Distribution Channels

Distribution channels of consumer goods in the U.S. consist of multiple hierarchical participants such as manufacturers, wholesalers, and retailers. Research in marketing and economics has studied the vertical structure of channels, i.e. the desirability and stability of vertical intermediation, in a single market (e.g. McGuire & Staelin, 1983). However, in this literature the impact of the *geographical* organization of distribution channels has not been studied.

An aspect of this geographical organization is the structure of retail trade areas. This structure is important to manufacturers because the retailers control the choice environment of consumers at the point of purchase to a large extent. It is therefore likely that observed spatial pricing policies have a component that reflects the geographic nature of the retail trade and that observed sales data have a component that reflects the unobserved retailer activity such as shelf-space allocations (see also Bronnenberg & Mahajan, 2001).

Another geographical aspect of the distribution channel is that the influence of a single retailer can extend beyond its own trade area. This is because retailers compete and often mimic each other's successful programs. To capture the influence of retailer competition, it is useful to look at how retail trade areas overlap. To exemplify this, Fig. 1 visualizes trade areas of a selection of U.S. retailers.¹ Panel (a) shows the trade area of Albertsons, a large U.S. chain of grocer stores. The trade area of retailer (b), Safeway, coincides largely with that of (a) Albertsons but not at all with that of retailer (d), Kroger. From a competitive perspective, it is therefore likely that for instance Albertsons and Safeway in Fig. 1 compete more directly than say Safeway and Kroger. We will subsequently use trade area overlap to define competitive "closeness" in a network of retailers (see also Baum & Singh, 1994).

The Geographical Organization of Media and Communication Channels

In addition to distribution channels, communication channels also have a distinct spatial organization. For instance, TV communication channels are organized in so-called advertising markets or Designated Market Areas (DMA's).

Nielsen Media Research constructs DMA's by grouping all counties whose largest viewing share is with the same TV stations. For instance, the New York TV advertising market or DMA consists of all counties where the New York TV stations attract the largest viewing share. DMA's are non-overlapping and cover all of the continental U.S., Hawaii and parts of Alaska. In total, the U.S. consists of 210 DMA's. The Nielsen company tracks viewing habits at the individual level for all of these 210 DMA's. Additionally, daily household level viewing data are collected for about 55 of the largest DMA's.

The geographical structure of DMA's is important to manufacturers because their TV advertising decisions are forcibly made at the DMA level. This creates dependence between two markets that are part of the same DMA. In sum, distribution and communication channels are controlled by manufacturers at different levels of spatial aggregation. For the purpose of delivering goods physically to the customer, a spatial control unit often is the trade area of a retail chain.² For the purpose of making the consumer aware of the product, an advertising market or DMA is a relevant spatial control unit. These units need not be (and usually are not) the same. Managerially, this causes an interesting control problem because these different units cause distribution and awareness creating policies to interact in a complicated way. Additionally, from an empirical modeling perspective, the differences in control units will need to be accounted for when modeling data from a cross-section of locations.

REPRESENTATION AND MEASUREMENT OF SPATIAL CONCENTRATION

In this section, we outline several empirical models to measure spatial concentration in brand-level market outcomes. These models combine data at the retailer, DMA, and market level.

The Geographical Concept of a Market

For empirical and economic purposes in the analysis of packaged goods, it is helpful to first define an elementary spatial unit of analysis that can be used in the empirical analysis of both the distribution as well as the communication channels.

