From Density to Destiny: Using Spatial Dimension of Sales Data for Early Prediction of New Product Success

Tal Garber, Jacob Goldenberg, Barak Libai, and Eitan Muller

November 2003

Tal Garber, School of Business Administration, the Hebrew University of Jerusalem, Jerusalem, Israel 91905 tal.garber@intel.com

Jacob Goldenberg, School of Business Administration, the Hebrew University of Jerusalem, Jerusalem, Israel 91905 msgolden@vered.huji.ac.il

Barak Libai, Davidson Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, and Leon Recanati Graduate School of Business Administration, Tel Aviv University, Tel Aviv, Israel 69978 libai@post.tau.ac.il

Eitan Muller, Leon Recanati Graduate School of Business Administration, Tel Aviv University, Tel Aviv, Israel 69978 muller@post.tau.ac.il Fax: +972 (3) 640-6321

The authors would like to thank Noa Golan, Oded Lowengart, Steven Shugan, an Area Editor, and two anonymous reviewers for their helpful comments and suggestions.

From Density to Destiny: Using Spatial Dimension of Sales Data for Early Prediction of New Product Success

Abstract

One of the main problems associated with early-period assessment of new product success is the lack of sufficient sales data to enable reliable predictions. We show that managers can use the spatial dimension of sales data to obtain a predictive assessment of the success of a new product shortly after launch time.

Based on diffusion theory, we expect that for many innovative products, word-of-mouth communications play a significant role in the success of an innovation. Because word-of-mouth spread is often associated with some level of geographical proximity between the parties involved, one can expect that "clusters" of adopters will begin to form. Alternatively, if the market general reaction is reluctance to adopt the new product, word-of-mouth effect is expected to be significantly smaller, leading to a more uniform pattern of sales. Hence, the product whose distribution is further away from a uniform distribution, or any other *known* distribution, will have a higher likelihood of generating a "contagion process" and therefore of being a success. This is also true if the underlying baseline distribution is non-uniform, as long as it is a known distribution.

To analyze this spatial phenomenon, we propose using a spatial divergence approach based on the Ali-Silvey class of divergence measures to determine the "distance" between two distribution functions. We apply this approach to both simulated and real-life data. Using two divergence measures, we find that the spatial divergence approach is capable of predicting success in the beginning of the process, which makes it appealing for use in marketing activity in general, and particularly for launches of new products. When applied to 17 actual product introductions, the method succeeded in correctly predicting the success or failure of the products in 15 cases.

1. Introduction

A marketing manager faces the following problem: S/he has three months' worth of bi-weekly sales data on two new products sold in the same geographical areas, targeted at roughly the same segments, using the same distribution channels. Both time series look similar, not counting random fluctuation. The manager, however, feels that one of the two products is not as successful as the other, and would like to back this hunch with quantitative data, as decisions must be made urgently regarding advertising, pricing, and promotion for the two products.

Such dilemmas are quite common, because although much of the research on new product introductions has traditionally focused on improving the initial go / no-go decision, one of the main problems associated with early-period assessments is the lack of data to enable further predictions. The few periods of aggregate sales of monthly data available to marketers render predictions relatively unreliable. Given the high failure rate of new products, as well as the often large investment in new product introduction, it is critical to assess the success potential of an innovation as early as possible in order to avert preventable financial losses and concentrate resources on the support of innovations that have a high chance of success. Although there is a school of research focused on early assessment of a product's likelihood of success, with few exceptions, there has been relatively little success in transforming early sales data into valid measures for predicting post-launch product success (Golder and Tellis 1997).

One of the problems associated with early-period assessments is the lack of data for enabling further prediction. For example, it has been suggested that using diffusion models such as the well-known Bass model, a stable estimate of the diffusion process parameters, requires sales data from introduction through peak of sales (Srinivasan and Mason 1986).

Therefore, it has been argued that diffusion models cannot serve as an effective predictive tool for the early years of a product's life (Mahajan, Muller, and Bass 1990; Kohli, Lehmann, and Pae 1999) unless combined with other information, from meta-analysis or market potential or multigenerational substitution (Sultan, Farley, and Lehmann 1990).

We propose a novel approach to overcome some of the problems associated with early prediction of new product success based solely on sales data. Our approach is based on the use of the spatial dimension of new product growth. Instead of using only time series of aggregate sales, we analyze the spatial dimension of the sales, i.e., not only *how much* the product sold, but also *where*. The rationale of our approach stems from diffusion theory. Specifically, our proposed approach is valid for products where the *internal influence* from previous adopters, namely, word of mouth and imitation, play a significant role in the success of an innovation (Mahajan, Muller, and Srivastava 1990). In order for a product to be successful, word-of-mouth and imitation effects must begin to take place. Because word-of-mouth spread is often associated with a certain level of geographical proximity between the parties, one can expect that "clusters" of adopters will begin to form (Allaway, Mason, and Black 1991; Rogers and Shoemaker 1971).

Conversely, if the overall market reaction is reluctance to adopt the new product, then word-of-mouth effect is expected to be significantly weaker, with product trials being the result of external efforts such as advertising, leading to a more uniform geographical distribution of adopters. Eventually, with the absence of a take-off, such a product might be declared a failure. Thus we can examine how "far" the adoption process is from a uniform geographical distribution. The product whose distribution is "further away" from a uniform distribution will have a higher likelihood of generating a contagion process, and therefore a take-off leading to success. As we show later, this scenario will also occur if the underlying baseline distribution is non-uniform, as long as it is a known distribution.

In this paper, we start with the relationship between the spatial and the temporal dimensions of new product growth and then explain and illustrate the methods used. Next, we utilize two measures of product success to determine the predictive ability of spatial divergence in both Cellular Automata and Small World environments. We then use spatial data on the diffusion of real products to determine the efficacy of our approach in the early periods of product introduction. Finally, managerial and methodological implications of the method, as well as its limitations, are discussed.

2. Pairing spatial and temporal dimensions of growth

When new products enter a market, they diffuse both in time and space. As compared with the wide interest in marketing in the temporal diffusion of new products, relatively scant attention has been paid to the spatial pattern of growth and its relationship to the temporal one. Few exceptions are Mahajan and Peterson (1979), Allaway, Mason, and Black (1991) and Bronnenberg and Mahajan (2001). Other studies have been conducted that examine the global diffusion of technologies focused on the proximity between countries and geographical regions to explain the temporal patterns of growth (Dekimpe, Parker, and Sarvary 2000; Putsis, Balasubramanian, Kaplan, and Sen 1997).

The rationale for the formation of adopter clusters is related to the role of word of mouth and imitation in the diffusion of innovations. Assuming the diffusion paradigm that views the communication process as the main driver of new product growth, one can see two types of communication effects: external and internal (Mahajan, Muller, and Bass 1990). While the external effect relates to the marketing efforts of the firm, the internal effect represents the influence of previous adopters, and provides two concurrent positive reinforcements: positive word of mouth, and a source for imitation and legitimation. To a large extent, internal effects represent the market reaction to the product. If the product is well received, then word of mouth and imitation will carry forth the message, followed by more and more adopters, further feeding the flow of internal influence. Eventually, when reaching a sufficient level of internal effects, the product will take off (for the rest of the paper we use the term “word of mouth” to mean all types of internal effects).

Internal effects have been found to be the underlying and driving force of innovation diffusion of many new products, exerting an influence exceeding that of external marketing efforts such as advertising (Rogers 1995; Goldenberg, Libai, and Muller 2001; Mahajan, Muller, and Srivastava 1990). However, for internal effects to take place and personal recommendations to begin circulating, adopters typically must share some form of physical proximity. Indeed, the diffusion literature reports a clear correlation between geographic proximity and the strength and speed of word-of-mouth spread, sometimes labeled the “neighborhood effect” (Baptista 2000; Case 1991; Mahajan and Peterson 1979). It is easy to see, even with a simulation as simple as that demonstrated below, that from a geographical point of view, word-of-mouth effects drive the formation of spatial clusters.

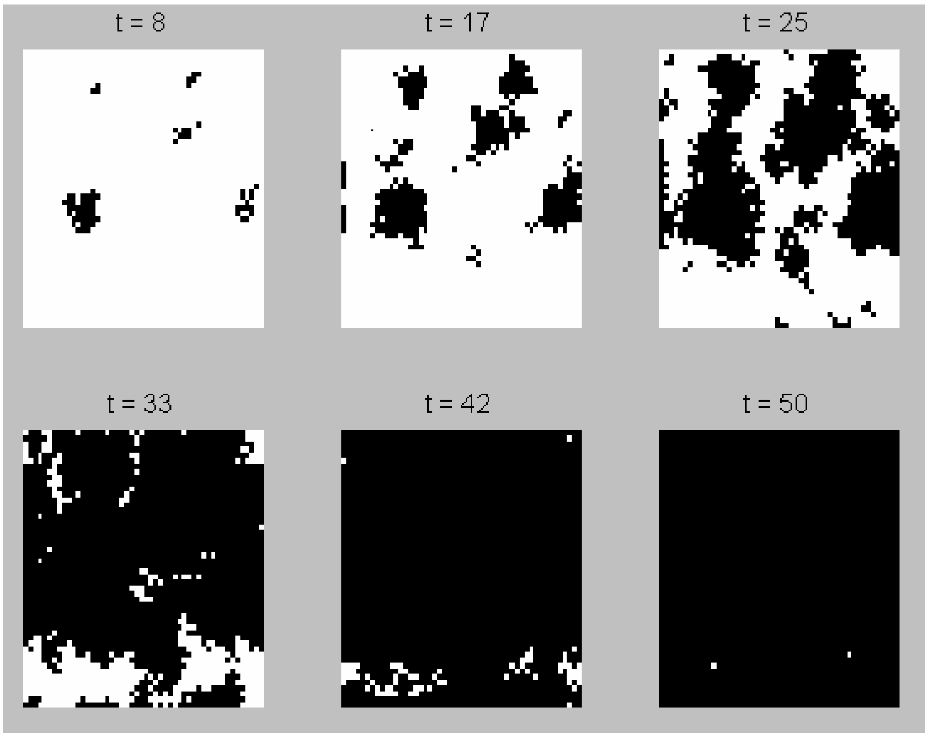
The above discussion concerned how—based on diffusion theory—one can expect a successful product to grow in a market. However, if the product in question is a “dud”, then one can expect the internal effects activity associated with it to be minimal, with no contiguous units buying it. While some consumers will adopt it, mainly as a result of external effects, the effect of their adoption on other consumers will be negligible: Clusters will not form, and spatial distribution will be mainly a result of external effects, i.e., adopters will be randomly distributed in space in what can be expected to be a geographical distribution close to a uniform distribution. Thus, a lack of contiguous units adopting the product may be a strong signal for the product’s failure.

