Physica A 592 (2022) 126784

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Applying Hurst Exponent in pair trading strategies on Nasdaq 100 index

Quynh Bui [a,1,2], Robert Ślepaczuk [b,*,1,2]

[a] *Quantitative Finance Research Group, Faculty of Economic Sciences, University of Warsaw, Poland*
[b] *Quantitative Finance Research Group, Department of Quantitative Finance, Faculty of Economic Sciences, University of Warsaw, Ul. Długa 44/50, 00-241 Warsaw, Poland*

A R T I C L E I N F O

Article history:
Received 23 June 2021
Received in revised form 30 October 2021
Available online 22 December 2021

JEL classification:
C4
C14
C15
C52
C53
C58
G11
G12
G14

Keywords:
Generalized Hurst Exponent
Algorithmic trading strategies
Portfolio choice
Mean-reversion strategy
Pair trading
Correlation & cointegration trading
Efficient market hypothesis

A B S T R A C T

This research aims to seek an alternative approach to stock selection for algorithmic investment strategy. We try to build an effective pair trading strategy based on 103 stocks listed in the NASDAQ 100 index. The dataset has a daily frequency and covers the period from 01/01/2000 to 31/12/2018 , and to 01/07/2021 as an additional out-of-time dataset. In this study, Generalized Hurst Exponent, Correlation, and Cointegration methods are employed to detect the mean-reverting pattern in the time series of a linear combination of each pair of stock. The result shows that the Hurst method cannot outperform the benchmark, which implies that the market is efficient. These results are quite sensitive to varying number of pairs traded and rebalancing period but they are less sensitive to financial leverage degree. Moreover, the Hurst method is better than the cointegration method but is not superior as compared to the correlation method.

1. Introduction

A trading strategy refers to a designed plan to achieve risk-adjusted returns by going long or short on an asset or portfolio of assets in the market. Trading strategies have become very popular amongst investors. Examples of them are: long/short equity, pair trading, momentum/contrarian trading, etc. Institutional investors usually build more complex trading models which are confidential and not disclosed to outsiders. A trading strategy can be based on fundamental analysis, technical analysis, or both fundamental and technical analysis and additionally any kind of other financial theory.

A trading strategy can be characterized by momentum or a mean-reverting pattern. Momentum refers to the fact that the prices are trending in the same direction. The slow spread of new information is the main source of the momentum

* Corresponding author.
E-mail address: rslepaczuk@wne.uw.edu.pl (R. Ślepaczuk).

[1] The views presented in this text are those of the authors and do not necessarily represent the institutions of their employers.
[2] This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

https://doi.org/10.1016/j.physa.2021.126784

effect. In other words, the price is pushed in the same direction as more people buy or sell a stock when certain news, such as quarterly earnings is announced. In contrast, mean-reverting assumes that asset prices will be reverted to a reference price (usually mean price), so if the price is currently high, it is expected to decrease (revert to the mean) and vice versa. Numerous research shows that stock prices are close to random walk, which makes it difficult to model. However, under some special conditions and depending on the time horizon, stock prices can exhibit some degree of trending and mean-reversion.

Determining the price patterns — mean-reverting or momentum, under certain conditions and for a certain period is the essential core of building a trading model. Hurst exponent, which was originally developed in hydrology for the practical matter of determining the optimum size of the dam for the Nile river by Harold Edwin Hurst, can help us classify the pattern of time series of prices under a certain time horizon [1]. Hence, the Hurst exponent can serve as a criterion to select either momentum or mean-reverting strategy. As such, this paper introduces a method based on the Hurst exponent to effectively select pairs from a portfolio of stocks for a pair trading strategy. As we know, pair trading takes advantage of the fact that the linear combination of 2 assets' prices may be stationary, hence we can apply a mean-reverting strategy on the price spread of such pairs to make a profit. Other popular methods to select pairs are correlation and cointegration, so we measure the correlation between the returns of 2 assets or check the cointegration between 2 time series of asset prices and decide which pairs should be selected. In this research, we compare the performance of all 3 methods to verify which method is superior to others.