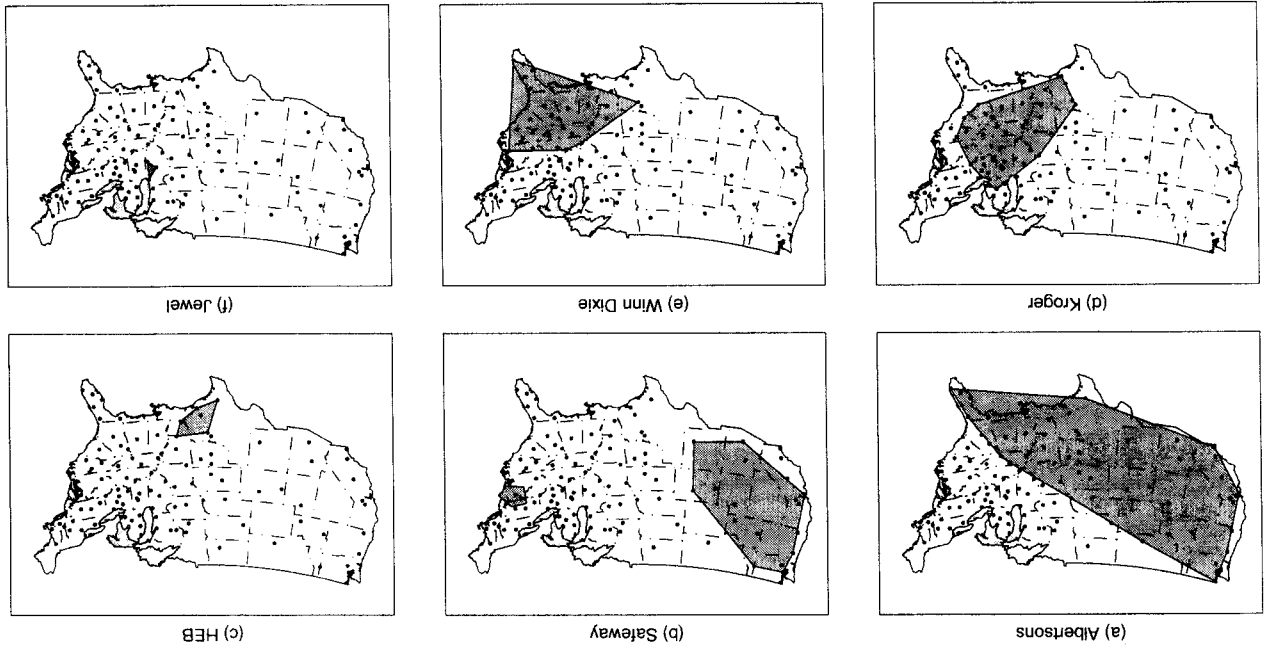


Fig. 1. Examples of Retailer Trade-Areas.

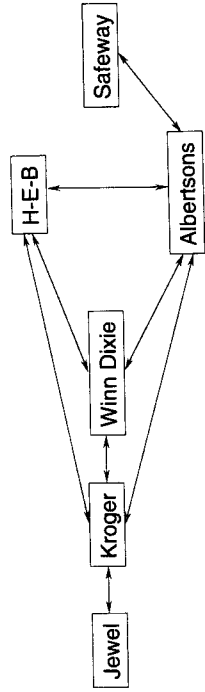


Fig. 2. Part of the U.S. Retail Network, with Linkages Based on Common Trade-Areas.

r' in the trade area of r to capture the influence of r' on r . Therefore, the influence of r' on r can be represented as

$$w_{r' \rightarrow r} = \begin{cases} \frac{\sum_{m \in T_r} ACV_{r'm}}{\sum_{r^m \neq r} \sum_{m \in T_r} ACV_{r'm}} & \text{if } r' \neq r \\ 0 & \text{if } r' = r \end{cases} \quad (1)$$

This measure sums to 1 across all direct competitors r' of retailer r . Using these weights, the representation of the complete retailer network is a sparse weight matrix W of dimension $K \times K$ whose elements are arranged as follows:

$$W = \begin{bmatrix} 0 & w_{2 \rightarrow 1} & \dots & w_{K \rightarrow 1} \\ w_{1 \rightarrow 2} & 0 & \dots & w_{K \rightarrow 2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1 \rightarrow K} & w_{2 \rightarrow K} & \dots & 0 \end{bmatrix} \quad (2)$$

This matrix is sparse because many pairs of retailers do not have overlapping trade areas. Further, the matrix W is asymmetric and can express that the influence of one retailer on the other is larger than vice versa. For any retailer, the definition of $w_{r' \rightarrow r}$ is sensitive to both the size of a given competitor, as well as to the number of markets in which they both meet. For instance, H-E-B in Texas competes in only a small part of the trade area of Albertsons. Albertsons, on the other hand, is present in the entire trade area of H-E-B. Therefore, all else equal and because of its limited scope, the influence of H-E-B on Albertsons, is modeled to be less than the influence of Albertsons on H-E-B. Alternative measures of $w_{r' \rightarrow r}$ can be formulated to account for interactions between the ACV of r' and r .