This is not to say that advertising is not important to the growth of new products.

First, it is critical at the early stages of the diffusion process. Second, it might be crucial to the consumer's selection of a specific brand. Third, there are products for which diffusion theory may be less relevant, as, for example, in the case of non-durable brand extensions, where consumers are mainly affected by advertising, promotions, and channel availability. However, for many new products, especially discontinuous innovations, marketers can expect diffusion theory to reasonably represent the process, and thus without market acceptance in the form of word of mouth and imitation, the product will not "take off". Indeed, as Golder and Tellis (1997) note, the absence of the takeoff is a decision criterion for managers to "kill" a product.

The following figures illustrate our main contention (the exact methods used will be discussed in the following sections). In Figure 1 we present a simulated adoption of a product (Product 1) in six discrete time periods (8, 17, 25, 33, 42, and 50) throughout a certain rectangular geographical area. The product is clearly adopted in clusters.

Figure 1: Spatial adoption of Product 1 (in clusters)



In comparison, in Figure 2 we present the adoption of the second product, whose distribution of adopters in the same geographical area is relatively uniform. The difference in the adoption curves of the two products is evident in Figure 3.

Figure 2: Spatial adoption of Product 2 (uniform)

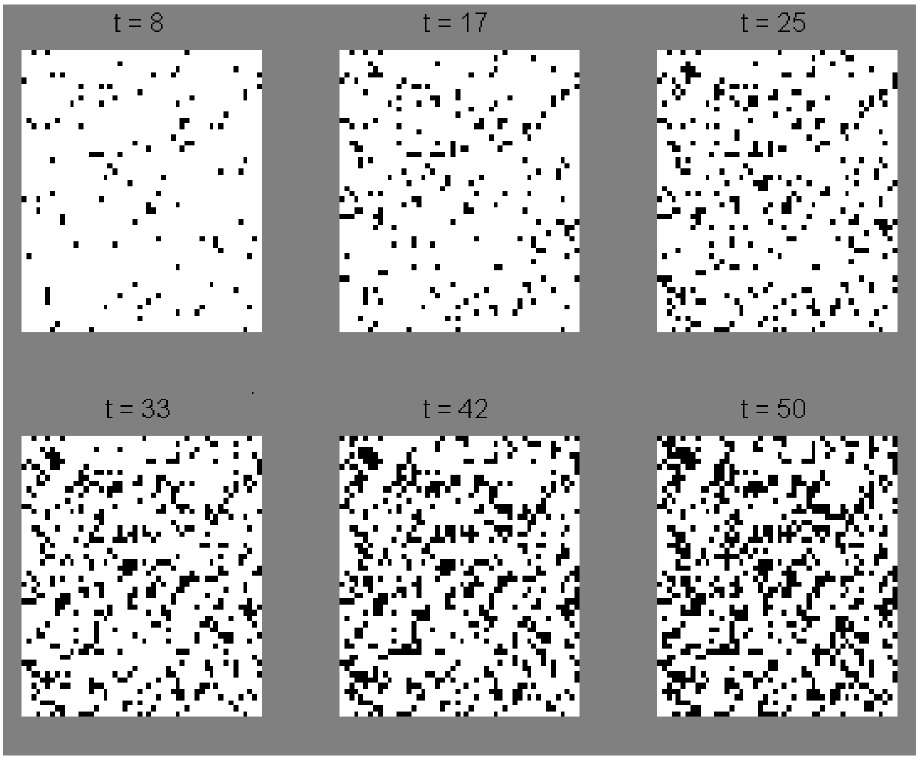
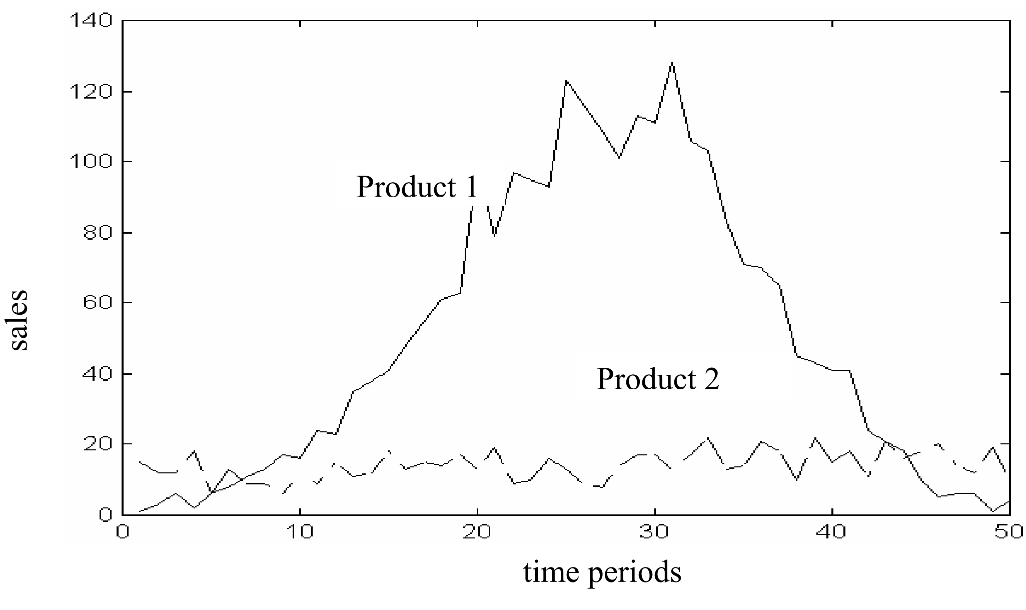


Figure 3: Temporal adoption of products 1 (high) and 2 (low)



Looking at Figures 1 and 2, one sees a pronounced difference in the spatial distribution of adopters, though the aggregate time series are more vague regarding the eventual success of the two products. Our claim is that as early as Period 8 after products' introductions, it is possible to predict the eventual success / failure of each based on their respective spatial diffusions, by comparing their spatial distribution to their uniform distribution (we will discuss the use of non-uniform baseline distribution in section 4).

3. Spatial divergence as a predictor of success: Method

3a. Measures of spatial divergence

As previously explained, we are interested in measuring the divergence between the spatial distribution of the product in question and its uniform distribution. There is a rich literature concerning divergence measures between two distributions (Lee 1999; Johnson and Sinanovic 2001). A widely used group of such measures is the Ali-Silvey class, which measures the expectations of (a function of) the likelihood ratio of the two distributions. Consider two probability functions p_1 and p_2. A (discrete) Ali-Silvey divergence measure is represented by:

$$(1) \qquad d(p_1, p_2) = \sum_x p_2(x)\, f[p_1(x)/p_2(x)]$$

Where f is a convex function (see Csiszar 1991). This divergence measure was constructed in an axiomatic approach so as to satisfy several rational statistical conditions ("axioms") such as consistency, continuity, and locality, where for example, the latter requirement is similar to the Independence of Irrelevant Alternatives condition in choice models. Moreover, it is the only class that satisfies these axioms (Csiszar 1991). Due to their rigorous statistical properties, measures of the Ali-Silvey class are being used in many applications such as machine learning, data mining, information theory, and statistical physics (Kullback 1997). Of this class, the most commonly used divergence measure is the *cross-entropy*, or *Kullback-Leibler* divergence (see Kullback 1997) in which the function f of Equation 1 is the following: $f(x) = x \log x$. Thus the cross-entropy divergence can be represented by the following distance:

$$(2) \qquad KL\,(p_1|\,p_2) = \sum_x \; p_1(x) \log[p_1(x)/p_2(x)]$$

As the logarithmic function $\log[p_1(x)/p_2(x)]$ in Equation 2 can be represented as $\log(p_1) - \log(p_2)$, *cross-entropy* sums up the distance between any two points in the two distributions weighted by the probability that these points could occur. Being concave, the log function smoothes out extreme values, while multiplication by $p(x)$ assures that higher-probability events receive a higher weight. While this function is not symmetric and thus is not a metric, its appeal lies both in its robustness and the fact that it captures the distance of patterns between distributions in a single parameter value.

