Regime Switching and Artificial Neural Network Forecasting of the Cyprus Stock Exchange Daily Returns

by

Eleni Constantinou[1][*], Robert Georgiades[1], Avo Kazandjian[2] and Georgios P. Kouretas[3][**]

June 2005

Abstract

This paper provides an analysis of regime switching in volatility and out-of-sample forecasting of the Cyprus Stock Exchange using daily data for the period 1996-2002. We first model volatility regime switching within a univariate Markov-Switching framework. Modelling stock returns within this context can be motivated by the fact that the change in regime should be considered as a random event and not predictable. The results show that linearity is rejected in favour of a MS specification, which forms statistically an adequate representation of the data. Two regimes are implied by the model; the high volatility regime and the low volatility one and they provide quite accurately the state of volatility associated with the presence of a rational bubble in the capital market of Cyprus. Another implication is that there is evidence of regime clustering. We then provide out-of-sample forecasts of the CSE daily returns using two competing non-linear models, the univariate Markov Switching model and the Artificial Neural Network Model. The comparison of the out-of-sample forecasts is done on the basis of forecast accuracy, using the Diebold and Mariano (1995) test and forecast encompassing, using the Clements and Hendry (1998) test. The results suggest that both non-linear models equivalent in forecasting accuracy and forecasting encompassing and therefore on forecasting performance.

Keywords: Regime switching, artificial neural networks, stock returns, forecast encompassing, forecasting accuracy, nonlinearity.

JEL Classification: G12

[1] Department of Accounting and Finance, The Philips College, 4-6 Lamias Street, CY-2100, Nicosia, Cyprus.
[2] Department of Business Studies, The Philips College, 4-6 Lamias Street, CY-2100, Nicosia, Cyprus.
[3] Department of Economics, University of Crete, University Campus, GR-74100, Rethymno, Greece.

1. Introduction

This paper addresses two important issues with respect to the behaviour of stock returns of a recently established emerging capital market the Cyprus Stock Exchange using daily data for the period 1996-2002. First, we explore the issue of volatility regime switching in the stock returns of the capital market of Cyprus and second, we compare the forecasting performance of two competing non-linear models the Markov-Switching (MS) model and the Artificial Neural Network (ANN) model.

The Cyprus Stock Exchange is the primary stock market in Cyprus. It is considered to be a small emerging capital market with a very short history since it was established in April 1993 when the inaugural Stock Exchange Law passed through the Cypriot House of Representatives. In July 1995 the Cypriot House of Representatives passed the laws for the stock exchange function and supervision, while additional laws led to the establishment of the Central Securities Depository. On 29 March 1996 the first day of transactions took place. The Cyprus Stock Exchange S.A. is supervised by the Ministry of Finance and the Minister of Finance is responsible for choosing the seven member executive committee that runs CSE. Furthermore, the Securities and Exchange Committee is mostly responsible for the well functioning of the capital market of Cyprus. Trading takes place electronically through the Automated Trade System. The main index is the CSE General Price Index that reflects, approximately, 93% of the trading activity and 96% of the overall capitalization. In November 2000 the FTSE/CySE 20 was constructed with the cooperation of CSE, the Financial Times and the London Stock Exchange in order to monitor closer the market. To highlight the increasing need for regional capital market integration the FTSE Med 100 was created in June 2003 with the cooperation

of CSE, ASE and the Tel-Aviv Stock Exchange. Figure 1shows the evolution of the CSE general price index and that of its returns.

During the last fifteen years there has been a growing interest among portfolio managers for the emerging capital markets as they provide opportunities for higher asset returns compared to those of the developed markets. However, in the aftermath of the financial crisis in Southeast Asia, Latin America and Russia in 1997-1998 it was also made clear that there is a need to understand and model these markets in a more coherent way. Modelling the stock returns volatility is of paramanount importance especially in the case that the stock market has undergone different structural changes.

In the present analysis, first we focus on nonlinear features in stock returns volatility of the CSE in the presence of regime switching. The motivation to model stock returns in Cyprus within a nonlinear framework lies on the fact that during the period under examination we observed a transition from low volatility regime to a high volatility regime due to the presence of the rational bubble in 1999-2000. We model nonlinearities in the Cyprus Stock Exchange by adopting regime switching models to analyze regime switching in the volatility of stock returns. Regime switching models make it possible to take into account the shifts between high and low volatility regimes and correlations due to changes in the economic and financial environment. In addition, regime switching models result to a reduction of the persistence in variance and therefore the problems of underestimation of the volatility in the high volatility regime or the overestimation of the volatility in the low volatility regime are minimized.

Evidence of regime switching in the volatility of stock returns have been found by Hamilton and Susmel (1994), Hamilton and Lin (1996), Edwards and

Susmel (2001), Coe (2002) and Kanas (2005). In addition, Engel (1994), Engel and Hamilton (1990), Engle and Kim (2001) and Kanas and Kouretas (2003) are examples of studies that they have estimated Markov-Switching regime models for currency markets. We incorporate the issue of univariate volatility regime switching to the stock returns of CSE. If the existence of different regimes is found then we are legitimate to search for behavioral differences across regimes. This would be in line with the Lucas (1976) critique that agents react differently across regimes. This is important for the Cypriot capital market, as it underwent several structural changes over the period under consideration, which may have caused volatility regime switching. Thus, the behaviour of market participants, according to the Lucas critique, can alter as a function of volatility. This is in line also with several studies which found that the volatility is a statistically significant factor affecting the conditional mean of the exchange rate, within a GARCH-in-mean type of model.

