Are stock markets really efficient? Evidence of the Adaptive Market Hypothesis

Abstract

This study examines the adaptive market hypothesis of the S&P500, FTSE100, NIKKEI225 and EURO STOXX 50 by testing for stock return predictability using daily data from January 1990 to May 2014. We apply three bootstrapped versions of the variance ratio test to the raw stock returns and also whiten the returns through an AR-GARCH process to study the nonlinear predictability after accounting for conditional heteroscedasticity through the BDS test. We evaluate the time-varying return predictability by applying these tests to fixed-length moving subsample windows and also examine whether there is a relationship between the level of predictability in stock returns and market conditions. The results show that there are periods of statistically significant return predictability, but also episodes of no statistically significant predictability in stock returns. We also find that certain market conditions are statistically significantly related to predictability in certain markets but each market interacts differently with the different market conditions. Therefore our findings suggest that return predictability in stock markets does vary over time in a manner consistent with the adaptive market hypothesis and that each market adapts differently to certain market conditions. Consequently our findings suggest that investors should view each market independently since different markets experience contrasting levels of predictability, which are related to market conditions.

JEL classification:

G12; G14; G15

Keywords:

Adaptive market hypothesis

Stock return predictability

Market conditions

Market efficiency

1. Introduction

This paper examines the predictability of stock returns over time and determines whether the Adaptive Market Hypothesis (AMH) can describe the predictability of major stock indices around the world. We also study whether the changing degree of stock return predictability can be linked to certain market conditions. The level of predictability of stock returns may depend on certain market conditions since market conditions can have strong consequences on the psychology of market participants and the way the market participants analyse information which in turn affects their decision-making.

This paper is motivated by the fact that stock returns have been found to have predictive power from past price and return information, which is in contrast to the weak-form efficient market hypothesis (EMH). The weak-form EMH states that stock prices fully reflect all available information, which is limited to past prices, and consequently stock returns are purely unpredictable from past prices (Fama 1970). However the literature has found that stock returns do have predictive power and that the EMH does not always hold, challenging the validity of the EMH. Also, most of the previous work on market efficiency has followed a conventional approach of testing weak-form efficiency over a specific time period which has two main problems. Firstly, as Campbell et al (1997) note, this approach determines whether a market is efficient over a whole period as an all-or-nothing condition and ignores the notion of relative efficiency which enables the efficiency of one market to be measured against another. Secondly, the conventional approach assumes that the level of market efficiency is constant over some pre-determined time period. However this is very unlikely as many factors will result in the degree of market efficiency varying over time (Lim and Brooks 2011). Further, Grossman and Stiglitz (1980) argue for the impossibility of perfectly efficient markets since traders would not have any incentive to acquire costly information if markets were not inefficient and profit-making opportunities available.

The AMH proposed by Lo (2004), enables market efficiency and inefficiencies to co-exist in an intellectually consistent manner. Under the AMH, market efficiency evolves over time instead of being subject to the conventional view of all-or-nothing efficiency. Natural selection ensures the survival of the fittest and determines the number and composition of market participants and trading strategies. As market participants adapt to an ever-changing environment, they rely on heuristics to make their investment choices. Therefore the return predictability can arise from time to time due to changing market conditions as demonstrated by Lo (2005) through a rolling first-order autocorrelation test on monthly returns of the S&P Composite Index. Therefore, convergence to market efficiency is neither guaranteed nor likely to occur and the level of efficiency depends on the market participants and the market conditions at that moment of time. There has been an explosion of studies in the recent literature that document strong evidence of the AMH in stock markets (see for example Kim et al 2011; Lim et al 2013; Urquhart and Hudson 2013; Manahov and Hudson 2014; Urquhart and McGroarty 2014; Ghazani and Aragli 2014).

We examine the predictability of returns for the four most economically important stock market indices globally (S&P500, FTSE100, NIKKEI225 and EURO STOXX 50) over time between January 1990 and May 2014. We use three versions of the variance ratio test with bootstrapped p-values to examine predictability, as well as the BDS test which enables us to examine the nonlinear predictability of stock returns. To ensure the BDS test examines the nonlinear predictability in stock returns, we whiten returns through an AR(q)-GARCH(1,1) model to remove all the linear correlation in returns and time-varying volatility since market efficiency has implications for and only for the conditional mean (Lim and Hooy 2013). We use a two-year fixed-length moving sub-sample framework similar to Kim et al (2011) where each window length is two-years long that rolls forward one month. From this, we also obtain monthly measures of the degree of return predictability and test whether they are related to different stock market conditions. We find that stock return predictability does vary over time for each of the markets studied and that the S&P500 is the most efficient of the four markets studied. We also show that levels of predictability are associated with certain conditions in certain markets, but there is no widespread consensus on the behaviour of markets during those market conditions.

The contributions of this paper to the literature are as follows. First, this is the first study to examine whether stock return predictability is related to market conditions on a multi-market basis. The literature so far on AMH has either studied just one market or has not linked the predictability to market conditions. Second, this is the first study to use three bootstrapped versions of the variance-ratio test and the nonlinear BDS test to examine the AMH through a fixed length moving sub-sample window framework. The majority of the literature of AMH ignores the importance of nonlinear predictability. Third, we use these methods to capture the main dynamics of stock returns in several dimensions while at the same time reducing the risk that a spurious result from one test may affect the conclusions. Fourth, we examine which market conditions are associated with high/low levels of stock market index return predictability for four largest stock markets in the world (by trading volume) during the period of our study, which has been largely ignored in the empirical literature.

The remainder of the paper is organized as follows. The next section discusses the recent literature on the AMH while Section 3 presents the methodology. Section 4 reports the data and the empirical results while Section 5 summarises the findings and provides conclusions.

2. Related Literature

The AMH has received increasing attention in the recent academic literature where there has been strong evidence of the adaptive behaviour of stock returns. Lim and Brooks (2006) examine the evolving efficiency of developing and developed stock markets using a rolling sample approach. Through the portmanteau bicorrelation test they find that the degree of market efficiency varies over time in a cyclical fashion. Lim (2007) uses the portmanteau bicorrelation test through a rolling sample framework on eleven emerging and two developed markets and find that each market's level of efficiency evolves over time in a way consistent with the AMH. Todea et al (2009) study the profitability of the moving average rule over windows using linear and nonlinear tests and show that returns are not constant over but episodic. Ito and Sugiyama (2009) investigate the time-varying autocorrelation of monthly S&P500 returns and show that the degree of market efficiency varies over time, with the market the most inefficient during the late 1980s and most efficient around the year 2000.

Kim et al (2011) use an automatic variance ratio test and automatic portmanteau test to examine stock return predictability of daily DJIA data over time from 1900 to 2009. They use a rolling window and find strong evidence of time-varying predictability which is driven by market conditions. They find that stock market crashes are associated with no significant return predictability while during economics and political crises, a high degree of return predictability with a moderate degree of uncertainty is observed while during economic or political crises they find that stock returns are highly predictable. Alvarez-Ramirez et al (2012) study the DJIA from 1929 to March 2012 using entropy concepts for proposing a degree of relative market efficiency and they find that market efficiency does vary over time in line with the AMH. Charles et al (2012) study the return predictability of major foreign exchange rates from 1975 to 2009 using daily and weekly nominal exchange rates. By applying the automatic variance ratio test, generalized spectral test and Dominguez-Lobato consistent tests, they show that return predictability does vary over time depending on changing market conditions, consistent with the AMH. Smith (2012) examines the changing efficiency of 15 European emerging stock markets and three developed markets. They use rolling window variance ratio tests and find that the return predictability varies widely, with the global financial market crisis of 2007-2008 coinciding with high return predictability in Croatia, Hungary, Poland, Portugal, Slovakia and the UK. Lim et al (2013) examine the return predictability for three major US stock indices using the automatic portmanteau Box-Pierce test as well as the wild bootstrapped automatic variance ratio test through a rolling estimation approach. They find evidence of time-varying return predictability and that those periods with significant return autocorrelations can be largely associated with major exogenous events, thus consistent with the AMH.

Urquhart and Hudson (2013) study whether the US, UK and Japanese stock markets conform to the AMH through linear and nonlinear tests for stock return independence. They show strong evidence in favour of the AMH, suggesting it provides a better explanation of the stock return behaviour than the EMH. Zhou and Lee (2013) study REIT data through the automatic variance ratio test and automatic portmanteau test and show that market efficiency changes over time depending on market conditions. Dyakova and Smith (2013a) study two Bulgarian stock prices indices and eight stock prices using variance ratio tests in a rolling window from October 2000 to August 2012 and show that changing level of predictability with supports the AMH. Dyakova and Smith (2013b) examine 40 Bulgarian stocks, two Bulgarian stock market indices and 13 other South East European stock market indices using three finite-sample variance ratio tests and show that the return predictability of both stocks and stock market indices varies widely over time. Niemczak and Smith (2013) study 11 Middle Eastern stock markets and show that most markets experience successive periods of efficiency and inefficiency, which is consistent with the AMH.