Mapping Retailer Networks to Consumer Markets

It is often of interest to analyze the performance of products at the market level. It would seem at first glance that the absence of consumer arbitrage across markets

We use the concept of a geographical "market." The term "market" is routinely used in the research and practice of the economic sciences, however it often lacks a formal definition. In the interest of modeling the potential strategic use of space in an economic context, we believe that a useful definition of a "geographic market" is implied by spatial limits on consumer arbitrage. In such a definition, two markets are separated if consumers are unwilling to invest time or resources in travel to benefit from potential price differences across geography. For instance, Los Angeles and New York are two different markets for consumer non-durable goods (e.g. food items), because consumers in Los Angeles do not travel to New York to benefit from deals on such products. On the other hand Los Angeles and New York can be part of the same market in the context of goods that are more expensive.

An interesting aspect of the U.S. geography is that it consists by and large of population centers with relatively empty space in between (see e.g. Greenhut, 1981). This obviously helps the geographic definition of markets. Large marketing research firms such as AC Nielsen and Information Resources Incorporated (IRI) sample selectively from such markets to provide sales and marketing data for consumers goods that cover the entire U.S. (see, e.g. Fig. 1 for an example of the spatial sample design that is used by such marketing research firms).

Modeling Distribution Networks

With consumer markets characterized as a set of locations, the influence of distribution and advertising decisions on the consumers in these markets can be represented using networks. For instance, consider a consumer product that is distributed through retail chains. The mere fact that manufacturers use retailers for the distribution of their brands causes the data to be related across markets in at least two ways. First, U.S. retailers are present in multiple markets. Second, in addition to multimarket presence, retailers influence each other. For example, retailers with overlapping trade areas compete for the same consumers.

To model the influence among retailers, we specify a network of retailers. In this network, retailers who's trade areas overlap are connected. Using Fig. 1 as an example, the subset of six retailers can thus be represented as a sociogram or a graph. Figure 2 shows this graph representation.

The arcs between the retailers can be modeled based on the context at hand. Bronnenberg and Sismeyro (2002) for instance use bi-directional arcs, and a measure based the importance of trade area overlap. Specifically, let any given retailer r have a trade area T_r , consisting of all markets in which r operates. The total dollar amount sold through a retailer r in a given market m is called "all commodity volume" of r in m or simply ACV_{rm} . We use the ACV share of retailer

allows researchers to analyze markets independently. However, it is easy to see that this is only efficient if the analyst observes all demand-relevant information about distribution and advertising. This is normally not the case. For instance, the analyst does not observe shelf-space allocations for consumer goods (such data are not collected on a frequent basis). To make efficient use of the available data, the analyst must therefore make reasonable assumptions about the behavior of each retailer $r = 1, \dots, K$. For example, it could be assumed that when setting shelf-space, each retailer acts in part independently and in part imitates those retailers with whom it competes. A formalization of such an assumption proceeds as follows. Denote unobserved retailer support or shelf space allocation for good j by retailer r by S_{jr} and array all such allocations into the $K \times 1$ vector S_j . Then,

$$S_j = \lambda W S_j + \eta_j \tag{3}$$

In this equation, retailer support S_j (e.g. shelf space allocation) is a linear function of the weighted average, $W S_j$ of retailer support at competing retailers. The coefficient λ measures the strength of the effect of competing retailers. The terms η_j represent the idiosyncratic component of retailer behavior. This model of retail support can be written as a reduced form of the idiosyncratic terms by taking $\lambda W S_j$ to the left-hand side and dividing through,

$$S_j = (I - \lambda W)^{-1} \eta_j \tag{4}$$

This model can be interpreted as a spatially-autoregressive model of retail support. The vector S_j is random from the perspective of the analyst because the idiosyncratic shocks η_j are not observed. However, if the shocks can be assumed to have a parametric distribution, the effects of S_j can be estimated. For instance, if the innovations η_j are normally distributed with mean 0 and variance σ_η^2 then the vector S_j is distributed multivariate normal with mean zero and variance covariance matrix equal to

$$E(S_j S_j') = \sigma_\eta^2 (I - \lambda W)^{-1} (I - \lambda W)^{-1'} \equiv \sigma_\eta^2 \Gamma \tag{5}$$