In order to study the dynamics of the regime switching and the stochastic processes evolved in the stock market of Cyprus, we adopt the univariate Markov Switching Model (MS) developed by Hamilton (1989). This model allows, in a univariate context, for shifts in the stochastic volatility regime driving the stock market. Thus, the change in regime should be considered as a random event and not predictable. In addition, the effect of these shifts must be considered when we investigate the stochastic properties of the stock market volatility.

The second stage of our analysis deals with the issue of forecasting stock returns of the CSE. To this end we use the estimated non-linear MS model of the first stage and a competing non-linear ANN model. The ANN methodology is preferred to alternative non-linear models because it is nonparametric, and thus appropriate in our case since we are not considering a specific nonlinear functional form between stock

3

prices and their lagged values. The input layer of the ANN contains one input variable, namely the lagged stock index returns. Then, from the estimation of the two empirical models we obtained 1-step ahead out-of-sample forecasts for the out-of-sample period we compare the forecasting performance of the two models by testing for forecasting accuracy using the Diebold and Mariano (1995) test and for forecast encompassing using the Clements and Hendry (1998) approach.

The main findings of the paper are summarized as follows. First, there is strong evidence in favour of volatility regime switching modelling of nonlinearities which affects the stock returns. Second, the estimation of the MS model accurately describes the two regimes based on the different pattern of adjustment of the stock returns due to the presence of a rational bubble; and the estimated model captures all the events that are responsible for the presence of nonlinear features in the CSE stock returns due to the presence of a rational bubble during the period 1999-2000. Third, there is a high probability for regime clustering with the likelihood that a low volatility regime will be followed by a low volatility regime greater that the likelihood a high volatility regime will be followed by a high volatility regime. Finally, the comparison of the out-of-sample short-run forecasts generated by the two models suggest that on the basis of the forecast accuracy and the forecast encompassing criteria, we conclude that the MS and the ANN models are equivalent in terms of their forecasting performance.

The rest of the paper is organized as follows. Section 2 presents the econometric methodology. Section 3 presents and discusses the empirical results. Section 4 provides the summary and the concluding remarks.

2. Econometric methodology

2.1. The Markov Switching Model

The Markov Switching Model (MS), is a single equation model developed by Hamilton (1989). The advantage of using a Markov chain to model the random discontinuous shift is that it allows conditional information to be used in the forecasting process. This allows us: (i) to model and explain time series dynamics; (ii) to demonstrate the presence of regime persistence (this is the well-known cluster effect, i.e. high volatility is usually followed by high volatility) and (iii) to provide better forecasts since switching regime models generate a time conditional forecasted distribution instead of an unconditional forecasted distribution.

In our case, the CSE general index stock returns are modeled on the basis of the following model MS model:

$$R_{it} = a_0(1 - S_t) + \sum_{i=1}^{4} a_i(S_t)R_{t-i} + [\sigma S_t]\varepsilon_{it} \tag{1}$$

where ε_{it} is the stochastic term distributed as an iid Gaussian variable. Equation (1) is a Markov two-regime switching process.[1] The random and unobservable variable S_t follows a first-order Markov chain, namely:

$$\Pr(S_t = 1 / S_{t-1} = 1) = q \tag{2a}$$

$$\Pr(S_t = 2 / S_{t-1} = 1) = 1 - q \tag{2b}$$

$$\Pr(S_t = 2 / S_{t-1} = 2) = p \tag{2c}$$

$$\Pr(S_t = 1 / S_{t-1} = 2) = 1 - p \tag{2d}$$

---

[1] A comprehensive discussion of the theory of Markov chains with application to Markov-switching models is given by Hamilton (1994).

Equations (2a) - (2d) indicate that the random variable S_t can be described as a two-regime Markov chain with transition probabilities p and q. The two regimes are $S_t = 1$ and $S_t = 2$. As shown in equation (1), regime switching is allowed both in the mean and variance. Statistical inference regarding the empirical validity of the two-regime switching process is carried out using nonstandard Likelihood Ratio (LR) tests [Davies (1987)]. The null hypothesis is that of no regime switching in the mean and variance. The alternative hypothesis of two-regime switching, either in mean ($\alpha_0 \neq \alpha_1$) or in variance ($\sigma_0^2 \neq \sigma_1^2$) or both in mean and variance, corresponds to equation (1). To explore the full richness of the MS modeling approach, we allow for regime-dependent autoregressive parameters.[2]

The parameter vector is estimated by maximum likelihood methods [Hamilton (1994)]. If the regime is assumed to be known, the sample likelihood function conditional on the given state S_t is:

$$f(\hat{e} / S_t; \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{\hat{\varepsilon}_{S_t}^{\ 2}}{2\sigma^2}] \qquad (3)$$

Maximum likelihood estimation of the model in (1) is based on the Expectation Maximisation (EM) algorithm discussed in Hamilton (1989), and Krolzig (1997).[3] As a byproduct of the maximum likelihood estimation, one can calculate the

