Information Content of the Trajectory-Domain Models

Shu-Heng Chen*and Chueh-Yung Tsao

AI-ECON Research Center
Department of Economics
National Chengchi University
and
Department of Finance
National Central University

*Corresponding author. E-mail: chchen@nccu.edu.tw

Abstract

In this paper, we examine the information content in the trajectory-domain model proposed by Chen and He (2003). The data to be tested are three American stock indies, including Dow Jones, Nasdaq, and S&P 500. We adopt two event study methods, the standardized-residual method (SRM) and the standardized cross-sectional method (SCSM), to test the abnormality of the aftermath return series. In addition, the GARCH-M plus MA(1) is considered as the benchmark to be compared with. It is found that some patterns of the models do transmit informative signals. However, the signals are not persistent. They would emerge during a period and then vanish, vice versa.

Keywords: financial modeling; self-organizing maps; event study methods; technical analysis

JEL Classification: C45; C51; G14

1 Introduction

Financial data mining concerns two general questions. First, define the financial patterns with appropriate data mining tools. Second, show the patterns derived are profitable or informative. In the literature, the first issue was largely addressed in the context of time series models, be they linear or non-linear. However, a recent progress in financial data mining starts to look at the alternative — the feature-domain approach.

In the time-domain models, e.g., ARIMA models, bilinear models, (G)ARCH models, etc., extrapolation of past values into the immediate future is based on correlations among lagged observations and error terms. The feature-based models, however, select relevant prior observations based on the symbolic or geometric characteristics of the time series in question, rather than their location in time. Then what will happen in the next time will depend on the current feature. Examples of feature-domain models include self-organizing maps (SOMs), decision trees, k-nearest neighborhood, etc. In a nutshell, feature-based models first identify or discover features, and then act accordingly by taking the advantage of these features. In a sense, this modeling strategies can be regarded as a change from the conventional *global modeling strategies* to *local modeling strategies*. The effectiveness of this modeling strategies is built upon the assumption that a global complex model can be effectively decomposed into many local simple models. To test the plausibility of this assumption, this paper attempts to examine whether the feature-domain models provide an effective representation of the financial time series data. In particular, we examine the feature-domain model proposed by Chen and He (2003).

Chen and He (2003) is the first to use SOMs to search for and identify price patterns. In their model, a geometric or trajectory pattern of the price series is considered as a feature. It is referred as to the *trajectory-domain* model. The motivation of Chen and He (2003) is based on the observation that in the financial market, chartists appear to have been good at doing pattern recognition for many decades, yet little academic research has been devoted to a systematic study of these kinds of activities.

They applied a 6×6 two-dimensional SOM to a time series data of TAIEX (Taiwan Stock Index), and hence 36 charts were derived automatically. Among the 36 charts, many are familiar, such as uptrends, downtrends, v-formations, rounding bottoms, rounding tops, double tops, and island reversal. Furthermore, many of these charts were able to transmit buying and selling signals. They also showed that trading strategies developed from these charts may have superior profitability performance.

As a follow-up of this research line, Chen and Tsao (2003) applied the same architecture to three American stock indices, including Dow Jones, Nasdaq, and S&P 500. In addition, they conducted a more rigorous statistical analysis of the discovered patterns. By using the one-sided studentized range test (Hayter, 1990), it was found that for the appearance of some charts, the aftermath equity curves established are either monotonically increasing or decreasing. This feature is hard to capture via ordinary econometric methods. However, after excluding unconditional mean return, such monotonicity disappears.

This paper provides a different approach to examine the SOM-discovered patterns, namely, the event study approach. We treat each pattern as an *event*. Every price trajectory classified to the same pattern is considered as the same event. The event study approach is then applied to estimate the impact of a pattern (event) on the aftermath return behavior, and, based on that, to examine the information content in the SOM-discovered patterns.

The empirical findings of econometricians suggest a general notion of "one model can not fit all." Many researches of asset pricing argue an issue of "whether beta is dead." Moreover, it is found that in the option pricing literature the Black-Scholes formula seem to provide reasonable accurate values during 1976 to 1978. However, since 1986 there has been a very marked and rapid deterioration (Rubinstein, 1994). In this paper, it is then interesting to see whether the informative patterns, if there are any, discovered by SOMs are consistently informative during the whole time horizon. We separate the data into two parts to examine this issue.

The rest of this paper is organized as follows. Section 2 will firstly give a brief review on Chen and He (2003)'s trajectory-domain model, and then describe the data and parameters considered. Section 3 contains a description on the event study approach and shows its relevance to our pattern analysis. Event study results are presented in Section 4. Section 5 concludes and gives several directions for future study.

2 The Trajectory-Domain Model

This section briefly reviews Chen and He (2003)'s trajectory-domain model. The model can be characterized as two parts. The first is the sliding window device that expresses a price trajectory as a chart, and the second is the SOMs that is used to charts clustering. We will firstly introduce SOMs in Section 2.1, and then show the sliding window device and the data used in this paper in Section 2.2.

2.1 Self-Organizing Maps

In contrast to the artificial neural networks (**ANNs**) which are used for *supervised learning*, SOMs are another special class of artificial neural networks. The SOMs are used for *unsupervised learning* to achieve *auto classification*, *data segmentation* or *vector quantification*. Unlike the supervised ANNs, SOMs do not require the user to know in advance the exact objects that they are looking for. This convenience is particularly important when one can only effectively recognize some patterns by *visual inspection* rather than by mathematical descriptions.

The SOMs adopt so-called *competitive learning* among all neurons. The output neurons that win the competition are called *winner-takes-all* neurons. In SOMs, the neurons are placed on the sites of an l-dimensional lattice. The value of l is usually 1 or 2. Through competitive learning, the neurons are tuned to represent a group of input vectors in an organized manner. The mapping from a continuous space to a discrete one or a two-dimensional space achieved by the SOMs reserves the spatial order.

Among a number of training algorithms for SOMs, Kohonen's learning algorithm is the most popular one (Kohonen, 1982; Haykin, 1994). Kohonen's learning algorithm adopts a heuristic approach. Each neuron on the lattice has a weight vector of w components attached. The w is the number of input variables of the input data sets. The winning neuron and its close neighbors in the lattice have their weight vectors adjusted towards the input pattern presented on each iteration. Unlike other clustering methods such as k-means clustering (Huang, 1997), Kohonen's SOMs have the advantage that the final training outcome is insensitive to the initial settings of weights. Therefore, Kohonen's SOMs have found a wide variety of applications in image processing, target detection, 3D dynamic modeling, the classification of pulse signals of the autonomic nervous system, speech processing, etc.

In the training process, for an input vector x, the weights of the winning neuron and its close neighbors are updated according to (1),

$$v_j(n+1) = v_j(n) + \eta(n)\pi_{j,i(x)}(n)[x - v_j(n)], \tag{1}$$

where $v_j(n)$ is the weight vector of the jth neuron at the nth iteration, $\pi_{j,i(x)}(n)$ is the *neighborhood function* (to be defined below) of node indices j and $i(x)$,

$$i(x) = arg \min_j \| x - v_j \|, j = 1, 2, ..., d^2, \tag{2}$$

and $\eta(n)$ is the *learning rate* at iteration n.

We take for the neighborhood function the *Gaussian* form,

$$\pi_{j,i(x)} = \exp\left(-\frac{d^2_{j,i(x)}}{2\sigma^2(n)}\right), \tag{3}$$

where $d_{j,i(x)}$ is the distance between node units j and $i(x)$ on the map grid, and $\sigma(n)$ is some suitably chosen, monotonically decreasing function of iteration times n. Here, the effective width σ decays with n linearly according to (4).

$$\sigma(n) = \sigma_0 + \frac{(\sigma_1 - \sigma_0)}{N-1}(n-1), \tag{4}$$

where σ_0 and σ_1 are constants ($\sigma_0 > \sigma_1$) and N is the total number of epoch. The learning rate decays in a *power* manner:

$$\eta(n) = \eta_0(0.005/\eta_0)^{(n-1)/N}, \tag{5}$$

where η_0 is constant.

The training takes a long time with almost all neurons initially having their weights updated. This training phase is called the *ordering phase*. During this phase, as the learning rate and effective width gradually decrease, the topological ordering of the weight vectors takes place. During this phase, the initial effective width assumes a large value and the weights of virtually all of the neurons are updated. Through competitive learning, the weight vectors gradually settle down to form a topological order. The weights then settle down gradually during the second phase of learning named the *convergence phase* where only the weights of the winning neuron and perhaps its nearest neighbors are updated according to the case presented.[1]

2.2 Model Design

In this paper we present the results of the application of the SOM to financial time series data. The data sets to be segmented are three empirical stock indices, which are the Dow Jones, Nasdaq, and S&P 500. The original data sets cover the daily closing prices from 1/1/80 to 7/10/02 and have 5688, 5682, and 5687 observations, respectively.

What we intend to do is to take a sliding window (Fig. 1) with different window width w moving from the first period to the last period of the whole data set indexed by t ($t = 1, ..., T$), so that all T observations will further subdivide into $T - w + 1$ subsamples, each with w observations of a time series. Each subsample represents a time series pattern. The SOM is then used to automatically divide all patterns into groups or clusters in such a way that members of the same group are *similar* (*close*) in the Euclidean metric space. The w observations of each subsample are normalized between 0 and 1. A two-dimensional 6×6 SOM is used to map the $T - w + 1$ records into 36 clusters.[2] The 6×6 lattice of SOMs is presented in Fig. 2. Here, we consider the hexagonal lattice.

[1] For more descriptions of SOMs and discussions on its mathematical properties, see, for example, Kohonen (1997). For the applications of SOMs to economics and finance, see Deboeck and Kohonen (1998).