In that context, the paper addresses two main hypotheses: (H1) *Does the Efficient Market Hypothesis work? i.e. whether Hurst Exponent can generate a trading strategy that can beat the market*; and (H2) *Does the method to select pairs based on Hurst exponent outperform the methods based on correlation or cointegration?* Additionally, based on these two main hypotheses, a few research questions are constructed: (RQ1*) Is the result obtained robust to a varying number of pairs selected?* (RQ2) *Is the result obtained robust to varying rebalancing periods?* (RQ3) *Is the result obtained robust to varying degrees of financial leverage?*

In order to verify the main hypotheses and answer the research questions mentioned above, empirical research is conducted based on 103 stocks listed in NASDAQ 100 index. The period of our daily dataset is ranging from 01/01/2000 to 31/12/2018. There are three pairs selection methods that are applied in this paper: Hurst exponent, correlation, and cointegration. The Hurst exponent is calibrated using Generalize Hurst Exponent (GHE) method. For examining correlation, the Pearson correlation coefficient is employed and for examining cointegration, the Engle–Granger two-step method is implemented. Furthermore, the technique to generate a buy/sell signal on each pair of stocks is the breakout volatility model (BVM). The upper and lower thresholds of BVM are calculated based on the exponential moving average and rolling standard deviation of the price spread for each pair. The performance of our strategy is evaluated based on five criteria: annualized return compounded, annualized standard deviation, maximum drawdown, information ratio, and adjusted information ratio.

We expect that Hurst exponent can help build a trading model that can outperform the market. Besides, it is expected to generate a superior pair selection model as compared to the correlation and cointegration method. On the other hand, as the pair trading strategy has a market neutrality feature, we expect its cumulative returns to be better than the market in the recession period but worse than the market in the expansion period but overall its behavior should be independent of market downward or upward movements. In addition, the returns gained by pair trading should be less volatile than benchmark returns i.e. its standard deviation and the maximum drawdown will be lower than those of the market index.

The paper is structured as follows: Section 2 is an overview of literature about Hurst exponent, correlation, cointegration and the Efficient Market Hypothesis. Section 3 describes all the details concerning the dataset. Section 4 explains the methodology of Hurst exponent calibration as well as the methods to select pairs based on Hurst exponent and also how the pair trading strategy is constructed. Section 5 discusses the empirical results. Section 6 conducts the sensitivity analysis to determine whether the result is robust to changes in initial assumptions. Section 7 draws conclusions and suggests some extensions for future research.

2. Literature review

2.1. Hurst exponent

Hurst exponent is a statistical measure which is used to study scaling properties in time series. The scaling property in time series refers to the patterns in financial asset prices which are repeated at different time scales and is studied in a number of past researches. In 1930s, the work of Elliot has pointed out the appearance of price patterns at different time horizon. The Brownian motion, originally proposed by Bachelier [2], was one of the most popular models to study the scaling properties and has been developed and modified by several researchers such as the fractional Brownian motion [3,4] or Lévy motion [5,6].

Hurst exponent measures the long-term memory and fractality of a time series. It is known [7] that a Hurst value of 0.5 implies a Brownian random process. The Hurst value between 0.5 and 1 indicates trending (or momentum) behavior and a Hurst value between 0 and 0.5 indicates anti-trending (or mean-reverting) behavior of the time series. There are a number of techniques which can be exploited to study scaling properties such as the re-scaled range analysis (R/S analysis) and its complement [8]; the modified R/S analysis [9]; the detrended fluctuation analysis (DFA) [10], generalized Hurst exponent, etc. Hurst initially applied the R/S analysis to estimate long-term dependence of water levels in rivers and reservoirs. According to Di Matteo et al. [11], this method can disclose the long-run correlation in random processes and help distinguish non-correlated from correlated time series. However, if the time series presents short memory, heteroskedasticity and multiple scale behavior, the original R/S approach proposed by Hurst may cause a problem and is also sensitive to outliers as it is calculated based on maxima and minima. Hence, several alternative approaches have been proposed, one of which is the modified R/S analysis by [9] that can detect long-term memory in the presence of short-term memory. Besides, the generalized Hurst exponent explores the scaling properties based on qth order moments of distribution, and also includes in the algorithm many types of dependence in the data.

There are a number of papers that present the application of Hurst exponent in finance. According to Di Matteo et al. [12], a powerful tool to describe and differentiate the scaling structure of different markets is the generalized Hurst exponent H(q). The scaling structure of markets is an evolving quantity which can assist in differentiating different development stage among markets and portraying the overall fluctuation of market conditions. Qian and Rasheed [13] analyzed the Hurst exponent in many rebalancing periods for Dow-Jones index. The results suggest that the periods with large Hurst exponent can be predicted more accurately than those with Hurst value closes to 0.5 (random series), which implies that stock markets are not always random — in some periods it can present strong persistence pattern which is predictable. Kopeliovich [14] uses Hurst exponent as a predictability measure for VIX index. It is observed that Hurst exponent of VIX index is stable and consistently less than 0.5 over the long period of time, which means that VIX index is a mean-reversion time series and can be predicted. More recent paper focusing on multiscaling property of financial time series was published by Antoniades et al. [15]. They applied the visual methodology to time-series comprising of daily close prices of four stock market indices: two major ones (S&P 500 and Tokyo-NIKKEI) and two peripheral ones (Athens Stock Exchange general Index and Bombay-SENSEX). Their results show that multiscaling varies greatly with time: time periods of strong multiscaling behavior and time periods of uniscaling behavior are interchanged while transitions from uniscaling to multiscaling behavior occur before critical market events, such as stock market bubbles.