---

[2] Hamilton (1996) discusses alternative specification tests in Markov-switching models.
[3] The EM algorithm was introduced by Dempser et al. (1977) for a class of models where the observed time series depend on the unobservable stochastic variable – such as the variable S_t. Each iteration of the EM algorithm involves two steps. The first step (the expectation step) involves the estimation of the unobserved states S_t by determining their smoothed probabilities $\Pr(S_t | \Delta Y_t)$. The second step, known as the maximisation step, involves estimating the parameter vector λ as a solution of the first-order conditions from maximising the likelihood function where the unknown conditional regime probabilities are replaced with the smoothed probabilities are replaced with the smoothed probabilities obtained from the expectation step. Having obtained the parameter λ, the filtered and the smoothed

unconditional probability that the system of the two currency is in regime i, $i = 1, \ldots m$, at any given date, $\Pr(s_t = i)$.[4] Also, a set of conditional probabilities can be obtained, namely the 'smoothed' probabilities, representing the ex-post inference about the system being in regime i at date t. Further, one could date the regime switches. For instance, for 2 regimes, an observation is assigned to the first regime if $Pr(s_t = 1 \mid \Delta Y_T) > 0.5$, and to the second regime if $Pr(s_t = 1 \mid \Delta Y_T) < 0.5$. Lastly, the average duration of each regime, d_i, is calculated from the formula $d_i = (1 - p_{ii})^{-1}$. The maximum likelihood estimates of the transition probabilities q and p satisfy

$$ q = \frac{\sum_{t=2}^{T} \Pr(S_t = 0, S_{t-1} = 0 / \hat{e}_t)}{\sum_{t=2}^{T} \Pr(S_{t-1} = 0 / \hat{e}_t)} \tag{4} $$

$$ p = \frac{\sum_{t=2}^{T} \Pr(S_t = 1, S_{t-1} = 1 / \hat{e}_t)}{\sum_{t=2}^{T} \Pr(S_{t-1} = 1 / \hat{e}_t)} \tag{5} $$

The usefulness of our estimated MS model will be further made clear with the evaluation of its forecasting performance. Given that the Markov switching regime model is a non-linear model its forecasting performance will be compared with that of another non-linear model.

2.2. Artificial Neural Network Model

An alternative nonparametric nonlinear approach for out-of-sample forecasting is the Artificial Neural Network (ANN) estimation. ANNs have been extensively used in out-of-sample forecasting of financial variables both in univariate

---

probabilities are updated and so on, guaranteeing an increase in the value of the likelihood function at each step.

[4] The formula is $\Pr(s_t = 1) = (1 - p_{22})/(2 - p_{11} - p_{22})$

and multivariate approaches [Swanson and White (1995)]. The specific type of ANN employed in this study is the Multilayer Perceptron (MLP).[5]

Following White (1992), Swanson and White (1995), Kanas (2001) and Kanas and Yannopoulos (2001) the architecture of the MLPs trained includes 1 hidden layer and 8 hidden units. Such an MLP is denoted as MLP(1, 8). The output variable is the stock returns series, and the input variable is the lagged stock returns (denoted by X_1).[6] We also set a second input variable X_2 which takes the value of 1, and plays the role of a constant in a regression setting. Given that there is no reliable method of specifying the optimal number of hidden layers, we specified one hidden layer following White's (1992) conclusion that single hidden layer MLPs do possess the universal approximation property, namely they can approximate any nonlinear function to an arbitrary degree of accuracy with a suitable number of hidden units.[7]

The graphical representation of the MLP(1, 8)'s architecture employed is given in Figure 3. The mathematical expression for the MLP(1,8) drawn in Figure 3 is given by equation (6), where the subscripts t from the output and input variables are suppressed to ease the exposition. Thus:

$$y = \sum_j a_j x_j + \sum_i b_i f(\sum_j c_{ji} x_j), \; \text{j=1, 2, and i} = 1,2,\ldots 8 \tag{6}$$

where $f(.)$ is the activation logistic cumulative distribution function[8], a_j are the weights for the direct signals from each of the two input variables to the output variable, b_i is the weight for the signal from each of the eight hidden units to the

---

[5] For a discussion of MLPs, see Campbell, Lo and MacKinlay (1998).
[6] Kanas and Yannopoulos (2001) point out that autoregressive models for stock index return are justified on the grounds that autocorrelation in index returns can appear due to nonsychronous trading.
[7] Moreover, as mentioned in Adya and Collopy (1998), 67% of the studies which carried out sensitivity analysis to determine the optimal number of hidden layers found that 1 hidden layer is generally effective in capturing nonlinear structures.
[8] The algebraic expression of this function is: $f(u) = 1/[(1+\exp(-u)]$.

output variable, and $c_{j,i}$ are the weights for the signals from each of the two input variables to the hidden units. The logistic function is used as the activation function in order to enhance the nonlinearity of the model.

The network interpretation of Figure 2 and equation (6) is as follows. Input variables, X_1 and X_2, send signals to each of the hidden units. The signal from the j-th input unit to the i-th hidden unit is weighted by some weight denoted by $c_{j,i}$ before it reaches the hidden unit number i. All signals arriving at the hidden units are first summed and then converted to a hidden unit activation by the operation of the hidden unit activation function $f(.)$. The next layer operates similarly with connections sent over to the output variable. As before, these signals are amplified by weights b_i and summed. Finally, signals are transmitted directly from the input variables to the output variable with weight a_j. The method of estimation adopted is nonlinear least squares and we were successful in obtaining convergence for the CSE general stock returns. The MLP reflected in Figure 2 is trained over the in-sample period (known as the 'training period'). Following Kanas and Yannopoulos (2001) in order to improve the in-sample-performance of the MLP, we take the estimated set of weights as the initial values for further training while at the same time we adopt stricter criteria for convergence.[9] To achieve this end we consider the following strategy: The ANN model is initially trained on a subset of the in-sample data 29 March 1996 to 31 December 2000. The estimated model is then used to generate forecasts for the remaining in-sample-period 1 January 2001 to 31 December 2001. We then compare the sample of original observations with the forecasted values calculating the root mean squared error (RMSE). This strategy is repeated until the RMSE reaches a minimum and then starts increasing which may imply model overfitting. Therefore,

---

[9] Kavalieris (1989) provides a thorough discussion for model selection procedures.

this training strategy was stopped at the *n*-th training if the RMSE of the (*n*+1)-th training was higher than the RMSE of the *n*-th training. We then used the rollover estimations to generate the 1-step ahead forecast for stock returns for the out-of-sample period.