Hull and McGroarty (2014) using the Hurst-Mandelbrot-Wallis rescaled range test as a measure of market efficiency on 22 emerging markets and show strong evidence in favour of the AMH. Ghazani and Aragli (2014) examine daily data from the Tehran stock exchange from 1999 to 2013 and show evidence of the AMH provides an appropriate evolutional perspective on market efficiency. Also, Manahov and Hudson (2014) develop artificial stock markets using a special adaptive form of the Strongly Typed Genetic Programming based learning algorithm and apply it to data from the FTSE100, S&P500 and Russell 3000. They show that the stock market dynamics are consistent with the evolutionary process of the AMH since trader population behave in an efficient adaptive system evolving over time. Further, Urquhart and McGroarty (2014) study the AMH through an examination of how well-known calendar anomalies behave over time. Using a rolling window analysis and a subsample analysis, they show that the four calendar anomalies studied support the AMH and that certain calendar anomalies are only present during certain market conditions. Rodriguez et al (2014) studies the AMH over weekly, monthly, quarterly and year time scales for the DJIA from 1929 to 2014 using the detrended fluctuation analysis and show that the interday and intraday returns are more serially correlated than overnight returns. They also show that the efficiency of the DJIA is not uniform over time thus providing evidence of the AMH. Smith and Dyakova (2014) use a rolling window analysis on three finite-sample variance ratio tests to examine the changing predictability of African stock market returns. They find that the stock markets go through successive periods of predictability and unpredictability with is consistent with the AMH. Also, Levich and Poti (2015) study the predictability in currency markets over the period 1972-2012 by constructing an upper bound on the explanatory power of predictive regressions of currency returns. They find that currency predictability exceeds this bound during recurring albeit short-lived episodes and that excess predictability is highest in the 1970s and tends to decrease over time but is still present in the final part of the sample period, thus providing evidence of the AMH. Urquhart et al (2015) study the simple moving average rule in the DJIA, FT30 and TOPIX and find that trading on anticipating signals generates superior profits for investors, suggesting that investors are anticipating signals in a way consistent with the AMH. Recently, Noda (2016) find that the degree of market efficiency in Japanese stock markets changes over time and thus supporting the AMH, while Ito et al (2016) develop a non-Bayesian time-varying model and show that the US stock market has evolved over time, consistent with the AMH.

3. Methodology

To examine the predictability of returns, we adopt three formulations of the variance ratio test, as well as the popular nonlinear BDS test. We include the BDS test due to the fact that linear tests may fail to pick up nonlinear predictability when nonlinear predictability is present (Amini et al 2010), and when the returns are whitened through an AR-GARCH process, any remaining nonlinear predictability cannot be attributed to conditional heteroscedasticity (Lim and Hooy 2013).

To obtain monthly measures of predictability, we employ the moving-subsample window of fixed length over the grid of months similar to Kim et al (2011) and Urquhart and McGroarty (2014). We use a two-year window and calculate the test statistics using data from the first trading day in January 1990 to the last trading day in December 1991 and then move the window forward one-month to cover the period February 1990 to January 1992. We continue this process to the end of the data and obtain measures for predictability of returns up to May 2014. We choose a two-year window to provide enough observations to generate reliable results, while at the same time providing enough results to analyse how the level of predictability has behaved over time.

3.1. Variance Ratio Test

Since the seminal work of Lo and MacKinlay (1988), the variance ratio (VR hereafter) test has emerged as a primary tool in examining whether stock returns are serially uncorrelated, with Hoque et al (2007) stating that it has become the most commonly used econometric tool for testing the random walk hypothesis[1].

The VR test is based on the statistical property that if a stock price follows a random walk, then the variance of the *k*-period return is equal to *k* times the variance of the one period return. Lo and MacKinlay (1988) provide a test for this hypothesis using the single VR, denoted by VR(k). Let r_t denote an asset return at time t, where t = 1,2,3….T. Then the variance ratio for r_t, with holding period k is;

$$VR(k) = \sigma_k^2 / k\sigma^2 \quad (1)$$

Where $\sigma_k^2 = Variance(r_t + r_{t-1} + ... + r_{t-k+1})$ is the variance of k-period return. It can be rewritten as;

$$VR(k) = 1 + 2\sum_{j=1}^{k-1}\left(1 - \frac{j}{k}\right)\rho(j) \quad (2)$$

Where $\rho(j)$is the autocorrelation of r_t of order j. That is, the variance ratio is one plus a weighted sum of autocorrelation coefficients for the asset returns with positive and declining

[1] For a survey on variance ratio tests see Charles and Darné (2009).

weights. The VR tests the null hypothesis that the variance ratio equals 1 for all k's since returns are serially uncorrelated with $\rho(j)$ = 0. Alternatively, values for VR(k) greater than 1 imply positive serial correlations while values less than 1 imply negative serial correlations or mean reversion.

Lo and MacKinlay (1988) determined the asymptotic distribution of VR(x; k) by assuming that k is fixed when T → ∞. They showed that if r_t is i.i.d., i.e. under the assumption of homoskedasticity, then under the null hypothesis that VR(k) = 1, the test statistic $M_1(k)$ is given by;

$$M_1(k) = \frac{VR(r;k) - 1}{\Phi(k)^{1/2}} \tag{3}$$

which follows the standard normal distribution asymptotically. The asymptotic variance, $\Phi(k)$, is given by;

$$\Phi(k) = \frac{2(2k-1)(k-1)}{3k} \tag{4}$$

To accommodate the returns exhibiting conditional heteroscedasticity, Lo and MacKinlay (1988) proposed the heteroscedasticity robust test statistic $M_2(k)$;

$$M_2(k) = \frac{VR(r;k) - 1}{\Phi^*(k)^{1/2}} \tag{5}$$

which follows the standard normal distribution asymptotically under the null hypothesis that VR(k) = 1, where;

$$\Phi^*(k) = \sum_{j=1}^{k-1} \left[\frac{2(k-j)}{k}\right]^2 \delta(j) \tag{6}$$

$$\delta(j) = \left\{ \sum_{t=j+1}^{T} (r_t - \hat{\mu})^2 (x_{t-j} - \hat{\mu})^2 \right\} \Big/ \left\{ \left[\sum_{t=1}^{T} (r_t - \hat{\mu})^2 \right]^2 \right\} \tag{7}$$

The $M_2(k)$ test is applicable to returns of a price series and this study utilises $M_2(k)$ due to the heteroscedastic property of the returns series' studied, as revealed in Table 1.

A problem with the traditional VR test is that under the random walk hypothesis, we must have $VR(k) = 1$ for all chosen values of k. Under the null hypothesis of no predictability, the VR value should be one for all values of k. Hence, the test for the null hypothesis should be conducted as a joint test for $VR(k) = 1$ for all k or multiple values of k. To account for this Chow and Denning (1993) propose a multiple VR test where only the maximum absolute value of $VR(k)$ in a set of m test statistics is considered[2]. The Chow-Denning test statistics is defined as;

$$CD_1 = \sqrt{T} \max_{1 \le j \le m} |M_2(k_j)| \tag{8}$$

And it follows the studentized maximum modulus (SMM) distribution with m and T degrees of freedom. One of the difficulties with the VR test is that the statistics are based on asymptotic theory making the statistical inference misleading in small samples (Richardson and Stock 1989). To overcome this problem, a wild bootstrap method for the VR test proposed by Kim (2006) is used for the Chow-Denning statistic which improves the small sample properties of the variance ratio test. This approach involves computing the individual and joint VR test statistics on samples of T observations formed by weighting the original data by mean 0 and variance 1 random variables, and using the results to form bootstrap distributions of the test statistics. The bootstrapped p-values are computed directly from the fraction of replications falling outside the bounds defined by the estimated statistics[3].

Further, Wright (2000) proposes a non-parametric alternative to the conventional VR test using ranks and signs that overcome the problems of biased and right-skewed samples. These two tests can be more powerful than the Lo-MacKinlay VR test since they have high power

[2] For all variance ratio tests, we use values of 2, 4, 8 and 16 for k.
[3] For more information on the wild bootstrap methodology, see Kim (2006).

against a wide range of models displaying serial correlation, the signs-based test is exact even under conditional heteroscedasticity and the ranks-based test displays low-size distortion under heteroscedasticity. Given log returns as r_t and r(r) be the rank of r(r) among (r_1, … , y_r) which, under the hypothesis that r_t is i.i.d, is just a random permutation of the numbers 1,2, …, T, each with equal probability. Define the rank based VR tests R_1 and R_2 as (for i = 1 or 2);

$$R_j(k) = \left(\frac{(Tk)^{-1}\sum_{t=k}^{T}(r_{jt} + \cdots + r_{jt-k+1})^2}{T^{-1}\sum_{t=1}^{T} r_{jt}^2} - 1\right)\left(\frac{2(2k-1)(k-1)}{3kT}\right)^{-1/2} \tag{9}$$

Where

$$r_{1t} = \frac{\left[r(r_t) - \left(T + \frac{1}{2}\right)\right]}{\sqrt{((T-1)(T+1))/12}} \tag{10}$$

$$r_{2t} = \frac{\varphi^{-1} r(r_t)}{T+1} \tag{11}$$

Where φ^{-1} is the inverse of the standard normal cumulative distribution function. The test based on signs of the first difference is given by;

$$S_j(k) = \left(\frac{(Tk)^{-1}\sum_{t=k}^{T}(s_{jt} + \cdots + s_{it-k+1})^2}{T^{-1}\sum_{t=1}^{T} s_{jt}^2} - 1\right)\left(\frac{2(2k-1)(k-1)}{3kT}\right)^{-1/2} \tag{12}$$

Where s_t = $2u(y_t, 0)$ and $u(y_t, 0)$ is ½ if y_t is positive and -1/2 otherwise. Under the assumption that r_t is generated from martingale difference sequence with no drift, s_t is an i.i.d. sequence with zero mean and unit variance and the critical values can be obtained by simulating its sampling distribution. We also conduct the joint variance ratio test for ranks and signs proposed by Belaire-Franch and Contreras (2004) and Kim and Shamsuddin (2008).