2.2. Correlation

The correlation is defined as a measure of the relationship between the changes of two or more financial variables over time. The most popular measure of correlation is Pearson correlation coefficient which measure the linear correlation between 2 variables and takes the value between -1 and $+1$, where -1 means perfectly negatively correlation; $+1$ means perfectly positively correlation and 0 means no correlation. Correlation plays a key role in Capital Asset Pricing Model (CAPM) as it indicates the potential for the diversification of the portfolio and such high diversification between assets improves the return/risk ratio.

The Correlation has a wide range of application, one of which is in portfolio management and risk assessment. Lee [16] applies Spurgin et al. [17] model, which is a model of correlation shifts with market return as independent variable, on 31 real estate market segments in the UK. The results show both significant negative and positive correlation shifts in different market segments. Depending on this piece of information, fund managers can effectively make decision on portfolio allocation based only on expected market volatility and can improve overall portfolio performance during different stages of real estate cycle. Kalotychou et al. [18] estimates the value of correlation dynamics in asset allocation using mean–variance analysis. They tested the correlation-timing framework models against industry correlation and found out how superior statistical properties of multivariate conditional correlation models account for different aspects of the correlation dynamics.

Another application of correlation is in trading strategy. Meissner [19] shows the analysis of opportunities and limitations of six correlation trading strategies which is most popular in financial practice, including Empirical Correlation Trading, Pairs Trading, Multi-asset Options, Structured Products, Correlation Swaps, and Dispersion trading. Most strategies associate with trading an underlying asset, so we have to control for the levels and volatilities of the underlying, except for pairs trading and variance dispersions which are pure correlation plays, hence they make correlation trading simpler. According to Riedinger [20], the two factors that influence profitability of pair trading strategy via the return per trade and trading frequency are pair volatility and correlation. There is a trade-off between the improvement of return per trade and trading frequency, that is, high pair volatility and low correlation are favorable for the return per trade but unfavorable for the trading frequency.

2.3. Cointegration

Cointegration, roughly speaking, describes the long-run relationship between two or more non-stationary but integrated at the same order time series, so we expect that any deviations from this long-run relationship are non-stationary. In a formal definition [21], time series x_t and y_t are cointegrated of order d; b (where $d \geq b > 0$) if: (1) both time series are integrated of order d and (2) there exists a linear combination of these variables, $a_1x_t + a_2y_t$, which is integrated of order d – b. It can be formally expressed as $x_t; y_t \sim CI(d; b)$ and vector $[a_1, a_2]$ is called cointegrating vector. A typical example for two cointegrated time series is a stock market index and its associated future contracts, both are integrated at order one I(1) but their linear combination forms a stationary time series I(0). Cointegration is one of the crucial characteristics in time series analysis.

Yule [22] first introduced and analyzed the concept of spurious or non-sense regression in 1926. Granger and Newbold [23] showed that using linear regression on non-stationary time series data is a dangerous approach that could lead to spurious correlation — untrue correlation. The paper by Granger and Engle in 1987 established the cointegrating vector approach and coined the term "cointegration". Granger and Newbold [23] presented that de-trending the integrated I(1) series does not solve the problem of spurious correlation and that checking for cointegration is a superior alternative as cointegration only exists if there exists a genuine relationship between two I(1) series. If we run a regression model on two I(1) time series to test their relationship, it is possible for the regression to generate spurious correlation.

There are 3 methods for cointegration test: Engle–Granger two-step method; Johansen test and Phillips–Ouliaris cointegration test. According to the Engle–Granger method, residuals are extracted from the static regression run on two non-stationary variables, then the residuals are tested for stationarity using ADF or KPSS test. We expect that the residuals will be stationary if the cointegration between two time series exists. The major issue in Engle–Granger method stays in the choice of dependent and independent variable when we run the regression to extract residuals as it may lead to inconsistent conclusions [24]. This issue may be addressed by alternative tests such as Phillips–Ouliaris and Johansen's. The Phillips and Ouliaris [25] residual-based unit root test is an enhancement over Engle–Granger test. Before the assumption underlying cointegration tests is that the regression residuals are independent from common variance, which is rarely true in practice [26]. Phillips–Oularis test resolves this issue by taking into account supplementary variability. The result generated from this test is also robust to the choice of dependent and independent variables. In this research, the Engle–Granger method is employed to recognize cointegration. The detail of this method is discussed later in the methodology section.