2.3. Out-of-sample forecasting criteria

Having estimated the two empirical models discussed above and obtained 1-step ahead out-of-sample forecasts for the out-of-sample period, we proceed to evaluating their relative forecast performance. We compare the out-of-sample forecasts using two different testing approaches. First, we examine forecast accuracy using the Diebold and Mariano (1995) procedure. We calculate the Root Mean Squared Error (RMSE) of forecasts from the two models, and then use the Diebold-Mariano (DM) statistic to test whether the difference between the two RMSEs.[10] For the DM test, the null hypothesis is that the difference between the RMSEs from two models is zero.

We next employ the forecast encompassing approach.[11] Following Donaldson and Kamstra (1997), and Clements and Hendry (1998), the formal forecast encompassing test is based on a set of OLS regressions. To illustrate, consider two competing models, namely M_1 and M_2. E_i denotes the forecast error for model M_i (i =1,2), and D denote the difference between the forecasts from the two models. Given forecasts from the two models, we can test the null hypothesis that neither model

---

[10] The DM test statistic is defined as $DM = \bar{d}/\sqrt{\omega}$, where $\bar{d}$ is the average loss differential, $\bar{d} = (1/m)\sum_{j=1}^{m} d_j$. d_j is the difference between the squared forecast errors of two competing models. For 1-step-ahead forecasts, ω is estimated using the variance of d_j. The DM statistic is asymptotically distributed as N(0,1).

[11] The forecast encompassing test has an easily derivable distribution when applied to out-of-sample data [Donaldson and Kamstra (1997)].

encompasses the other by running two regressions: the first involves regressing the forecast error from M_1 on the difference of forecasts, i.e.

$$E_{1,t} = b_0 + b_1 D_t + w_t \tag{7}$$

and estimating the coefficient b_1. The second involves regressing the forecast error from M_2 on the difference of forecasts, i.e.:

$$E_{2,t} = d_0 + d_1 D_t + u_t \tag{8}$$

and estimating the coefficient d_1. If $\hat{b}_1$ is not statistically significant and $\hat{d}_1$ is statistically significant, then we reject the null hypothesis that neither model encompasses the other in favour of the alternative hypothesis that M_1 encompasses M_2. If $\hat{b}_1$ is significant and $\hat{d}_1$ is not significant, then M_2 encompasses M_1. If both $\hat{b}_1$ and $\hat{d}_1$ are significant or if both are not significant, then we fail to reject the null hypothesis that neither model encompasses the other.[12]

3. Empirical results

The data consist of daily observations of the CSE general price index and the sample period is 29 March 1996 (first day of trading at CSE) until 19 April 2002, excluding all weekends as well as holidays giving a total of 1444 observations. The data for the general price index were taken from the CSE database. The series are taken in natural logarithms. For both the MS and the ANN models, the in-sample

---

[12] Clemen (1989) and Donaldson and Kamstra (1996) provide a detailed analysis of combining forecasts.

period was 29 March 1996 – 31 December 2001. The out-of-sample period was taken to be 1 January 2002 – 19 April 2002.

Given that the stock price index is a nonstationary variable we only consider the first differences of the general price index:

$$\Delta p_t = 100*(p_t - p_{t-1})$$

which corresponds to the approximate percentage nominal return on the stock price series obtained from time t to $t\text{-}1$.[13]

To address the issue of volatility regime switching and to discriminate between low and high volatility regime in the CSE general price index, we estimate and test for an MS given by (1). Table 1 reports the estimated coefficients of the proposed MS along with the necessary test statistics for evaluation of the adequacy of the estimated model. The Likelihood Ratio test for the null hypothesis of linearity is statistically significant and this suggests that linearity is strongly rejected. This is a nonstandard LR test due to Davies (1987). This outcome is reinforced from the AIC and HQIC criteria. Therefore, we can conclude that the returns of the CSE general price index are characterized by two stochastic regimes in the mean and in the variance of the distribution of the returns.

Table 1 also reports several other diagnostics which further highlight the use of regime switching induced nonlinearities in the relationship of the two exchange rates. First, the standard deviation σ takes the values of 0.014 and 0.052 for regime 1 and regime 2 respectively; these values help us to identify regime 1 as the low volatility regime and regime 2 as the high volatility regime. Second, the duration measure shows that the low volatility regime lasts approximately 60 weeks, whereas

---

[13] We have applied the standard ADF, the Elliott et al. (1996) and Elliott (1999) GLS-ADF and the Ng and Perron (2001) GLS versions of the modified Phillips-Perron unit root tests as well as the Kwiatkowski et al. (1992) KPSS stationarity test and we have found conclusive evidence that both exchange rate series are I(1) processes. The results are available upon request.

the high volatility regime lasts approximately 105 months. This outcome is expected, given that during the period under examination we had the emergence of a rational bubble which had long lasting effects. Third, the calculated unconditional probability shows that there is a probability of 33 percent that a low volatility regime occurs and 67 percent that a high volatility regime takes place. Furthermore, the transition probabilities p_{11} and p_{22} explain the possibility of regime clustering and it is shown that there is 99.05 percent probability that a high volatility regime will be followed by a high volatility regime, while there is a 98.32 percent probability that a low volatility regime will be followed by a low volatility regime. This implies that both regimes exhibit high persistence and this outcome leads to the argument that once the CSE finds itself in one regime there is a great likelihood to remain in this regime for the rest time period.