In order to overcome the lack of symmetry of the above measure, another measure in the Ali-Silvey class can be formed (Lee 2001): The *Jensen-Shannon* measure is obtained by assigning the function f of Equation 1 as: $f(x) = -x \log[(1+1/x)/2] - \log[(1+x)/2)]$.
If $KL(p_1|p_2)$ represents cross-entropy as in equation 2, then following is the *Jensen-Shannon (JS)* divergence measure:

$$(3) \qquad JS\,(p_1, p_2) = KL(p_1|(p_1+p_2)/2) + KL(p_2|(p_1+p_2)/2)$$

In this paper we have performed the calculations on our models using both measures of cross-entropy and the Jensen-Shannon divergence. We found virtually no difference between the predictive abilities of the two measures. The correlation between the results using the two divergence measures was above 98%. We report most results using both measures.

In order to operationalize the divergence measure, we must first determine the geographical units according to which we examine the data. In actual marketing situations, marketers have to base their decisions on whatever division is available in their data source. For example, the geographical unit might be based on retail establishment sales, regional sales, statewide sales, and so forth. Yet assuming that geographical databases can be finely constructed, we wish to discover the division that allows for efficient analysis of divergence. We therefore divide the geographical area under consideration into smaller subsets, called "windows." The density of adoptions for each window would be the number of adopters in the window divided by the total market potential.

Two opposing considerations should be taken into account when determining the window's size. The first is the adequacy of the representation of the distribution by the complete set of windows. For this purpose, a high resolution is needed and a smaller window is required. However, by decreasing the window's size, the number of individuals inside the window decreases accordingly, causing possible decrease in estimation accuracy. This possibility leads to the second consideration: In order to reduce the noise of the estimated distribution curve, a larger window is required. These two demands are contradictory, and thus it is important to define an efficient window size. One recommended approach is to employ *Parzen Windows* of equal size as the area unit of analysis, an important and widely used technique considered an efficient method for the construction of density functions (Parzen 1962). Note that even with equal-size windows, their number still remains to be determined. An efficient solution as suggested by Parzen (1962) is to set up windows of equal size that are proportional to the square root of the potential market. In the Appendix we show how the forecasting accuracy changes when we move from coarse partition to finer partitions.

In the first three studies, we will use a simulated environment to evaluate the model. In order to create this environment, we utilize two different complex system tools called *stochastic cellular automata* and *small world*. Then, in three more studies we will test (or validate) the proposed approach on field data of different cases. We will first explain the nature of the cellular automata and the small world environments, as well as the measure we use for a new product's success in this environment.

3b. Cellular Automata and Small World

In order to calibrate and validate the proposed method, one must create dyadic sets of data of successes and failures with a reasonable variance within these sets. A convenient and efficient way to achieve this objective is to utilize a *Complex Systems* approach in which a system is analyzed through a simulated "would-be world" that allows testing of a wide range of scenarios. In this case, a simple yet appropriate technique is *Stochastic Cellular Automata*, a computational tool aimed at analyzing problems of complex systems. Cellular automata have been extensively applied in disciplines such as physics, computer science, and biology. Recently, an increasing interest in using complex systems techniques such as cellular automata has arisen in business disciplines such as economics and marketing (Krider and Weinberg 1997; Goldenberg, Libai, and Muller 2001).

We follow previous approaches (Goldenberg, Libai, and Muller 2001, 2002) to let the elements of the matrix component of the model represent adoption of individuals comprised of various discrete cells, where the location of each cell is determined and taken into account in order to provide a spatially meaningful process. Each cell interacts with its neighboring cells, and this interaction evolves in time and may produce complex behaviors. Each of the cells can exist in more than one state, and the value of the cell in the next time unit depends on that cell's current state and the states of its neighboring cells.

We define a cellular automata-based model as consisting of three components:

i) A matrix representing the behavior of individuals

ii) The relationships among individuals

iii) The transition rules of the probabilities of change between periods

i) The matrix representing the behavior of individuals is a two-dimensional array of cells. In our case, each cell, representing a potential consumer, can accept one of two states: "0", denoting a potential consumer who did *not* adopt the innovation, and "1", denoting a consumer who *has* adopted the new product. In addition, irreversibility of transition is assumed, so that consumers cannot "un-adopt" after they have adopted.

ii) The individual maintains relationships with all individuals in his / her immediate neighborhood. We will discuss other possible modes of relationships presently.

iii) The rules that define transitions of potential adopters from state "0" to state "1" are classified into two types: *External Factors*: probability p exists, such that in a certain time period, an individual will be influenced by external influence factors such as advertising to adopt the innovative product. *Internal Factors*: a probability q exists, such that during a single time period, an individual will be affected by an interaction with *a single* other individual who has already adopted the product.

The time-dependent (non-cumulative) individual probability of adoption, *prob(t)*, given that the individual has not yet adopted, is based on the following binomial formula:

$$prob(t) = 1 - (1 - p)(1 - q)^{k(t)} \qquad (4)$$

Where $k(t)$ is the number of previous adopters with whom the individual maintains interactions in the vicinity of the cell under consideration.

Cellular Automata's popularity stems from the fact that despite its parsimony, it generates a wide range of dynamics and growth patterns. However, techniques that model the connection among agents that are not necessarily neighbors, can also be used to examine the evolution of markets.

Among them a promising technique that draws considerable attention is *Small World Networks* (see Watts and Strogatz 1998, and for recent marketing implications, Balakrishnan, Shaikh, and Rangaswami 2002)[1].

---

[1] One of the more colorful examples of a small world setting is the **Erdös number** project. Paul Erdös (1913–1996), a widely traveled and prolific Hungarian mathematician, wrote hundreds of research

An advantage of the small world system is that it enables us to describe a social system with a flexible connection structure among networks. According to the Small World Network approach, nodes are uniformly distributed on a circle, each connected to its neighbors up to some pre-specified range. In addition, a limited number of non-neighboring nodes are allowed for interaction through shortcuts, appearing as strings inside the circle. This allowance enables conditioning growth patterns and firms' strategies on a single parameter of the degree of randomness in the network.

Thus small world brings more complexity into the network generated by cellular automata by modeling—in addition to communication among neighboring cells—some random communications between cells that are not in proximity to each other. Such an environment corresponds to one in which individuals communicate with friends who do not necessarily live in their geographical proximity.

3c. Measures of success

Common measures of new product success among market researchers include factors such as achievement of financial objectives, achievement of market share objectives, and achievement of technical objectives. Estimation of these factors is often subjective and based on managers' input and extensive data collection.

The situation becomes more complicated if one wants to base one's judgment on sales data only, since a universally accepted sales-based measure of success does not exist. However, a possible important and indicative sales-based measure is the eventuality of

---

papers in collaboration with others. His Erdös number is 0. Erdös's co-authors have Erdös number 1. Co-authors of individuals with Erdös number 1 have Erdös number 2, and so on. For example, one of the authors of this paper has an Erdös number 4, and thus the Erdös number of the other three co-authors is at most 5. The latter might be smaller if there exists a chain of links fewer than 5 connecting them to Erdös. An interesting point to observe—found in the Erdos Number Project homepage—is that almost everyone with a finite Erdös number has a number less than 8.

reaching the *takeoff* point of the sales curve (Golder and Tellis 1997). This point in time, in which a rapid change takes place in the rate of the new product's adoption, is associated with the end of the Introduction Stage and the start of the Growth Stage in the product's life cycle. This visual sales pattern—sometimes labeled by practitioners as the "hockey stick"—is easy to perceive due to the often long left tail of new product growth and the steep growth that follows. Practitioners who often use the takeoff as an indication of a product's future might terminate a new product if it does not reach what they perceive as a "takeoff" (Golder and Tellis 1997). An exact definition of what constitutes a "takeoff" is not straightforward. One approach is to base the definition on outside judges who determine when the takeoff took place. This method can be carried out only *a posteriori*, when the judges are able to observe multiple points of sales data.

Another approach is to estimate the percentage of adopters associated with a takeoff. Rogers (1995) suggests that takeoff typically occurs when about 16% of the market potential adopts. Indeed, 16%, representing the cumulative number of innovators and early adopters according to Rogers' adopter categories, is sometimes an accepted number with regard to the anticipated end of the Introduction Stage (Moore and Pessemier 1993). Thus, if an innovation is accepted by 16% of the expected market potential, it can be an indication of acceptance. It should be noted that this might be a conservative estimate: Golder and Tellis (1997) report lower percentages in their calculation of the takeoff. In the following studies we will use visual identification that does not require knowledge of market potential, as well as the 16% criterion respectively, to identify the success of a new product.

