

Geography and Marketing Strategy in Consumer Packaged Goods

by

Bart J. Bronnenberg and Paulo Albuquerque*

December 2002

Third and final version

Submitted to:

Advances in Strategic Management, Vol 20.

Joel Baum and Olav Sorenson, editors,

Elzevier

* Thanks to Joel Baum and Olav Sorenson for excellent comments on an earlier draft. Bart Bronnenberg is an Associate Professor and Paulo Albuquerque is a PhD student both at the John E. Anderson Graduate School of Management at UCLA. Correspondence to bart.bronnenberg@anderson.ucla.edu or paulo.albuquerque@anderson.ucla.edu.

Contents

1	Introduction	4
2	Geographical aspects of marketing strategy	5
3	Representation and measurement of spatial concentration	7
3.1	The geographical concept of a market.	7
3.2	Modeling distribution networks	8
3.3	Mapping retailer networks to consumer markets	9
3.4	Direct measures of spatial concentration across markets	11
3.5	An empirical example	12
4	Path dependent growth processes: the interaction of geography (space) and history (time)	14
4.1	Spatial and network diffusion in retail distribution	15
4.2	Order of entry and consumer learning	17
5	Marketing strategy and sustenance of spatial concentration in brand shares	20
5.1	Spatial distributions of consumer tastes and path-dependence	21
5.2	Multi-market contact	21
6	Conclusions	23
7	References	24

Geography and Marketing Strategy in Consumer Packaged Goods

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 US 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.

Keywords: Multi-market competition, retailing, vertical channel competition, spatial analysis, network analysis.

1 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 and 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 and 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. Section 4, focuses on the first question above and outlines two path-dependent processes that create spatial concentration of outcomes. Section 5 focuses on the second question and discusses

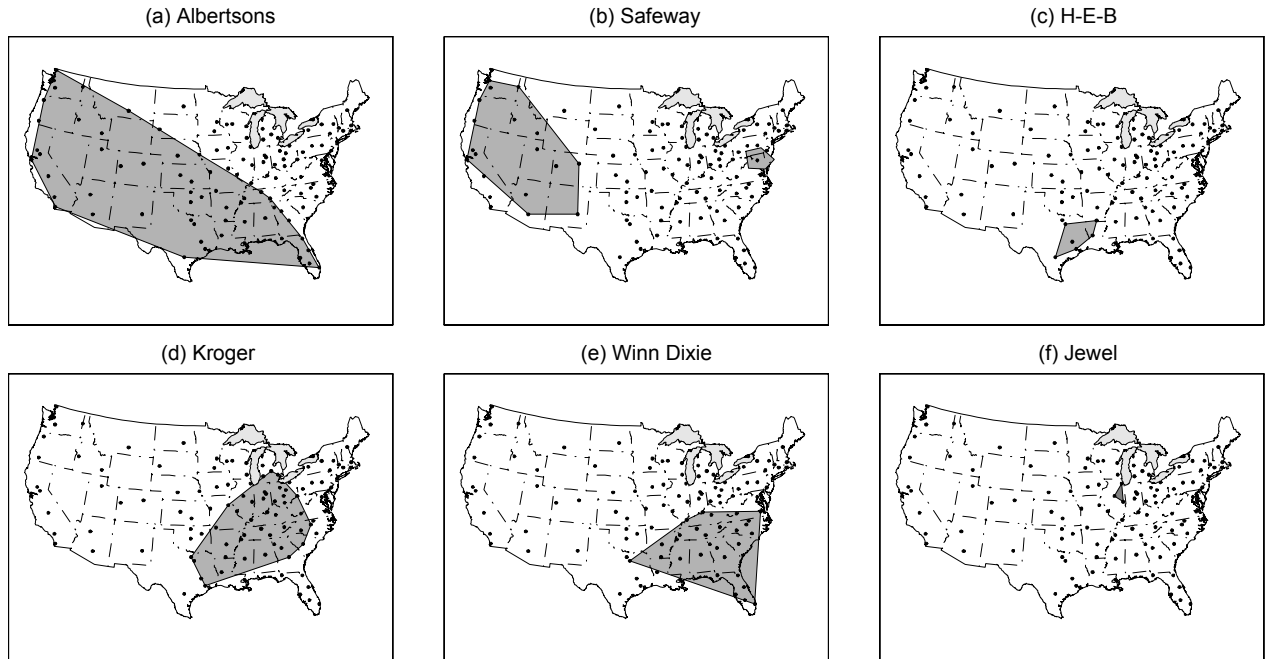


Figure 1: Examples of retailer trade-areas

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.

2 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 United States 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 and Staelin 1983). However, in this literature the impact of the *geographical* organization of distribution channels has not been studied.