3.2. BDS Test

The BDS test, proposed by Brock, Dechert and Scheinkman (1987), is a popular non-parametric test for serial dependence (or a nonlinear structure) in stock returns. The null hypothesis of this test is that the data generating processes are i.i.d., while the alternative hypothesis is "an indication that the model is misspecified" (Brock et al 1996). Given a sample of i.i.d. observations, $\{x_t: t = 1,2, . . , n\}$, Brock et al (1996) show;

$$W_{m,n}(\varepsilon) = \sqrt{n}\frac{T_{m,n}(\varepsilon)}{V_{m,n}(\varepsilon)} \tag{13}$$

Where $W_{m,n}(\varepsilon)$ is the BDS statistic, n is the sample size, m is the embedding dimension and the metric bound (ε) is the maximum difference between pairs of observations counted in computing the correlation integral. $T_{m,n}(\varepsilon)$ measures the difference between the dispersion of the observed data series in a number of spaces with the dispersion that an i.i.d. process would generate in these spaces ($C_{m,n}(\varepsilon) - C_{1,n}(\varepsilon)^m$) and has an asymptotic normal distribution with zero mean and variance $V^2_m(\varepsilon)$. The asymptotic distribution of the BDS test does not depend on the existence of higher-order unconditional moments. As Hsieh (1991) points out, structural changes in the data series can cause a rejection of the null hypothesis of i.i.d. on the basis of the BDS test. Thus it is rational to break up the sample period and examine subsamples separately. The choice of ε and m values can be problematic since too small a ε will capture too few points so we follow the common approach in the literature by setting ε as a proportion of the standard deviation of the data. With regards to m, we again follow the literature by setting m from 2 to 5 because the small sample properties of the BDS test degrade as m increases (Patterson and Ashley 2000). To determine how the predictability from the BDS test changes over time, we take the mean of the p-values generated from the m values.

To examine the nonlinear dependence in returns, the returns need to be whitened to remove all linear correlations. An AR(p) model is fitted to the data to remove the linear correlations with the optimal lag length determined when the standardised residuals are no longer correlated through the Ljung-Box Q-statistic up to 10 lags. However as Lim and Hooy (2013) note, it is generally accepted that most of the nonlinear dependence in financial returns are due to conditional heteroscedasticity that can be captured by an ARCH-type model. As the EMH does not impose any restrictions on the dynamics of the conditional variance, any nonlinear dependence found due to conditional heteroscedasticity is not a violation of the

EMH. As the BDS has been proven to have high power against ARCH and GARCH models where nonlinearity enters through the conditional variance, we fit an AR-GARCH(1,1) to the returns and its standardised residuals are then tested for i.i.d. using the BDS test such that;

$$
\begin{gathered}
r_t = \beta_0 + \sum_{t=1}^{p} \beta_i r_{t-i} + \varepsilon_t \\
\varepsilon_t \sim N(0, h_t) \\
h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 \varepsilon_{t-1}^2
\end{gathered} \tag{14}
$$

Where r_t is the return series, ε_t is the residual of the mean equation and h_t standards for the conditional variance of the residual. The natural logarithm of the squared standardised residuals, $log(\zeta_t^2)$ where $\zeta_t = \varepsilon_t/\sqrt{h_t}$, are then subject to the BDS test. Thus if the BDS test finds the AR-GARCH filtered returns[4] do have significant dependence, there is nonlinear dependence in stock returns and a violation of the EMH.

4. Data and Results

In this section, we present the data details and their descriptive statistics. We also present the empirical returns of the tests discussed in the previous section.

4.1. Data

The sample data consists of daily closing prices for the S&P500, FTSE100, NIKKEI225 and EURO STOXX 50 in order to cover the main stock indices in the world. The data spans 1st January 1990 to 30th May 2014. The S&P500, FTSE100 and NIKKEI225 are three of the most important and well established world markets, while the EURO STOXX provides a representation of Blue-chip stocks within the Eurozone area. The data is obtained from Bloomberg and Thomson Financial Datastream. Figure 1 presents the time-plots of the four markets studied while descriptive statistics of the three indices are presented in Table 1, where the daily returns for each index is calculated by;

$$r_t = In(P_t) - In(P_{t-1}) \tag{15}$$

[4] The models required to whiten each series is AR(5) for the S&P500, FTSE100 and EURO STOXX 50, while an AR(1) was required for the NIKKEI225.

Table 1 shows that the S&P500 the highest mean return and that the NIKKEI has a negative mean return. The standard deviation values show that the NIKKEI has the highest volatility while the FTSE100 is the least volatile. All four return series' indicate negative skewness indicating a longer left tail. Excess kurtosis is observed for all returns series, showing that their distributions are leptokurtic. The Jarque-Bera test statistics is significant at the 1% level for all series indicating the non-normal nature of their returns. The LM test is applied to the residuals of a fitted ARMA model to each series to test for conditional heteroscedasticity in returns[5]. All four returns series provide significant evidence at the 1% level for conditional heteroscedasticity in returns.

(Insert Table 1 and Figure 1 here)

4.2. Time-varying returns predictability

Figures 2-5 presents the three different variance ratio p-values over time through a two-year fixed-length moving subsample window analysis for the S&P500, FTSE100, NIKKEI225 and EURO STOXX 50. The statistical significance of the three variance ratio tests is evaluated using p-values where if the p-value is less than or equal to 0.05, the p-value is deemed to reject the null hypothesis and therefore indicates significant evidence in support of the alternative hypothesis.

Figure 2 presents the variance ratio p-values over time for the S&P500 and it is clear to see that in some periods the three variance ratio tests generate very different p-values, reflecting the difference between these formulations of the variance ratio tests. Throughout the sample the p-values generated vary over time, with some periods generating statistical significant p-values and some periods generating quite high p-values. From the start of the sample to January 2003, nearly all of the p-values for each of the three variance ratio tests are insignificant, suggesting the independence and unpredictability of stock returns. However from January 2003 to April 2008 the p-values for each test fluctuate between being statistically significant and insignificant, indicating a varying behaviour of stock returns. From April 2008 to April 2010 all of the p-values for the three tests are statistically significant indicating the predictability of stock returns. After April 2010 to the end of the

[5] Where the lag length is selected based on the Akaike information criterion. The lag lengths required are ARMA(0,0) for the S&P500, ARMA(1,0) for the FTSE100 and EURO STOXX 50, and ARMA(5,2) for the NIKKEI225.

sample, nearly all of the p-values are insignificant, suggesting the unpredictable nature of stock returns. Therefore the variance ratio test results for the S&P500 support the AMH as the stock market goes through periods of predictability and unpredictability.

Figure 3 shows the variance ratio p-values over time through a two-year rolling window analysis on the FTSE100. At the start of the sample until May 1998, nearly all of the p-values are insignificant, indicating the unpredictable nature of the stock returns. However from May 1998 to May 2010 nearly all of the p-values are statistically significant or very close to being significant at the 5% level indicating that stock returns were predictable during this period. Nevertheless, from this period onwards, only one of the p-values is significant indicating that stock returns were generally unpredictable during this period. Therefore the varying behaviour of the stock returns is supportive of the AMH.

In Figure 4, the variance ratio p-values for the NIKKEI225 are also shown to vary considerably over time. There is evidence of strong predictability from June 1997 to June 1999 where nearly all of the p-values are statistically significant. However during the rest of the sample period, there are only a few p-values that are statistically significant indicating the unpredictable behaviour of NIKKEI225 returns. Nevertheless, there is clear evidence of the changing level of predictability of stock returns which is consistent with the AMH.

Figure 5 presents the three variance ratio p-values over time for the EURO STOXX 50 and there is again clear evidence of time-varying behaviour of stock return predictability. There is significance evidence of predictability from October 1994 to November 1995, January 1999 to March 2000, January 2005 to December 2008 and July 2013 to May 2014 for the joint-rank and joint-sign tests only. Therefore each variance ratio test shows that the EURO STOXX 50 has gone through periods of predictability and unpredictability consistent with the AMH.

(Insert Figures 2-5 here)

Figures 6-9 report the average p-values of the BDS test from the four dimension sizes through a two-year rolling window analysis. The statistical significance is evaluated using the 95% confidence interval based on the bootstrap methodology described earlier.