4. Spatial Divergence as a predictor of success

4a. Study 1a: Predicting success using a visual takeoff measurement

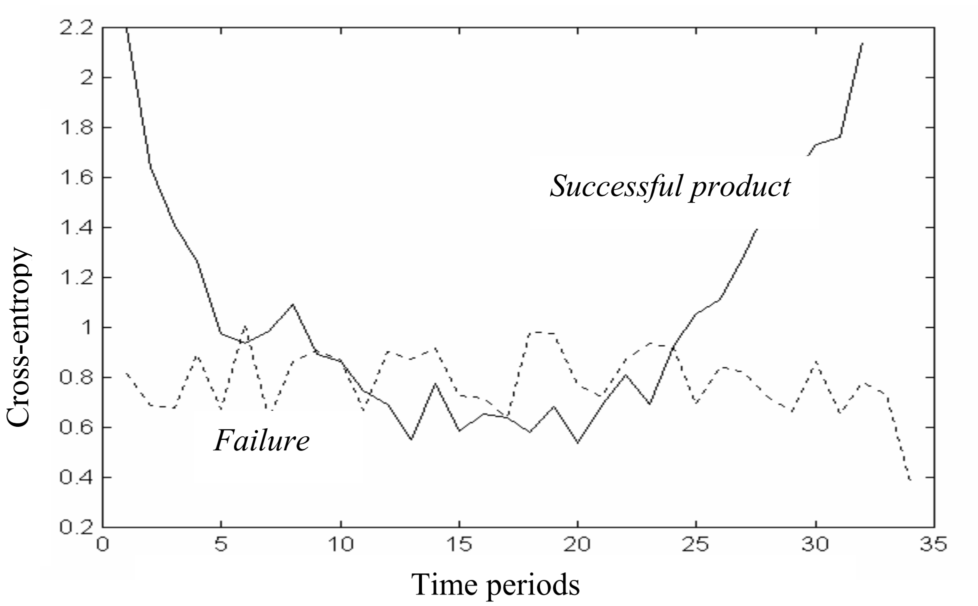
The aim of Study 1a is to examine the capability of spatial divergence of predicting a new product's success when the success measure is a visually observed takeoff in sales.

We ran the cellular automata process described earlier using a Matlab application customized specifically for our purpose. We ran 44 cellular automata simulations of 2,500 potential adopters each, with varying p and q values so as to cover various possible ranges of adoption curves (from complete failures to a strong noticeable takeoff). The parameters p and q were chosen to comply with findings on values of aggregate diffusion, transformed to an individual-level grid (see Sultan, Farley, and Lehmann 1990 for aggregate diffusion modeling results and standard diffusion parameters; Goldenberg, Libai, and Muller 2001 and 2002 for a discussion of the transformation of parameters to individual-level cellular automata) where p ranged from 0.0001 to 0.04, and q from 0.0001 to 0.03. The lower values of q were necessary to generate failure processes for the binary comparison.

The sales curves generated for the various parameter levels were shown to three expert judges. All three are familiar with both the theory and the practice of new product growth; two of them have extensive experience in firms at various levels of management of new product development and marketing. For each graph, the judges were asked to assign a score ranging from 1 to 4, where 4 indicates a definite takeoff, and 1 indicates a definite failure. For each process, the divergence measure was calculated. Logistic regressions were performed in order to assess the predictive ability of the divergence measure.

How would the two processes differ in terms of the divergence measure? Figure 5 illustrates the time-dependent cross-entropy calculation for the two products represented in Figures 1 and 2 of the previous section. As can be seen, the difference in cross-entropy is noticeable very early in the growth process. As the process continues, however, the difference between the cross-entropies decreases, and in this example, becomes similar as sales approach the peak level for the two cases of success and failure.

Figure 5: Cross-entropy measures of the two products



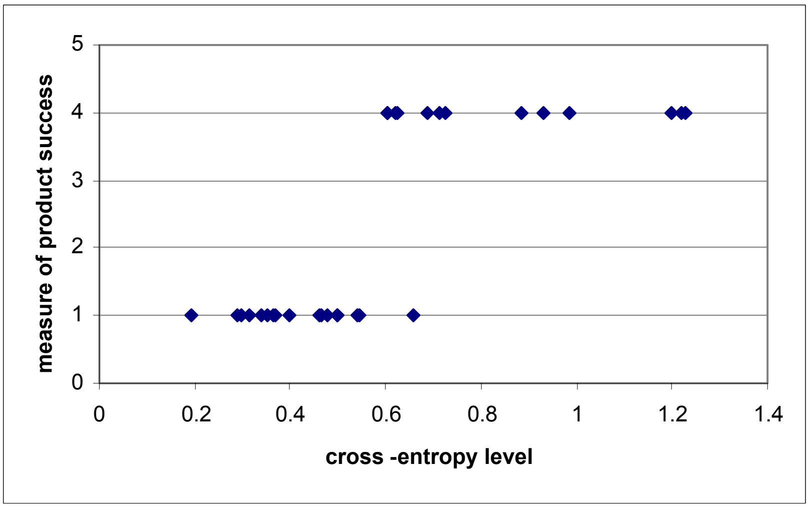
The phenomenon depicted in Figure 5 may look counter-intuitive: Typically one can expect that as more sales data are accumulated, better predictions regarding a product's success will be generated. Here, however, better separation is observed between successful and non-successful processes at earlier stages. This characteristic of cross-entropy has important implications for the ability of the approach to predict success at very early stages of a product's life.

The reason for this behavior lies in the intrinsic development of clusters through word of mouth. For a successful product, initially (as can be observed in Figure 1) there are small kernels of adopters. Then the kernels grow to form clusters, and later on clusters start to merge. At a certain point in time, it becomes harder to distinguish between cluster-based and non-cluster-based growth processes, as the distribution of buyers reverts to a spatially near-uniform distribution. In fact, toward the end of the process, there are clusters of non-adopters (as opposed to a uniform distribution of non-adopters in the failure case), and the difference between the cross-entropies of the two processes is once again distinguishable.

In order to quantify the difference in divergence measure between the successful and non-successful products, three judges determined the existence of a takeoff on a four-point scale for the 44 cellular automata processes. The judges' results yielded seventeen 1s, three 2s, seven 3s, and twelve 4s. Inter-judge reliability tests showed high agreement between judges (alpha = 0.94). In order to perform a regression, a dependent variable of the judges' decision was composed by taking a majority decision. In five cases, there was no agreement between the judges; these cases were excluded from the analysis, leaving us with 39 valid cases. For each growth process, we measured cross-entropy (or Jensen-Shannon) very early in the growth process (after 0.4% of the population had adopted).

We performed an ordinal logit analysis on all four groups using LIMDEP. McFadden R^2 was found to be 0.59. The correct predictions of the ordinal probit were 75%. As one of the intermediate categories had only three observations, we followed the practice of deleting sparse cell classes and were left with three categories. Indeed, the correct predictions increased to 86%. Judgments of 2 and 3 may indicate cases for which success and failure are not clear. In addition, the number of 2 and 3 grades assigned was small (10 cases) compared to the number of 1 and 4 grades (29 cases). Hence, we performed a logistic regression, in which the 2- and 3-graded cases were excluded (leaving us with 29 cases) to allow a regression between the cross-entropy and clear successes and failures. In this case the logistic regression correctly predicted 96.7% of the cases (see Figure 6). The logistic regression performed with the Jensen-Shannon measure as an independent variable yielded the exact same percentages (96.7%).

Figure 6 – Cross-entropy of clear cases of successes and failures



4b. Study 1b: Predicting success using time of 16% adoption measure

Study 1b differs from Study 1a in its definition of success for the new product. In Study 1b, instead of relying on the visual identification of a takeoff, we follow Rogers' (conservative) suggestion as to when in a product's life a takeoff should take place. Thus, a growth process is deemed a success if 16% of the market is obtained before a specified time. A new set of 44 growth processes was generated in the same parameter range as before. A deterministic measure of $t_{16\%}$ (the time in which 16% of the potential adopters have adopted) was computed for each process. We consider failure all cases in which one of the following two events takes place:

i) 16% of the population has not adopted within ten periods (T = 10)
ii) The peak in sales occurs prior to the time at which 16% of the potential adopters have adopted the product

Note that the determination of the time T may differ for different growth processes. However, for the durables studied in the meta-analysis reported by Sultan, Farley, and Lehmann (1990), takeoff was typically reported prior to ten periods. Thus, we define success / failure measure *S* as:

$$S = \begin{cases} 1 & if\ t_{16\%} > T\ and\ t_{peak} > t_{16\%} \\ 0 & otherwise \end{cases}$$

For each process, the cross-entropy (or Jensen-Shannon) was calculated, and a logistic regression performed on the success measure *S* and the cross-entropy (or J-S).

The logistic regression based on the cross-entropy measure correctly classified about 91% of the cases, as can be seen in Table 1. The regression based on the Jensen-Shannon measure correctly classified about 86% of the cases.

Table 1: Logistic Regression Confusion Table:
Cellular Automata data, uniform baseline distribution, cross-entropy measure

	Predicted success	Predicted failure
Observed success	90%	10%
Observed failure	8.3%	91.7%

Average correct predictions - 90.8%

4c. Study 1c: Comparison of results to benchmark models

In order to validate our model, we compare the cross-entropy to a naïve criterion of cumulative sales in a specific period. Although this is not a precise criterion, it is nonetheless simple to test and can serve as a benchmark for comparison of new methods for predicting success. The 44 processes were randomly divided into two equal groups, with the same ratio of successes and failures in each group. The first group was used to establish a criterion for prediction of success by cumulative sales, and the second was a holdout sample used to test the reliability of this criterion. Each adoption process in the first group was classified as success or failure using the same classification used for the cross-entropy test. The average sales in periods 2, 3, and 4 were calculated for both groups, with one average calculated for the successful processes and one for the failures. A cumulative sales measure of each of the 22 processes of the holdout group was then calculated in the same time periods (2, 3, and 4).