A geographical aspect of this 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 and 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, Figure (1) visualizes trade areas of a selection of United States retailers.¹ Panel (a) shows the trade area of Albertsons, a large US chain of grocery 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 Figure (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 and 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 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 United States, Hawaii and parts of Alaska. In total, the US 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.

¹Figure 1 visualizes the trade areas of chains, but not of their subsidiaries.

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.

3 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.

3.1 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. 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

²During the introduction of new products, firms are often additionally interested in retailer adoption at the market level. The same holds for retailers that have very large trade areas. Some of these larger retailers have spatial control units themselves, e.g., the Albertsons supermarket chain is organized in various geographical clusters.

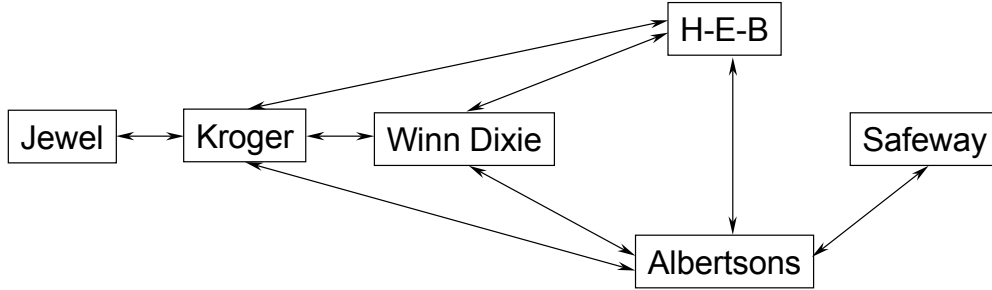


Figure 2: Part of the U.S. retail network, with linkages based on common trade-areas

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 United States (see, e.g., Figure 1 for an example of the spatial sample design that is used by such marketing research firms).

3.2 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, United States 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 Figure 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 Sismeiro (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 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

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

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 \mathbf{W} of dimension $K \times K$ whose elements are arranged as follows:

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

This matrix is sparse because many pairs of retailers do not have overlapping trade areas. Further, the matrix \mathbf{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 .

3.3 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 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 \mathbf{S}_j . Then,

$$\mathbf{S}_j \underset{(K \times 1)}{=} \lambda \mathbf{W} \mathbf{S}_j + \boldsymbol{\eta}_j. \quad (3)$$

In this equation, retailer support \mathbf{S}_j (e.g., shelf space allocation) is a linear function of the weighted average, $\mathbf{W} \mathbf{S}_j$, of retailer support at competing retailers. The coefficient λ measures the strength of the effect of competing retailers. The terms $\boldsymbol{\eta}_j$ 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 $\lambda \mathbf{W} \mathbf{S}_j$ to the left hand side and dividing through,

$$\mathbf{S}_j \underset{(K \times 1)}{=} (\mathbf{I}_K - \lambda \mathbf{W})^{-1} \boldsymbol{\eta}_j. \quad (4)$$

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

$$E(\mathbf{S}_j \mathbf{S}_j') \underset{(K \times K)}{=} \sigma_\eta^2 (\mathbf{I}_K - \lambda \mathbf{W})^{-1} (\mathbf{I}_K - \lambda \mathbf{W})^{-1'} \equiv \sigma_\eta^2 \boldsymbol{\Gamma} \quad (5)$$

The random effects \mathbf{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 \mathbf{x}_{jm} , $m = 1, \dots, M$ and the random effects \mathbf{S}_j . To translate the \mathbf{S}_j to the market level define a retail-structure matrix \mathbf{H} of size $M \times K$ which lists the ACV based market share of retailer r in market m (\mathbf{H} is sparse). Stacking over markets, we model

$$\mathbf{v}_j \underset{(M \times 1)}{=} \mathbf{x}_j \boldsymbol{\alpha} + \beta \mathbf{H} \mathbf{S}_j + \mathbf{e}_j \quad (6)$$

where the effects $\boldsymbol{\alpha}$ are responses to the exogenous variables (it is possible to estimate other effects than common-effects $\boldsymbol{\alpha}$ but we do not discuss such elaborations here) and the scalar β is the effect of the unobserved retail variables such as shelf-space. The $M \times 1$ vector $\mathbf{H} \mathbf{S}_j$ contains the market averages of the unobserved retailer variables. We assume that \mathbf{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 \mathbf{S}_j . We can rearrange the last equation to

$$\mathbf{v}_j - \mathbf{x}_j \boldsymbol{\alpha} \stackrel{(M \times 1)}{=} \beta \mathbf{H} \mathbf{S}_j + \mathbf{e}_j. \quad (7)$$

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 $\beta^2 \sigma_\eta^2 \mathbf{H} \mathbf{H}' + \sigma_e^2 \mathbf{I}_M$. We usually define $\sigma_\eta^2 = 1$ to set a metric (σ_η^2 and β 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 $\beta^2 \sigma_\eta^2 \mathbf{H} \mathbf{H}' + \sigma_e^2 \mathbf{I}_M$. In this case, the off-diagonals in $\mathbf{H} \mathbf{H}'$ 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 dependences in unobserved retailer behavior.