[2] Chen and He (2003) and Chen and Tsao (2003) give some intuition to why SOM is a suitable tool for geometric pattern recognition. In addition, the choice of *two*-dimensional

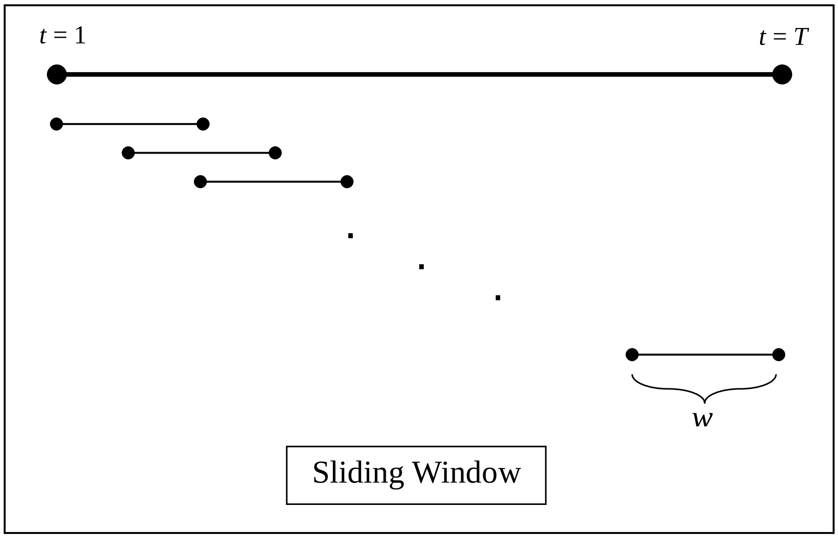

Figure 1: Sliding Window of the Trajectory-Domain Model.

In this paper, we consider six different w, their being 10, 30, 60, 90, 120, and 150. Then the results can be compared with different window widths. The control parameters used to conduct this experiment are given in Table 1.

3 Event Study Approach

The event study approach is widely used in finance as a quantitative tool to examine the aftermath of an event. Early event studies are primarily concerned with the impact of firm-specific events, such as the dividends payout, on stock returns. Focus generally lies on how stock prices adjust to the release of relevant information around certain events or announcements. Binder (1998) and MacKinlay (1997) provide nice survey of the literature on the firm-specific event studies. In the following subsections, the event study approaches are fine-tuned to match the need of this study.

lattice model is justified in those studies.

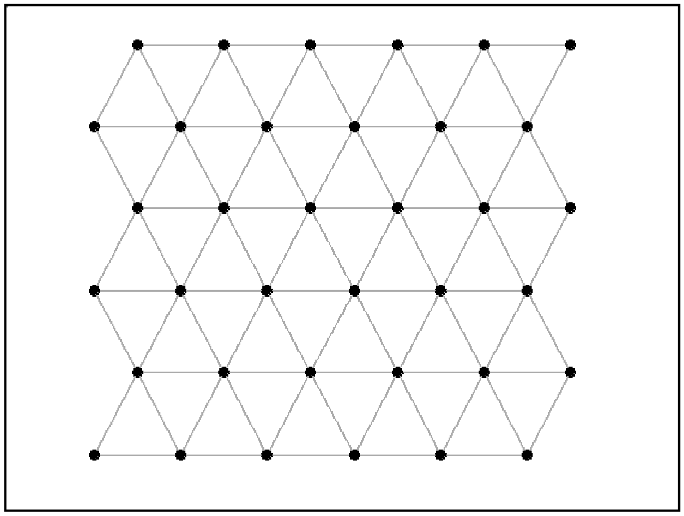

Figure 2: Map Structure of SOMs (Hexagonal Grid).

3.1 Event and Estimation Periods

For the design of sliding window of the trajectory-domain model, there is a *persistence* of a pattern. For example, if at period t_2, pattern j is observed, it may continue to appear for the next $m_f - 1$ days. In this case, we count the appearance of pattern j only once but attach to it a duration of m_f days. The index word f is the appearance index of the pattern under such modification, $f = 1, 2, \ldots, F_j$. Drawing this fact into the event study approach, we regard time t_2 as event date, and the event period is determined by the duration of the pattern on question. Since what interested is whether there is any abnormal after the pattern has been observed, we regard $[t_2 + 1, t_2 + m_f]$ as event period. Fig. 3 depicts the time horizon of the event study approach.

Fig. 4 introduces an example. In this case, the length of the series (T) is 11 and the window width (w) is 3. Then there are totally 9 ($T - w + 1 = 9$) charts for the trajectory-domain model. The number attached to each 3-period segment indicates the pattern recognized. At time 5, 6, 7, 9, and 10, the charts were recognized as Pattern 1. Then the first event period for the

Table 1: Parameter setup for the implementation of the 2-dimenisonal d $\times$ d SOM.

Window width (w)	10, 30, 60, 90, 120, 150
Dimensionality of SOM (l)	2
Number of neurons ($d \times d$)	6×6
Ordering phase initial radius (σ_0)	6.00
Ordering phase final radius (σ_1)	1.00
Ordering phase initial learning rate (η_0)	0.90
Ordering epoch (N)	1000
Convergence phase initial radius (σ_0)	1.00
Convergence phase final radius (σ_1)	0.10
Convergence phase initial learning rate (η_0)	0.10
Convergence epoch (N)	1000

Pattern 1 event is [6, 8] and the second is [10, 11]. If there are significant positive (or negative) returns during these two periods, the Pattern 1 reveals the signal of future price rising (or falling).

Some problems may raise under such construction of analysis. Firstly, the length of event period (m_f) is, of course, not constant, contrast to the general uses of event study approach of constant event period. However, it is *deterministic* after the SOMs have been trained. Then the test statistics of the event study approach can still be obtained using central limit theorem.[3] Secondly, from Section 2 we know that the learning process of SOMs is iterative and it uses all of the sample to train the map. Then one chart classified into some specific pattern will depend on charts before and after that chart. Hence, this is an *in-sample analysis*, but, there is no *in-sample problem*, i.e., the evidence, if there are any, of the abnormal returns will not been overemphasized. Notice the unsupervised learning properties of SOMs. The purpose of the learning process is not to find the pattern that could induce any aftermath return behavior, i.e, , the determinant of the pattern is *independent* of the abnormal returns. On the contrary, it just clusters the

[3]For details, see Section 3.3.

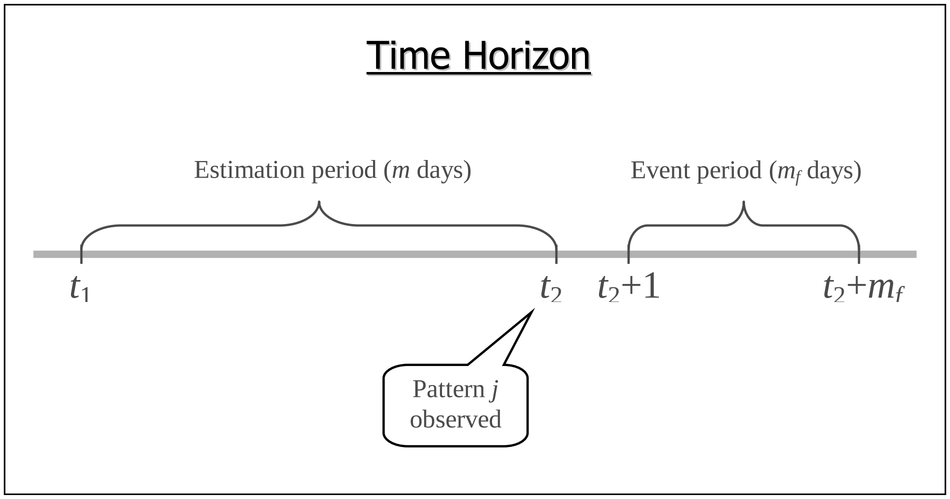

Figure 3: Time Horizon of Event Study Approach.

charts based on the similarity in the Euclidean space. Therefore, although the events (patterns) considered here are endogenous, the event studies can show us the impact of the events with the same reliable as any other exogenous events such as tax policy of the markets or the macroeconomic circumstances.

Another essential element of the event study approach is to define the so-called abnormal return, such that it can correctly measure the impact of a specific event. To define what is abnormal, one has to define what is normal. Alternatively speaking, under the event study framework, the criterion used to distinguish the informative patterns from non-informative patterns is the abnormal return, which mainly is a statistic of the first-order moment. In event study approach, the normal return comes from the prediction of *benchmark model*. The difference between the actual return (r_t) and the predicted return ($E_b(r_t)$) is then called the abnormal return (AR_t), i.e.,

$$AR_t = r_t - \mathrm{E}_b(r_t).$$

In practice, the benchmark model is estimated from *estimation period* $[t_1, t_2]$ (Fig. 3). The length of estimation period ($m = t_2 - t_1 + 1$) is arbitrarily set

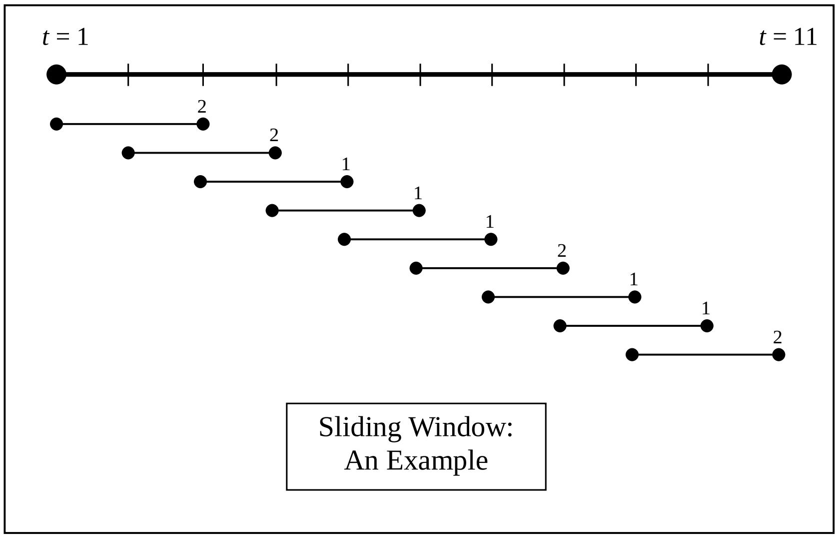

Figure 4: Sliding Window: An Example.

as 200 days in all of the cases.