In order to predict success or failure for a process in early periods, its cumulative sales measure was compared to the two values—or success and failure averages—of the first sample. If the cumulative sales measure was closer to the success average, it was denoted a success, and if it was closer to the failure average, it was denoted a failure. A confusion table of the prediction of success and failure revealed that even in period 4, the correct predictions are less than 50%. One can improve the results of this method by making the prediction in later periods. This option, however, is not in line with the purpose of our work, which attempts to predict success at an early stage of the product's life cycle.

In order to further validate our model, we also compared the classification provided by the spatial divergence measures to predictions made using diffusion theory that is exclusively based on time series. We used the Bass model on the 44 processes reported above. A Bass model requires three parameters: external coefficient - p, internal coefficient - q, and the market potential N. Following Srinivasan and Mason (1986) for estimation using only early-period data, the known market potential N is used (2,500 in our case). Thus we estimated the two parameters of p and q utilizing data in four early time periods, and using the $t_{16\%}$ measure of success discussed earlier, determined if the resultant growth process would be labeled a success or a failure (note that this is a conservative approach, as only two parameters were estimated). We then compared the results with the true success or failure of the process. The proportion of successful predictions was around 58%. Note that this number is not only much lower than our 91% successful prediction based on the cross-entropy measure, but also not significantly different from the 50% that one would achieve in a prediction based on random drawing.

4d. Study 1d: Small World

A small world model was programmed with the same market potential as in the previous two studies, in order to allow for similar conditions for the adoption processes. In keeping with this requirement, the generated adoption processes were in the same range of parameters as the previous process. In order to use the small world model efficiently, one must decide on the proportion of the distant links from the entire set of links. We have chosen the classic level of 5% as an upper limit, as beyond this level, the social system becomes similar to a random network and less comparable to real-life social systems (Amaral et al. 2000). Because the distant links are random, the expected variance of the dynamics may be larger than those of the cellular automata. Therefore, the number of generated processes was increased to 100.

Two logistic regressions were performed, one with cross-entropy and another with the Jensen-Shannon measure. The spatial divergence performances as success predictors were compared to the two benchmark models that were used in study 1b (cumulative sales and the Bass model). In addition, we used a third naïve model, where we checked the derivative at the beginning of the process, and compared it, using the same type of logistic regression, to the success or failure measure. This test was based on the fact that if one looks at Figure 3, the high-growth product (1) is characterized by an upward trend in adoption from period 0, whereas sales are flat during the first four periods for the low-growth product (2).

The logistic regression based on the cross-entropy measure correctly classified 90% of the cases, as can be seen in Table 2. The regression based on the Jensen-Shannon measure correctly classified 90% of the cases as well. Once again, the correlation between the two spatial divergence measures was found to be high (99%).

Table 2: Logistic Regression Confusion Table:
Small World data, uniform baseline distribution, cross-entropy measure

	Predicted success	Predicted failure
Observed success	93.1%	6.9%
Observed failure	14.3%	85.7%

Average correct predictions - 90.0%

The rate of successful predictions using the new naïve model was low: Only 57% of the processes were successfully predicted. When adding the naïve model to the spatial divergence as a second predictor in the logistic regression, there was no change in the percentage of correct predictions, and its coefficient was not significant.

Overall, we established the capacity of the spatial divergence method for predicting a success when just 0.4% of the market has been successfully demonstrated in Studies 1 through 3 on simulated processes that cover a wide range of "clean" scenarios. In the next studies, we extend the analysis to cases in which a uniform distribution of consumers does not hold.

5. Non-uniform baseline distributions

While uniform distribution might be a reasonable assumption in some cases, there are many reasons why the baseline distribution might be other than uniform. This fact might have an effect on predictions, particularly false positives.

The most common cause for this phenomenon might be a distribution of individuals in the various geographical areas that might not be uniform. Another cause for (illusory) formations of clusters that do not necessarily imply potential success of a new product is the case of non-uniform marketing. Such efforts might include local advertising, promotions in specific points of sale, and other localized, idiosyncratic efforts. In such a case, calculation of the cross-entropy must be modified by changing p_2 in Equation 2 from a uniform distribution to a normalized distribution of the non-uniform investments of the marketing efforts, measured in monetary values.

A related effect that can create a cluster-like phenomenon is a deliberate policy on the part of marketers of introducing innovations—at least in an initial phase—in restricted geographical areas. In such cases, areas in which the innovation is actually introduced may appear as clusters when compared to other areas. Varying structures of distribution channels and marked differences in marketing mixes among regions—such as deep price cuts in only a few areas—can create a similar phenomenon.

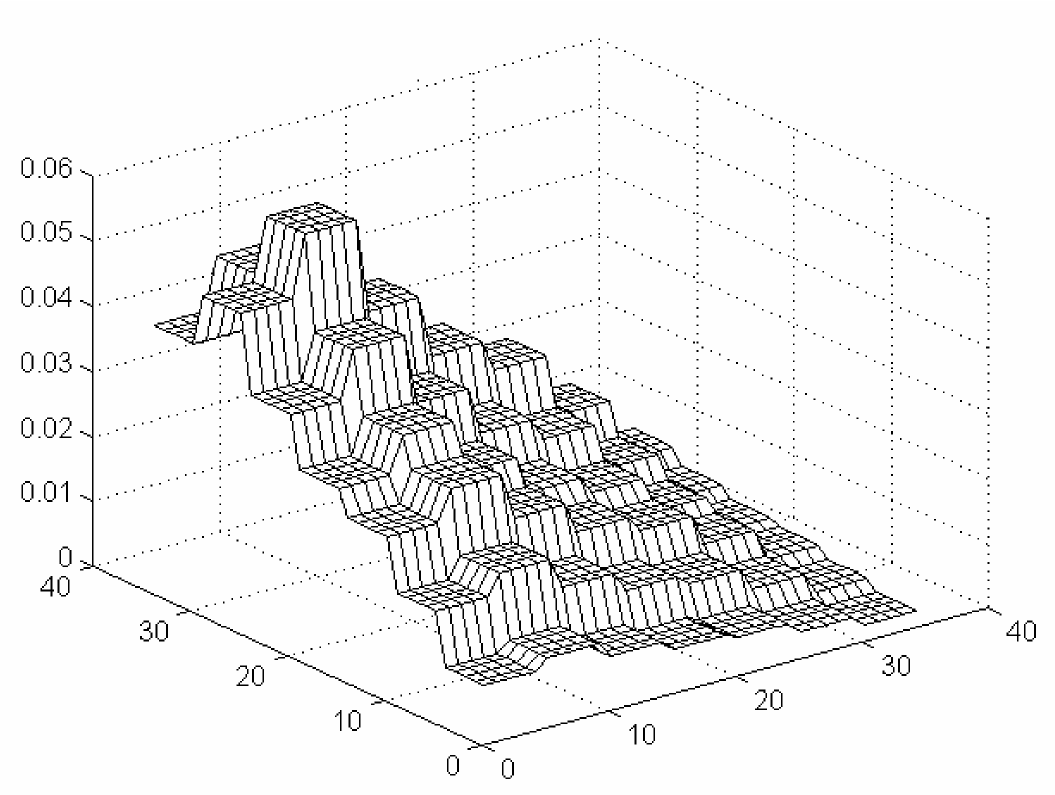
In order to use the spatial divergence approach as an indicator of success in markets where organic clustering occurs, marketers can take one of two approaches: One possibility is to conduct the analysis in relatively homogeneous areas for which "natural" clustering is less of a problem. The second approach is not to compare the distribution of actual penetration to the uniform distribution, but rather to another baseline distribution, which represents marketing executives' information on possible clustering based on drivers that are external to the communication process such as demographics or reasons related to the nature of the marketing plan in specific areas.

In order to examine the usefulness of the spatial divergence approach even when the distribution of consumers is not uniform, we performed two additional studies as follows: in Study 2a we compare the resultant sales distribution of the new product to a non-uniform (Beta) distribution that represents the non-uniform external marketing effect, and in Study 2b, we present an empirical test for which the underlying distribution is known to be non-uniform due to variances in the number of users in each region.

5a. Study 2a: Non-uniform distribution; simulated test

First, we created a Beta-type baseline distribution of population. As seen in Figure 7, the density of potential adoption varies considerably between geographical areas. The spatial divergence approach could now be used while comparing the resultant distribution of adopters to the given Beta baseline distribution, instead of to the uniform one.

Figure 7: Beta-type baseline distribution of consumers in an area of 49 different regions
The vertical axis represents population density.



In order to do so, we repeated the calculations for both the cellular automata processes and the small world processes, but this time the population is spatially distributed according to the above Beta function. We used both cross-entropy and the Jensen-Shannon measure in order to determine the distance of the two distributions.

The results of the logistic regressions show that both measures correctly classified the majority of the cases. The results for the cross-entropy measure on small world data are presented in Table 3. The results for the other studies of the Jensen-Shannon measure and cellular automata data are similar, the highest correct predictions average being 81.8%.

Table 3: Logistic Regression Confusion Table:
Small world data, non-uniform baseline distribution, cross-entropy measure

	Predicted success	Predicted failure
Observed success	86.2%	13.8%
Observed failure	28.6%	71.4%

Average correct predictions - 80%

5b. Study 2b: Non-uniform distribution: Hybrid corn revisited

In the previous study, we showed that the spatial approach could be used by using simulated data when the underlying baseline distribution is non-uniform. The purpose of this study is to apply the spatial divergence approach to an actual field application. The well-known case of hybrid corn was selected, as the actual distribution of the examined variable, namely the corn acreage planted, is known. Detailed descriptions and insights from this case can be found in Griliches' seminal article of (1957).

