PROPOSAL FOR THE COMPOSITION OF A SOLVENT PORTFOLIO WITH CHAOS THEORY AND DATA ANALYSIS

Dimitris Karapistolis [1], Costas Siriopoulos [2], Iannis Papadimitriou [3], Raphael Markellos[4]

[1] Technological Educational Institute of Thessaloniki,
Thessaloniki GR 54101, Greece

[2] University of Macedonia, Dept. of Economics,
156 Egnatia str., P.O.Box 1591, Thessaloniki GR 540 06, Greece

[3] University of Macedonia, Dept. of Informatics,
156 Egnatia str., P.O.Box 1591, Thessaloniki GR 540 06, Greece

[4] Loughborough University, Dept. of Economics,
Loughborough, Leics LE11 3TU, UK

Abstract: This paper deals with the structure and dynamics of the Athens Stock Exchange (ASE) in Greece. Chaos Theory and Data Analysis methods are applied and produce evidence of a reasonably low-dimensional system, in both the phase and data space domain respectively. Based on the determined the concept of the solvent firm is identified and the solvent portfolio is constructed and traded according to a passive and active strategy. While the solvent portfolio returns of ASE for the period 1/1/1993 - 31/12/1993, clearly outperforms the market return, it is shown that it should be used as an investment tool, rather than for speculation.

Key words: Greek Stock Market, Chaos Theory, Data Analysis, Portfolio Management

1. Introduction

One of the fundamental issues of empirical finance in the study of stock markets is: given knowledge about the system and its past behaviour, what can be said about its future evolution. As shown in Figure 1. two basic approaches exist and may be classified into the econometric or model driven approach [1] and the non parametric or data driven approach.

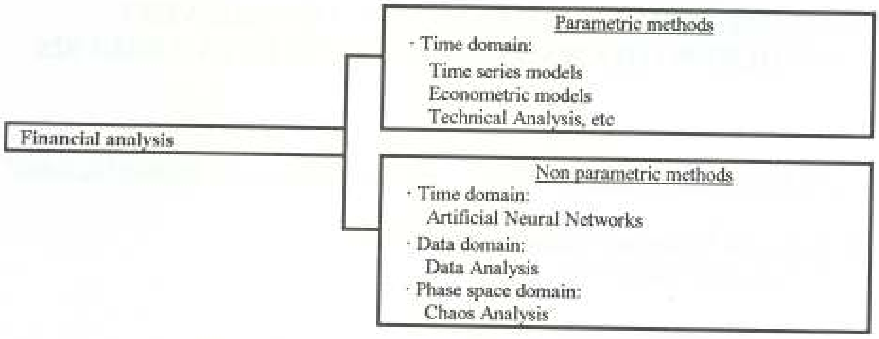

Figure 1: Approaches in Financial Analysis

The first approach attempts to analyse the sequence of observations produced by the underlying mechanism directly. From the statistics obtained from the observation sequence one hopes to be able to infer some knowledge about the future observation of the observation sequence. The strict statistical restrictions imposed by parametric model driven methods have often proven themselves to be unrealistic since properties such as noise, non stationary, nonlinearies and non-normality have been found to dominate stock market returns [2], [3].

The second approach postulates that no a priori assumption can be made about the structure of the stock market and the interaction of its components and that a data driven methodology should be adopted in order to estimate both interactions and components. Such methodology includes Data analysis, Artificial Neural Networks, Chaos analysis etc [21]. In this paper the perspective of interest is based on nonparametric methods of Chaos theory and Data Analysis. According to Chaos theory fluctuations are endogenous to the system and reflect the presence of important nonlinearities in the behavioural relationships of the system.