3.4 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 equation (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 and Liljeros 2003). In the latter approach, a contiguity matrix \mathbf{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., e.g. Okabe et al. 2000). These polygons use the (irregular, i.e., non-lattice) location of markets to exhaustively divide the US 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 \mathbf{C} add to 1, the matrix

\mathbf{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 \mathbf{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

$$\begin{aligned}\mathbf{v}_j &= \mathbf{x}_j\boldsymbol{\alpha} + \boldsymbol{\xi}_j\beta + \mathbf{e}_j, \\ \boldsymbol{\xi}_j &= \lambda\mathbf{C}\boldsymbol{\xi}_j + \boldsymbol{\eta}_j\end{aligned}\tag{8}$$

with both \mathbf{e}_j and $\boldsymbol{\eta}_j$ 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 \mathbf{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 $\boldsymbol{\xi}_j$ on the left hand side, we obtain that $\boldsymbol{\xi}_j = (\mathbf{I}_M - \lambda\mathbf{C})^{-1}\boldsymbol{\eta}_j$. The model above can therefore be statistically formulated as

$$\mathbf{v}_j - \mathbf{x}_j\boldsymbol{\alpha} = \boldsymbol{\xi}_j\beta + \mathbf{e}_j,\tag{9}$$

where the right hand side is distributed Multivariate Normal with mean 0 and variance covariance matrix equal to $\beta^2 (\mathbf{I}_M - \lambda\mathbf{C})^{-1} (\mathbf{I}_M - \lambda\mathbf{C})^{-1'} + \sigma_e^2\mathbf{I}_M$. Whereas this model has the same number of parameters as the model in equation (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.

3.5 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 United States. 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 \mathbf{W} , we need data on the total volume (ACV_{rm}) 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 \mathbf{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 equation (9) that is based on market contiguity. Finally, the third model (CHAIN) is the model in equation (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 US 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 \mathbf{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.

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)
σ_e^2	0.32	(11.03)	0.24	(1.5)	0.06	(1.8)
loglikelihood	-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.

4 Path dependent 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.

4.1 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 on *temporal* patterns of sales growth (see e.g., Mahajan, Muller and Bass, 1995). Recently, spatial and spatiotemporal patterns of diffusion have become the subject of empirical study (e.g., Bronnenberg and Mela 2002, Vandenburg and Lilien 2001). In addition to empirical methods, an other way to study spatial diffusion is by using differential equations derived from theoretical models (Edling and 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} 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 \mathbf{C} from the previous section, and arraying the market entry variables of $t - 1$ across markets into the $M \times 1$ vector \mathbf{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 $\mathbf{C} \cdot \mathbf{y}_{it-1}$. Denoting the m th row of \mathbf{C} by \mathbf{c}_m , the weighted average of past entry in neighboring markets is thus $\mathbf{c}_m \cdot \mathbf{y}_{it-1}$.

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 \mathbf{H} (defined previously as the M by K matrix \mathbf{H} containing the ACV share of chain k in market m). Write the m th row of \mathbf{H} by \mathbf{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 across chains to obtain a $K \times 1$ vector \mathbf{z}_{it} . Then, the total share of chains on market m that are already carrying the brand in other markets $m' \neq m$ is equal to the m th element of $\mathbf{H} \cdot \mathbf{z}_{it-1}$, which is equal to $\mathbf{h}_m \cdot \mathbf{z}_{it-1}$. To summarize, the adoption function U_{imt} above contains (potentially among other variables) the following components

$$U_{imt} = \alpha_i + \gamma_1 \mathbf{c}_m \cdot \mathbf{y}_{it-1} + \gamma_2 \mathbf{h}_m \cdot \mathbf{z}_{it-1}$$