The spatial distribution is not uniform in the hybrid corn case, so the classic Parzen windows approach had to be modified. Consistent with study 2a, different mapping and regional treatments were utilized. The windows density was constrained to fit the shape and area of the states in the US (data on the penetration of hybrid corn total cornfield acreage were obtained from USDA agricultural statistics).

Due to the lack of reliable data prior to 1933, the spatial divergence was calculated for that year. This is a stringent test for the validity of cross-entropy, as studies1 and 2 demonstrated that the best separation and predictions of success are produced at earlier stages (e.g., 1930 in our case). As opposed to the uniform case, in the hybrid corn case, each state was considered a window with a unique size. In the first stage, the acreage for each state was calculated, and the relative acreage percentage in each state compared to the total US acreage was calculated. A probability function representing the adoption of hybrid corn in the US was calculated from this data. Thus, when calculating the sum of p(x)*log(p(x)/q(x)) for Illinois, q(x) was equated to 0.0754, or the corn acreage in Illinois divided by the total corn acreage in the US. The cross-entropy value of the hybrid corn spatial adoption was calculated accordingly.

In order to use this measure to predict success, a comparison had to be performed to the cross-entropy and the Jensen-Shannon values obtained in studies 1, 2, and 3. Consequently, the simulations should take into account the fact that the area and the windows distribution are different. In other words, in order to test the probability of hybrid corn being adopted successfully, one has to simulate the US geographically on the computer. Consequently, the simulated space world was divided into windows similar to the respective states' sizes. Both cellular automata and small world simulations were performed again, and both spatial divergence measures for each process were calculated taking into account the new windows sizes. Because the cellular automata and small world data were now calibrated to represent the world of the hybrid corn, the cross-entropies of both the cellular automata and the real case can be matched into the same graph. Hence we can "plug" the divergence measure value into the logistic curve and use it as the independent variable to read the value of the dependent variable, or the probability of hybrid corn becoming a success. In the same way as in Study 2a, *mutatis mutandis*, the following table is obtained:

Table 4: Divergence measures calculations for hybrid corn

Comparison	Cross-entropy	Jensen-Shannon	Probability of success*
Cellular Automata	1.63	.35	93%
Small World	1.63	.35	99.9%

* Computed from the logistic regression whose independent variable is the cross-entropy and whose dependent variable is the outcome (success/failure)

6. Field tests

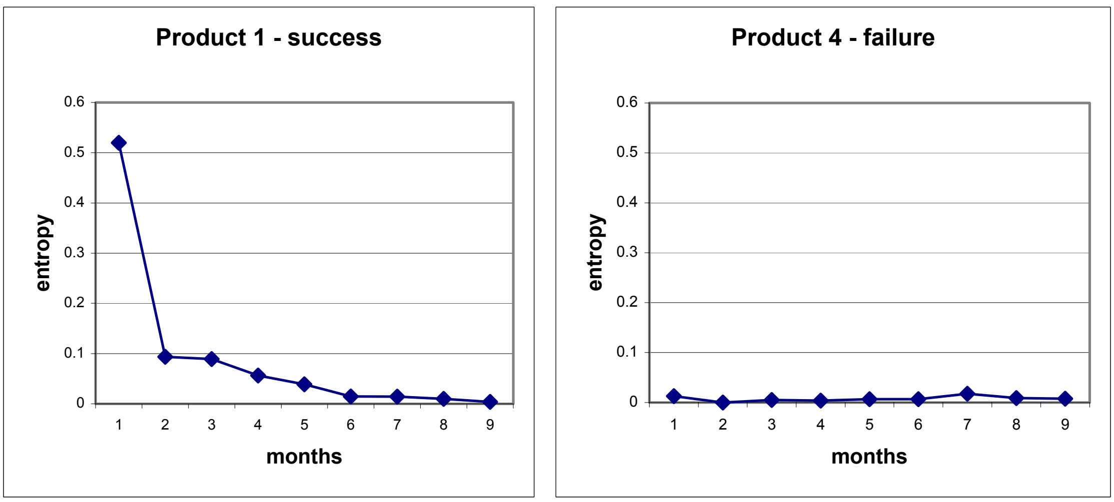
6a. Study 3a: Supermarket products

Obtaining the data needed to test the divergence method in a real-life application is not simple. First, one needs to obtain growth data of a product in a number of geographic locations, as opposed to only aggregate data. Second, for discriminating validity purposes, one needs to obtain spatial sales data on failures, which are generally hard to obtain. We were able to obtain data on the sales of eight new health and personal hygiene products sold in a dominant supermarket chain in a Mediterranean country. The data include monthly sales during the first years. While the general product categories are relatively mature, as is the case with most supermarkets products, the products themselves were considered innovative relative to the category. Each product was launched simultaneously in each of the chain's locations. The successful products had a pattern of a rapid, monotonic increase in sales. In contrast, the failed products did not sell well based on the chain's standards, and very moderate growth was observed. At the time, these products were under consideration for removal from the shelves and cessation of their distribution.

Because we were limited to data provided by the supermarket chain, data could not be broken down to optimal window (Parzen windows) size. Rather, it was coded based on the retailer's data into 12 regions. Consequently, window size had to be constrained by the weight and area of the retailer distribution regions. Each window was weighted according to the corresponding area's relative population size. For cross-entropy calculation purposes, each region was designated as a window. Before comparing the probabilities of success in a manner similar to the previous study, we can examine the cross-entropy differences between the successful and failed products to see if a difference is sufficiently great to discriminate between the cases. What we find is a pattern—similar to that in Figure 3—in which the successful products have a *declining* cross-entropy measure, while the failures have a *consistently low* cross-entropy measure. Figure 8 presents the cross-entropy measure for each month for one successful product and one failed product. Both graphs use the same scale. The difference between failure and success is indeed pronounced in the first periods (about 50 times higher in the first period for the successful product). As can be seen, the early-period cross-entropy pattern of these cases is similar to the cross-entropy pattern for the cellular automata case in Figure 5, obtaining its maximum level at the beginning of the process.

This figure demonstrates how a clear difference in divergence helps to differentiate between successes and failures by qualitative visual assessment. Yet a more rigorous approach should call for a quantitative tool that would help to convert cross-entropy results into probabilities of success in an unambiguous manner. Thus we return to cellular automata simulation, this time as a tool to translate the empirical results into success probabilities. We ran cellular automata and connected success or failure status (as the dependent variable) to cross-entropy divergence measure (as an independent variable) through logistic regression. As will be explained next, using the resultant logistic regression function, we can determine the probability of success for each cross-entropy value, including real-life cases.

Figure 8. Cross-entropy, calculated monthly for the two products



We must, however, take into account that unlike the simulations in which the area unit was based on the Parzen windows, real-life data may present itself as geographical areas with varying sizes, market potentials, and concentrations of adopters. Hence, the cellular automata analysis should use spatial units calibrated to take into account actual size and potential. Consequently, since the actual market was divided into twelve geographical areas, the simulated space was divided into twelve windows of relative sizes to match the actual areas of distribution. We calculated cross-entropy for each simulated process again, taking into account the new windows sizes. We performed a logistic regression that uses the new results with a dependent variable of the $t_{16\%}$ to produce a function that describes the success / failure probability as a function of the cross-entropy in the field case at hand. Thus, when a marketing manager wishes to estimate the probability for success for a specific case, s/he should perform the following procedure:

1. Calculate the cross entropy of the spatial distribution of the sales, based on the windows made possible by the data available.
2. Next one needs to form the function that will translate the cross-entropy values into probabilities of success. In order to do this one has to follow these three steps:

   2a. Use the set of the cellular automata simulations with the following modification. The windows used to calculate cross-entropy are now changed from optimal (Parzen) size to windows that match the window numbers and sizes in the real-life case of stage 1. The result is the same set of processes but with new cross-entropies values.

   2b. Create a data set for the logistic regression: Each point is actually one cellular automata process, the independent variable is the cross entropy measures early in the process and the dependent variable indicates whether the product is a success or failure.

   2c. Run the logistic regression to form the relationship between cross entropy values and probability of success.
3. Now, going back to the cross entropy of the real-life case, and using the relationship determined in stage 2c, find the probability of success for this specific case.

The calculation of the cross-entropy of the cellular automata simulation for the 12 regions of the supermarket chain again revealed a phase transition-like phenomenon in which there was clear distinction between success and failure. Again, high values of cross-entropy indicate a high probability of success, and low values indicate a low probability of success. The transition between these two regimes is abrupt, with the threshold value that discriminates between success and failure found to be around 0.3. The fact that this measure is sufficiently robust, even when the window sizes are far from optimal, suggests that cluster formation may be an eminent factor of success. The correct predictions of the logistic regression (based on the cellular automata cases) are 80%. This value is naturally not as high compared to the values achieved when using optimal window size in the simulation study (see The Effect of Number of Windows Reduction in the Appendix). Table 5 presents the predictions of the logistic regression.

Table 5: Supermarket product category:
Cross-entropy calculations for successful and failed products

Product	Outcome	Cross-entropy	Jensen-Shannon	Probability of success*
1	Success	.52	.30	99.9%
2	Success	.51	.30	99.9%
3	Success	.31	.19	99.8%
4	Success	.30	.17	99.8%
5	Success**	.03	.01	0.6%
6	Failure	.01	.006	0.4%
7	Failure	.02	.001	0.4%
8	Failure	.08	.05	6.5%

* Computed from the logistic regression whose independent variable is the cross-entropy of the cellular automata, and whose dependent variable is success / failure.