We applied chaos analysis to calculate Hurst's exponent in an attempt to determine whether our system is of long term memory. Next, we estimated the system's fractal dimension - which is a number that quantitatively describes how the system fills its space - and that lead us to the number of components least needed to represent the stock market [18]

Since we cannot determine those components through chaos analysis, we apply data analysis (DA) in the study of the system's behavioural structures and relationships and the isolation of critical properties by specifying its components. The isolated properties and relationships resulting from DA, are used by the critical construction of the "solvent portfolio" [4]. The solvent portfolio is based on the complex interaction of many qualitative and quantitative criteria it is a concept much wider than the efficient portfolio. The empirical findings of Data analysis indicate that it can be defined in terms of stock Corporate validity, Acceptability and Economic vigour. It is found that the ASE does not confirm to the strict assumptions made by traditional portfolio management and parametric modelling techniques, it of a low dimensional-system characterised by complex nonlinear regularities. The proposed 3-aspect solvent portfolio, produced by DA, is empirically justified and validated as a powerful investment tool and a satisfactory alternative to other methods [17]. Additionally, the solvent portfolio performance is consistent to Rossenberg's portfolio theory of extra market covariance, while its theoretical construction is based on an extension of Larrain's nonlinear model [8] It must be noted that the philosophy of Chaos and Data analysis is very similar: in Data analysis we reduce the degrees of freedom of a system described by a large number of variables, while in Chaos analysis we assess the degrees of freedom that are needed to reconstruct the system, but in a topological sense.

The paper is organised as follows: in the next section, a statistical description of the data is given. In section 3, the basic concepts of Chaos theory are introduced along with empirical evidence on the dimensionality and properties of the ASE returns. In section 4, Data analysis is applied in order to decrease the degrees of freedom, in a quantitative and qualitative data space of criteria concerning stocks listed in the ASE. The concept of the solvent portfolio and firm is then defined and determined via Data analysis. In section 5, the effectiveness of the proposed solvent portfolio is examined and in section 6, the major findings are presented along with a route for future research.

2. The data

The data analysed in section 3 consist of closing prices for the Athens Stock Exchange General Index (GIASE) for the period October 1986 to February 1994, a total of 1810 daily observations. To reduce the effects of non-stationarity and serial correlation, the raw prices are transformed to logarithmic returns. In figure 2 the descriptive statistics of the data are presented. The kurtosis and skewness measures indicate that the distribution of the GIASE returns has heavy tails and is skewed towards the right while the Jarque-Bera test strongly rejects the assumption of normality. The ARCH heteroskedasticity and McLeod - Li tests detect substantial non-linearities in the variance and mean respectively.

Average	0.0014
StDev	0.0221
Skewness	0.4
Kurtosis	17.6
Jarque Berra test	14315.5
ARCH test	176.4
McLeod-Li test	541.3

Figure 2: Descriptive statistics of the GIASE returns

The data used in section 4 of the paper, consists of 15 items for 240 stocks listed in the ASE for the year 1992. These items are: Company size, Stock market value, Capitalisation ratio, Financial position progress Marketability, Traded shares per day, Transaction value, Flow ratio, capital gain, dividend profits, Dept to equity, Equity to assets, Current ratio, P/E and Equity earnings. As shown in figure 3, these 15 aspects represent the 3 aspects of solvency: Corporate validity, Acceptability and Economic vigour.