The estimation of MS model using all the available data allows us to make inference of being in one of the two volatility regimes for each week of the sample. Figure 3 shows the resulting smoothed probabilities of being in the high and the low volatility regimes along with the evolution of the CSE general price index. As shown in this figure, several weeks of the sample are characterized by high volatility and this may be attributed mainly to the existence of the rational bubble during the period under investigation. Table 2 provides a full account of the data falling in the two regimes, regime 1 (low volatility) and regime 2 (high volatility). It is very interesting to note that the estimated MS model provides an accurate account of the distinction between regimes of high volatility and low volatility. Specifically, the low volatility regime lasts from July 1996 to January 1999. Since mid January 1999 we moved to the high volatility regime which lasts until the end of the sample. This regime switching identifies to a great extent the period during which a rational bubble

characterized the behaviour of the general price index of CSE. Thus, the period when the rational bubble was present corresponds to the high volatility regime and this phenomenon is captured extremely well from our model.

Another important result derived from our estimations is the statistical significance of the autoregressive parameters from one regime to the other. As, it is clear from Table 1 during the low volatility regime none of the autoregressive coefficients is statistically significant at the 5% level of significance and thus we can argue that we can not reject the weak form of market efficiency hypothesis for the stock market of Cyprus. In contrast, during the high volatility regime the coefficients with one and two periods lag are statistically significant. This implies that during this period that coincides with the presence of the rational bubble the weak form of the market efficient hypothesis is rejected and therefore we can argue that during that period the capital market of Cyprus was not efficient.

The estimated coefficients from training the ANN model for the training period (29 March 1996 to 31 December 2001) are reported in Table 3. Along with the estimates of the weights a_j, b_i and $c_{j,i}$ we also report the RMSE of the generated forecasts for the sub-sample of the in-sample period.

Tables 4 and 5 reports the results from the application of the forecasting criteria discussed in Section 2. Specifically, Table 4 shows that the ANN technique has the smallest RMSE (0.15291), with the Markov regime switching model coming second (RMSE = 0.15610). The DM statistic indicates that the null hypothesis cannot be rejected. Thus, we conclude that the out-of-sample forecasts obtained from the MS model and the ANN models are not statistically significant different in terms of forecast accuracy.

Table 5 reports the results from the forecast encompassing tests. We report the heteroscedasticity-robust t-statistics of the estimated coefficients $\hat{b}_1$ and $\hat{d}_1$ from regressions (7) and (8) and the corresponding p-values. As shown in Table 5, the ANN model (M_2) does not encompass the MS model (M_1), and the MS model does not encompass the ANN Model, since both regression coefficients $\hat{b}_1$ and $\hat{d}_1$ are statistically insignificant. Therefore, on the basis of the forecast accuracy and the forecast encompassing criteria, we could conclude that the MS and the ANN models are equivalent in terms of their forecasting performance.

4. Summary and Concluding Remarks

The purpose of this paper is twofold. First, it aims to model volatility regime switching for the stock returns of a recently established emerging capital market, the Cyprus Stock Exchange using daily data for the period 1996-2002. To conduct our study, we adopt the univariate MS developed by Hamilton (1989). There are several important findings that stem from the present analysis. First, there is strong evidence in favour of regime switching modelling of nonlinearities in the stock returns of the CSE. Second, the estimation of the MS accurately describes the two regimes based on the different pattern of adjustment of the stock returns; and the estimated model captures all the events that are responsible for the presence of nonlinear features in the CSE stock returns due to the presence of a rational bubble during the period 1999-2000. Third, there is a high probability for regime clustering with the likelihood that a low volatility regime will be followed by a low volatility regime greater that the likelihood a high volatility regime will be followed by a high volatility regime. Fourth, we consider two competing non-linear models to conduct forecasting of the stock returns. These models are the estimated MS univariate model and the Artificial

Neural Network model. By obtaining the 1-step ahead forecast for stock returns for the out-of-sample period, we compare the out-of-sample performance of the two competing models on the basis of forecasting accuracy by applying the Diebold and Mariano (1995) test and forecast encompassing using the Clements and Hendry approach. The results suggest that on the basis of the forecast accuracy and the forecast encompassing criteria, we could conclude that the MS and the ANN models are equivalent in terms of their forecasting performance.

References

Adya, M. and F. Collopy, 1998, How effective are neural networks at forecasting and prediction? A review and evaluation, Journal of Forecasting, 17, 481-495.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay, 1998, *The econometrics of financial markets*. Princeton, NJ: Princeton University Press.

Clemen, R.T., 1989, Combining forecasts: A review and annotated bibliography, International Journal of Forecasting, 5, 671-690.

Clements M.P. and D.F. Hendry, 1998. *Forecasting economic time series*, Cambridge University Press: Cambridge, UK.

Coe, P.J., 2002, Financial crisis and the great depression: A regime switching approach, Journal of Money, Credit and Banking, 34, 76-93.

Davies, R.B, 1987, Hypothesis testing when the nuisance parameter is present only under the alternative, Biometrika, 74, 33-43.

Dempster, M.P., N.M Laird and D.B. Rubin, 1977, Maximum likelihood estimation from incomplete data via the EM algorithm, Journal of the Royal Statistical Society, 39, series B, 1-38.