The random effects S_j (which are at the retailer level) can help in measuring spatial concentration of brand performance across markets by mapping the retailer trade areas to the markets. To exemplify this, suppose we are interested in modeling market shares v_{jm} of product j in market m , as a function of a $1 \times P$ vector of exogenous variables x_{jm} , $m = 1, \dots, M$ and the random effects S_j . To translate the S_j to the market level, define a retail-structure matrix H of size $M \times K$ which lists the ACV based market share of retailer r in market m (H is sparse). Stacking

over markets, we model

$$v_j = x_j \alpha + \beta H S_j + e_j \tag{6}$$

where the effects α are responses to the exogenous variables (it is possible to estimate other effects than common-effects α but we do not discuss such elaborations here) and the scalar β is the effect of the unobserved retail variables such as shelf-space allocations. The $M \times 1$ vector $H S_j$ contains the market averages of the unobserved retailer variables. We assume that e_j is a set of IID residuals that are normally distributed with mean 0 and variance σ_e^2 . These residuals are also independent of the S_j . We can rearrange the last equation to

$$v_j - x_j \alpha = \beta H S_j + e_j \tag{7}$$

Estimation of this model proceeds by realizing that the right-hand side is a Normally distributed random term with mean 0 and variance-covariance matrix equal to $\beta^2 \sigma_\eta^2 H \Gamma H' + \sigma_e^2 I_m$. We usually define $\sigma_\eta^2 = 1$ to set a metric (β^2 -and σ_e^2 can not be identified separately).

It is instructive to observe that two sources of spatial dependence are present in this model. First, the contagion among retailers, λ , creates that the influence of a given retailer spreads beyond its own territory. Second, when this contagion is absent, $\lambda = 0$, the variance covariance matrix in the model reduces to $\beta^2 \sigma_\eta^2 H H' + \sigma_e^2 I_m$. In this case, the off-diagonals in $H H'$ will account for spatial dependence due to the multimarket presence of – independent – retailers.

This discussion implies that in the analysis of multimarket data, even when consumers do not travel from market to market, dependencies across markets will often emerge because of spatial dependencies in unobserved retailer behavior.

Direct Measures of Spatial Concentration Across Markets

Another often used model to express the dependence of data across markets relies on a direct measurement of spatial dependence (see, e.g. Anselin, 1988). Rather than using a factor model such as Eq. (3) to build the spatial dependence matrix from the areas over which retailers exercise direct control, one can take a more statistical perspective and, analogous to the temporally autoregressive model, directly model spatial dependence based on for instance distance or contiguity (see also Edling & Lijerros, 2003). In the latter approach, a contiguity matrix C of size $M \times M$ is defined (M is the number of markets). Each row m of this matrix identifies which markets $m' \neq m$ are neighbors of market m . Various definitions of neighborhood or contiguity exist. The definition of contiguity that most frequently used empirically

with irregularly spaced data is based on so-called Voronoi polygons (e.g. Okabe et al., 2000). These polygons use the (irregular, i.e. non-lattice) location of markets to exhaustively divide the U.S. geography into mutually exclusive market areas. A contiguity-set for a given market is then constructed by the set of all markets areas that are adjacent to the area of the market under study. The contiguity-set of a market is called its *spatial lag operator* (in analogy to approaches in time series analysis). If the rows of the matrix C add to 1, the matrix C is said to be standardized. Denote the number of neighbors of market m by N_m . In this paper, we use a standardized matrix C , with $C(m, m') = 0$ if the two markets are not neighbors, and with $C(m, m') = 1/N_m$ if m and m' are adjacent.

A model of spatially dependent market shares for brand j is then defined by the following variance components model

$$\begin{aligned} v_j &= x_j \alpha + \beta \xi_j + e_j \\ \xi_j &= \lambda C \xi_j + \eta_j \end{aligned} \quad (8)$$

with both e_j and η_j are $M \times 1$ vectors of independently normally distributed variables with mean 0 and variance σ_e^2 and 1 respectively. This model is known as a spatially autoregressive model with autoregression parameter λ . For various technical properties of this model see, e.g. Lesage (2000).