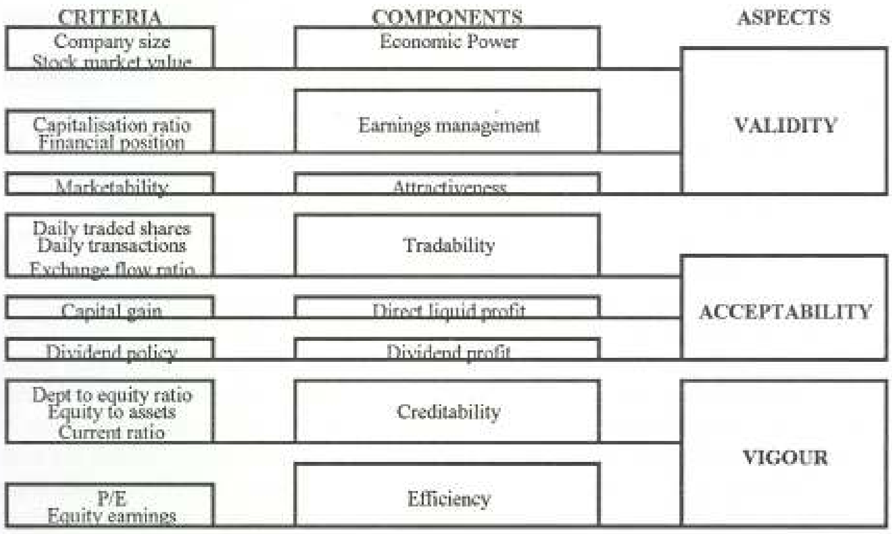

Figure 3: Aspects of corporate solvency

3. Chaos Analysis

Rather than assuming that the observation sequence may be considered as one specific realisation of a random process - where randomness arises from the many independent degrees of freedom interacting linearly - an emerging view exists in finance that postulates that apparently random behaviour may be generated in the long term by chaotic deterministic systems with only a few degrees of freedom that interact nonlinearly. A necessary, though not sufficient, condition for the occurrence of chaos is the presence of non-linearity. Chaotic systems are described by fractal dimensions and have strange attractors. In a nonlinear dynamic series, an attractor, is a definition of the equilibrium level of the system. Also, chaotic systems have some very interesting characteristics: due to their sensible "dependence on initial conditions", it is possible to make only very short-term predictions.

By plotting one variable - in our case stock returns - with different lags in different embedding dimensions (m), one can represent a system in the so-called phase space domain and treat it as a geometric object with invariant properties. A phase space is a graph that allows all possible states of a system. In this graph, the value of a variable is plotted against possible values of the other variables at the same time. Due to a theorem by Takens [5] one can fully reconstruct the original, unknown phase space with only one dynamic observal variable and obtain the attractor of the system, using the so called time delay method. Takens showed that a topologically equivalent picture of the attractor in phase space can be constructed by the time delay method, which consists of choosing a proper delay time and reconstructing a set of n-dimensional vectors, where n is non known a priori. The reconstructed phase space refers to the true dynamical system that generated the series and gives us information on the possibilities of the system.

Using this reconstructed phase space we can calculate the fractal dimension which measures how much m-dimensional space is occupied by and object. The most commonly used method to estimate the fractal dimension is the Grassberger Procaccia method [6], that uses the correlation dimension (CD). The Grassberger-Procaccia method offers a reliable, relatively simple method for estimating the fractal dimension when only one dynamical observal variable is known, as it is in our case. The CD measures the probability that two points chosen at random in phase space, will be within a certain distance of each other and examines how this probability changes as the distance is increased. The CD can be interpreted as a lower bound to the significant degrees of freedom of a dynamical system. Although only logarithmic returns are analysed it must be clear that according to the Takens theorem the estimated degrees of freedom refer to the stock market system as a whole and not to the return series alone. The CD has been also used to differentiate between deterministic, stochastic and chaotic systems. If chaos is present, a strange attractor can be identified that only occupies a small fraction of the available phase space. The computation of the CD allows us to find the dimension of this attractor. If the value of the CD does not change further with embeddings, it is assumed that the CD has converged to its correct value. That is, if chaos is present in the data the correlation dimension saturates with increasing embedding dimensions of the phase space. If this stabilisation does not occur, the system is considered high-dimensional or stochastic. The data are generated from some deterministic process when the correlation dimension remains well below the value of the embedding dimension and keeps increasing with increasing embedding without ever saturating. The estimated CD, for embedding dimensions from 2 to 10 and initial distance 0.0405 increased 10% each time, are presented in figure 4. The CD shows a strong saturating tendency for increasing number of dimensions in about 2,35. We can presume that the ASE dynamics are possibly chaotic and that at least 3 variables are needed to represent the system [17].