The BDS statistic p-values from the rolling window analysis are reported in Figure 6 where there is clear evidence of no significant BDS statistics from the start of the sample to May 1998. From this point to August 1999 all p-values are statistically significant at the 5% level indicating the nonlinear predictability in stock returns. After this point the returns are deemed unpredictable since the p-values are all insignificant. However the p-values are significant from June 2002 to October 2004 suggesting returns are predictable during this period. Nevertheless, the p-values from November 2004 to October 2007 are all insignificant indicating the unpredictable nature of the returns. From November 2007 to the end of the sample, nearly all of the p-values are statistically significant thus suggesting that the returns are predictable during this period. Therefore the BDS test results indicate that the stock returns go through periods of predictability and unpredictability, consistent with the AMH.

Figure 7 reports the FTSE100 BDS test p-values over time and from the start of the sample to August 1997 nearly all the p-values are insignificant indicating the unpredictability of stock returns. The FTSE100 then goes through a period of predictable BDS test statistics until April 2000 and then goes through a period of unpredictability until August 2001 when the p-values are statistically significant. From August 2000 to the end of the sample, all BDS p-values are statistically significant except two periods between April 2005 and April 2006, and November 2010 to July 2011. These results are again consistent with the AMH.

Figure 8 presents the NIKKEI BDS test statistic p-values over time and show that from the start of the sample to August 2000, all but 4 of the p-values are statistically significant at the 5% level indicating strong evidence of the predictable nature of stock returns during this period. From September 2000 to June 2005 all the p-values are insignificant indicating the unpredictable behaviour of stock returns. However from July 2005 to October 2010, all but one of the p-values is statistically significant indicating that stock returns are predictable once again. After this point to the end of the sample, stock return behaviour fluctuates between predictable and unpredictable. Therefore the NIKKEI225 results support the AMH since stock return predictability fluctuates over time.

The EURO STOXX 50 BDS test statistic p-values are presented in Figure 9 and show initially there is strong evidence of predictability however from October 1993 to June 1997 all p-values are insignificant indicating the unpredictable nature of stock returns. After this point until February 2005, all but two of the p-values are statistically significant at the 5%

level indicating that stock returns were predictable during this period. Apart from two periods between March 2005 to June 2006 and October 2010 to August 2011 when the p-values insignificant, the results suggest that stock returns are predictable. Hence the EURO STOXX 50 results also provide strong evidence of the AMH.

(Insert Figures 6-9 here)

Therefore our results show that each market has experienced differing periods of statistically significant predictability and periods of no statistically significant predictability. These changing levels of predictability depend on the tests employed, since differing testing procedures will capture differing aspects of return predictability. Therefore a direct comparison of markets can only be made when the testing procedure is consistent.

To enable a comparison of the four markets, we study the relative efficiency according to the tests for predictability. In Table 2 we report the percentage of each test statistic which does not reject the null hypothesis at the 5% significance level, similar to Smith (2012). The S&P500 is deemed the most efficient of the four markets since 76.21% of the p-values over the four tests fail to reject the null hypothesis of efficiency, while the EURO STOXX 50 is the least efficient with only 66.95% of the p-values failing to reject the null hypothesis. The BDS test fails to reject the null hypothesis of efficiency a lot less of the time than the three formulations of the variance ratio test, indicating that the levels of nonlinear predictability are high and that there can be high levels of nonlinear predictability even in the absence of linear predictability[6].

(Insert Table 2 here)

4.3. Market Conditions and Return Predictability

Lo (2004) states that the degree of predictability of a market varies over time with changes in market conditions but gives no suggestion to specific indicators of market conditions or predictions about the relationship between predictability and market conditions (Kim et al 2011). Hence we examine the relationship between certain market conditions and the level of stock return predictability from the variance ratio tests and the BDS test. Therefore we

[6] As documented by Amini et al (2010).

regress the monthly measures of return predictability against a number of market conditions. The measures of return predictability we use are the p-values of the variance ratio tests and BDS test. As discussed in Section 3, the p-values of the variance ratio tests and BDS tests indicate significant predictability in stock returns.

Since we generate a measure of predictability as the predictability over the previous 24 months, we calculate the market conditions on data for the previous 24 months. Firstly, we separate our data into bull and bear markets similar to Fabozzi and Francis (1977). We define each period as either Up or Down (UD) periods, where an UP period is one where the average return was non-negative and DOWN periods are when the average return was negative. This procedure yields a mutually exclusive and exhaustive division of the total sample into two subsets, however it ignores trends in the market and views each month independently. We also use Klein and Rosenfeld (1987) definition of bull and bear markets, which is a modified version of the methodology of Fabozzi and Francis (1977). A period is defined as a 'substantial market mover' when the periods market's return is greater than one-half of the market's standard deviation of that periods returns. Following Klein and Rosenfeld (1987), we place all the moving sample period in the sample either into bull, normal, or bear categories based on trend. For example, if the market index rises substantially in one period while the surrounding periods are average, this period is classified as average. If the market either rises or is normal in one period while the surrounding periods are deemed bearish, this period is classified as bearish. Thus each type of market must contain a minimum of two consecutive substantial movements. As a measure of market risk, we calculate the realised volatility of each stock market as the square root of the sum of daily return squares over the previous 24 months, which is a purely non-parametric measure of total market risk[7] (Anderson et al., 2003).

Table 3 reports the number of periods designated with a certain market condition. The S&P500, FTSE100 and EURO STOXX 50 all have more UP periods than DOWN periods, while the NIKKEI225 have more DOWN periods which reflects the markets negative mean return. Each market has more normal periods than bull or bear periods, while only the NIKKEI225 has more bear periods than bull periods.

[7] We do not classify the volatility measure but regress the level of volatility against the levels of predictability.

(Insert Table 3 here)

Table 4 reports the regression results for the market conditions and the different measures stock return predictability. The results are mixed for the markets studied, with certain market conditions associated with significant predictability in certain stock returns but not for other measures and in the other markets studied. More specifically, the S&P500 results show a significant negative relationship between the BDS p-values of the S&P500 and DOWN periods, indicating high levels of predictability during DOWN periods. We also find that during Bull periods there is significantly low level of predictability according the joint-sign test, but significantly high levels of predictability during bear markets according the BDS test. Further, we also find that there are significantly high levels of predictability during normal periods according to the joint-rank and joint-sign test, while the BDS test suggests significantly low levels of predictability during these normal periods. Also the BDS test shows significant high predictability during high volatility periods. The FTSE100 results show that during bull markets there is significantly low predictability according the Chow-Denning test, while there is significantly high levels of predictability during normal periods according the joint-rank and joint-sign tests. Similar to the S&P500 results, there is significantly evidence of high levels of predictability during high volatility periods according to the BDS test. The NIKKEI225 results show that according the joint-rank test, there are high levels of predictability during bear periods. However according the BDS test, there are significantly low levels of predictability during bear periods. Also, the BDS test suggests significantly high level of predictability during normal periods, and the joint-sign test suggests significantly low levels of predictability during periods of high volatility. The EURO STOXX results suggest significantly high levels of predictability during UP and bull periods according to the joint-sign test, and significantly low levels of predictability during DOWN periods. All other periods generate insignificant coefficients indicating no significant relationship between stock return predictability and market conditions.

(Insert Table 4 here)

5. Conclusions

This paper examines the return predictability of the returns in four of the world's largest stock markets, namely the S&P500, FTSE100, NIKKEI225 and EURO STOXX 50, by testing the AMH using daily data from January 1990 to May 2014. As measures of the degree of return predictability, we use three formulations of the variance ratio test, as well as the nonlinear BDS test on pre-whitened returns. We evaluate the time-varying return predictability by applying these tests to fixed-length moving sub-sample windows of two years, which move forward one month at a time. A regression analysis is also conducted to determine how these measures of return predictability are related to changing market conditions. Therefore we add to the growing literature on the AMH and also link the levels of stock return predictability to market conditions.

We find evidence that return predictability fluctuates over time in each market, with each return series going through periods of significant predictability and periods where no predictability is found. This is found for each variation of the variance ratio test, as well as for the BDS test for nonlinear predictability. This suggests that the linear and nonlinear stock return predictability does vary over time and that market efficiency is not an all-or-nothing condition. We also can see that different markets experience significant predictability at different periods of time, suggesting that each market evolves differently over time and the predictability in the markets is not very correlated.

We also show that certain market conditions are more favourable to produce periods of significant predictability, however this varies with each market. This suggests that markets adapt differently over time and interact differently to varying market conditions. A bull market may represent a period of significant stock return predictability according to the Chow-Denning test in the FTSE100, but that does not necessarily mean that a bull market in the S&P500 will be associated with significant predictability. Therefore even though we find evidence of the AMH in each market, each market must be viewed as an individual entity as they interact differently to market conditions. Thus investors need to view each market independently since the predictability of these markets vary over time along with their market conditions.

References

Alvarez-Ramirez, J., Rodriguez, E., Espinosa-Paredes, G. (2012). Is the US stock market becoming weakly efficient over time? Evidence from 80-year-long data. *Physica A: Statistical Mechanics and its Applications*, 391(22), 5643-5647.

Amini, S., Hudson, R., Keasey, K. (2010). Stock return predictability despite low autocorrelation. *Economics Letters*, 108, 101-103.

Anderson, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579-625.