One of the usual benchmark models in firm-specific event studies is the market model. However, from the data we considered is stock indices, the market models or other benchmark models in firm-specific event studies are not suitable. We must then search for another model that is usually used to capture stock indices progress. Conventionally, there are couple of econometric model can help us predict what the normal return of stock index is. This paper consider GARCH-M plus MA(1) as the benchmark. Two questions arise from this choice. Why GARCH-M plus MA(1)? What are the consequences of misspecification? We try to justify such choice in the next subsection. In addition, once the benchmark model might be misspecified, the bootstrap method is considered as a remedy to consolidate the test results. We detail this in Section 4.2.

3.2 GARCH-M plus MA(1)

The ARCH process introduced by Engle (1982) is the one econometric tool that can capture the volatility clustering phenomenon in financial time series data. The basic idea of ARCH model is to allow the conditional variance to change over time. Bollerslev (1986) proposes GARCH model in which on one hand allow for a more flexible conditional variance structure and on the other hand convert a high-order ARCH model into a more parsimonious GARCH representation that is much easier to identify and estimate, while in empirical applications of the ARCH model a relatively long lag in the conditional variance equation is often required. Bollerslev et al. (1992) found that the GARCH(1,1) is most identified in financial time series data. The GARCH(1,1) model can be written as

$$\begin{aligned} r_t &= c + \epsilon_t \\ \epsilon_t | \Omega_{t-1} &\sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \end{aligned} \tag{6}$$

where Ω_{t-1} is the information set available at time $t-1$.

Based on the idea that the risk-averse investors require compensation for holding risky assets, ARCH model is extended by Engle et al. (1987) to allow for the variance to be a determinant of the mean and is called ARCH-M. Thus as the risk of an asset changes over time, the risk premium changes accordingly, and also, the expected return. It is straightforward to enlarge ARCH-M to having GARCH-M model. Consider a GARCH(1,1)-M model

$$\begin{aligned} r_t &= c + \delta h(\sigma_t) + \epsilon_t \\ \epsilon_t | \Omega_{t-1} &\sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \end{aligned} \tag{7}$$

The choice of $h(\sigma_t) = \sigma_t$ ($h(\sigma_t) = \sigma_t^2$) represents the assumption that the conditional expected return is proportion to the conditional standard deviation (variance). In French et al. (1987)'s research, it is found that the specification of $h(\sigma_t) = \sigma_t$ fits the data slightly better than that of $h(\sigma_t) = \sigma_t^2$, however, the evidence is not strong. Engle et al. (1987) states that empirically, $h(\sigma_t) = \log \sigma_t$ is found to be a better choice. In this paper, we consider both the specifications of σ_t and σ_t^2. That is

$$\begin{aligned} r_t &= c + \delta\sigma_t + \epsilon_t \\ (\text{or } r_t &= c + \delta\sigma_t^2 + \epsilon_t) \\ \epsilon_t|\Omega_{t-1} &\sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2. \end{aligned} \tag{8}$$

The final forecasting model is chosen between them via the Akaike information criterion (AIC).

Spurious first-order autocorrelation can be usually found in an asset return due to two possible reasons: nonsynchronous trading and bid-ask spread. Most of the financial asset tradings, such as the individual stocks on the NYSE, do not occur in a synchronous manner. For daily stock returns, nonsynchronous trading can induce lag-1 cross-correlation between stock returns and, thus, lag-1 serial correlation in a portfolio return.[4] Another financial issue that can cause spurious lag-1 correlation is the bid-ask spreads, which exist in the stock exchanges with market makers. The Market makers are individuals who stand ready to buy or sell whenever the public wishes to sell or buy. They buy from the public at the bid price and sell at the ask price. The difference between these two prices is called the bid-ask spread. The realized price thus jumps between the bid and ask price, which introduces a negative lag-1 serial correlation in the return series (Roll, 1984). Not only in individual stock, but also the effect of bid-ask spread continues to exist in portfolio returns.

In order to capture the first-order autocorrelation induced by the bid-ask spread and nonsynchronous trading, a MA(1) term is included in the mean equation of (8). We obtain

$$\begin{aligned} r_t &= c + \delta\sigma_t + \epsilon_t - \theta\epsilon_{t-1} \\ (\text{or } r_t &= c + \delta\sigma_t^2 + \epsilon_t - \theta\epsilon_{t-1}) \\ \epsilon_t|\Omega_{t-1} &\sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2. \end{aligned} \tag{9}$$

French et al. (1987) applied Model (9) to the Standard and Poor's composite portfolio to examine the relation between stock returns and stock market

[4] See, for example, Campbell et al. (1997) for more detail discussion.

volatility. In this paper, we apply Model (9) to Dow Jones, Nasdaq, and S&P 500 indices, but take it as a benchmark model in event study to examine the information content of the SOMs-discovered patterns.

3.3 Test Statistics

For pattern j ($j = 1, ..., 36$), we are interested in the null hypothesis:

H_0: There is no abnormal return after pattern j has been observed.

Due to the construction for the analysis in this paper described in Section 3.1, the null hypothesis was rewritten as:

H_0': There is no cumulative abnormal return after pattern j has been observed.

Two statistical tests have been frequently used to test the significance of events in the event study approach. One is the standardized-residual method (SRM) proposed by Pattel (1976), and the other is the standardized cross-sectional method (SCSM). The former assume that there is no event-induced variance, whereas the latter assume there is.

The standardized-residual method assumes that the variance structure of return is the same in both estimation and event period. The test statistic is as follow:[5]

$$t_{SRM} \equiv \frac{\sum_{f=1}^{F_j} \sum_{i=1}^{m_f} SAR_{f,t_2+i}/\sqrt{m_f}}{\sqrt{F_j(m-5)/(m-7)}} \xrightarrow{d} N(0,1). \tag{10}$$

where $\{SAR_{f,t_2+i}\}_{i=1}^{m_f}$ is the set of standardized abnormal return during the event period. [6]

If during the event period the variance increase or decrease, the t_{SRM} seems not to be a good test statistic. It may reject the null too often or seldom. Cowan and Sergeant (1996) point out that three commonly used

[5] This test statistic is a little different from Pattel (1976) due to that the event length here is not constant over each occurrence of event. Appendix 1 details the derivation of the statistic.

[6] For rigid definition, see Appendix 1.

tests are potential solutions to the problem of event-induced variance: the generalized sign test proposed by Cowan (1992), Corrado's rank test, and a standardized cross-sectional test proposed by Boehmer et al. (1991). The second test statistic used in this paper is the modified version of Boehmer et al. (1991)'s standardized cross-sectional test to allow for nonconstant event length. The test statistic is as follow:[7]

$$t_{SCSM} \equiv \frac{\sqrt{F_j}\overline{SCAR}}{S_{F_j}} \xrightarrow{d} N(0,1). \tag{11}$$

where $\overline{SCAR}$ and S_{F_j} is the sample mean and sample standard deviation of standardized cumulative abnormal return.[8]

Before examining the magnitudes of $t_{SRM,j}$ and $t_{SCSM,j}$, $j = 1, 2, \ldots, 36$, to judge which pattern is informative, we first emphasize that it would be desirable to conduct a joint test of the 36 patterns together rather than 36 tests for each individual pattern. It is clear that joint test can give us a better control than the individual tests regarding the general conclusion: SOM can discover informative patterns. Then the null hypothesis we are interested is:

H_0'': SOM can not discover informative patterns.

On the other hand, from the statistical point of view, the result from the joint test is robust because it avoids the problem of test size diminishing which happened when conducting many tests together. So the two chi-square tests are consider before going through each patterns:[9]

$$q_K^2 \equiv \sum_{j=1}^{36} t_{K,j}^2 \xrightarrow{d} \chi_{36}^2, \quad K = SRM, SCSM \tag{12}$$

In summary, the experiment design and analysis are depicted in the flowchart displayed Fig. 5.

---

[7] Here, we make two regular assumptions about the return series in order to apply central limit theorem and the Slutsky theorem. Appendix 2 details the assumptions and the derivation of the statistic.

[8] For rigid definition, see Appendix 2.

[9] If the intertemporal dependency of the return series is perfectly captured by the benchmark model, the $\{t_{K,j}^2\}_{j=1}^{36}$, $K = SRM, SCSM$, is independent identical distributed (iid). Then the asymptotic distribution of the following statistics can be obtained.

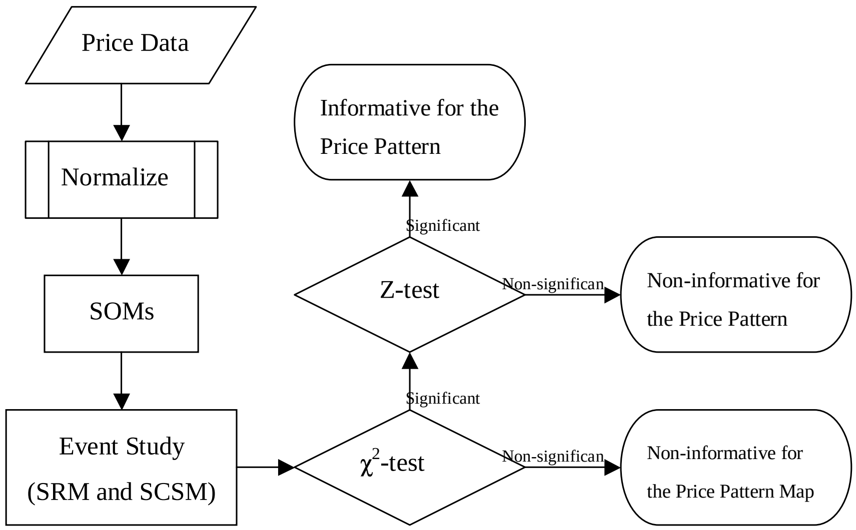

Figure 5: Flowchart of the Analysis.