Embedding dimension	Correlation dimension
2	0.61917
3	0.95229
4	1.33443
5	1.73523
6	2.17384
7	2.33678*
8	2.44838
9	2.74870
10	2.21977

Figure 4: Correlation dimension of the GIASE returns

Another method of analysis in the phase space domain involves the calculation of the Largest Lyapunov exponents (LLE). These exponents show the average* exponential rates of divergence or convergence of nearby points in the phase space. The LLE can be interpreted as the average loss of predictive power in terms of bits of information. By applying the Wolf method [7], the average loss of information was estimated to be 0.0013, that is 0.0013 bits of information lost per day. So, the whole information set is lost in 1/0.0013 days which is about 37 months. A bit is a measure of information.

While various theoretical market models have been proposed by those who study chaos, their common point is that nonlinearities and determinism results from the interaction of long term fundamental and short term technical analysis or sentimental factors. Fundamental factors help to determine overall trends while technical factors help to determine near-term volatility. Such a model has been proposed by Larrain [8] and is called the K-Z map. Larrain shows that since fundamental factors (Z-map) alone cannot produce a true distribution of prices the nonlinear dynamics are produced by the effect of sentimental factors and short-term speculators (K-map) that use technical analysis. Erratic stock market behaviour occurs when the K-map overpowers the Z map. Larrain also argues that financial markets do not respond instantaneously to fundamental events as a result of the time and costs involved in acquiring, processing and interpreting information and the availability of such information.

From fractal geometry and topology we know that although a chaotic system is totally unpredictable in the long term it is bounded in a certain area of the attractor. In the case of the stock exchange these bounds are set by fundamental information and the structure of the market: speculation, sentimental and technical factors can move the price of a stock in a nonlinear fashion but this movement is restricted by the non-systematic fundamental boundaries of the stock, in the long term. An attractor exists for every stock and although we can assume that the structure of these attractors will be in general similar to the overall attractor of the market, their fundamental boundaries differ substantially.

The results of Chaos analysis indicate that predicting the ASE is essentially impossible in the long term. The evidence of nonlinearity and nonnormality of returns do not confirm to the strict statistical assumptions of the EMH and parametric modelling and to mean/variance portfolio optimisations. In the next section it is attempted to extract, discriminate and extrapolate the fundamental boundaries of each stock in order to exploit their differences in the determination of the solvent portfolio.

4. Data Analysis

Many studies that make use of Data analysis have been reported in the financial literature [19]. Recurrence plot analysis (RPA) [9] uses the same techniques as in Chaos analysis, particularly in the reconstruction of the phase space. With RPA we try to find the characteristics of a time series in terms of geometric criteria. The idea is to identify the similarities of the behaviour of points in time. Similarities between Chaos and Data analysis methods go beyond surface since they both are nonparametric methods and attempt to reduce the dimensionality of a system outside the time series domain.

The statistical properties of the returns justify the adoption of a nonparametric methodology such as Data analysis, since no a priori hypotheses are needed. The time series of returns has been found extremely difficult to model, thus one should focus on other fundamental qualitative and quantitative information about the stocks.

The concept of portfolio solvency receives special attention in this study. The solvent portfolio of firms is fundamentally different from the efficient portfolio but does confirm to Rossenberg's [10] portfolio theory. Rossenberg reformed Merkowitz's and Sharpe's ideas in an enriched and more applicable form. He introduced the concept of extra market covariance, which means that many stocks move together independently of what the market does as a whole. For example, stocks of companies in the same industry, stocks that are small in size or, as in our case, solvent stocks may move independently of the market. The term corporate solvency is defined as the competence of a firm listed in the stock exchange to fulfil its obligations in the long term. Thus solvency is strongly related to the reliability of a firm.