Retailer adoption The second stage of the model focuses on the retailer's decision to adopt the brand in its assortment. As before, this decision can be represented as a probit model. Adoption can only occur if the brand is made available by the manufacturer in at least one market that belongs to chain k 's territory. Defining the moment of earliest entry into the trade area of retailer k by t_k^{avail} , and the moment of first time adoption by the retailer by t_k^{adopt} , we define

$$\Pr(z_{ikt} = 1) = \begin{cases} 0 & \text{if } t < t_k^{\text{avail}} \\ \Phi(V_{ikt}) & \text{if } t_k^{\text{avail}} \leq t \leq t_k^{\text{adopt}} \\ 1 & \text{if } t > t_k^{\text{adopt}} \end{cases}, \quad (11)$$

in which the terms V_{ikt} capture the attractiveness of brand i to retailer k at week t , and the function Φ is the again the cumulative standard Normal distribution. Of interest in this model is whether retailer adoption decisions depend on similar decisions made by its direct rivals. As outlined in the previous section, such an effect can be introduced as a network effect. Implementation in the adoption model proceeds by making attractiveness V_{ikt} dependent on past adoption by rival retailers. Rival retailers are identified by the $K \times K$ matrix \mathbf{W} (defined previously) whose rows add to one, and whose entries $[k, k']$ are 0 if k and k' do not compete in the same geographic markets and positive if they *do* compete directly. Also define the k th row of \mathbf{W} as \mathbf{w}_k . To define the diffusion variable of retailer adoption, array the K distribution variables z_{ikt-1} at $t - 1$ across markets into the $K \times 1$ vector \mathbf{z}_{it-1} . Next, the value of the diffusion variable is the k th element of the spatially and temporally lagged chain adoption variables $\mathbf{W} \cdot \mathbf{z}_{it-1}$. For each retailer k this variable assumes the value $\mathbf{w}_k \cdot \mathbf{z}_{it-1}$. These variables can be interpreted as weighted averages of past adoptions by competing retailers. The weights capture the degree of influence by each direct competitor in one's trade area. Thus, a model for V_{ikt} would contain (among other components)

$$V_{ikt} = \theta_i + \gamma_3 \mathbf{w}_k \cdot \mathbf{z}_{it-1}$$

Bronnenberg and Mela (2002) use chain and market level data from the Frozen Pizza industry and find evidence for the spatial (geographic), selection, and network (retailer) effects that are implied by the effects $\gamma_1 - \gamma_3$ respectively. Further, it was found that retailer adoption and manufacturer roll-out reinforce each other. This means that lead-market selection is non-trivial in the sense that brands diffuse faster from some markets than others. Bronnenberg and Mela (2002) find that attractive lead markets are those that are on a common edge of multiple large retailer trade areas.

Obviously, this work does not stand alone, but is a part of an existing stream of empirical studies in network and spatial diffusion. For instance, the seminal paper by Strang and Tuma (1993) provides alternative measures of spatial and social contagion. Wasserman and Faust (1994) give a very complete overview of social contagion variables. VandenBulte and Lilien (2001) argue that it is important to test for rival explanations for social contagion. In a reanalysis of the famous data from Coleman, Katz and Menzel (1966), they show that interpretations of contagion can be confounded with marketing mix activity such as sales-calls or advertising. In marketing, other studies have found that market characteristics, culture and demographic details, number of urban conglomerations and similarities between countries and size or importance of the old technology influence international diffusion. (Dekimpe, Parker and Sarvary, 2000). Network diffusion, which started with research on innovations (e.g., Valente, 1995) and on sociology (e.g., Wasserman and Faust, 1994), attempts to formalize the links between the different participants in the network and explain the diffusion process.

4.2 Order of entry and consumer learning

Spatial concentration can emerge from the combination of consumer learning processes and local order-of-entry (the latter is implied by the model above). That is, order-of-entry in a certain market influences consumer preferences if such preferences follow a learning process that is based on past experience. For instance, in product categories in which consumer preferences are initially diffuse (e.g., high tech products, discontinuous innovations), several studies found that consumer preferences are not exogenous but are formed on the basis of an anchoring-and-adjustment process (Kahneman, Slovic and Tversky, 1982; Kahneman and Snell 1990). In this process, consumers learn about their

own preferences from the available choice options. In a similar context, Carpenter and Nakomoto (1989) find that, over time, the ideal point of the consumer (i.e., what the consumer wants) tends to shift toward the pioneer’s location in perceptual space. In effect, the pioneer becomes the prototype for the category and an asymmetric product comparison process emerges between the pioneer and later entrants (see also Tversky 1977).

An effective model of path dependent preferences is given by Pólya (1931). In this model, a consumer’s choice history is represented by an urn with different brands represented by balls of different colors (say two for simplicity). For discussion, suppose the balls are either red or green. At time $t = 0$ the urn contains G_0 green and R_0 red balls.