Belaire-Franch, J., Contreras, D. (2004). Ranks and signs-based multiple variance ratio tests, Working Paper, University of Valencia, Valencia.

Borges, M. R. (2010. Efficient market hypothesis in European stock markets. *The European Journal of Finance*, 16, 711-726.

Brock, W. A., Dechert, W. D., Schieinkman, J. A., (1987). A test for independence based on correlation dimension, SSRI Working Paper No. 8702, Department of Economics, University of Wisconsin, Madison, WI.

Brock, W. A., Dechert, W. D., Schieinkman, J. A., LeBaron (1996). A test for independence based on correlation dimension. *Econometric Reviews*, 15, 197-235.

Campbell, J. Y., Lo, A. W., MacKinlay, A. C. (1997). The Econometrics of Financial Markets. Princeton University Press, Princeton.

Charles, A., Darné, O. Variance-ratio tests of random walk: An overview. *Journal of Economic Surveys,* 23(3), 503-527.

Charles, A., Darné, O, Kim, J. H. (2012). Exchange-rate return predictability and the adaptive markets hypothesis: Evidence from major foreign exchange rates. *Journal of International Money and Finance*, 31, 1607-1626.

Chow, K., Denning, K. (1993). A simple multiple variance ratio test. *Journal of Econometrics,* 58, 386-401.

Dyakova, A., Smith, G. (2013a). The evolution of stock market predictability in Bulgaria. *Applied Financial Economics*, 23(9), 8005-816.

Dyakova, A., Smith, G. (2013b). Bulgarian stock market relative predictability: BSE-Sofia stocks and South East European markets. *Applied Financial Economics,* 23(15), 1257-1271.

Fabozzi, F. J., Francis, J. C. (1977). Stability Tests for Alphas and Betas Over Bull and Bear Market Conditions. *Journal of Finance*, 32(4), 1093-1099.

Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance*, 25, 383-417.

Ghazani, M. M., Araghi, M. K. (2014). Evaluation of the adaptive market hypothesis as an evolutionary perspective on market efficiency: Evidence from the Tehran stock exchange. *Research in International Business and Finance*, 32, pp. 50-59.

Grossman, S. J., Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3), 393-408.

Hsieh, D. A. (1991). Chaos and nonlinear dynamics: Application to financial markets. *Journal of Finance*, 46, 488-502.

Hoque, H. A. A. B., Kim, J. H., Pyun, C. S. (2007). A comparison of variance ratio tests of random walk: a case of Asian emerging stock markets. *International Review of Economics and Finance*, 16, 488-502.

Hull, M., McGroarty, F. (2014). Do emerging markets become more efficient as they develop? Long memory persistence in equity indices. *Emerging Markets Review*, 18, 45-61.

Ito, M., Noda, A., Wada, T. (2016). The evolution of stock market efficiency in the US: a non-Bayesian time-varying model approach. *Applied Economics*, 48(7), 621-635.

Ito, M., Sugiyama, S. (2009). Measuring the degree of time varying market inefficiency. *Economics Letters*, 103(1), 62-64.

Kim, J. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, 92, 38-43.

Kim, J., Shamsuddin, A. (2008). Are Asian stock markets efficient? Evidence from new multiple variance ratio test. *Journal of Empirical Finance,* 15, 518-532.

Kim, J. H., Shamsuddin, A., Lim, K. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data. *Journal of Empirical Finance*, 18, 868-879.

Klein, A., Rosenfeld, J. (1987). The Influence of Market Conditions on Event-Study Residuals. *Journal of Financial And Quantitative Analysis,* 22(3), 345-351.

Levich, R. M., Poti, V. (2015). Predictability and 'good deals' in currency markets. *International Journal of Forecasting*, forthcoming.

Lim, K-P., Brooks, R. D. (2006). The evolving and relative efficiencies of stock markets: Empirical evidence from rolling bicorrelation test statistics. *SSRN Working Paper* (available at http://ssrn.com/abstract=931071)

Lim, K-P. (2007) Ranking market efficiency for stock markets: A nonlinear perspective. *Physica A*, 376, pp. 445-454

Lim, K-P., Brooks, R. (2011). The evolution of stock market efficiency over time: A survey of the empirical literature. *Journal of Economic Surveys*, 25, 69-108.

Lim, K-P., Hooy, C. (2013). Non-linear predictability in G7 stock index returns. *The Manchester School*, 81, 620-637.

Lim, K-P., Luo, W. L., Kim, J. H. (2013). Are US stock index returns predictable? Evidence from automatic autocorrelation-based tests. *Applied Economics*, 45(8), 953-962.

Lo, A., MacKinlay, A. (1988). Stock market prices do not follow random walks: evidence from simple specification test. *The Review of Financial Studies,* 1, 41-66.

Lo, A.W. (2004) The Adaptive Markets Hypothesis. *Journal of Portfolio Management*, 30, pp. 15-29.

Lo, A. W. (2005). Reconciling efficient markets with behavioural finance: the adaptive markets hypothesis. *Journal of Investment Consulting*, 7, 21-44.

Manahov, V., Hudson, R. (2014). A note on the relationship between market efficiency and adaptability – New evidence from artificial stock markets. *Expert Systems with Applications*. 41(16), 7436-7454.

Niemczak, K., Smith, G. (2013). Middle Eastern stock markets: absolute, evolving and relative efficiency. *Applied Financial Economics*, 23(3), 181-198.

Noda, A. (2016). A test of the adaptive market hypothesis using a time-varying AR model in Japan. *Finance Research Letters*, 17, 66-71.

Patterson, D. M., Ashley, R. A. (2000). A Nonlinear Time Series Workshop: a Toolkit for Detecting and Identifying Nonlinear Serial Dependence, Kluwer Academic, Boston, MA.

Rodriguez, E., Aguilar-Cornejo, M., Alvarez-Ramirez, J. (2014). US stock market efficiency over weekly, monthly, quarterly and yearly time scales. *Physica A: Statistical Mechanics and its Applications.* 413, 554-564.

Todea, A., Ulici, M., Silaghi, S. (2009). Adaptive market hypothesis: Evidence from Asia-Pacific financial markets. *The Review of Finance and Banking*, 1(1), 7-13.

Smith, G. (2012). The changing and relative efficiency of European emerging stock markets. *The European Journal of Finance*, 18(8), 689-708.

Smith, G., Dyakova, A. (2014). African Stock Markets: Efficiency and Relative Predictability. *South African Journal of Economics*, 82(2), 258-275.

Urquhart, A., Gebka, B., Hudson, R. (2015). How exactly do markets adapt? Evidence from the moving average rule in three developed markets. *Journal of International Financial Markets, Institutions and Money*, 38, 127-147.

Urquhart, A., Hudson, R. (2013). Efficient or adaptive markets? Evidence from major stock markets using very long run historic data. *International Review of Financial Analysis,* 28, 130-142.

Urquhart, A., McGroarty, F. (2014). Calendar effects, market conditions and the Adaptive Market Hypothesis: Evidence from long-run U.S. data. *International Review of Financial Analysis*, 35, 154-166.

Wright, J. (2000). Alternative variance-ratio tests using ranks and signs. *Journal of Business and Economic Statistics*, 18, 1-9.

Zhou, J., Lee, J. M. (2013). Adaptive market hypothesis: evidence from the REIT market. *Applied Financial Economics*, 23(21), 1649-1662

Figures and Tables

Figure 1: Time plots of the S&P500, FTSE100, NIKEEI225 and EURO STOXX 50.