4 Results

4.1 General Description

Before a formal presentation of our testing results, it would be useful to have a general picture of the patterns we discovered via the SOM. They are depicted in Figs. 6-8. Each figure stands for different stock index, and each map in the figure stands for different window width. There are 36 patterns in the map, in which the relative position of the patterns match the hexagonal lattice. From these figures, it is worthy noting that some patterns are very similar due to exogenous setting of the size of SOM. [10] From these diagrams, it is also clear that the patterns who are neighbors to each other behave similar. This is also what one can expect from a full-spanned SOM. [11]

Frequencies of each patterns are not uniformly distributed. Some pat-

[10] There must be 36 clusters to form, no more and no less, regardless of how they are similar or dissimilar.

[11] The SOM algorithm does not guarantee the full-span of the web.

Table 2: Event Study: Joint Tests for the Patterns Discovered by SOMs.

	Dow Jones		Nasdaq		S&P 500	
w	SRM	SCSM	SRM	SCSM	SRM	SCSM
10	***(0.05)	**(0.10)	***(0.07)	***(0.00)	***(0.04)	*(0.12)
30	**(0.35)	(0.40)	***(0.36)	**(0.26)	*(0.33)	(0.46)
60	**(0.22)	*(0.24)	***(0.15)	***(0.11)	***(0.05)	*(0.16)
90	***(0.06)	(0.73)	***(0.07)	***(0.08)	*(0.37)	(0.84)
120	***(0.04)	(0.76)	***(0.25)	(0.76)	**(0.10)	***(0.01)
150	**(0.07)	(0.90)	***(0.30)	**(0.31)	*(0.34)	(0.79)

*, **, and *** denote 10%, 5%, and 1% significant levels, respectively, based on regular χ^2 tests. The numbers in the parentheses indicate the p-value obtained via bootstrap method.

terns were found more prevalent than others. This can be seen from the display of Figs. 9-11. In these figures, the size of the black hexagons indicate the amount the charts clustered. The larger the size is, the more widespread the pattern is in the whole time series. There is a general finding in term of window width (w) that larger window width has less uniformly distributed frequency of each pattern. However, as we shall see in the next subsection, most of these patterns are not informative from the perspective of the event study approach.

4.2 Event Studies

From Table 2, we first notice that the testing results are sensitive to the test methods and the window widths. The null hypothesis is rejected more frequently in the method of SRM. Moreover, the results among different indices are also different. Among all possible combinations, the one looks particular impressive is the Nasdaq, whose SOM-discovered patterns are significant in almost all window widths by using either the SRM or the SCSM. Why is SOM so powerful for the Nasdaq index? Is that powerfulness real or spurious?

To answer this question, one have to first notice that the creditability of our tests may crucially depend upon the benchmark from which the abnormal return is derived. One way to gauge the effect of the misspecification of the benchmark upon our testing results is to use the bootstrap approach. By that approach, we shuffled the sequence of the patterns discovered by the SOM. Ideally, by this doing, the information revealed by the original patterns shall all gone. In other words, the "patterns" after shuffling shall no longer guide us to see the abnormal return.

The shuffling procedure was repeated 100 times. For the shuffled sequence b, the event study is applied and then $\left\{(t^{(b)}_{SRM,j}, t^{(b)}_{SCSM,j})\right\}_{j=1}^{36}$ is obtained and, also, $\left(q^{2(b)}_{SRM}, q^{2(b)}_{SCSM}\right)$, $b = 1, 2, \ldots, 100$, does. The p-value is calculated via:

$$\text{p-value} = \frac{\sharp\left(q^{2(b)}_{K} > q^{2}_{K}\right)}{100}, \quad K = SRM, SCSM. \tag{13}$$

The numbers in the parentheses in Table 2 indicate the p-value obtained via such bootstrap approach. It is obviously that the bootstrap methods give more conservative results. Take Nasdaq as example, there are only two window widths ($w = 10, 90$) indicating that the patterns discovered by SOMs are informative under 10% significant level, in terms of both event study methods.

Based on our results, there are some evidences to support the relevance of the SOM to pattern discovery. Following the event-study tests, we found that some charts discovered by SOM can in effect transmit signals of abnormal returns. For example, in terms of bootstrap p-values of both test statistics, there are four maps disclosing informative signal under 10% significant level. They are the maps with window width $w = 10$ for Dow Jones and Nasdaq, with $w = 90$ for Nasdaq, and with $w = 120$ for S&P 500. There are, however, two remarks made for this finding. First, the evidences are not consistent for different window widths. Second, the evidences are also not consistent for different markets. The first remark is not entirely surprising considering that the real charts used by chartists also do not have a fixed window width.[12] The second remark indicates that charts may

[12] An interesting issue left for the future is to extend the SOM to deal with window-size-free patterns.

be more informative in some markets. This is also familiar because many chartists believe that technical analysis may be more supported by some specific markets. It would then be the next pursuit to understand what factors may cause the emergence of informative charts.

To show how the informative patterns in the four informative maps look like, Figs. 12-15 depict the charts with aftermath abnormal returns. The thick lines in the figures demonstrate that the aftermath return of the pattern has positive abnormal behavior. The patterns with dot line indicate there being negative abnormal return. Although we do not presume the patterns beforehand to be some specific types of chart, in contrary, the patterns emerge themselves from the data, some of the informative patterns discovered by SOMs can roughly be given a name in chartist' eyes, such as uptrends (Pattern (5,2) in Fig. 13 and (2,4) in Fig. 14), downtrend (Pattern (4,6) in Fig. 13), V-formations (Pattern (4,2) in Fig. 12 and (5,4) in Fig. 14), rounding bottom (Pattern (1,5) in Fig. 13), flat (Pattern (3,4) in Fig. 12), wedge (Pattern (5,2) in Fig. 14), and single zigzag (Pattern (4,5) in Fig. 12 and (3,3) and (6,6) in Fig. 13).

The portions of the data which was found to have significant patterns are 14.44%, 25.37%, 16.18%, and 4.35% for the cases of Fig. 12-15, respectively.

4.3 Life of Informative Patterns

There are tremendous evidences indicating that the pattern has a life. It can emerge, and will die as well. With this background, it is questionable whether there are indeed any pattern which can signal abnormal returns and are not found for 20 years. The second analysis of this paper is to examine the life of patterns. A simple device to do so is to divide the whole data set into two parts, and then check whether the patterns found significant in previous section survived in both sub-periods or just one of the them. Table 3 is the joint test results for the patterns' life. There are six panels in Table 3, each stands for different window widths. In each panel, the first row is the test results using all data, which is the same as in Table 2. The second and the third rows are the results using the first and the second half of the data, respectively. In terms of bootstrap p-values of both test statistics under 10%

Table 3: Event Study: Joint Tests for the Patterns' Life.

	Dow Jones		Nasdaq		S&P 500	
w	SRM	SCSM	SRM	SCSM	SRM	SCSM
10	***(0.05)	**(0.10)	***(0.07)	***(0.00)	***(0.04)	*(0.12)
	**(0.20)	(0.38)	***(0.08)	***(0.00)	(0.47)	(0.42)
	(0.58)	(0.74)	***(0.07)	***(0.07)	**(0.06)	(0.22)
30	**(0.35)	(0.40)	***(0.36)	**(0.26)	*(0.33)	(0.46)
	**(0.37)	(0.56)	***(0.37)	(0.34)	(0.54)	(0.81)
	(0.65)	(0.76)	***(0.49)	**(0.36)	(0.38)	*(0.17)
60	**(0.22)	*(0.24)	***(0.15)	***(0.11)	***(0.05)	*(0.16)
	*(0.40)	***(0.00)	***(0.44)	**(0.13)	***(0.03)	***(0.02)
	(0.33)	(0.83)	***(0.02)	***(0.09)	(0.73)	(0.83)
90	***(0.06)	(0.73)	***(0.07)	***(0.08)	*(0.37)	(0.84)
	***(0.05)	(0.50)	***(0.09)	**(0.30)	***(0.19)	*(0.39)
	**(0.10)	*(0.50)	***(0.05)	***(0.18)	***(0.01)	**(0.29)
120	***(0.04)	(0.76)	***(0.25)	(0.76)	**(0.10)	***(0.01)
	***(0.03)	(0.47)	***(0.34)	(0.65)	(0.78)	(0.44)
	(0.53)	(0.72)	***(0.06)	***(0.12)	**(0.04)	**(0.45)
150	**(0.07)	(0.90)	***(0.30)	**(0.31)	*(0.34)	(0.79)
	***(0.03)	(0.38)	***(0.13)	(0.47)	***(0.00)	(0.58)
	(0.12)	(0.76)	***(0.03)	***(0.14)	(0.72)	(0.99)

*, **, and *** denote 10%, 5%, and 1% significant levels, respectively, based on regular χ^2 tests. The numbers in the parentheses indicate the p-value obtained via bootstrap method.

significant level, there is one case (Nasdaq with $w = 10$) the patterns are informative in both subsamples and, of course, in the whole sample, there are three cases (Dow Jones with $w = 10$, Nasdaq with $w = 90$, and S&P 500 with $w = 120$) the informative signals are not significant in both periods but are in the whole sample, and there are two cases (Nasdaq with $w = 60$ and S&P 500 with $w = 60$) the informative signals only appear in one of the two periods.