** Product 5 was predicted to be a failure, while its actual status is a success.

The first column of Table 5 represents the product, while the second column indicates the field observations based on the growth of the product's sales during the nine-month period. The third column shows cross-entropy value for the first month. It can be seen that the cross-entropy of a successful product is in order of magnitude larger than that of a failure.

The last column shows the prediction of success probability, calculated through the logistic regression. Thus, for example, for product 1, a success, the cross-entropy measure of 0.52 yields a 99.9% probability of success in the logistic regression function.

Note that the fifth case (denoted with a double star) is a successful product that was predicted to be a failure. In a discussion with the executives of the distribution chain, the possibility was raised that the success of the product was so phenomenal, that after one month the diffusion covered too large an area, causing the cross-entropy to drop sooner than the other products. If correct, this means that regarding a rapid diffusion of a highly successful product, the cross-entropy value must be measured earlier than the first month.

6b. Study 3b: Home furniture sets

The purpose of this study is twofold; the first is to replicate the results of the previous study with products from another, different, category, the second is to extend the analysis of spatial divergence and test it in an environment with strong non-uniform effects.

In order to use a different product category and a different distribution channel, the home furniture category was selected. The advantages of this category are the availability of regional data, frequent introduction of new sets (models), and, as was confirmed by the experts in this field, an apparent word-of-mouth effect. A large home furniture chain was asked to select four successful new furniture sets and four failures. A failure was defined by the managers as a set whose sales have not reached a predefined level and whose first stock were not sold out. A successful product was defined as a set all of whose entire first run units have been sold, and more have been ordered from the manufacturers. The chain supplied data of monthly regional sales by individual store.

The regional data were rich, including dozens of locales, in contrast to the 12 windows of the previous study. This richness enabled a more accurate measure of spatial divergence, which may be sufficiently robust to the non-uniform effects. Indeed, each new furniture set was distributed in a different way. The main difference is that not all points of sale received the new set. In addition, each point of sale differed in its volume and in the local promotion invested. These two parameters represent an actual situation, in contrast to the simulated data studies in which the windows can be fixed at the same size. The comparison between failures and successes was obtained by computing the success probability for each case in the same way that this was performed in the previous study *mutatis, mutandis*.

Because the spatial grid is different for each case, the cellular automata simulation, based on the specific grid, was repeated eight times (with the relevant grid). This procedure is identical to the procedure that was applied in the hybrid corn case, and included weighing each point of sales by its size, to factor out the non-uniform effect of size. It was also assumed that local marketing efforts are correlated to store size. While this is not necessarily an accurate assumption, it is a conservative one. If, for example, small stores invest more in promotion, the spatial measure performances are expected to decline. As in the hybrid corn study, the computed cross-entropy marked the success probability based on the logistic regression curve that was applied on the 100 Cellular Automata simulations.

As in the previous studies, the cross-entropy and J-S measures produced similar values that were found to be largest in the first period. Both spatial divergence measures declined with time. Table 6 presents the computed prediction of success probability. It can be seen that the divergence measure of a successful product is in order of magnitude larger than that of a failure. Note that unlike in Table 5, a monotonic relationship between cross-entropy and probability of success is not found in Table 6. The reason is that the analyzed areas for the furniture sets were different for different sets, and thus the windows used for the analysis were not identical. Thus, for each product we used a different logistic regression as explained earlier. Had the windows been equal, a single logistic regression could be used and a monotonic relationship between cross-entropy and probability of success would have been expected.

Table 6: Home Furniture Sets product category:
Divergence measures calculations for successful and failed products

Product	Outcome	Cross-entropy	Jensen-Shannon	Probability of success*
1	Success	.21	.11	91%
2	Success	.27	.15	96%
3	Success	.21	.12	70%
4	Success	.26	.13	98%
5	Failure	.12	.07	0.1%
6	Failure**	.37	.21	92%
7	Failure	.07	.04	23%
8	Failure	.19	.11	1%

*Computed from the logistic regression whose independent variable is the cross-entropy and whose dependent variable is the outcome (success / failure)

** Product 6 was predicted to be a success, while its actual status is that of a failure.

7. Discussion and limitations

In this paper we demonstrated the ability of the spatial divergence approach to serve as a robust, early-period prediction tool, in both simulated and actual field studies. This demonstration adds to the potential contribution of spatial analysis to new product theory and practice. Often, marketers aggregate data on new product growth from diverse regions and areas of operations. However, these rich data are not used when the focus of attention is primarily on the temporal aspect of product growth. This research suggests one way in which spatial growth data can be used to better understand and predict aggregate temporal growth, further emphasizing marketers' need to collect and analyze spatial data that has recently been shown to aid in assessing important issues such as brand performance (Bronnenberg and Sismeiro 2002).

While the spatial divergence method might be a promising tool for new-product analysis, care should be taken in its application. A number of limitations should especially be noted:

Type of products susceptible to word-of-mouth communications

Following diffusion theory, our analysis assumes that word of mouth and imitation play a dominant role in the adoption process of consumers. Word-of-mouth communications intensify as resistance to adoption is higher, purchase is less frequent, and product use has a higher uncertainty, a higher perceived risk, and a higher consumer involvement level (Rogers 1995). As an example, consider the case of the Personal Digital Assistant (PDA) for which well known brand names include Apple's Newton and the Palm Pilot. Such an innovation is discontinuous, as it changes the consumption patterns of those who use it. Clearly, at least in the early years of the innovation, the move to electronic organizers of these types is not trivial, is characterized by uncertainty, and involves the outlay of considerable resources. Thus we can expect that word of mouth and imitation will play an important role in the adoption of such a product, especially at the early stages of its product life cycle. On the other hand, a minor brand extension in the personal toiletry category such as a new fragrance of a deodorant will command considerably less word-of-mouth activity for the opposite reasons.

Negative word of mouth

Negative word of mouth is a latent component of any diffusion process that does not appear at all in sales data, yet its effect on sales might be notable. It can be argued that in cases of strong negative word-of-mouth activity, the product might be a failure even though the researcher might observe an artificial formation of adoption clusters, which are actually the result of large areas of *resistance* to the product.

However, we expect that pockets of adopters, formed by negative word of mouth, might emerge only after a certain time period. In order for large and homogeneous areas of resistance to be formed, negative word of mouth has to propagate and cover this area. This propagation takes time, which will probably be longer than the introductory stage. This case is somehow less relevant to the early predictive ability of our approach. Also, in order for large clusters to evolve, one must assume that consumers affected by negative word of mouth will spread negative word of mouth themselves without buying the product, a scenario less likely to occur. Hence we believe that the existence of some negative word of mouth does not represent a threat to the validity of our approach, though a further exploration of its exact interaction with the positive one warrants future research.

Geographical proximity and electronic commerce

Following past research, we assumed that word-of-mouth communications spread faster for nearby geographic neighbors and thus help develop the formation of geographic clusters. While this assumption should be robust in most cases, electronic methods of data transfer may make word of mouth less relevant for some products. Specifically, Internet-based word of mouth (sometimes called "word of mouse") does not demand geographical proximity, and thus in places where Internet communications drive the diffusion process, looking at geographical clusters may not be useful.

We suggest, however, that even for many products sold in electronic commerce environments, geography-based word of mouth and imitation can still be a major driver of the diffusion process. For example, after seeing it or hearing about it from a friend, a consumer may purchase a book or a CD over the Internet because he or she finds a better price there than at the local retailer. Such a purchase does not affect the geographical proximity assumption. This is especially true when observing others using the product, or even trying it out personally plays an important role in the adoption process. Indeed, recent findings suggest that geographical clustering does occur in the growth of Web-based grocery shopping

(Song and Bell 2002).

The case may be different regarding Web-based products such as instant messaging software that can be evaluated and spread based on Internet contacts. In fact, even here, one can argue that many users will adopt the software after actually seeing it somewhere nearby, and that an individual's friends typically live or work in her or his vicinity. There is not enough empirical evidence to edify us on how exactly the Internet influences internal effects, yet future research will probably help managers decide for which products geographical proximity is less relevant.

Some anecdotal evidence suggests not only geographical spread but also an accelerated rate of diffusion for Internet-purchased products ("word of mouth on steroids"). Such is the case, for example, with Hotmail service: the two venture capitalists that supplied the seed money for the service noted in 1996 that: "We would notice the first user from a university town or from India, and then the number of subscribers from the region would rapidly proliferate" (Rosen 2000).

References

Allaway, Arthur W., Barry L. Mason and William C. Black (1991), "The Dynamics of Spatial and Temporal Diffusion in a Retail Setting," in Avijit Ghosh and Charles A. Ingene (eds.) *Spatial Analysis in Marketing*, New York: JAI Press, Inc.

Amaral, Luis A., Marisa A. Scala, Marc Barthelemy, and Eugene H. Stanley (2000), "Classes of Small- World Networks," *Proceedings of the National Academy of Sciences of the United States of America*, (97), 21, pp. 11149-52.

Balakrishnan, Anant, Narzul I. Shaikh and Arvind Rangaswami (2002), "Small-World Markets and Their Implications for the Diffusion of Innovations," Marketing Science Conference, University of Alberta, Edmonton.