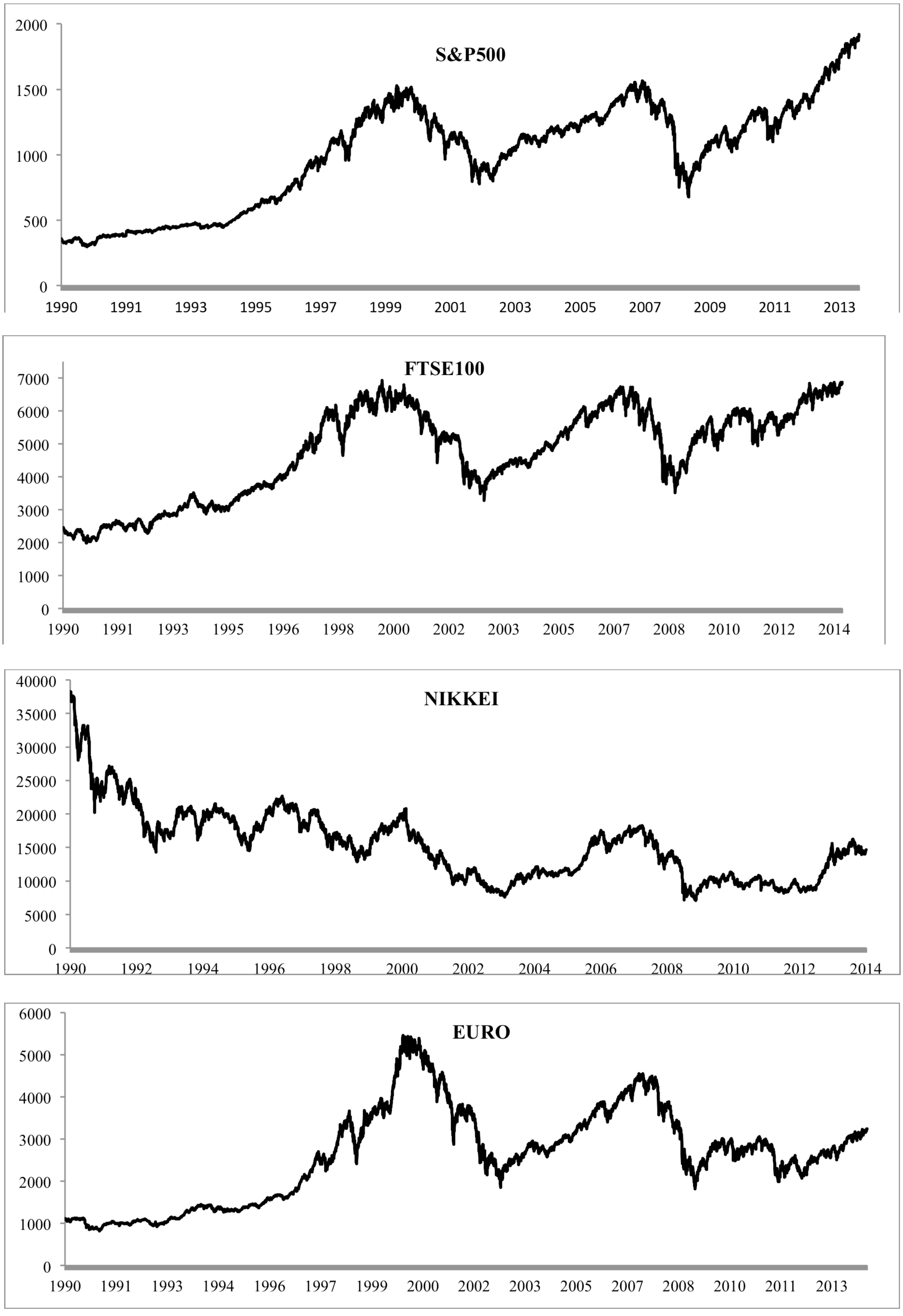

Figure 2: The three different variance ratio joint-test statistic p-values over time for the S&P500 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

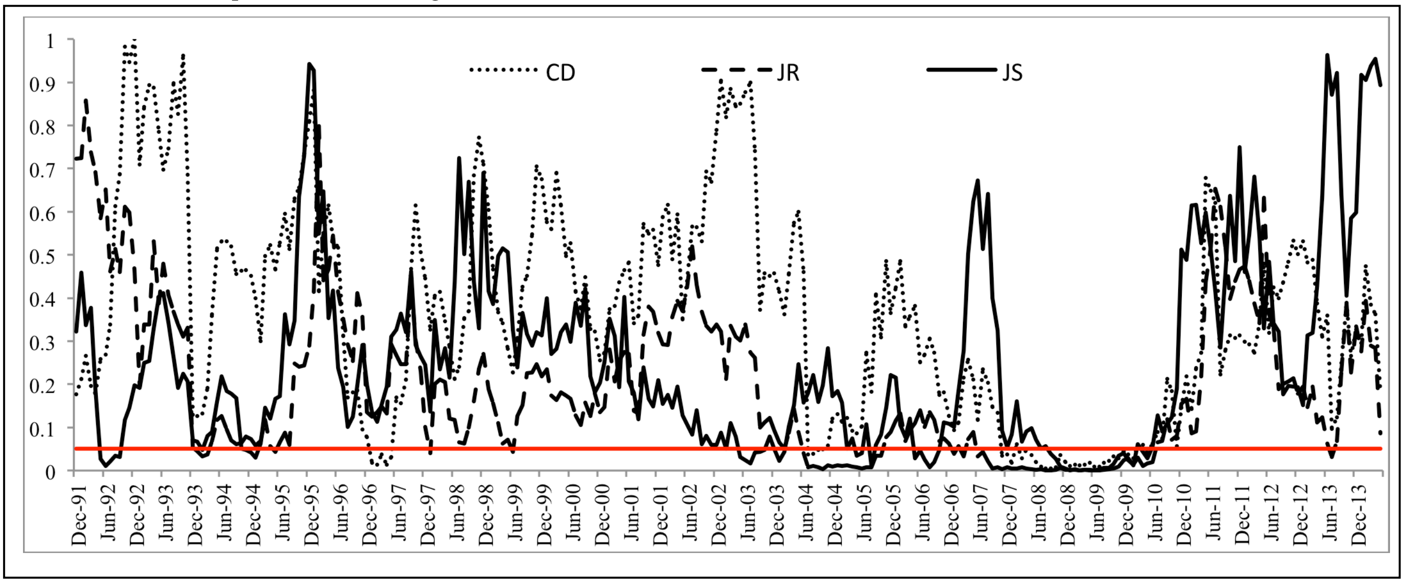

Figure 3: The three different variance ratio joint-test statistic p-values over time for the FTSE100 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

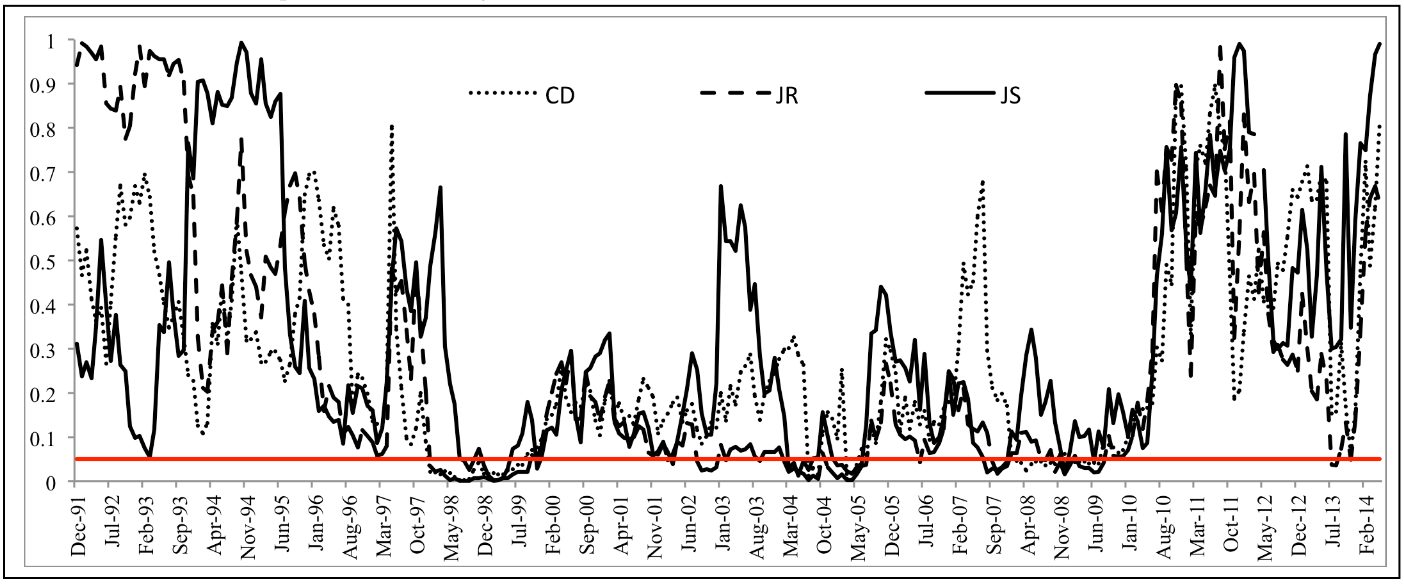

Figure 4: The three different variance ratio joint-test statistic p-values over time for the NIKKEI225 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

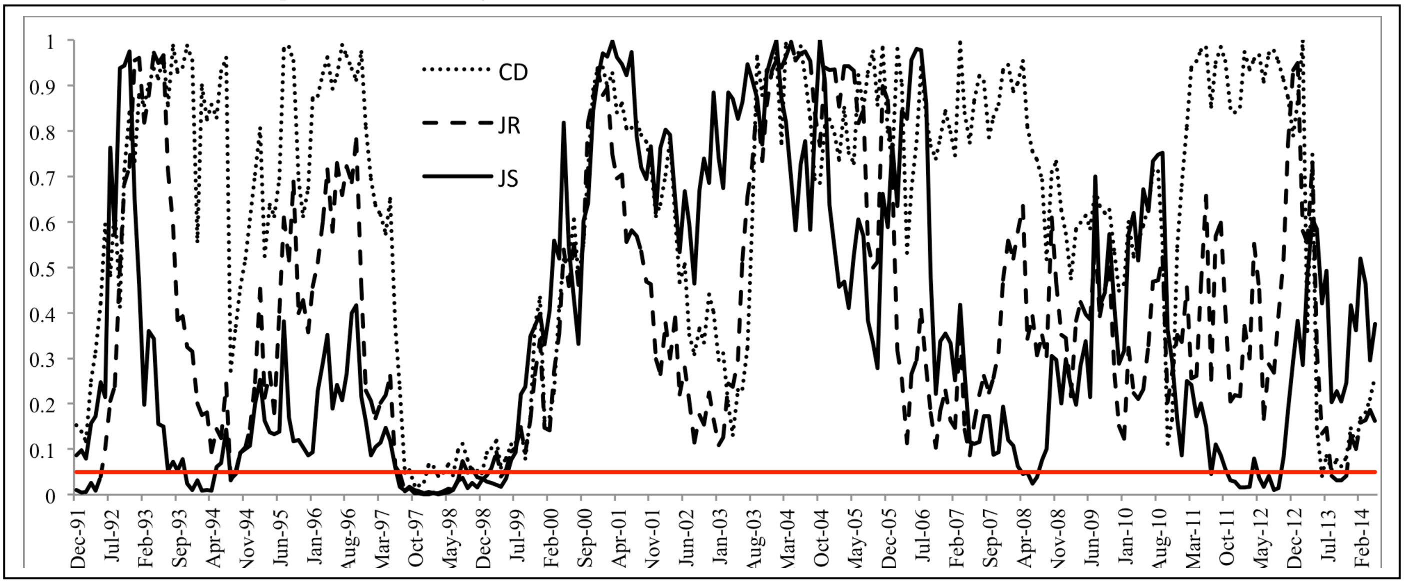

Figure 5: The three different variance ratio joint-test statistic p-values over time for the EURO STOXX 50 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