The selection of the fifteen criteria analysed in this section, was done in order to succeed the best possible representation of a firm's fundamental status and obtain the maximum quantity of information with minimum covariation of items. It is assumed that this data set contains sufficient information to define the boundaries of each stock. An additional assumption made is that the fundamental information contained in the criteria is not absolved immediately and that influences the market until it is drastically altered, which is consistent with the long term information decay of 37 months found by the LLE [17].

By applying methods of Data analysis the original 15 criteria are organised in three groups of five criteria each, as shown in figure 3.

The Validity of a firm is determined by the following 5 criteria: Company size, Stock market value, Capitalisation ratio, Financial position progress and Marketability. The first two criteria form the component of Economic power while the next two form the component of Earnings management. The last criteria form the attractiveness component of corporate validity. In order to decrease the large variations observed in the values of the above five criteria, they were divided in quartiles. As a result the initial quantitative criteria were transformed in qualitative binary values. The above transformation justified the use of Correspondence analysis [11] in the investigation of the data.

The Acceptability aspect is determined by the following 5 criteria: Traded shares per day, Transaction value, Exchange Flow ratio, Capital gain, and Dividend profits. The first three criteria form the component of Tradability while the Capital gain and the Dividend profits criteria form the Direct liquid profit and Dividend policy components respectively. Since the above 5 criteria concern ratios and we are interested in both their relations and factors, we apply Component analysis [12].

Finally, the aspect of Economic vigour is determined by the following 5 criteria: Dept to equity, Equity to assets, Current ratio, P/E and Equity earnings. The first three criteria form the Creditability component while the last two form the Efficiency component. Since Economic vigour is determined according to the classification of firms according to the above 5 criteria, the Range Analysis is used.

After performing the above analysis and classifying the stocks, it is possible to mark each one of the three aspects of a firm's solvency in a discrete scale of 1 to 5 according to each stock's integrity at the respective aspect. We then form a (240x3) table containing marks for each solvency aspect of 240 stocks. By summing the 3 marks of a firm we can obtain an overall measure which constitutes the solvency mark of the respective stock.

Mark	Group
[12,15]	Solvent portfolio
[9,11]	Potential alternatives
[3,8]	Uninteresting firms

Figure 5: Ranks of Corporate Solvency

As shown in figure 5, the 240 stocks can then be ranked in three groups, according to their solvency mark. Descriminant Analysis is then applied on the 240x3 table and after the possible recontructing emerging from analysis, the first group constitutes the solvent portfolio. The stocks that belong to the first group determine the solvent portfolio. The stocks ranked in the second group are acceptable in terms of solvency and can be used if needed for diversification. The stocks that form the third group are characterised for the time being as not interesting and will be evaluated in the next inflow of fundamental information. After applying the above analysis it was found that on 1/1/1993 the solvent portfolio of the ASE was constituted by nine stocks [20].

It is very interesting that the dimension estimated by Data analysis in the data domain we used is three, since Chaos analysis found that at least three components are required in order to model the Athens Stock Exchange in the phase space of the GIASE returns.

5. Evaluating the solvent portfolio *via* technical analysis indicators.

In this section we test the performance of the proposed solvent portfolio using an active strategy based on a mechanical trading rule, and a passive buy and hold strategy. The use of mechanical trading rules has been recently proposed by different authors in testing market efficiency or portfolio returns. We use the Stochastic Momentum Index (SMI) developed by Blau [13], since it is better adopted by ASE structures and magnitudes formulated as follows:

$$SMI\,(q,r,s) = 100\,E(s)\,\{\,E(r)\,[SM(q)]\,\}\,/\,0.5\,\{\,E(s)\,[E(r)\,[HH{:}\,q - LL{:}q]\,]\,\}$$