Bronnenberg, Bart and Vijay Mahajan (2001), "Unobserved Retailer Behavior in Multimarket Data: Joint Spatial Dependence in Market Shares and Promotion Variables," *Marketing Science*, (20), pp. 284-299.

Bronnenberg, Bart and Catarina Sismeiro (2002), "Using Multi-Market Data to Predict Brand Performance in Markets for Which No or Poor Data Exist," *Journal of Marketing Research*, (39), pp. 1-17.

Case, Anne C. (1991), "Spatial Patterns in Household Demand," *Econometrica*, 59(4), 953-965.

Csiszar, Imre (1991), "Why Least Square, and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems," *The Annals of Statistics*, (19), pp. 2032-2066.

Dekimpe, Marnik G., Philip M. Parker and Miklos Sarvary (2000), "Multimarket and Global Diffusion," in Vijay Mahajan, Eitan Muller, and Yoram Wind (eds.) *New Product Diffusion Models*, Boston: Kluwer Academic Publications.

Goldenberg, Jacob, Barak Libai and Eitan Muller (2001), "Talk of the Network: A Complex System Look at the Underlying Process of Word of Mouth," *Marketing Letters*, (12), pp. 209-221.

Goldenberg, Jacob, Barak Libai and Eitan Muller (2002), "Riding the Saddle: How Cross-Market Communications Can Create a Major Slump in Sales," *Journal of Marketing*, (66), 2, 1-16.

Golder, Peter N. and Gerard T. Tellis (1997), "Will it Ever Fly? Modeling the Takeoff of Really New Consumer Durables," *Marketing Science*, (16), pp. 256-70.

Griliches, Zvi (1957), "Hybrid Corn: An exploration in the Economics of Technological Change," *Econometrica*, 25, pp. 501-522.

Johnson, Don H. and Sinan Sinanovic (2001), "Symmetrizing the Kullback-Leibler Distance," Rice University Working Paper.

Kohli, Rajeev, Donald R. Lehmann and Jae Pae (1999), "Extent and Impact of Time in New Product Diffusion," *Journal of Product Innovation Management*, (16), pp. 134-44.

Krider, Robert E. and Charles B. Weinberg (1997), "Spatial Competition and Bounded Rationality: Retailing at the Edge of Chaos," *Geographical Analysis*, (29), pp. 16-63.

Kullback, Solomon (1997), *Information Theory and Statistics*, New York: Dover.

Lee, Lillian (1999), "Measures of Distributional Similarity," *Proceedings of the 37th Association for Computational Linguistics*, pp. 25-32.

Mahajan, Vijay, Eitan Muller and Frank M. Bass (1990), "New Product Diffusion Models in

Marketing: A Review and Directions for Research," *Journal of Marketing*, (54), pp. 1-26.

Mahajan, Vijay, Eitan Muller and Yoram Wind (2000), *New Product Diffusion Models*, Boston: Kluwer Academic Publications.

Mahajan, Vijay, Eitan Muller and Rajendra K. Srivastava, (1990), "Determination of Adopter Categories Using Innovation Diffusion Models," *Journal of Marketing Research*, (27), pp. 37-50.

Mahajan, Vijay and Robert A. Peterson (1979), "Integrating Time and Space in Technological Substitution Models," *Technological Forecasting and Social Change*, (14), pp. 231-241.

Moore, William L. and Edgar A. Pessemier (1993), *Product Planning and Management*, New York: McGraw-Hill.

Parzen, Emanuel (1962), "On Estimation of a Probability Density Function and Model," *Annals of Mathematical Statistics*, (33), pp. 1065-1076.

Putsis, William, Sridhar Balasubramanian, Edward H. Kaplan and Subrata K. Sen (1997), "Mixing Behavior in Cross-Country Diffusion," *Marketing Science*, 16(4), 354-369.

Rogers, Everette M. (1995), *The Diffusion of Innovations*, 4th, New York: Free Press.

Rogers, Everette M. and Floyd F. Shoemaker (1971), *Communication of Innovation*, New York: Free Press.

Rosen, Emmanuel, (2000), *The Anatomy of Buzz: How to create Word-of-Mouth Marketing,* Doubleday Currency: New York, pp. 21.

Song, Sangyoung and David Bell (2002), "Space-Time Adoption: The Case of Internet Grocery Shopping," Marketing Science Conference, University of Alberta, Edmonton.

Srinivasan V. and Charlotte Mason (1986), "Nonlinear Least Square Estimation Procedure for Innovation Diffusion Models of New Product Acceptance," *Marketing Science*, 169-178.

Sultan Fareena, John U. Farley, and Donald R. Lehman (1990), "A Meta-Analysis of Application of Diffusion Models," *Journal of Marketing Research*, (25), pp. 331-41.

Watts, Duncan. J. and Steven. H. Strogatz (1998), "Collective Dynamics of 'Small-World' Networks," *Nature* (393) pp. 440-442.

Appendix

This Appendix demonstrates the comparison of forecast accuracy as we move from coarse partitions to finer ones.

This test was first performed on field data (the hybrid corn case, see section 5b). Instead of using the countries as windows, we used two other classifications (census regions and divisions of the US), with four and nine windows, respectively.

The four windows are the standard division of *Northeast*, *Midwest*, *South*, *and West*, and the nine windows are a finer division such as, for example, the subdivision of the South into *South Atlantic*, *East South Central*, *and West South Central*. We simulated the USA as a Cellular Automata representation, and calculated the predicted probability for success. In nine windows, the result again was 99% for success. This is probably related to the fact that hybrid corn is indeed a successful innovation. However, in four windows classifications, the predicted probability dropped to 45%, an indication of failure.

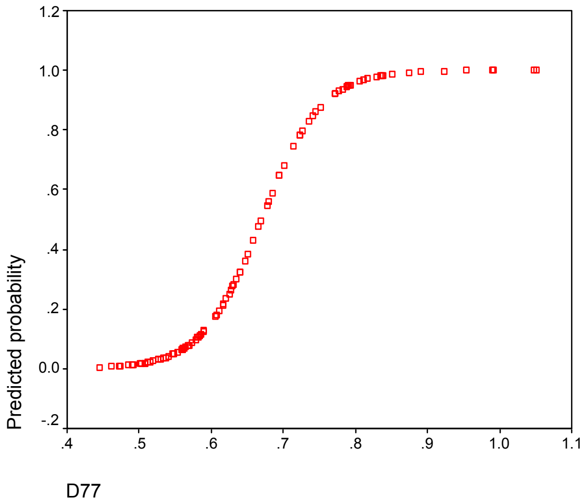
The test was also performed in synthetic data using cellular automata. As suggested by Parzen, the optimal windows number is the square root of the number of individuals (2,500); thus in our case it is 50. In order to stay within a symmetrical environment, we performed all analyses of the simulated data on squares, and thus the near-optimal number is 49 windows. We varied this number above and below 49. Three effects were noticed:

1) Indeed, the number of windows larger than Parzen's leads to a decrease in the correct predictions: 83% as opposed to 90%.

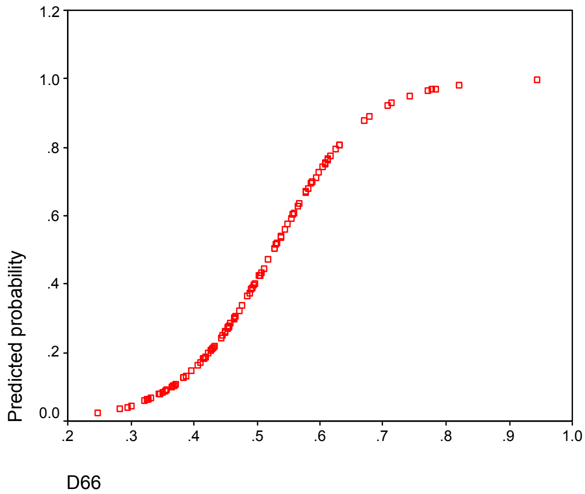
2) Decreasing the number of windows leads to a noticeable decrease in correct predictions: from 90% for 49 windows, to 80% for 36 windows, 80% for 25 windows, 72% for 16 windows, 65% for nine windows, and 57% for three windows (this last figure is not significantly different than random draw).

3) The clear step function of the logistic curve whose purpose is to separate success from failure slowly metamorphosed from an S-shaped function to a straight line. To illustrate this change, consider the graphs below:

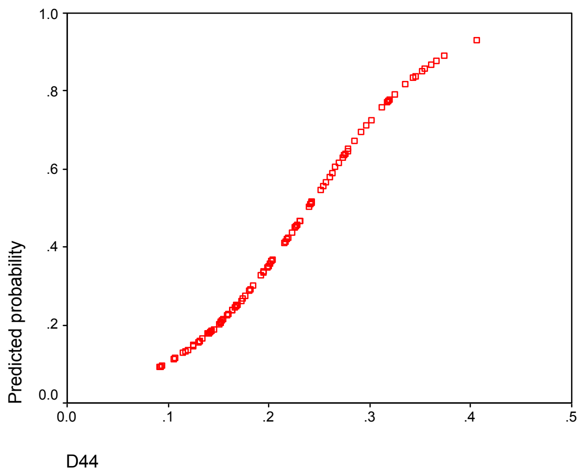
Parzen-recommended 49 windows:



36 windows



16 windows:



9 windows