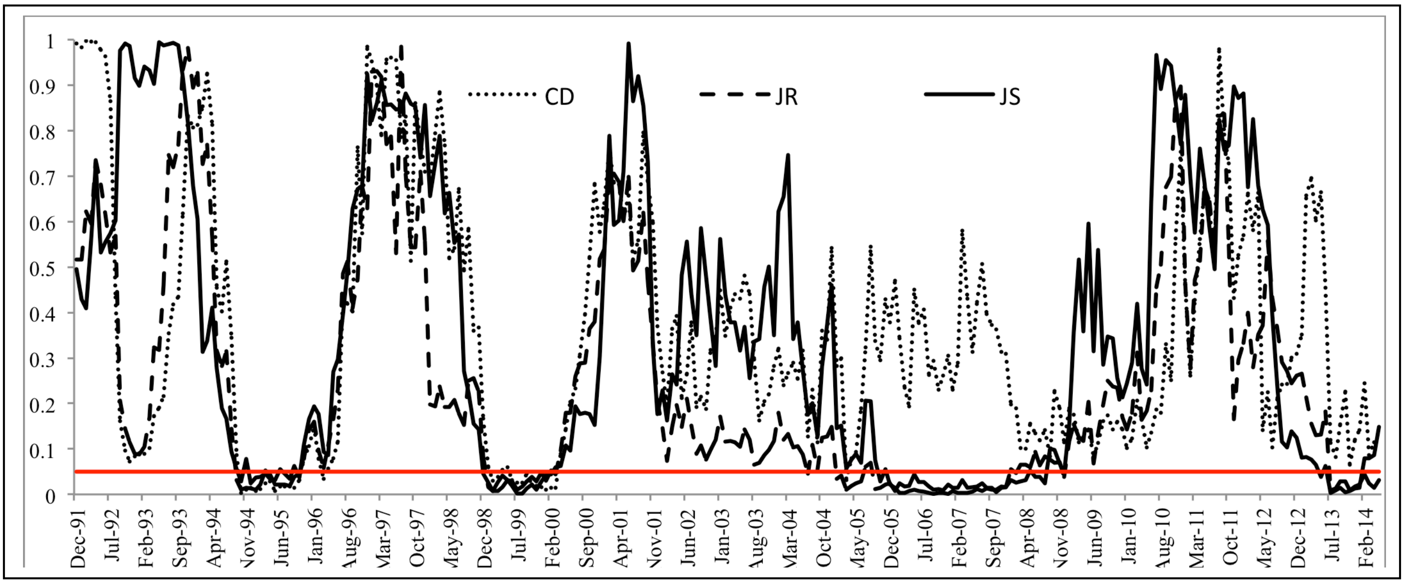

Figure 6: The average BDS statistic p-values over time for the S&P500 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

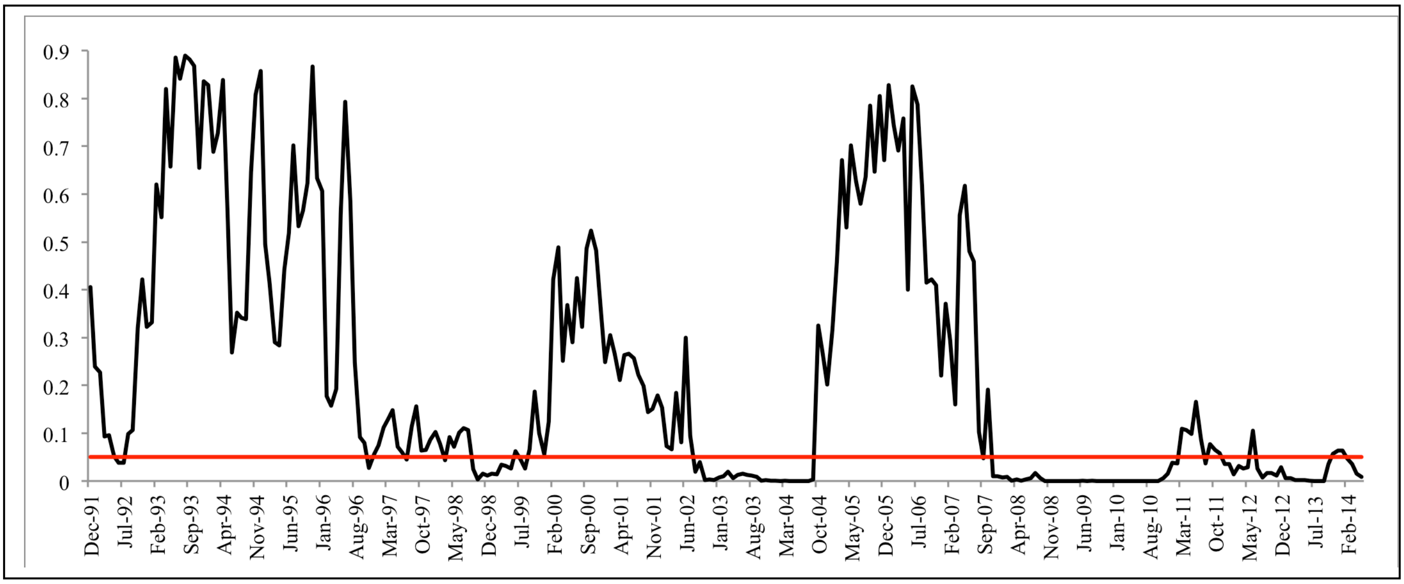

Figure 7: The average BDS statistic p-values over time for the FTSE100 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

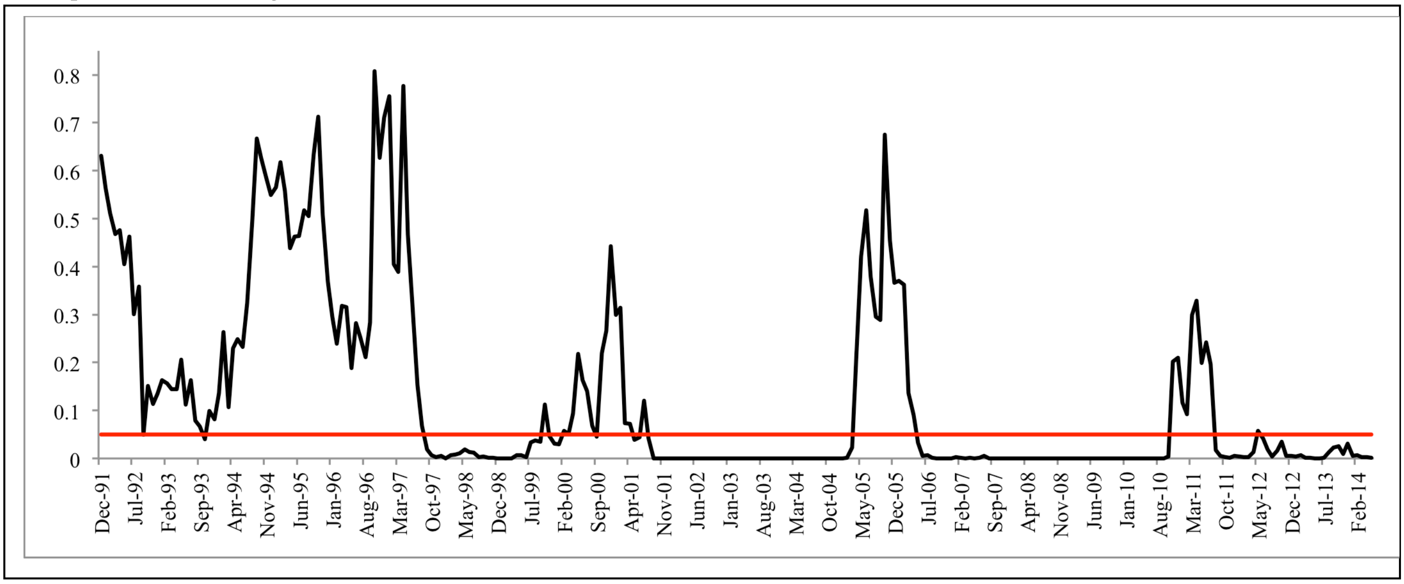

Figure 8: The average BDS statistic p-values over time for the NIKKEI225 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