Cointegration is usually employed in trading strategy. Rad et al. [27] examines the performance of three pair trading strategies – the distance, cointegration and copula methods – on the US equity market in the period 1962–2014 with dynamic trading costs. They find that the monthly return of distance method is slightly higher than that of cointegration method, but the Sharpe ratio of cointegration method is a bit higher than that of distance method. Caldeira and Moura [28] proposed pair trading strategy for stocks of Sao Paulo stock exchange. The strategy is expected to make profit based on the cointegration between each pair of stocks. They apply cointegration test on all possible pair combinations to select the pairs that have long-term relationship, as the spread of those pairs will have mean-reverting pattern – a key factor to make pair trading strategy profitable.

2.4. *Efficient market hypothesis*

Efficient Market Hypothesis [29] states that markets are efficient in the sense that the current stock prices reflect completely all currently available information that could anticipate future prices, i.e. there is no information hidden that could be used to predict future market development. This model is based on the assumption that market changes can be represented by a normal distribution. The underlying implication of Efficient Market Hypothesis is that investors cannot consistently outperform the market on a risk-adjusted basis by employing any complex stock selection or market timing process. Investors can only achieve higher returns by chance or by investing in riskier assets. Only when the advancement in technology allows computers to work faster and the comparison of prices of hundreds of stocks became quick and efficient in 1960s, the efficient market hypothesis started to be popular.

The random walk hypothesis, which claims that stock market prices is a random walk process and cannot be predicted, has a close connection with EMH. There are several empirical researches that indicated that US stock prices and other financial time series followed random walk process in the short-terms although it may be predicted in the long-terms. However, the reason for this predictability is due to rational time-varying risk premia or behavioral reason still remains in debate. Moreover, professional investors in general cannot beat the market and keep earning excess returns [30].

There are three forms of market hypothesis: weak-form efficiency, semi-strong-form efficiency and strong-form efficiency. The main difference between those forms lies in the assumptions on how market works. In the weak-form efficiency, we cannot forecast future prices based on historical prices. In semi-strong form efficiency, it states that share prices quickly and unbiasedly reflect the publicly available information, so no one can earn excess return by trading on this piece of information. In strong-form efficiency, all public and private information are reflected in share prices, so no one can make profit from that information.

More recent attempts showing that the market is inefficient analyze the topic from the perspective of algorithmic investment systems. In such approaches, the buy/sell signals were based on both classic investment techniques enriched with new asset classes [31,32] and machine learning methods [33] showing that it is possible to build investment systems with abnormal risk-weighted returns.

3. Data description

The dataset includes 103 stocks of companies listed in NASDAQ 100 index and the NASDAQ 100 index, which is used as the benchmark for our strategy. The data frequency is daily. The period is from 01/01/2000 to 31/12/2018, so the sample size for each stock is 4778 observations, in total making 496 912 observations for the whole dataset. The data for each stock was carefully investigated and cleaned for outliers if any. The reason to select stocks pairs from NASDAQ 100 index was that it was one of the most volatile, diversified and well performing index during the last 30 years. We could base our research on the selection of pairs from the spectrum of equity indices, commodities, currencies, bond, cryptocurrencies, etc. but due to the fact that the research literature focused mainly on stock pairs and that we wanted to compare our research with other works we finally selected stock pairs. In sensitivity analysis we added the out-of-time data until 01/07/2021 which were not present in the time when research was prepared but were able to gather before the final paper was published.

3.1. Problem with survivorship bias

Our dataset includes only stocks which are listed in NASDAQ 100 index on 31/12/2018 because we cannot obtain the full constituents of NASDAQ 100 during the period 2000–2018 easily or it is too costly to get access to such information. All stocks which are delisted during the period 2000–2018 are excluded from the dataset, which may induce a problem called survivorship bias. We cannot fix this problem due to the lack of data and information regarding the constituents of NASDAQ 100 index, therefore we discuss the problem here and evaluate its impact on our result.

A historical database suffers from survivorship bias if it does not include stocks that disappeared because of bankruptcy, delisting, merger or acquisition. Hence, only surviving stocks stay in the database, leading to the overestimation of historical performance when we backtest a strategy based on such database. In particular, testing of mean-reverting strategy type such as pair trading can be vulnerable to the effects of survivorship bias. This is because mean-reverting strategy would short very-high-price stocks and long very-low-price stocks, but these stocks are likely to be either acquired (for high valued stocks) or go bankrupt (for low valued stocks), hence leading to loss in both cases. However, if the historical database is characterized by survivorship bias, these stocks may not be included in that database, thus overestimating the test performance.