The characteristic process that gives rise to Pólya’s urn is that balls are randomly drawn from the urn with replacement of B additional balls of the last drawn color. As an example, if at $t = 0$, $G_0 = R_0 = B = 1$, then at $t = 1$, we replace a red draw with 2 red balls and a green draw with 2 green balls. At $t = 1$, both these events happen with equal likelihood. However, at $t = 2$, the likelihood of drawing either a red or a green ball depends on the previous draw and favors the color that was drawn at $t = 1$. As more and more balls are added to the urn, the odds of drawing either red or green keep changing depending on all past draws. However, it is readily verified that as the number of the balls in the urn increases, the proportion of green (or red) balls in the urn will become constant. In other words, there exists a stable distribution of the long-run share of green balls in the urn. Pólya (1931) proved that this distribution is a Beta distribution with parameters G_0/B and R_0/B .

Figure 3 illustrates. Each panel in this figure gives the distribution density of the long term proportion of green balls in the urn (between 0 and 1). Moving across panels horizontally, the expected proportion for green remains constant at $G_0/(G_0 + R_0)$, i.e., 0.5 in the top graphs and 0.33 in the bottom graphs.

The growth rate B increases across the panels from left to right. The associated distribution of the equilibrium proportion for $G/(G + R)$ goes from unimodal (suggesting a tendency to stay close to the initial conditions) to U-shaped (suggesting a tendency for one color to dominate). *Ex ante*, the expectations for the share of green are identical. However, the variance of these expectations is higher when the growth rate is high compared to the size of entry (the initial conditions). Moreover, when the growth rate is high enough, the urn becomes “tippy” in the sense that one color tends to

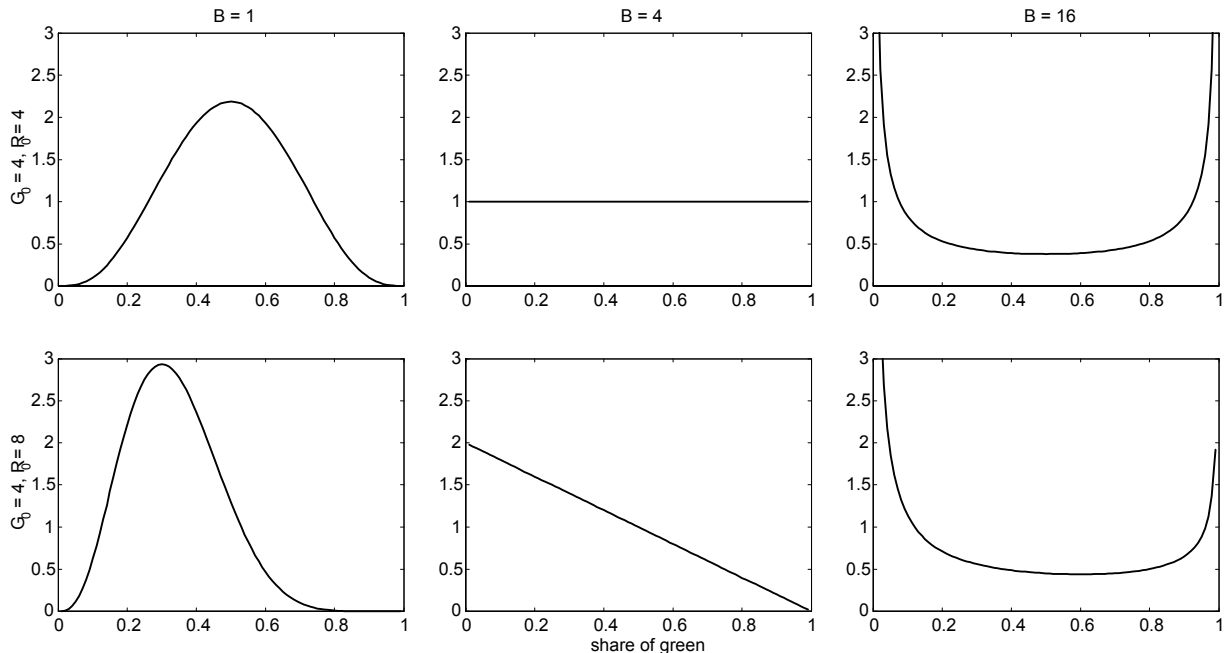


Figure 3: Density of the long-term market share of the “green product” in the Pólya urn.

dominate (shares of 0 or 1 are most likely).

As stated, this process can operate as a representation of path-dependent consumer preferences, especially when buying behavior is based on past choices. The contents in the urn substitutes for experience of the consumer in the category at hand. The parameter B can be seen as a learning parameter which controls the speed of updating preferences for the brands that have been purchased in the past. The steady-state distributions now represent brand preferences. The model captures both those consumers who repurchase out of inertia (those that update “fast” so that they either favor one brand or another) or consumers who consider more brands (those that update “slower”). Another appealing characteristic of this interpretation is that updating of preferences occurs most when the consumer is inexperienced. Purchase feedback becomes less informative when the consumer gains experience.