Diebold, F.X., and R.S. Mariano, 1995. Comparing predictive accuracy, Journal of Business and Economic Statistics, 13, 253-263.

Donaldson, R.G and M. Kamstra,1996, Forecast combining with Neural Networks, Journal of Forecasting, 15, 49-61.

Donaldson, R.G and M. Kamstra, 1997, An artificial neural network-GARCH model for international stock return volatility, Journal of Empirical Finance, 4, 17-46.

Edwards, S. and R. Susmel, 2001, Volatility dependence and contagion in emerging equities markets, Journal of Development Economics, 66, 505-532.

Elliott, G., 1999, Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution, International Economic Review, 44, 767-784.

Elliott, G., J.T. Rothenberg and J.H. Stock, 1996, Efficient tests for an autoregressive unit root, Econometrica, 64, 813-836.

Engel, C., 1994, Can the Markov Switching Model Forecast Exchange Rates? Journal of International Economics, 36, 151-165.

Engel, C. and J.D. Hamilton, 1990, "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" *American Economic Review*, 80, 689-713.

Engle C., and C. J. Kim, 2001, The long-run US/UK real exchange rate, Journal of Money Credit and Banking, 31, part 1, 335-356.

17

Hamilton J. D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica, 57, 357-384.

Hamilton, J. D., 1994. *Time Series Analysis*, Princeton University Press: Princeton, NJ.

Hamilton J.D., 1996. Specification tests in Markov-switching time-series models, Journal of Econometrics, 70, 127-157.

Hamilton, J.D. and R. Susmel, 1994, Autoregressive conditional heteroskedasticity and changes in regime, Journal of Econometrics, 64, 307-333.

Hamilton, J.D. and G. Lin, 1996, Stock market volatility and the business cycle, Journal of Applied Econometrics, 11, 573-593.

Kanas, A. and A. Yannopoulos, 2001, Comparing linear and nonlinear forecasts for stock returns, International Review of Economics and Finance, 10, 383-398.

Kanas, A., 2001, Neural networks linear forecasts for stock returns, International Journal of Finance and Economics, 6, 245-254.

Kanas, A., 2005, Regime linkages between the Mexican currency market and emerging equity markets, Economic Modelling, 22, 109-125.

Kanas, A. and G.P. Kouretas, 2003, Regime dependence between the official and parallel foreign currency markets for Dollars in Greece", Discussion Paper 03-15, Department of Economics, University of Crete.

Kavalieris, L., 1989, The estimation of the order of an autoregression using recursive residuals and crossvalidation, Journal of Time Series Analysis, 10, 178-271.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, 1992, Journal of Econometrics, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, Journal of Econometrics, 54, 159-178.

Lucas, R., 1976. Econometric policy evaluation: A critique. Carnegie-Rochester Conference on Public Policy, 1, 19-46.

Ng, S. and P. Perron, 2001, Lag length selection and the construction of unit root tests with good size and power, Econometrica, 69, 1519-1554.

Swanson N. and H. White, 1995. A model selection approach to assessing the information in the term structure using linear models and artificial neural networks, Journal of Business and Economics Statistics, 13, 265-275.

White H. 1992. *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell: Oxford.

Table 1. Estimation of the MS Model

Parameters	Regime 1	Regime 2
a_0	0.0005 (0.3760)	-0.0001 (-0.0378)
a_1	0.1837 (1.8662)	0.2529* (3.4198)
a_2	0.0930 (0.9769)	0.1777* (2.3255)
a_3	-0.0201 (-0.2229)	0.0510 (0.6670)
a_4	-0.0306 (-0.3501)	-0.0306 (-0.3501)
σ	0.014	0.052
Duration (in weeks)	60	105
Unconditional probability	0.3614	0.6386
p_{11}	0.9832	
p_{22}	0.9905	
Null hypothesis: The variance and autoregressive parameters are equal across regimes (Linear model) Alternative: The variance and autoregressive parameters are different across regimes (MS model)		
Likelihood	634.95{551.52}	
LR	166.86 [0.000]*	
AIC	-3.98 {-3.49}	
HQIC	-3.91 {-3.46}	

Notes:
1. LR denotes the likelihood ratio test for the null of a linear single equation model. The value in squared brackets next to LR is the marginal significance level of this test, based on Davies (1987).
2. The values in curly brackets report the respective values from the linear model.
3. t-statistics are reported in parentheses. (*) denotes statistical significance at the 5% level.

Table 2. Regime classification.

Period	Regime 1 (Low Volatility) or Regime 2 (High Volatility)
1996.18 – 1996.30	Regime 2
1996.31 – 1999.02	Regime 1
1999.03 – 2002.17	Regime 2

Notes: Regime chronology (in weeks) according to MS model for the Cyprus Stock Exchange general price index.

Table 3. ANN model estimation

ANN parameters	CSE index
a_1	-0.018
b_1	2.32
b_2	-1.46
b_3	-2.01
b_4	4.27
b_5	-2.11
b_6	-1.99
b_7	2.89
b_8	-0.68
$c_{1,1}$	-0.91
$c_{1,2}$	-1.02
$c_{1,3}$	-3.42
$c_{1,4}$	-4.77
$c_{1,5}$	3.21
$c_{1,6}$	-4.32
$c_{1,7}$	-2.11
$c_{1,8}$	-1.01
RMSE	0.41

Notes: a_j are the weights for the direct signals for each of the two input variables to the input variable, b_i are the weights for the signal from each of the eight hidden units to the output variable and $c_{j,i}$ are the weights for the signals from the input variable, j, to the hidden units, i, where $j = 1$ and $i = 1,2,3,4,5,6,7,8$

Table 4: Tests of forecast accuracy

Models	RMSE
MS model	0.121
ANN model	0.110
Diebold and Mariano (DM) test statistic Null: The RMSE of MS is equal to the RMSE of ANN Alternative: The RMSE of MS is higher than the RMSE of ANN Test statistic = 0.86 p-value = 0.80	

Notes:

1. * denotes rejection of the null at the 5% level of significance.
2. RMSE stands for Root Mean Squared Error.