There are a number of papers analyzing the impact of survivorship bias, especially on hedge fund performance. Lo [34] shows the analysis of impact of survivorship bias, stating that entire cross section of funds is influenced and the impact is magnified over time, hence it can result to significant misleading for investors that are not aware of this fact when building an optimal portfolio. Brown et al. [35] estimated the impact of survivorship bias on offshore hedge funds and recorded an annual attrition rate of 14% over the period from 1989 to 1995. The result shows that survivorship bias leads to 3% higher in annual returns. On the other hand, Ackerman et al. [36] evaluated only 0.2% survivorship bias. The gap in these estimates may be due to the difference in structure of databases and time horizon [37].

In the context of this paper, the exact effect of survivorship bias on the strategy performance cannot be characterized in details because it can affect the results of our market neutral strategy negatively and positively as well. However, when we compare the result with the benchmark, we can approximately evaluate whether the strategy really outperforms the benchmark by looking at the difference in the risk-adjusted returns.

4. Methodology

4.1. GHE methodology for calibration of Hurst exponent

GHE, a generalization of original method introduced by Hurst [1], is one of the most popular technique for Hurst exponent calibration. It is an effective tool that applies the scaling of qth-order moments of the increments distribution in order to investigate the scaling properties of time series [3,38]. Compared to Rescaled Range (R/S) analysis which takes into account the maxima/minima of data, it is less sensitive to the outliers. The Generalized Hurst Exponent verifies whether some statistical properties of data adjust with both number of observations and the time resolution.

To do this, we first define a time series $X(t)$ with $t = 1, 2, \ldots, k$. The statistic $K_q(\tau)$ is calculated according to the following formula:

$$K_q(\tau) = \frac{Mean(|X(t+\tau) - X(t)|^q)}{Mean(|X(t)|^q)} \tag{1}$$

where τ can vary between 1 and τ_{max}, while τ_{max} is usually chosen as a quarter of the length of the series.

The GHE characterizes the scaling properties of time series $X(t)$ and hence is associated with the scaling behavior of the statistic $K_q(\tau)$. Given the power-law, the GHE is calculated as:

$$K_q(\tau) \propto \tau^{qH(q)} \tag{2}$$

where $\propto$ denotes direct proportionality. Given two variables x and y, y is directly proportional to x if there is a non-zero constant k such that $y = kx$.

The GHE method is appealing for the fact that all of the information about scaling properties are reflected in the only one measure $H(q)$, which makes the analysis and interpretation of Hurst exponent simpler. The Hurst exponent is involved with different characteristics for each value of q. In specific, $H(q = 1)$, which should be close to the original Hurst exponent H, illustrates the scaling properties of absolute deviation of time series. On the other hand, the value of $H(q = 2)$ scales with the autocorrelation function of the increments $(C(t, \tau) = \langle X(t + \tau)X(t)\rangle)$ and has connection with the power spectrum and is crucial if we are interested in investigating long-range dependence [39]

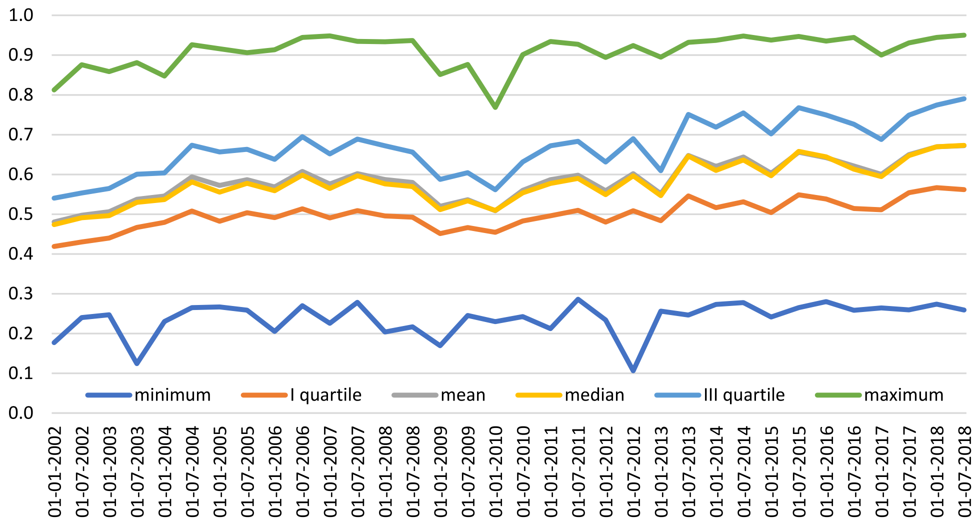

Fig. 1. Fluctuations of descriptive statistics for all Hurst exponents values estimated for all stock pairs every half year period.
Note: Each line represents the fluctuation of descriptive statistics (minimum, I quartile, mean, median, III quartile, maximum) for all Hurst exponents estimated individually for all pairs of stocks of every half year period during the research period (01.01.2002–01.07.2018).