where:
$SM(q)$ = closing price - 0.5 (HH: q + LL:q)
$E(x)$ is an exponential x-day moving average
HH:q is the highest high value over q periods
LL: q is the lowest low value over q periods
and $ASM(q) = E(s)\,\{E(r)\,[SM(q)\,]\,\}$ is the exponential moving average of period s of the exponential moving average of period r of the quantity $SM(q), s>r$. That is the average stochastic momentum, i.e. a double smoothing of stochastic momentum which is a known technical indicator. In our case q is chosen to be equal to 5 (a week), and after optimisation with MetaStock v.5.1, $r = 5$ and $s = 20$. We would choose other values for the parameters q,s and r but the results will be quite similar.
This technical indicator is based on the location of the closing price between the highest high and the lowest low value over q periods (days, weeks,...). The idea is that averaging this formula produces a relatively smooth indicator with a fast response giving buy and sell signals (trades). The SMI index is one of the followed technical indicators by portfolio managers for active trading (speculation) in stock markets. Other technical indicators are discussed in [14] and [16].

We take a long (L) and a short (S)position, respectively, if:
L: Close(t-1) < SMI(t-1) and Close(t) > SMI(t)
S: Close(t-1) > SMI(t-1) and Close(t) < SMI(t)
Where close(t) denotes the closing price at time t.

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	Total
LT	10	8	8	7	10	11	8	7	10	
ST	11	9	8	8	11	11	7	7	10	
%LT	60	63	75	57	20	28	63	57	40	
%ST	28	11	25	25	37	37	43	43	30	
%Total	80	75	46	35	38	7	72	25	2	47
%B&H of the solvent portfolio	220	210	120	70	-5	-6	110	70	30	89
%Market										40

Si , I=1,...,9 are the nine solvent stocks, LT: # of long trades, ST: # of short trades, and %LT, %ST: profitable long and short trades respectively.

Figure 6: Solvent portfolio returns for the period 1/1/93 - 31/12/93.

The results of this strategy are shown in figure 6 for the equally weighted (11.1%) solvent portfolio of 9 stocks. These results are lower than a simple buy and hold strategy (47% vs. 89%), but greater than the return of the market portfolio (40%), for the same period. It is obvious that the proposed Solvent portfolio serves as an long term investment tool rather than for speculation since the passive buy and hold strategy clearly outperforms the speculative strategy, using technical analysis tools such as the SMI trading indicator. This is true, since the Return of the Solvent Portfolio (SP%) over the one year period is greater than the Return of a speculative strategy (SMI%) and even greater than the Return of the Market (M%): SP% > SMI% > M%.

6. Conclusions

In this paper two distinct nonparametric approaches were adopted in the study of the ASE in Greece. The statistical properties of the GIASE index returns for the period October 1986 to February 1994, were found to violate the assumptions made by traditional parametric modelling and portfolio management.

The application of Chaos analysis methods produced evidence that the ASE dynamics are not purely stochastic and can be modelled with at least three variables. Based on the 37-month average decay of information period found by Chaos analysis and on a model by Larrain Data analysis is applied on 15 criteria derived from fundamental analysis, concerning 240 stocks listed in the ASE. The results of Data analysis confirm to those of Chaos analysis showing that three significant factors can be extracted from the data. These factors are translatable in economic terms and can be used in the determination of a portfolio of stocks, that is defined as solvent. The estimated solvent portfolio return of the ASE for the period 1/1/1993 to 31/12/1993 outperformed the market return. It was also found that the it performed better when used as a long term investment tool rather than for speculation.

Future research is concentrating on the technology of Artificial Neural Networks (ANN's) and Expert systems [14] where the ratings of solvency aspects can be used as input variables. Solvency combines a wide set of information in only three aspects or measures and can provide ANN's a condensed set of three inputs that contain the maximum possible information. This is essential to the training process of ANN's since a small number of input variables results in acceleration of the training process and better generalisation without limiting the width of information to be considered. The basic problem of finding the optimal architecture of an Artificial Neural Network can be solved based on the findings of Chaos and Data analysis [15]. The estimated dimension of three can serve as a lower bound to the number of distinct features (nodes in the so-called hidden layer) that the network must recognise. Previous research has shown empirically that the ASE is best described by ANN's with 3 nodes in the hidden layer [16].