Furthermore, to examine individual patterns, it is evidenced that most patterns that are found significant could in effect only successfully transmitted profitable signals in one of the two periods. For example, in the case of Nasdaq with $w = 10$, there are only 3 among 13 informative patterns the informative signal is significant in both periods. This is a highly interesting result. It shows that most of the charts which were informative in the first period died in the second period. However, chartists did not die, because in the next period there were new charts appearing and waited to be found. The finding lends support to the recent simulation study of agent-based artificial financial markets. The finding is further strengthened by the evidence that some non-informative patterns were actually found to be significant in one of the sub-periods. Take Nasdaq with $w = 10$ as example again, there are 7 patterns were judged to be informative in one of the periods but not in the whole sample. This implies that a number of patterns which were not found significant is because they could no longer transmitting the profitable signal in the second period. In other words, they died in the second period.

5 Conclusions

From the chartists' view, chart is an event. What scenario of the market in the future will depend on what pattern has been recognized today. Thus, it is straightforward to apply the event study approach to the analysis of information content of the trajectory-domain model, which is proposed by Chen and He (2003) based on the motivation of chart analysis. In this paper there are some evidences to support the relevance of the trajectory-domain model to informative pattern discovery. Some patterns discovered by SOM can in effect transmit signals of abnormal returns. However, the signals are not persistent. They would emerge during a period and then vanish, vice versa.

There are several directions that could extend this study, include:

- Multivariate model.

  Constructing multivariate models might be a direct extension of the study. The variables might be adopted with the same attributes, e.g., two markets prices, or with different attributes, e.g., the price and volume.

- Profitability.

  The purpose of this paper is not to test whether the trajectory-domain models can help us to make money. However, the profitability might be an alternative interesting issue to be examined. Then the results of the empirical analysis can be compared with theoretical financial issues, such as the efficient market hypothesis.

- The effect of model parameters.

  For example, what is the effect of the size of the SOM? Should the high-dimensional SOM make one easier to find patterns, if there are any?

- The effect of event study setting.

  For example, how is the event period discerned? Would the length of the event play a role in its significance?

Appendix 1

Consider the f-th appearance of pattern j,

$$AR_{f,t_2+i} = r_{t_2+i} - \hat{r}_{t_2+i}, \quad i = 1, 2, ..., m_f,$$

where $\hat{r}_{t_2+i}$ is the estimated predicted return from the benchmark model. Suppose there is no event-induced variance and H_0 is true, then

$$SAR_{f,t_2+i} \equiv \frac{AR_{f,t_2+i}}{SE_{f,t_2+i}} \overset{iid}{\sim} t(m-5),$$

where SE_{f,t_2+i} is estimated predicted error. The average standardized cumulative abnormal return can be obtained as

$$\overline{SCAR}_f \equiv \frac{\sum_{i=1}^{m_f} SAR_{f,t_2+i}}{m_f} \sim D\left(0, \frac{m-5}{(m-7)m_f}\right),$$

where $D(a,b)$ denote some distribution with mean a and variance b. Normalizing $\overline{SCAR}_f$ we have

$$K_f \equiv \frac{\overline{SCAR}_f}{\sqrt{\frac{m-5}{(m-7)m_f}}} \overset{iid}{\sim} D'(0,1), \quad f = 1, 2, ..., F_j.$$

Applying central limit theorem we get

$$\sqrt{F_j}\bar{K} \xrightarrow{d} N(0,1).$$

I.e.,

$$\frac{\sum_{f=1}^{F_j} \sum_{i=1}^{m_f} SAR_{f,t_2+i}/\sqrt{m_f}}{\sqrt{F_j(m-5)/(m-7)}} \xrightarrow{d} N(0,1).$$

Appendix 2

We now consider the case in which there is event-induced variance, then

$$SAR_{f,t_2+i} \equiv \frac{AR_{f,t_2+i}}{SE_{f,t_2+i}} \overset{iid}{\sim} D''(0, \sigma^2_{t_2+i}), \quad i = 1, 2, ..., m_f,$$

Thus, the standardized cumulative abnormal return will be

$$SCAR_f \equiv \sum_{i=1}^{m_f} SAR_{f,t_2+i} \sim D'''(0, \sigma^2_f),$$

where

$$\sigma_f^2 = \sum_{i=1}^{m_f} \sigma_{t_2+i}^2.$$

Let

$$\bar{\sigma}_{F_j}^2 = \sum_{f=1}^{F_j} \sigma_f^2 / F_j,$$

and

$$S_{F_j}^2 = \frac{\sum_{f=1}^{F_j} (SCAR_f - \overline{SCAR})^2}{F_j - 1}.$$

We make the first assumption here: Suppose $\lim_{F_j \to \infty} \max(\sigma_f)/(F_j \bar{\sigma}_{F_j}) = 0$ and $\bar{\sigma}^2 = \lim_{F_j \to \infty} \bar{\sigma}_{F_j}^2$ exists. Then apply central limit theorem (Lindgerg-Feller) we get

$$\frac{\sqrt{F_j}\overline{SCAR}}{\bar{\sigma}} \xrightarrow{d} N(0,1). \tag{14}$$

What left is to show $S_{F_j}^2 \xrightarrow{p} \bar{\sigma}^2$, then using Slutsky theorem

$$\frac{\sqrt{F_j}\overline{SCAR}}{S_{F_j}} \xrightarrow{d} N(0,1). \tag{15}$$

Two things need to be verified: $S_{F_j}^2$ is asymptotically unbiased and the variance of $S_{F_j}^2$ converges to zero.

1.

$$\begin{aligned}
\mathrm{E}(S_{F_j}^2) &= \mathrm{E}\left[\frac{\sum_{f=1}^{F_j} (SCAR_f - \overline{SCAR})^2}{F_j - 1}\right] \\
&= \frac{1}{F_j - 1}\left[\sum_{f=1}^{F_j} \mathrm{E}(SCAR_f^2) - F_j \mathrm{E}(\overline{SCAR}^2)\right] \\
&= \frac{1}{F_j - 1}\left[\sum_{f=1}^{F_j} \sigma_f^2 - \frac{\sum_{f=1}^{F_j} \sigma_f^2}{F_j}\right] \\
&= \frac{\sum_{f=1}^{F_j} \sigma_f^2}{F_j} \to \bar{\sigma}^2
\end{aligned}$$

2.

$$\mathrm{var}(S_{F_j}^2) = \mathrm{var}\left[\frac{\sum_{f=1}^{F_j}(SCAR_f - \overline{SCAR})^2}{F_j - 1}\right]$$

Then

$$\lim_{F_j \to \infty} \frac{\sum_{f=1}^{F_j} \mathrm{E}(SCAR_f^4)}{F_j} < \infty \tag{16}$$

is a sufficient condition for $\mathrm{var}(S_{F_j}^2) \to 0$ as $F_j \to \infty$. Here, we make the second assumption that Eq. (16) hold.

References

Binder, J. (1998), The event study methodology since 1969, *Review of Quantitative Finance and Accounting*, **11**, 111-137.

Boehmer, E., J. Musumeci, and A.B. Poulsen (1991), Event-study methodology under conditions of event-induced variance, *Journal of Financial Economics*, **30**, 253-272.

Bollerslev, T., (1986), Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307-327.

Bollerslev, T., R.Y. Chou, and K.F. Kroner (1992), ARCH modeling on finance: a review of the theory and empirical evidence, *Journal of Econometrics*, **52**, 5-59.

Campbell, J.Y., A.W. Lo, and A.C. MacKinglay (1997), The econometrics of financial markets, Princeton University Press.

Chen, S.-H. and H. He (2003), Searching financial patterns with self-organizing maps, in S.-H. Chen ed., *Computational Intelligence in Economics and Finance*, Springer (forthcoming).

Chen, S.-H. and C.-Y Tsao (2003), Self-organizing maps as a foundation for charting or geometric pattern recognition in financial time series, in *Proceedings of 2003 International Conference on Computational Intelligence for Financial Engineering (CIFEr2003)*, Hong Kong, March 20-23, 2003 (forthcoming).

Cowan, A.R. (1992), Nonparametric event study tests, *Review of Quantitative Finance and Accounting*, **2**, 343-358.

Cowan, A.R. and A.M.A. Sergeant (1996), Trading frequency and event study test specification, *Journal of Banking and Finance* **20**, 1731-1757.

Deboeck, G. and T. Kohonen (1998), *Visual Explorations in Finance with Self-Organizing Maps*, Springer.

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, *Econometrica*, **50**, 987-1007.

Engle, R.F., D.M. Lilien, and R.P. Robins (1987), Estimating time-varying risk premia in the term structure: the arch-m model, *Econometrica*, **55**, 391-407.

French, K.R., G.W. Schwert, and R.F. Stambaugh (1987), Expected stock returns and volatility, *Journal of Financial Economics*, **19**, 3-29.

Haykin, S.S. (1994), *Neural Networks: A Comprehensive Foundation*, New York: MacMillan.

Hayter, A.J. (1990), A one-sided studentized range test for testing against a simple ordered alternative, *Journal of the American Statistical Association*, **85**, 778-785.

Huang, Z. (1997), A fast clustering algorithm to cluster very large categorical data sets in data mining, *First Asia Pacific Conference on Knowledge Discovery and Data Mining*, Singapore, World Scientific, February.

Kohonen, T. (1982), Self-organized foundation of topologically correct feature maps, *Biological Cybernetics*, **43**, 59-69.

MacKinlay, A.C. (1997), Event studies in economics and finance. *Journal of Economic Literature*, **35**, 13-39.

Pattel, J.M. (1976), Corporate forecasts of earnings per share and stock price behavior: empirical tests, *Journal of Accounting Research*, **14**, 246-76.

Roll, R. (1984), A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance*, **39**, 1127-1139.

Rubinstein, M. (1994), Implied binomial trees, *Journal of Finance*, **49**, 771-818.