In this paper, we choose $q = 1$ to calculate the Generalized Hurst Exponent as we are not concerned about the multi-fractal feature of GHE. In order to describe the typical values obtained for all Hurst exponents values estimated between all stock pairs every half year period Fig. 1 presents the fluctuations of descriptive statistics for such Hurst exponents. We can see that mean and median values are above 0.5 and that additionally all descriptive statistics are slightly increasing between the beginning and the end of our research period.

4.2. Pair selection based on Hurst exponent

From the 103 stocks of NASDAQ-100, we form pairs of stock, so the total number of possible pairs is 5253 pairs. For each possible pair, we calculate the spread according to the formula:

$$\text{Spread} = \ln(\text{P\_A}) - h * \ln(\text{P\_B}) \tag{3}$$

where P_A and P_B are the prices of stock A and B; and

$$h = \text{std}(\ln(\text{ret\_A}))/\text{std}(\ln(\text{ret\_B}))$$

where std is standard deviation and ret_A and ret_B are the log returns of stock A and B.

When we buy the pair, it means that for each share of A that we buy, we short sell h shares of B. According to the above formula to estimate h, we will have the same volatility for both the position in A and in B.

We then calculate the Generalize Hurst Exponent at q = 1 for each spread and select N pairs of stock whose spread has the lowest Hurst Exponent, which indicates the strong mean reversion characteristic of the spread - a key condition for our pair trading strategy to make profit. We can vary the number of pairs parameters (N) to estimate the robustness of our result. The selection of pair will be made every 6 months, so at the beginning of each 6 month, we conduct the calculation of GHE based on the data of past one-year (past 252 observations) and select 10 (N = 10) pairs with lowest Hurst exponent to be traded during that 6 months. This process will be repeated from the beginning of 2002 until the end of 2018.

4.3. Pair selection based on correlation

We employ the Pearson correlation to gain the correlation for the two series of log returns. The Greek letter ρ (rho) is commonly used to represent the Pearson's correlation coefficient of a population. Given a pair of random variables (X, Y), the formula for ρ is:

$$\rho_{X,Y} = \frac{cov(X, Y)}{\sigma_x \sigma_Y} \tag{4}$$

or

$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_x \sigma_Y} \tag{5}$$

where μ_X and μ_Y are the mean of X and Y and σ_x and σ_y are their standard deviation.

The Pearson sample correlation coefficient is commonly represented as r_{xy}. We can get a formula for r_{xy} by replacing the estimates of sample covariances and variances into the formula above. Given paired data $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ consisting of n pairs, r_{XY} is defined as:

$$r_{X,Y} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{6}$$

In this context, we would like to use the correlation formula for sample. Similar as pair selection based on Hurst exponent, we will conduct calculation of correlation based on the data of past one year (past 252 observations) at the beginning of each 6 months during the period 2002–2018. The 10 pairs that have highest correlation would be used in pair trading strategy during that 6 months.

4.4. Pair selection based on cointegration

Similar to Hurst exponent and Correlation, we select pair based on Cointegration between 2 time series of prices. We employ Engle–Granger two-step approach to test for cointegration. The pair selection is conducted every 6 months and 10 pairs with the lowest KPSS test statistics, which indicates higher chance of cointegration between 2 time series, would be used to trade during that 6 months. The detail of Engle–Granger approach is as follows:
Step 1 check whether 2 time series is integrated at the same order d using Augmented Dickey–Fuller test. For stock prices, they are normally integrated at order 1.

$$X, Y \sim I(d) \tag{7}$$

The ADF test is applied to the model:

$$\Delta y_{\mathrm{t}} = \alpha + \beta t + \omega . y_{\mathrm{t}-1} + \sum_{i=1}^{k} \delta_i \Delta y_{t-i} + \varepsilon_{\mathrm{t}} \tag{8}$$

The null hypothesis is that the series is non-stationary and alternative hypothesis is that the series is stationary:

$$H_0 : \omega = 0 \text{ (non-stationary)}$$
$$H_1 : \omega < 0 \text{ (stationary)}$$

The test statistic is calculated as follow:

$$DF_\tau = \frac{\hat{\omega}}{SE(\hat{\omega})} \tag{9}$$

As this is a lower-tail test, we reject the null when test statistic is lower than critical value (which is normally −2.86 for 5% significance level)
Step 2: Estimation of the cointegrating vector using standard OLS regression and testing whether residuals obtained in the regression are stationary using KPSS test.

$$Y_{\mathrm{t}} = \alpha + \beta X_{\mathrm{t}} + \varepsilon_{\mathrm{t}} \rightarrow \varepsilon_{\mathrm{t}} = Y\mathrm{t} - \alpha - \beta X_{\mathrm{t}} \tag{10}$$

$$\text{Check} : \varepsilon_{\mathrm{t}} \sim I(0) \tag{11}$$

KPSS stationary test is quite similar to ADF test but more powerful. The most important difference between KPSS and ADF test stays in the null hypothesis. For KPSS test, the null hypothesis is that the series is stationary and the alternative hypothesis is that the series is non-stationary. Hence, it is an upper-tail test and we reject the null when our test statistic is higher than the critical value, which is normally 0.463 for 5% significance level.