This model predicts that in a market with “Pólya consumers,” early entrants will generally end up with larger market shares than later entrants. This is the case because initial choices are reinforced in this process. Furthermore, successful entry and influencing consumer preferences for new brands becomes harder after a critical amount of learning has taken place. This is because preferences change less and less as experience grows. Implicitly, the Pólya urn implies that there is



Figure 4: Spatial variability of market shares for undifferentiated goods

an opportunity window outside of which it is difficult and more expensive to enter the market.

Together, the feedback model of retailer distribution and market rollout, and the model of path-dependent consumer learning are consistent with the emergence of spatial concentration of market shares. The following section addresses two mechanisms of why spatial concentration of market shares tends to persist.

5 Marketing strategy and sustenance of spatial concentration in brand shares

Spatial concentration of market shares once established often persists. For instance, Figure 4 visualizes local market shares (averaged over two years of weekly data) for four brands of Mexican Salsa across sixty-four different geographical markets in the United States. The weekly market shares are stable across time. It is an interesting puzzle that in the face of this apparent lack of product differentiation, the observed market share differences can be sustained. Below, we discuss two broad theories that may help to explain this puzzle.

5.1 Spatial distributions of consumer tastes and path-dependence

Consumers may not be homogeneously distributed across space (in either quantity or type). If consumers are immobile (i.e., if intermarket distances are large enough), the Pólya process leads to local preferences that reflect the entry decisions by brands at the market level. If the Pólya process becomes a representation of the market, the ex ante prediction of long term share of a brand would be a random draw from the Beta distribution with parameters based on initial conditions. Note from Figure 3 that therefore market shares can stabilize around different values in different locations. In this explanation, the variation in market shares across markets is caused by the fact that the growth process takes different (sample)-paths in markets with different order-of-entry patterns.

The stability of the market structure or the persistence of concentration is caused by the fact that the Polya process will “lock in” a certain division of market shares after a growth process during which the markets are in flux. A defining characteristic of this explanation (at least in its pure form) for spatial concentration is that firms can not change the market structure once it has locked in.

Although consumer mobility can be used to explain the differences in shares across large distances, to a lesser degree consumer mobility even impacts retailer price-discrimination strategies at the neighborhood level as well. For instance, retailers charge higher prices in neighborhoods that have more consumers with higher travel cost or lower mobility (Hoch et al 1995).

5.2 Multi-market contact

In addition to the lock-in of market shares in path-dependent models, another reason for why spatial concentration may persist is that it is beneficial for the manufacturers to sustain it. In this interpretation, spatial concentration is the outcome of manufacturer competition when consumers are immobile. Especially if firms compete in many markets, it is a priori not clear whether they are better off dividing the universal market geographically into local markets with low and high market power, or, conversely, having symmetric market shares in all markets. Anderson, de Palma and Thisse (1992) show that within-market competition becomes more and more fierce as the differentiation of brands becomes less in the eyes of consumers. In such cases, multi-market contact among the same set of firms could achieve that firms maintain a pre-existing differentiation on the basis of geography (i.e., exploit the lack of consumer arbitrage across markets). This mutual forbearance hypothesis was introduced Bernheim and Whinston (1990) and has since received much attention in

the literature on economics and strategy (e.g., Baum and Greve 2001).

Directly related to the data in Figure 4 is a proposition by Karnani and Wernerfelt (1985). They introduce a so-called “mutual foothold” equilibrium in which firms take a large lead in some geographic markets but maintain a small position in other markets. This small position (the foothold) allows the locally small firm to inflict damage on attackers in its large markets. Mutual footholds then suffice to keep all players from attacking each other in the markets where they are large. For the top three brands in the Mexican Salsa category this seems a feasible explanation for why the brands do not exit the markets in which they have sometimes very small market shares.

Another strategic yet rather different reason for asymmetric market power in local markets is to allow that some product-unrelated source of differentiation is under control of firms. Yarrow (1989), using a duopoly model of logit demand, shows the existence of three candidate equilibria when firms first set advertising and then prices. One of these candidates is a symmetric equilibrium, while the two remaining candidates are mirror images of an asymmetric market outcome in which one firm advertises more than the other and has a higher profit margin. Yarrow (1989) characterizes the existence conditions for these candidate equilibria, and finds that the asymmetric equilibrium is unique when the product category is undifferentiated whereas both the existence and the uniqueness of the symmetric equilibrium requires a lower-threshold of product differentiation. This means that as long as product categories are well differentiated symmetric firms will compete with symmetric outputs (in each market). However, when the danger of ruinous price competition looms large in cases of undifferentiated goods, symmetric firms may compete by creating differentiation based on advertising investments. A surprising aspect of Yarrow’s analysis is that the asymmetric equilibrium can be sustained even in a single market.