Table 5: Forecast encompassing tests

Model 1	Model 2	$\hat{b}_1$	$\hat{d}_1$	Conclusion
MS	ANN	-0.66	0.34	Neither model
		(-1.34)	(0.70)	encompasses the
				Other

Notes:

1. * denotes statistically significant coefficient at the 5% level of significance.

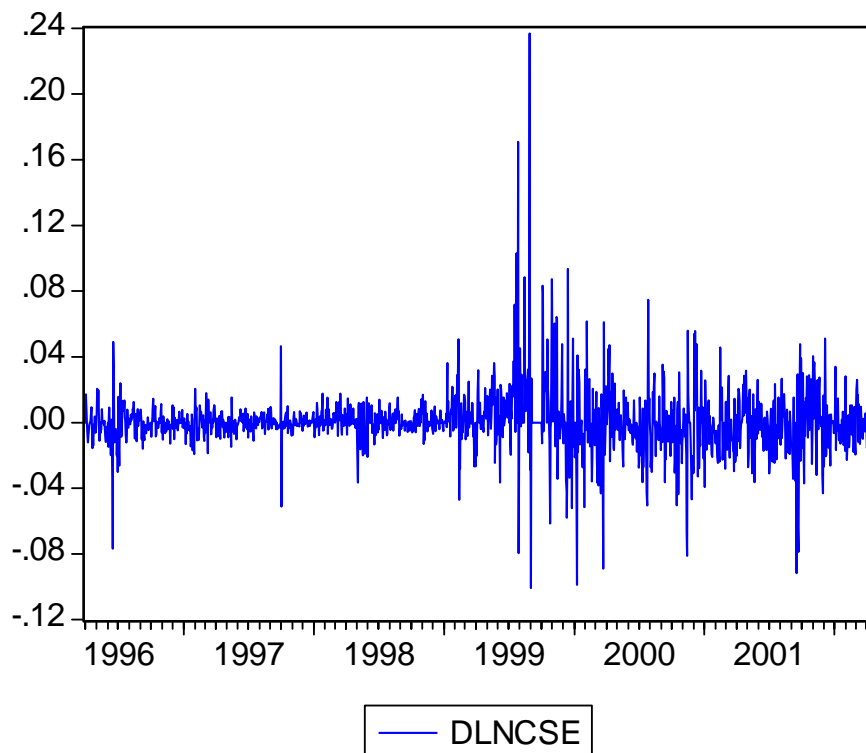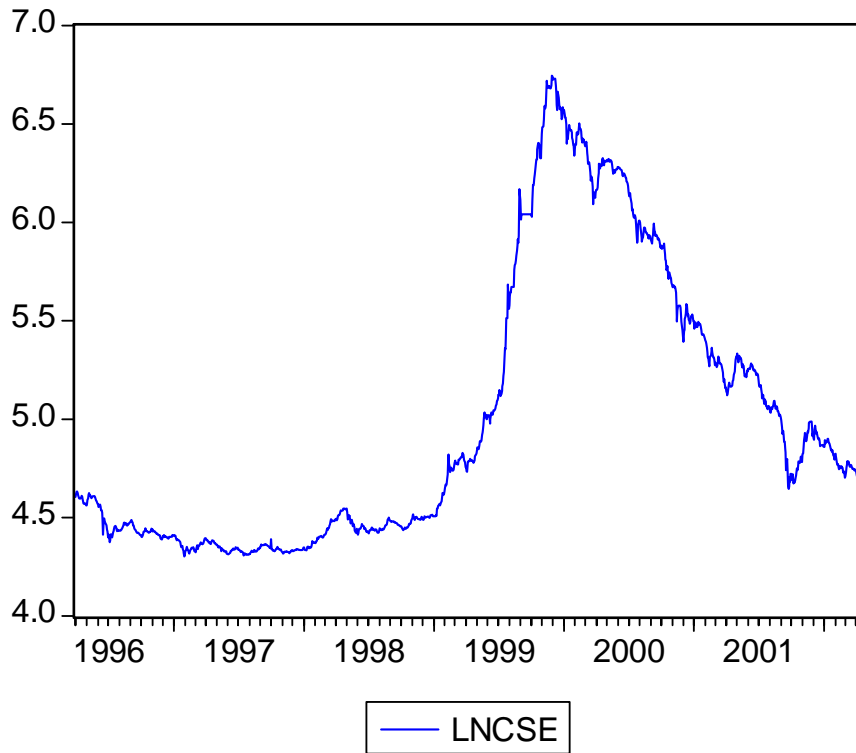2. Heteroscedasticity-robust t-statistics in parentheses.
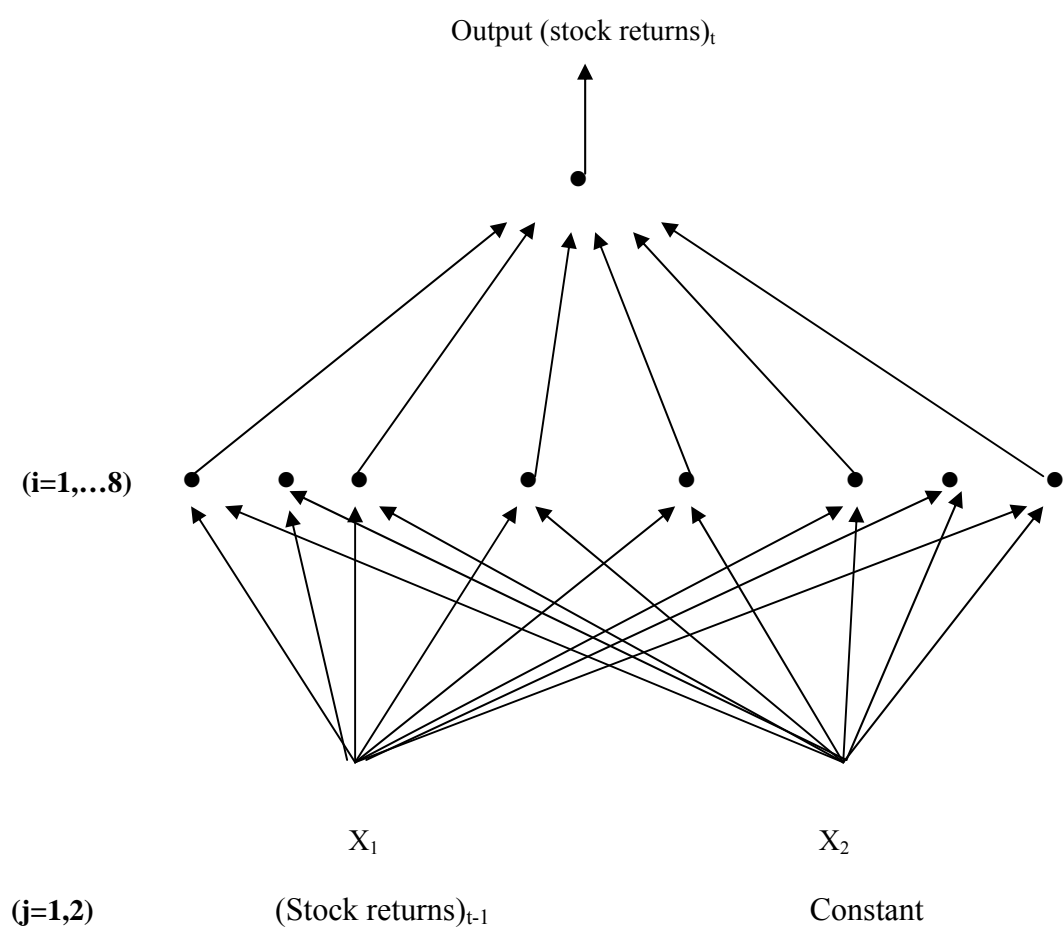
Figure 1. CSE General Price Index and Returns