4.5. Trading strategies criteria

We use breakout volatility for each series of spread to generate signal to go long or short the pair. The upper and lower threshold are calculated as follow:

$$\text{Upper threshold} = \text{EMA}_{\text{n}} \text{ (spread)} + \text{m} * \text{rolling\_std}_{\text{k}} \text{ (spread)} \tag{12}$$

$$\text{Lower threshold} = \text{EMA}_{\text{n}} \text{ (spread)} - \text{m} * \text{rolling\_std}_{\text{k}} \text{ (spread)} \tag{13}$$

where EMA_n is the rolling Exponential Moving Averages of the spread of each pair with certain window size n ranging from 10 to 180 (days), $\text{rolling\_std}_\text{k}$ is the rolling standard deviation of the spread with certain window size k ranging from 10 to 120 (days), multiplier m takes a certain value ranging from 0.5 to 3.

Note that the window size of EMA (n), the window size of rolling standard deviation (k) and multiplier m are 3 parameters that will be optimized i.e. the set of parameters that give the best performance on historical data up to the moment of selection will be chosen. This optimization method would prevent our results from suffering from forward

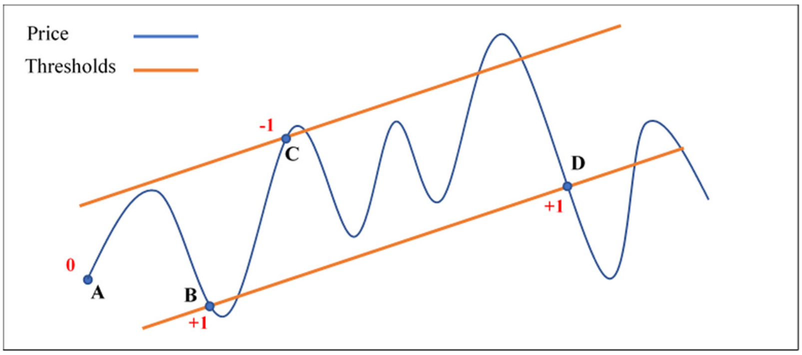

Fig. 2. Breakout volatility model illustration. Note: Point C and D indicate sell and buy signal based on the breakout volatility channel rule.

looking or look-ahead bias. Forward looking bias refers to the fact that we use the data or information from the future to backtest or analyze the current situation. Forward looking bias may artificially inflate our test result of trading strategy. In the context of this research, if we optimize parameters for breakout volatility model based on the whole sample of the dataset and then apply the optimized parameters to deduct the performance of our strategy, our result will be inaccurate as it suffers from look-ahead bias. To avoid this situation, we will use the data available up to the moment of selection to optimize parameters i.e. we use data of the first rebalancing period to get optimized parameters and apply these parameters for calculating performance statistics of second rebalancing period, then we use data of the first and second rebalancing period to get optimized parameters and apply these parameters for the third period, and so on.

For pair trading strategy, we will short the recent winners (the price is increasing) and long the recent losers (the price is decreasing) as we believe that their prices will behave according to mean-reverting pattern. It means that the current high price will go down and the current low price will go up in the future. In general, the rule for the entry and exit is as follow: if our position is being flat at time $t-1$, we will short the pair if the spread reaches upper threshold or long the pair if the spread reaches lower threshold at time t; if our position is being long at time $t-1$, we only switch to short position if the spread reaches the upper threshold at time t; if our position is being short at time $t-1$, we only switch to long position if the spread reaches the lower threshold at time t. Fig. 2 illustrates the rule of breakout volatility model as described above. We can see that initially at point A, our position is neutral (0), then we switch to long position (+1) when the price reaches lower threshold at point B, then we switch to short position (−1) when the price reaches upper threshold at point C, and then we keep short position and switch to long position when the price reaches lower threshold at point D.

The initial investment is assumed to be $10 000 at the beginning of rebalancing period. This investment is divided equally for N pairs, so each pair takes up $10 000/N. In our case, we have 10 pairs, so the amount of $1000 is invested in each pair. The transaction fee is assumed to be 0.01% of value of the pair. So, the transaction cost for trading 1 pair is 0.01% * (Price_A + h * Price_B). Moreover, we can also take advantage of financial leverage to improve the returns of pair trading strategy. So, instead of investing $10 000/N, we can invest $10 000/(N/2) or $10 000/(N/4) in each pair by borrowing money. The financial leverage can help us magnify the positive returns, but it also magnifies the negative returns if we make loss. Moreover, the double or quadruple times of degree of financial leverage will definitely make the equity curve more volatile.