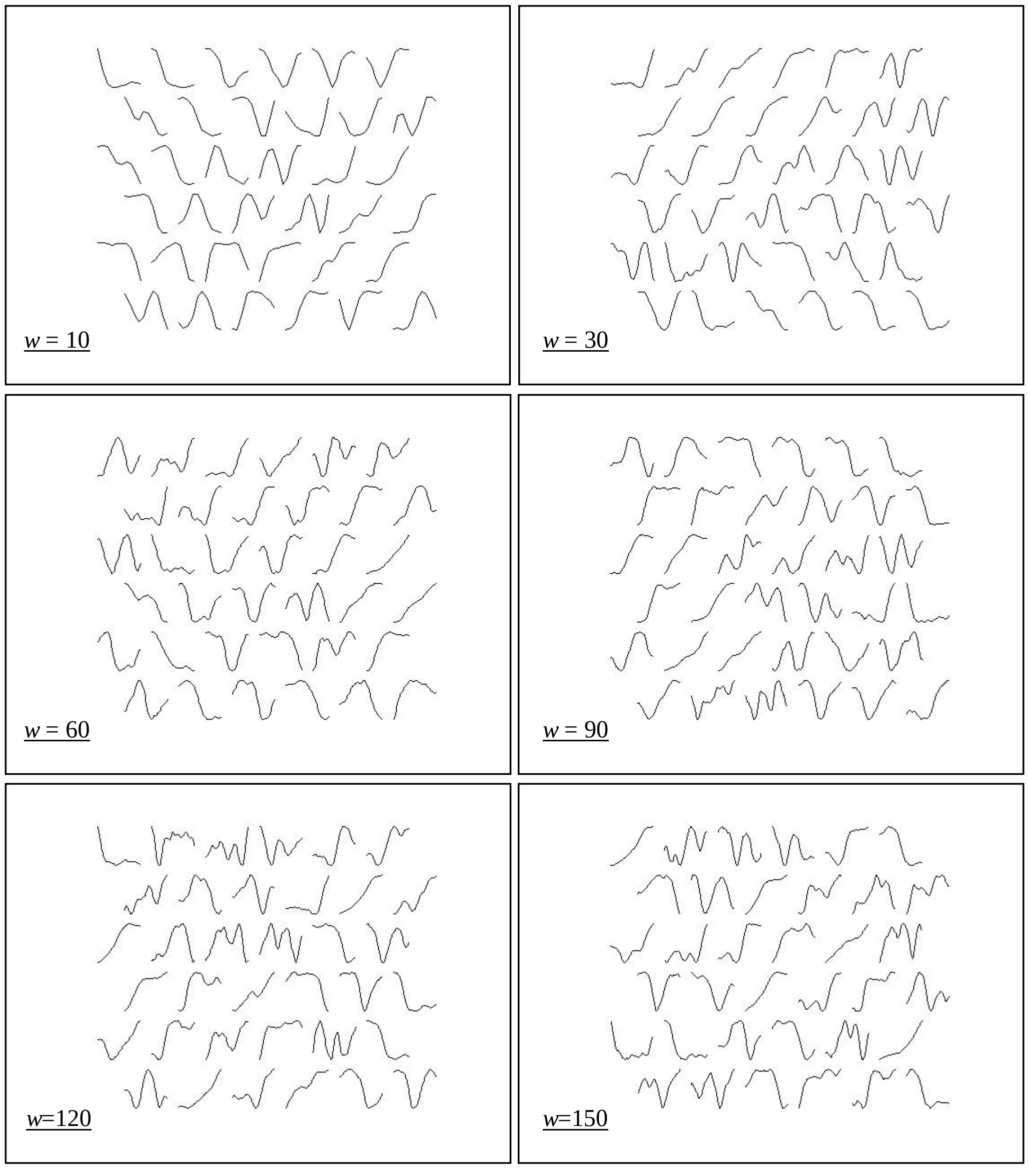

Figure 6: 6 × 6 Patterns Discovered by SOMs (Dow Jones).

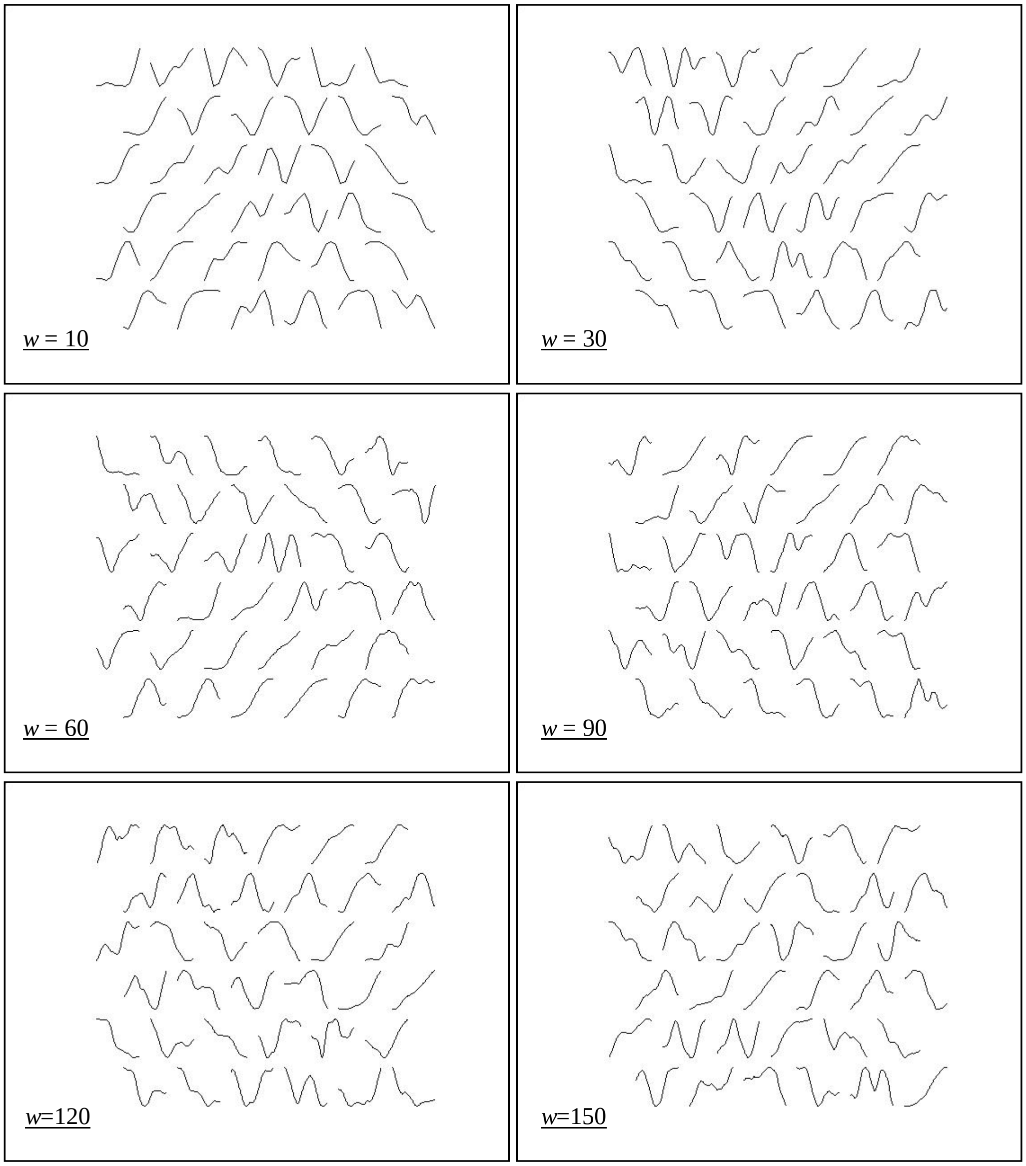

Figure 7: 6 × 6 Patterns Discovered by SOMs (Nasdaq).

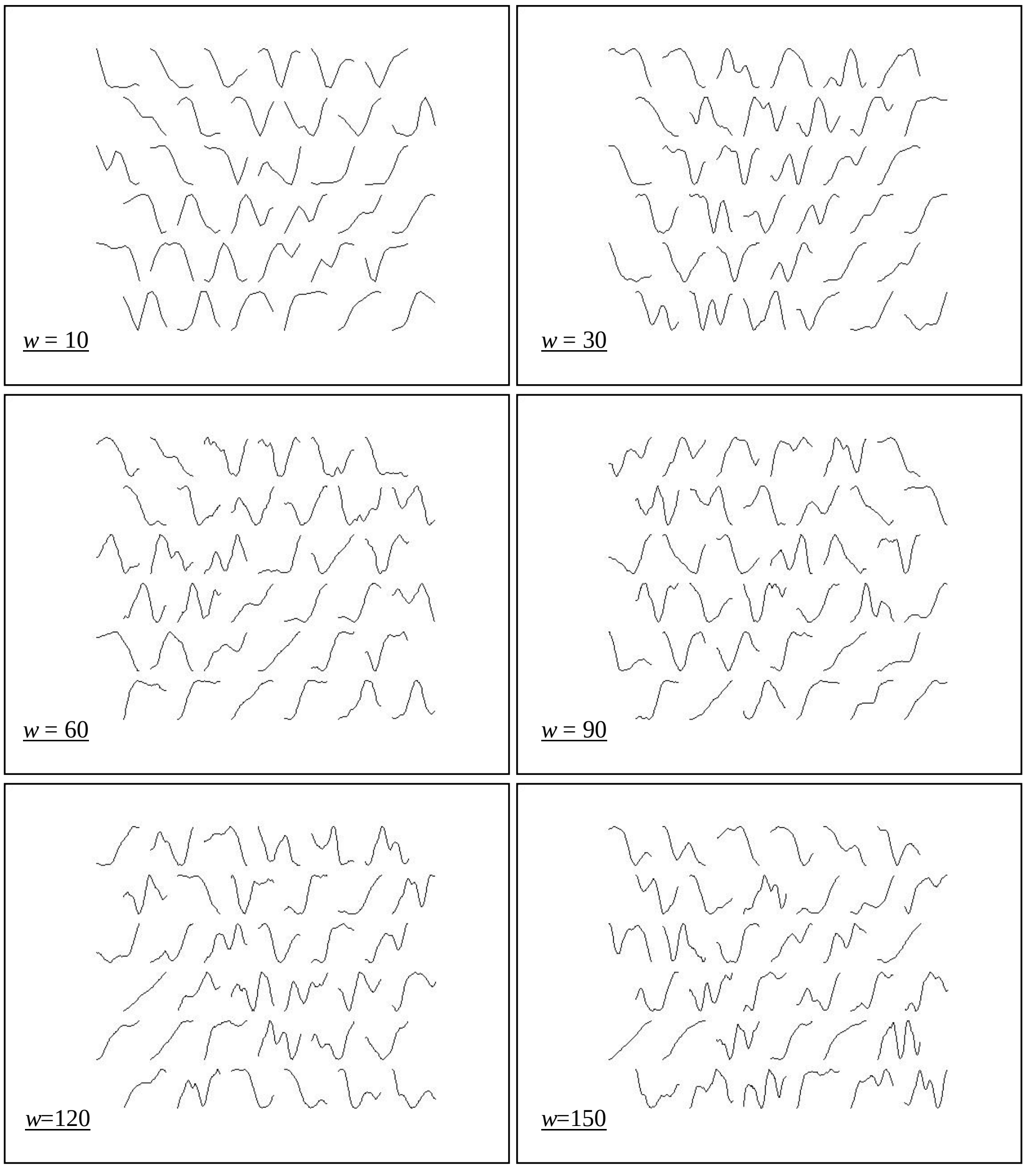

Figure 8: 6 × 6 Patterns Discovered by SOMs (S&P 500).