References

[1] Mills TC, *The Econometric Modelling of Financial Time Series*, Cambridge University Press, 1993.

[2] Hsieh DA, Chaos and Nonlinear Dynamics: Applications to Financial Markets, Journal of Finance, 1991;5:pp. 1839-1877.

[3] Peters, EE, *Fractal Market Analysis*, New York: John Wiley, 1994.

[4] Karapistolis D, Papadimitriou I, Construction of a Solvent Portfolio. Proceedings of 4th International Conference of the Thessaloniki Economic Society, 1994 : Thessaloniki, Greece, Vassiliadis S (ed), University Studio Press, 1995: 269-296.

[5] Takens F, "*Detecting Strange Attractors in turbulence.*" In Dynamical Systems and Turbulence, Rand D, and Young LS, eds. Berlin: Spinger-Verlag, 1980.

[6] Grassberger P, Procaccia I, Characterisation of strange attractors, Physics Review Letters, 1983;50: 3465-3490.

[7] Wolf A, Swift JB, Swinney HL, Vastano J, Determining Lyapunov Exponents From a Time Series. Physica D, 1985;16: 285-317.

[8] Larrain M, Testing Chaos and Non-Linearities in T-Bill Rates. Financial Analysts Journal, 1991;Sept.-Oct: 51-62.

[9] Eckmann JP, Kamphrost O, Ruelle D, Recurrence plot of dynamical systems. Europhysics Letters, 1986: 973-977.

[10] Deboeck GL,"The Impact of Technology on Financial Markets" in *Trading on the Edge*, ed. Deboeck, GL, 1994.

[11] Benzecri JP, *Correspondence Analysis Handbook*, New York: Marcel Dekker, 1985.

[12] Lebart L, Morineau A, Fenelon JP, *Traitement des donees Statistiques*, Dunod, Paris, 1979.

[13] Blau W, Stochastic Momentum. Technical Analysis of Stocks and Commodities, 1990;11 (1):26-32.

[14] Siriopoulos C, Doukidis G, Karakoulos G, Perantonis S, Varoufakis S, Applications of Neural Networks and Knowledge Based Systems is Stock Investment Management: A comparison of performances. Journal of Neural Network World, 1992; 6:785-795.

[15] Markellos RN, Siriopoulos C, Sirlantzis K, Testing Non- linearities and Chaos in Emerging Stock Markets: Implications for Financial Management. Proceedings (forthcoming) of the 4th Annual Meeting of the European Financial Management Association, June 1992, London, UK.

[16] Siriopoulos C, Markellos RN, Sirlantzis K, Applications of Artificial Neural Networks in Emerging Financial Markets. Proceedings of the 3rd International Conference of Neural Networks in Capital Markets, (ed) World Scientific Publishing (forthcoming), October 1995; London, UK.

[17] Sirlantzis C, Siriopoulos C, Deterministic chaos in stock market: Empirical Results from monthly returns. Journal of Neutral Network World, 1993; vol. 3, 6: 855-864.

[18] Brock W, Sayers C, Is the business cycle characterised by deterministic chaos? Journal of Monetary Economics, 1988; 22: 71-90.

[19] Benzecri JP, "L'analyse des donnees. Tome 1: La Taxinomie. Tome 2: L'analyse des correspondances, Dunod, Paris, 1973.

[20] Karapistolis D, Katos A, Papadimitriou I, Proposal for solvent portfolio selection with data analysis methods, Proceedings of the 7th National Conference of the Greek Statistical Institute, 1994 May 5-7: Nicosia. Cyprus,(ed) Hellenic Statistical Institute 1995: 90-98.

[21] Siriopoulos C, Markellos M, Semi-chaotic financial time series and neural network forecasting: evidence from the European stock markets. Yugoslav Journal of Operations Research (forthcoming), vol. 6, 2.