In sum, while some geographic markets are similar in aspects such as size, prices, consumer characteristics, etc., the associated market structures can be different. For example, a market may be highly concentrated, with one brand having a large share, while other markets may have numerous brands fighting aggressively. It is important to understand the reasons why markets evolve like they do and what makes one brand so predominant in one region but less significant in others. In this context, it is fortuitous that empirical data are becoming available to test alternative models of product-growth and market-structure.

6 Conclusions

Geographical space is an important ingredient of marketing strategy and marketing practice. Consumer immobility, transportation cost of the firm, advertising “markets,” retailer trade areas, distribution channels, etc. are all ingredients that make a case for the relevance of physical space in marketing and strategy. Spatial price discrimination, sustenance of asymmetric market power, etc., are likely an outcome of using geographical space as a source of differentiation in competition even when product differentiation is not enough to sustain profits. Despite this, currently, geographical space is not an important ingredient in the academic tradition of theory building in marketing or economics. Indeed, much theory building in marketing concentrates on within-market research questions. We hope that this chapter is instructive in suggesting ways in which spatial growth of new products and spatially concentrated outcomes of these growth processes can be modeled.

At least three avenues for future empirical research seem important. The first should focus on descriptive models of spatial growth. Research that combines both temporal and spatial data for the study of such models is scarce, but the data have recently become available in packaged goods. Second, not much work has been done to analyse the observed differences in within firm marketing strategy across markets. Indeed, multimarket data provide a great opportunity to study firm decision making with respect to advertising and pricing decisions within and across markets. A final area in which spatial analysis can play a major role in theory building is work on positioning new products in the attribute space. The Defender model (Hauser and Shugan, 1983) is one of the most used approaches to position new products and defend incumbents in marketing. It makes use of a perceptual map where each brand is defined by the location of two attributes and consumers have a preference distribution on those attributes. Elrod (1988) developed the model to identify the positions of the brands in a perceptual map from panel data. The implications of such and other “address” models of product positioning are only currently being uncovered (see e.g., Berry and Pakes 2001).

7 References

- Anderson, Simon P., André de Palma (1988), "Spatial Price Discrimination with Heterogenous Products," *the Review of Economic Studies*, vol 55, 573-592.
- , —————, and Jacques-François Thisse (1992), "*Discrete Choice Theory of Product Differentiation*," MIT press, Cambridge MA.
- Anselin, Luc (1988), "*Spatial Econometrics: Methods and Models*," Kluwer Academic Publishers, Boston.
- Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science*, 15(5), 215-227
- Baum, Joel A.C. and Henrich R. Greve (2001), "Multiunit Organization and Multimarket Strategy," *Advances in Strategic Management*, vol 18, JAI, Amsterdam.
- and Jitendra V. Singh (1994), "Organizational Niches and the Dynamics of Organizational Mortality," *American Journal of Sociology*, vol 100, p. 346-380
- Bernheim, B. Douglas and Michael D. Whinston (1990), "Multimarket Contact and Collusive Behavior," *Rand Journal of Economics*, vol 21:1 (Spring), 1-26.
- Berry, Steven and Ariel Pakes (2001), "The Pure Characteristics Discrete Choice Model with an Application to Price Indices," NBER working paper.
- Bonanno, G. (1990), "General equilibrium theory with imperfect competition," *Journal of Economic Surveys*, 4, 297-328.
- Bronnenberg, Bart J. and Vijay Mahajan (2001), "Unobserved Retailer Behavior in Multimarket Data" Joint Spatial Dependence in Market Shares and Promotion Variables," *Marketing Science*, 20:3, 284:299
- and Catarina Sismeiro (2002), "Using Multimarket Data to Predict Brand Performance in Markets for which No or Poor Data Exist," *Journal of Marketing Research*, vol 39, February, 1-17
- and C. Mela (2002), "Market Roll-out and Retail Adoption for New Brands of Non-durable Goods," *working paper*, no. 372, Anderson Graduate School of Management, UCLA
- Carpenter, G. and K. Nakamoto (1989), "Consumer Preference Formation and Pionnering Advantage," *Journal of Marketing Research*, 26 (Aug.), 285-98.
- Coleman, J.S., Katz, E. & Mentzel, H. (1966), "*Medical Innovation*," New York: Bobbs-Merrill.
- Dekimpe, M., P.M. Parker and M. Sarvary (2000), "Globalization: Modeling Technology Adoption Timing Across Countries", *Technological Forecasting and Social Change*, 63, 25-42
- Dixit, A. K., and J. E. Stiglitz (1977), "Monopolistic Competition and Optimum product Diversity", *American Economic Review*, 67 (3), pp. 297-308.
- Donthu, N. and A. Garcia (1999), "The Internet Shopper", *Journal of Advertising Research*, 39(3): 52-58.