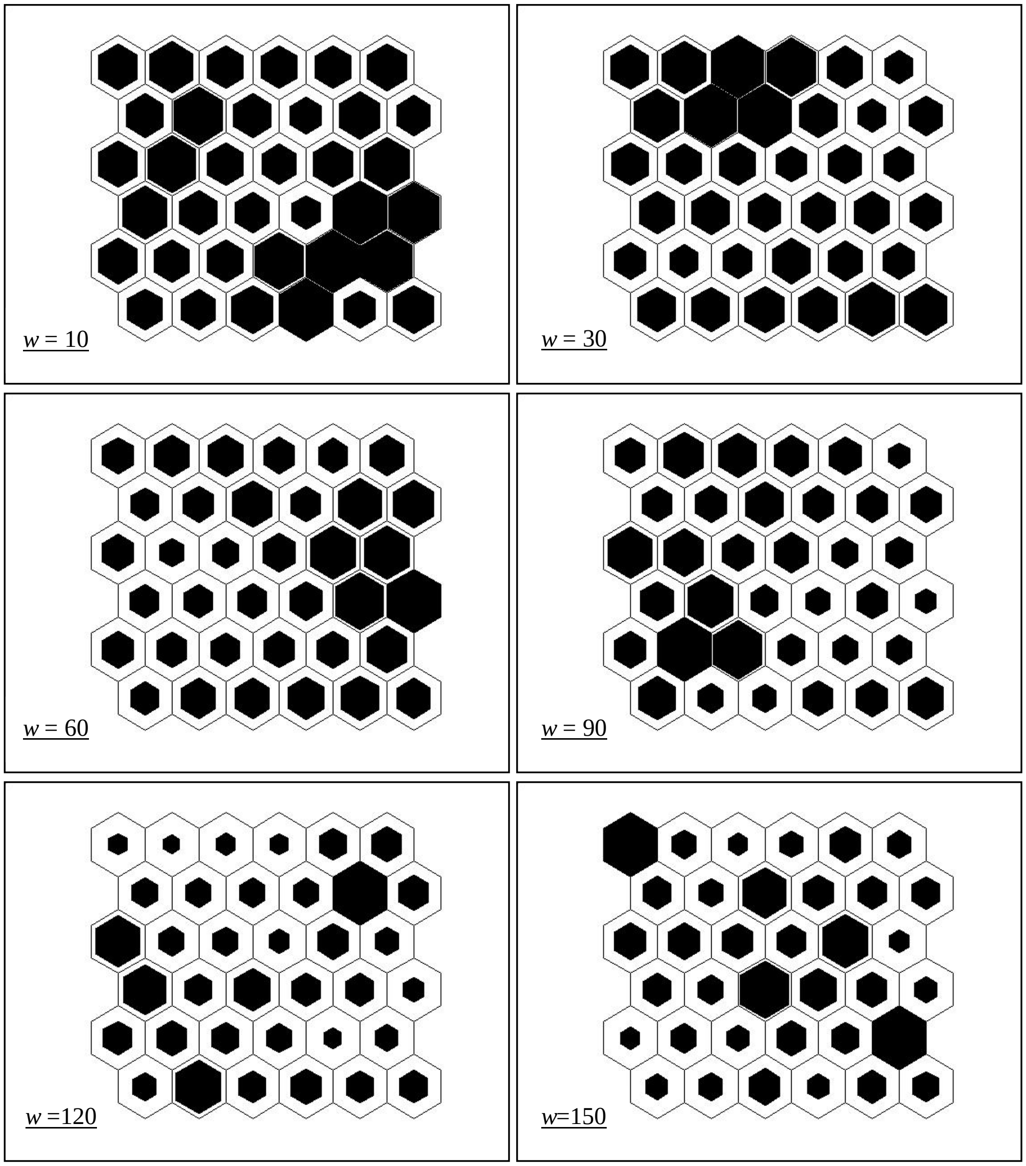

Figure 9: 6 $\times$ 6 Patterns Hits Clustered by SOMs (Dow Jones).

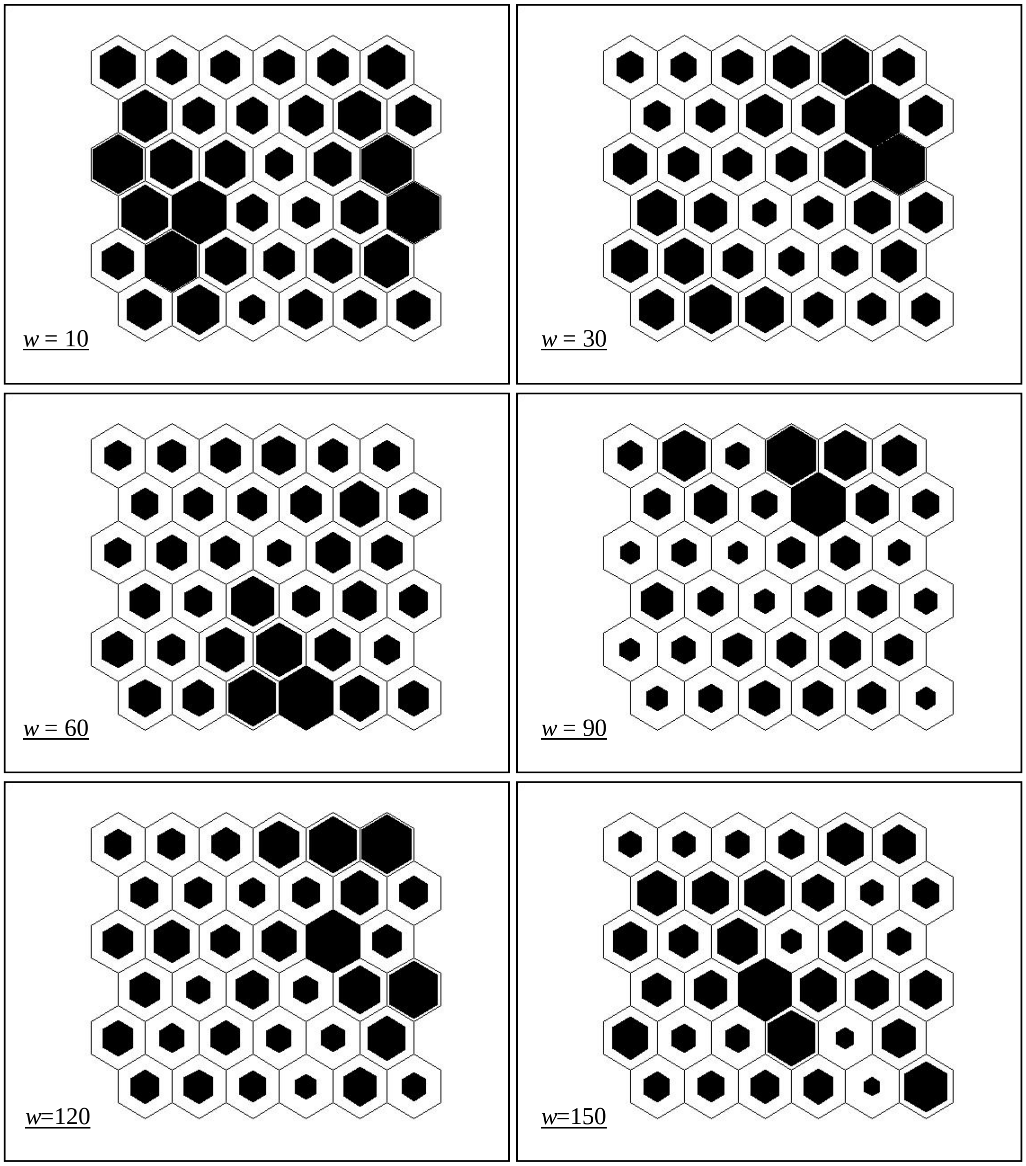

Figure 10: 6 × 6 Patterns Hits Clustered by SOMs (Nasdaq).