Using a standardized matrix C , the spatial lag of a given observation can be interpreted as the (weighted) average of the observations at neighboring locations. The model thus basically allows for the possibility that the average of neighboring observations is informative about the observation under investigation. Turning back to the model, and taking ξ_j on the left-hand side, we obtain that $\xi_j = (I_m - \lambda C)^{-1} \eta_j$. The model above can therefore be statistically formulated as

$$v_j - x_j \alpha = \beta \xi_j + e_j \quad (9)$$

where the right-hand side is distributed Multivariate Normal with mean 0 and variance covariance matrix equal to $\beta^2 (I_m - \lambda C)^{-1} (I_m - \lambda C)^{-1} + \sigma_e^2 I_m$. Whereas this model has the same number of parameters as the model in Eq. (7) it implies a different type of spatial dependence. Specifically, the model based on retailer networks accounts for the geographical constellation of retailer trade areas, whereas the market-contiguity model is purely based on proximity

An Empirical Example

The models (7) and (9) can be estimated from multimarket data. To provide a simple empirical example of their performance, we use Information Resources Inc. (IRI) optical-scanner supermarket data from 64 local markets, sampled from the entire continental U.S. Markets are defined by IRI as a metropolitan area (e.g.

Los Angeles) or a combination of metropolitan areas (e.g. Raleigh-Durham). In all cases, markets are sufficiently distant from each other that the assumption of absence of arbitrage is very reasonable in the case of consumer packaged goods. The data that we have at our disposal are at the market level and cover sales, prices, and indicators of the presence of promotion displays and feature ads (store flyer ads). For illustration purposes, we calibrate our models on a cross-sectional sample dating from 1995 of 64 observations of market shares, prices, promotion display intensity, and feature intensity (computed as the fraction of time and market volume that a given brand is on display or is featured). We transformed the data by taking natural logs so that regression constants may be interpreted as elasticities. The data analyzed herein are from the largest brand of Mexican Salsas in the United States, Pace.

To estimate the model, we also need data on retailer trade-areas and location of markets. Specifically, to compute the matrix W , we need data on the total volume, ACV_m , of all retailers in the 64 IRI markets. These data were obtained from TradeDimension in New York, who maintains a data base of retail-chains, that includes their location and local size of operation. To compute the matrix C we used the latitude and longitude data of the locations of the IRI markets, and a MATLAB function to compute the Voronoi tessellation of space on which contiguity is defined.

To estimate the models, we maximized the log of the normal likelihood under three different models. The first model, BASE, is a base model for which the coefficient β is constrained to be 0. This creates a standard regression model with IID residuals. The second model (MKT) is the model in Eq. (9) that is based on market contiguity. Finally, the third model (CHAIN) is the model in Eq. (7) and is based on chain level random effects and contagion across chains. The results of the three models are in Table 1.

The parameters in the BASE model have the intuitive pattern. The price elasticity is negative, while the promotion effects are positive. The MKT model shows a high autoregression constant λ . This implies that local averages are informative about the process at the location under investigation and suggests that the data are spatially dependent. However, the importance of the spatial component is relatively low ($\beta = 0.11$). Note the effects of price and promotion are estimated to be lower when spatial dependence is accounted for. Within the confines of this single example, the improvement in loglikelihood over the BASE model is modest.

Finally, when accounting for the geographical structure of the U.S. retail industry through the CHAIN model, we find that the spatial component in the data becomes quite important ($\beta = 0.41$). The parameter λ is lower than in the MKT model, because part of the spatial dependence is already accounted for through the matrix H which lists the market share of each retailer in each market. The loglikelihood of the CHAIN model is better than the two other models.

on temporal patterns of sales growth (see e.g. Mahajan, Muller & Bass, 1995). Recently, spatial and spatiotemporal patterns of diffusion have become the subject of empirical study (e.g. Bronnenberg & Mela, 2002; Van den Bulte & Lilien, 2001). In addition to empirical methods, an other way to study spatial diffusion is by using differential equations derived from theoretical models (Edling & Liljeros, 2003). Recently, also simulation studies using aggregations of micro-level agents or decision makers have been used to model spatial diffusion (see e.g. Lomi et al., 2003 for additional references). However, we focus on empirical models. Bronnenberg and Mela (2002) develop a two stage model of new product assortment-adoption by retailers. The first stage captures how manufacturers roll out the new product and enter local markets. The second stage models how retailers adopt a brand given that it is available in at least one market that is part of its territory. A basic version of this model can be stated as follows.