Output (stock returns)$_t$

(i=1,...8)

X$_1$ X$_2$

(j=1,2)   (Stock returns)$_{t-1}$ Constant
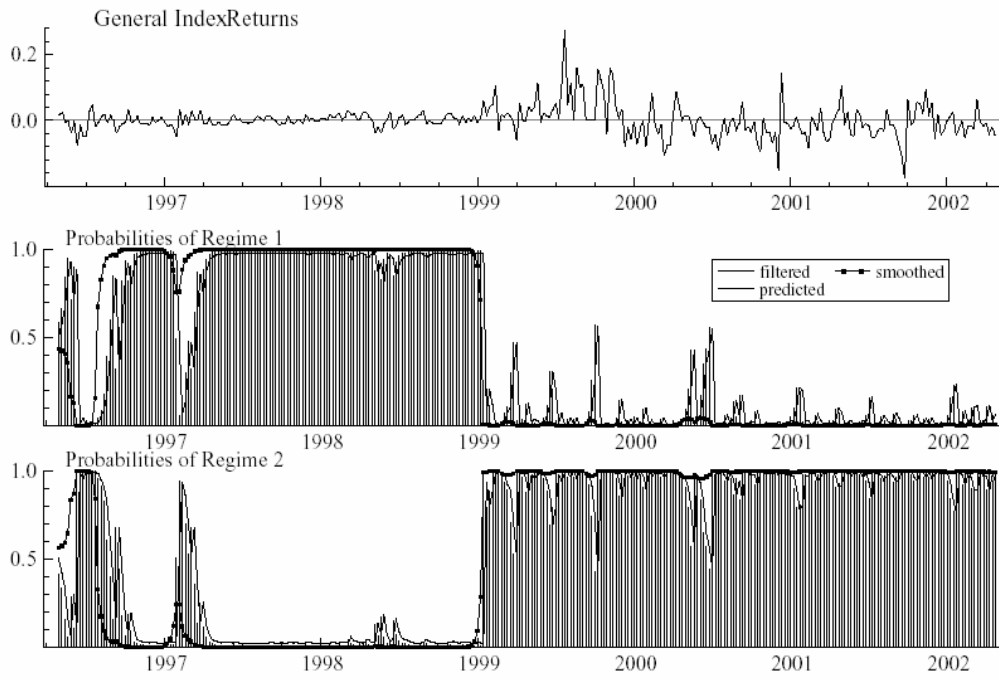
Figure 2. An MLP(1,8) for stock returns.

Figure 3. MS Regime Classification and Smoothed Probabilities