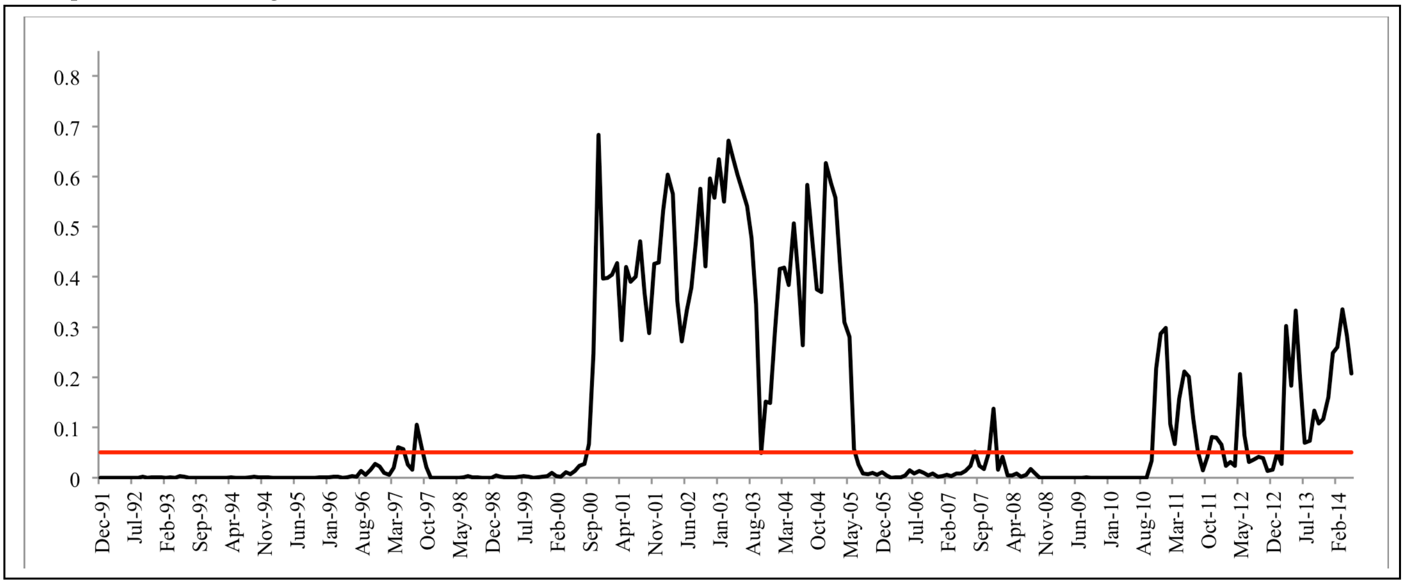

Figure 9: The average BDS statistic p-values over time for the EURO STOXX 50 (daily, two-year window). The horizontal line corresponds to the 5% significance level.

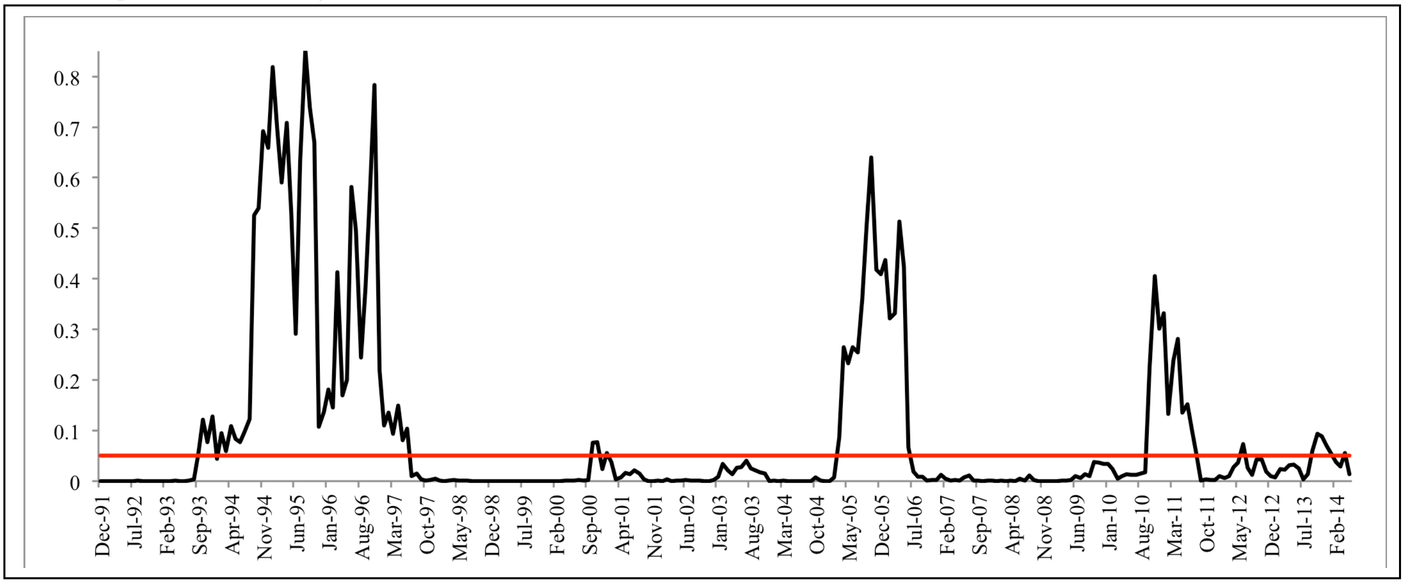

Table 1: Descriptive statistics of the daily returns of the S&P500, FTSE100 and EURO STOXX 50. ***, ** and * indicate significance at 1%, 5% and 10% respectively.

	Obs	Mean	SD	Skewness	Kurtosis	JB	ARCH(10)
S&P500	6151	0.000273	0.011503	-0.24	11.68	19376.9***	203.25***
FTSE100	6370	0.000163	0.011086	-0.12	9.44	11024.93***	181.96***
NIKKEI225	6007	-0.000162	0.015516	-0.13	8.26	6929.17***	130.83***
EURO STOXX 50	6278	0.000173	0.013534	-0.08	8.38	7575.74***	140.47***

Table 2: Relative efficiency. The percentage of test statistics that fails to reject the null hypothesis of market efficiency at the 5% significance level.

	CD	JR	JS	BDS	Average
S&P500	86.30%	76.30%	82.96%	59.26%	76.21%
FTSE100	84.44%	75.56%	88.89%	39.26%	72.04%
NIKKEI225	96.30%	88.52%	83.70%	34.81%	75.83%
EURO STOXX 50	90.00%	72.22%	75.19%	30.37%	66.95%

Table 3: Number of months characterized as market conditions.

	UP	DOWN	BULL	BEAR	NORMAL
S&P500	166	104	93	56	121
FTSE100	153	117	103	59	108
NIKKEI225	133	137	62	97	111
EURO STOXX 50	151	119	82	60	128

Table 4: Regression results of the p-values for the predictability tests and dummy variables for the states of the market. P-values are denoted in parentheses. ***, **, * indicate significance at 1%, 5% and 10% respectively.

		CD	JR	JS	BDS
S&P500	UP	-0.01328	-0.02283	0.00761	0.06580**
	DOWN	0.01328	0.02283	-0.00761	-0.06580**
	BULL	0.03004	0.03183	0.06858**	0.00355
	BEAR	0.00831	0.03765	0.00930	-0.11101***
	NORMAL	-0.03411	-0.05657***	-0.07101***	0.07532**
	Volatility	-101.30100	-58.13383	5.23938	-250.58***
FTSE100	UP	0.02939	0.01084	0.02088	-0.01154
	DOWN	-0.02939	-0.01084	-0.02088	0.01154
	BULL	0.06813***	0.03216	0.00459	0.03344
	BEAR	0.05380*	0.07527*	0.04974	-0.03518
	NORMAL	-0.10527***	-0.08518**	-0.03952	-0.00785
	Volatility	51.92574	72.24016	87.29663	-160.09670***
NIKKEI225	UP	0.02501	0.01180	0.01931	0.00652
	DOWN	-0.02501	-0.01180	-0.01931	-0.00652
	BULL	0.07442*	0.07335*	-0.02690	-0.01368
	BEAR	-0.07163*	-0.08703**	0.04537	0.06514***
	NORMAL	0.01373	0.02915	-0.02348	-0.05194**
	Volatility	-55.77214	10.39429	120.86**	-4.38824
EURO STOXX 50	UP	-0.04407	-0.04544	-0.10072***	0.01966
	DOWN	0.04407	0.04544	0.10072***	-0.01966
	BULL	-0.04070	-0.04254	-0.11403***	0.01928
	BEAR	-0.01178	0.01861	0.07355	-0.03577
	NORMAL	0.04268	0.02318	0.04573	0.00845
	Volatility	-113.93*	18.05372	82.30	-76.36*