Manufacturer's Market-Entry

Denote the presence of the brand in a market by a dummy variable y_{imt} , where $i = 1, \dots, I$ indexes brands, $m = 1, \dots, M$ = indexes markets, and $t = 1, \dots, T$ indexes time. Entry into market m by manufacturer i in week t can be formalized as a probit model, i.e.

$$\Pr(y_{imt} = 1) = \begin{cases} \Phi(U_{imt}) & \text{if } y_{imt-1} = 0 \\ 1 & \text{else} \end{cases} \quad (10)$$

in which U_{imt} is a deterministic function and Φ is the cumulative standard Normal distribution. Spatial dependence of manufacturer rollout can be introduced in this model by making U_{imt} a function of whether i 's brand was launched in neighboring markets m' in the past time periods. Using the definition of the matrix C from the previous section, and arraying the market entry variables of $t - 1$ across markets into the $M \times 1$ vector y_{it-1} , a spatial effect on the local entry decisions can be operationalized as the m th element of the spatially and temporally lagged market entry variables Cy_{it-1} . Denoting the m th row of C by c_m the weighted average of past entry in neighboring markets is thus $c_m y_{it-1}$.

Another variable that influences spatial concentration and affects market-entry is the sum of market shares in market m of chains who adopted manufacturer i 's new brand in any market $m' \neq m$ prior to t . This variable captures the degree to which retailers on a given market already carry the new brand in other markets. This variable can be defined on the basis of the matrix H (defined previously as the M by K matrix containing the ACV share of chain k in market m). Write the m th row of H by h_m . Denote the distribution status of brand i by $z_{ikt} = 1$ if chain k adopted before or in week t , and $z_{ikt} = 0$ if the chain did not adopt up until week t . Array

Table 1. Maximum Likelihood Estimates (*t*-statistics).

	Model		
	BASE	MKT	CHAIN
α_0	1.79 (3.4)	0.83 (1.5)	0.82 (1.8)
α_{price}	-3.20 (-3.8)	-2.33 (-3.7)	-2.37 (-4.1)
α_{display}	0.21 (3.7)	0.09 (1.4)	0.10 (2.4)
α_{feature}	0.14 (3.8)	0.12 (1.4)	0.06 (1.6)
λ		0.90 (8.5)	0.67 (4.0)
β		0.11 (7.8)	0.41 (6.9)
c_e^{-2}	0.32 (11.0)	0.24 (1.5)	0.06 (1.8)
Log likelihood	-16.94	-14.40	-1.42

We have illustrated that spatial concentration exists and outlined two methods through which it can be measured. Within the confines of our data, it seems (1) that spatial concentration in these data is substantial, (2) that the spatial component in the data seems consistent with unobserved retailer conduct and (3) that it is necessary to account for this structure when analyzing multimarket data. Especially the second finding is interesting. Essentially, the second point states that after accounting for price, display and feature effects, the unobserved components left in the data are mostly consistent with retailer level variation.

The following sections discuss theoretical perspectives that help to explain why spatial concentration emerges and why it generally persists.

**PATH DEPENDANT GROWTH PROCESSES:
THE INTERACTION OF GEOGRAPHY (SPACE)
AND HISTORY (TIME)**

In this section, we discuss two path-dependent processes of growth. Both processes partly explain the emergence of spatial concentration of market share data. The first process offers a spatial and network diffusion perspective on how retailers adopt new products (leading to local rollouts), while the second process concentrates on how consumers learn about new products based on past experiences.

Spatial and Network Diffusion in Retail Distribution

New product diffusion research has been important in marketing (see, e.g. Bass, 1969). However, the diffusion literature in marketing has almost uniquely focused