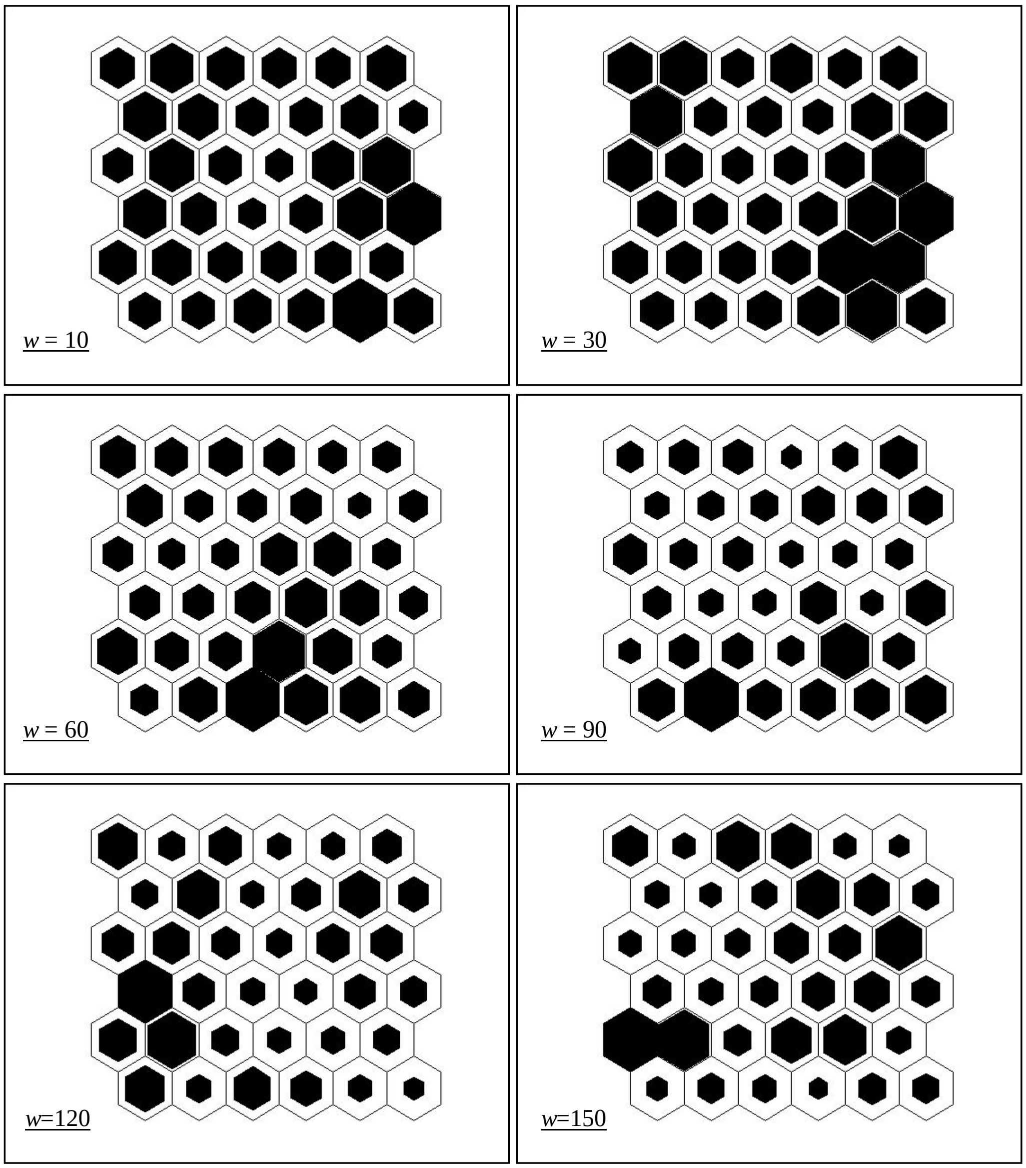

Figure 11: 6 × 6 Patterns Hits Clustered by SOMs (S&P 500).

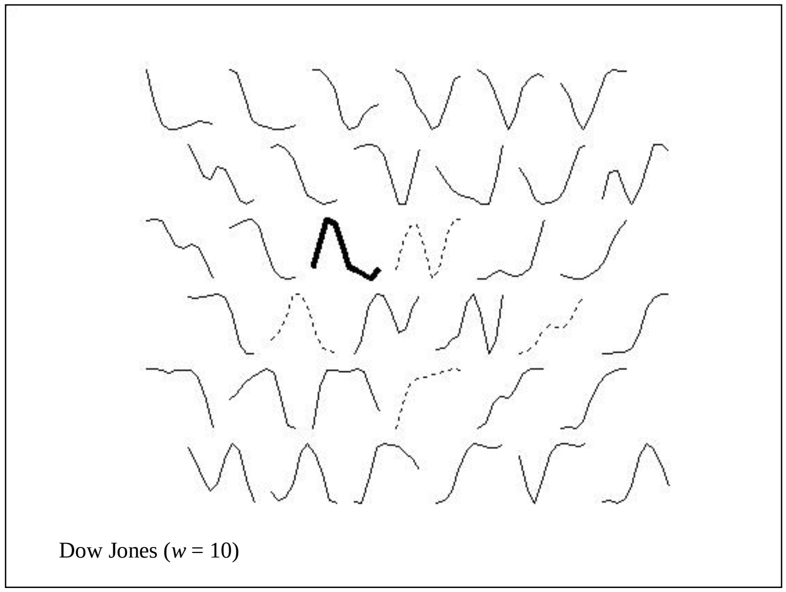

Figure 12: The Informative Patterns for Dow Jones ($w = 10$).

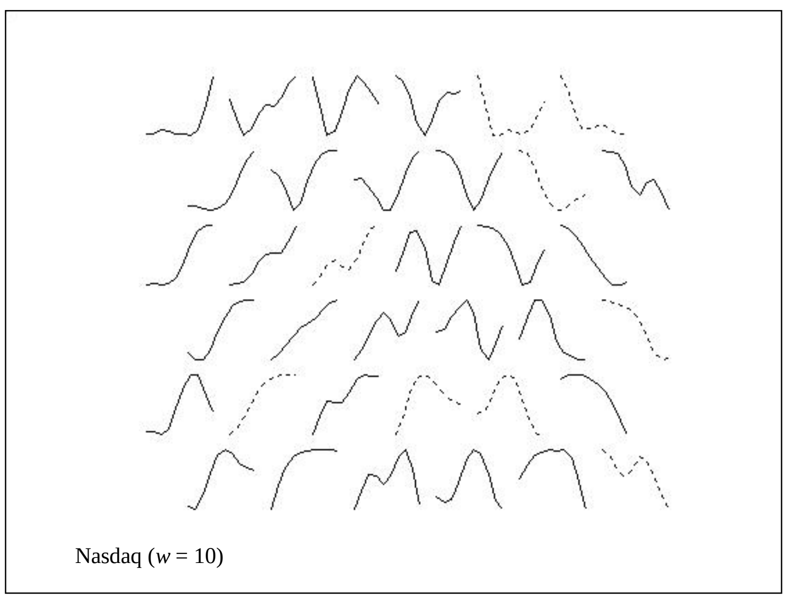

Figure 13: The Informative Patterns for Nasdaq ($w = 10$).

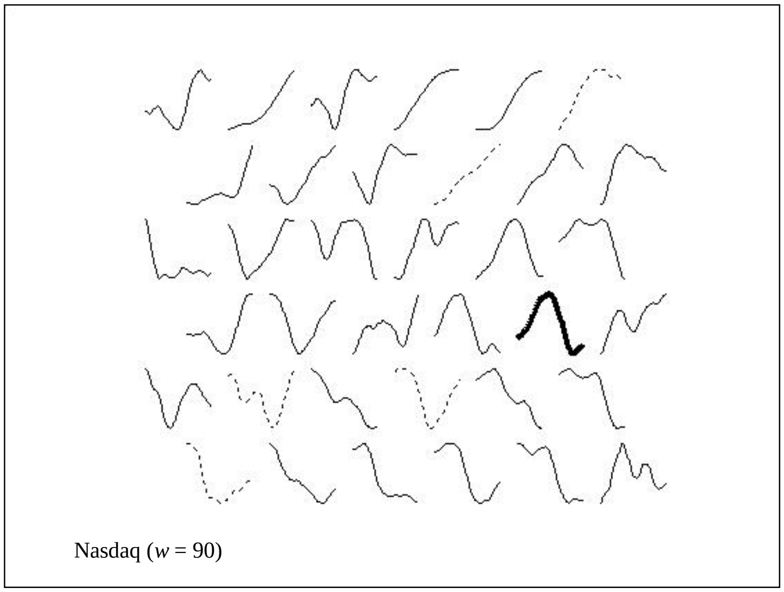

Figure 14: The Informative Patterns for Nasdaq ($w = 90$).

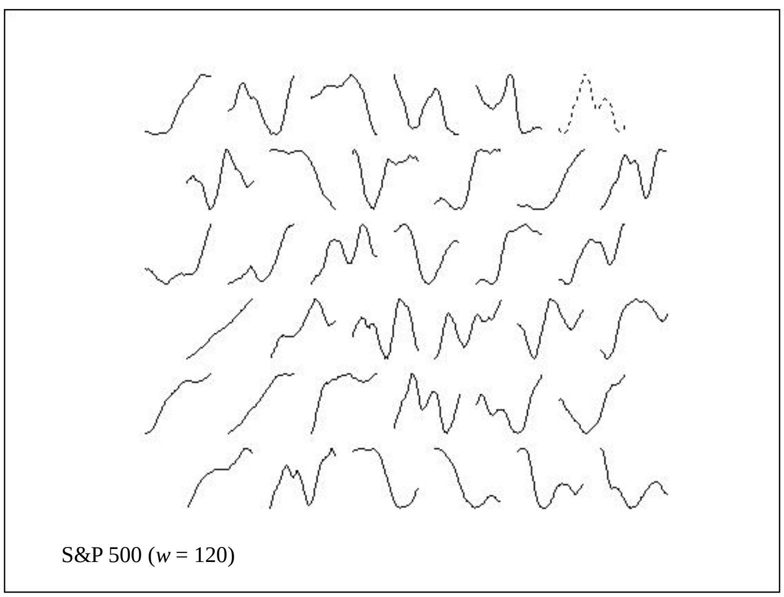

Figure 15: The Informative Patterns for S&P 500 ($w = 120$).