- Edling, Christopher, and Fredrik Liljeros (2003), "Spatial Diffusion of Social Organizing: Modeling Trade Union Growth in Sweden 1890-1940," *this volume*
- Elrod, T. (1988), "Choice Map: Inferring a Product-Market Map from Panel Data", *Marketing Science*, Vol. 7, No.1, Winter, pp. 21-40.
- EMarketer (1999). *The eGlobal Report*, July, <http://www.emarketer.com>
- Fujita, M., P. Krugman and A. J. Venables (1999), "*The Spatial Economy*", The MIT Press, Cambridge, Massachusetts.
- Fujita, M., Thisse J.-F. (2002), *Economics of Agglomeration*, Cambridge University Press.
- Greenhut, M.L (1981), "Spatial Pricing in the United States, West Germany, and Japan," *Economica*, vol 48, 79-86
- Hauser, J. R. and S. M. Shugan (1980), "Intensity Measures of Consumer Preference", *Operations Research*, Vol.28, No. 2, March-April, pp.278-320.
- Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi (1995), "Determinants of Store-Level Price Elasticity", *Journal of Marketing Research*, vol 32 (February), 17-29.
- Hofstede F.T., M. Wedel and J. B. Steenkamp (2002). "Identifying Spatial Segments in International Markets", *Marketing Science*, forthcoming.
- Kahneman, D. , P. Slovic and A. Tversky, eds. (1982), *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge University Press.
- and J. Snell (1990), "Predictin Utility", in *Insights in Decision Making: A tribute to Hillel J. Einhorn*, Robin M. Hogarth, ed. Chicago: University of Chicago Press, 295-310
- Kalyanaram G., W. T. Robinson and G. L.Urban (1995). "Order of market entry: established empirical generalizations, emerging empirical generalizations and future research", *Marketing Science*, Vol. 14, No. 3.
- Kardes, F. and G. Kalyanaram (1992), "Order-of-Entry Effects on Consumer Memory and Judgment: An Information Integration Perspective", *Journal of Marketing Research*, August, pp. 343-357
- Karnani, Aneel and Birger Wernerfelt (1985), "Multiple Point Competition," *Strategic Management Journal*, vol 6, 87-96
- Lesage, James, (2000), "Bayesian Estimation of Limited Dependent variable Spatial Autoregressive Models", *Geographical Analysis*, Volume 32:1, pp. 19-35. 1999
- Lomi, Alessandro, Erik R. Larsen, and Ann van Ackere (2003), "Organization, Evolution, and Performance in Neighborhood-based Systems," *this volume*.
- Lynch, P. and J. Beck (2001), "Profiles of Internet Buyers in 20 Countries: Evidence for Region-Specific Strategies", *Journal of International Business Studies*, 32, 4, 725-748
- Okabe, Atsuyuki, Barry Boots, Kokichi Sugihara, and Sung Nok Chiu (2000), "*Spatial Tessellations*," Wiley and Sons, New York, NY.
- Pólya, George (1931), "Sur Quelques Points de la Théorie de la Probabilités", *Ann. Inst. Henri Pointcaré*, vol 1, 117-161

- Rossi, P. and Allenby, G. (forthcoming), "Statistics and Marketing", *Journal of American Statistics Association*.
- Strang, David, and Nancy Brandon Tuma (1993), "Spatial and Temporal Heterogeneity in Diffusion," *American Journal of Sociology*, 99:3, 614-639
- Tversky, A. (1977), "Features of Similarity", *Psychological Review*, 84 (July), 327-52.
- Urban, Glen L., T. Carter, S. Gaskin and Z. Mucha (1986), "Market Share Rewards to Pioneering Brands: An Empirical Analysis and Strategic Implications," *Management Science*, 32 (June), 645-659
- Valente, Thomas W. (1995), "*Network Models of the Diffusion of Innovations*," Hampton Press, Cresskill, NJ
- Van den Bulte, Christophe and Gary L. Lilien (2001), "Medical Innovation Revisited: Social Contagion versus Marketing Effort," *American Journal of Sociology*, 106 (5) 1409-1435
- Wasserman, S. and K. Faust (1994), "*Social Network Analysis - Methods and Applications*", Cambridge University Press
- Yarrow, George (1989), "*The Kellogg's Cornflakes Equilibrium and Related Issues*," Working paper, Hertford College, Oxford.