For the benchmark, we just use buy & hold strategy on NASDAQ 100 index in the period 2002–2018 and compare our strategy's performance with this benchmark.

4.6. Summary of assumptions made for the purpose of trading strategy

Before investigating the result, we would like to describe the list of assumptions made in our methodology. It will give us a clear picture of the factors that may have impacts on our result and based on that we can then test the sensitivity of our results against the change of these factors (see Table 1),

4.7. Performance statistics

Based on the methodology presented in [40] we calculated the following performance statistics enabling us to evaluate tested algorithmic strategies.

Table 1
The list of initial assumptions and optimized parameters.

Parameters	Assumptions/parameters
Window size of EMA	To be optimized: {10, 20, 45, 60, 90, 120, 150, 180}
Window size of rolling std	To be optimized: {10, 20, 45, 60, 90, 120}
Multiplier m	To be optimized: {0.5, 1, 1.5, 2, 2.5, 3}
Rebalancing period	6 months
Number of pairs (N)	10 pairs
Initial investment	$ 10 000
Degree of Financial Leverage	100% (investment in each pair = Initial investment/N)
Spread of each pair	ln(P_A) – h * ln(P_B) (refer to Section 4.2)
Transaction cost	0.01% * (Price_A + h * Price_B) for trading 1 pair

Note: All optimized parameters were set on in-sample window. *Window size of EMA* — the number of days used for EMAn in calculation of Upper and Lower threshold of breakout volatility. Window size of rolling std — the number of days used rolling std in calculation of Upper and Lower threshold of breakout volatility. Multiplier m — the magnitude which will be multiplied by rolling std in calculation of Upper and Lower threshold of breakout volatility. Rebalancing period — the number of months after which we change the parameters of our models. Number of pairs (N) — how may pairs of stocks are selected for the trading purposes. Initial investment — the value of initial investment in USD for all pairs of stocks. Degree of Financial Leverage — what is the level of financial leverage used in the process of investment. Spread of each pair — how the spread was calculated for the trading purposes. Transaction cost — the level of transaction costs in percentage terms.

4.7.1. Annualized return compounded (ARC)

The Annualized Return Compounded is expressed as percentage (%) and calculated as:

$$ARC = \left(\prod_{t=1}^{N}(1+R_t)\right)^{\frac{252}{N}} - 1 \tag{14}$$

where:

- R_t is the percentage rate of return
- N is the sample size

4.7.2. Annualized standard deviation (ASD)

The Annualized Standard Deviation is expressed in the percentage terms (%) and calculated in the standard way:

$$ASD = \sqrt{252} \times \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}\left(R_t - \overline{R}\right)^2} \tag{15}$$

where:

- R_t is the percentage rate of return
- $\overline{R}$ is the average rate of return
- N is the sample size

4.7.3. Maximum drawdown (MD)

The Maximum Drawdown is the difference between the global maximum and the global minimum of the equity curve. Time order is important here, which means the global minimum must occur after the global maximum. It is expressed as percentage (%) and calculated as:

$$MD\,(S)_{t_1}^{t_2} = \max_{(x,y)\in\left\{[t_1,t_2]^2 : x\le y\right\}} \frac{S_x - S_y}{S_x} \tag{16}$$

where:

- S is the price process
- $[t_1, t_2]$ is the period between time t_1 and t_2

4.7.4. Information ratio (IR*)

Information Ratio is simply the ratio of ARC and ASD:

$$IR^* = \frac{ARC}{ASD} \tag{17}$$

Table 2
Performance of all methods and benchmark.

Method	ARC (%)	ASD (%)	IR*	MD (%)	IR**
Hurst method	3.14	8.37	0.374	15.02	0.078
Correlation method	2.97	**5.60**	0.531	**7.99**	**0.197**
Cointegration method	0.06	8.44	0.075	36.66	0.001
Buy & Hold (NASDAQ-100)	**13.36**	21.96	**0.609**	53.71	0.151

Note: ARC: Annualized Return Compounded; ASD: Annualized Standard Deviation; MD: Maximum Drawdown; IR* = ARC/ASD: Information Ratio; IR** = (ARC^2 * sign(ARC))/(ASD * MD): adjusted Information Ratio. Bold font indicates the best results in case of each performance statistics.

*4.7.5. Adjusted information ratio (IR**)*

Adjusted Information Ratio is similar to Information Ratio, but it also takes into account Maximum Drawdown as one of the risk factors:

$$IR^{**} = \frac{ARC^2 \times signARC}{ASD \times MD} \tag{18}$$

where:

− sign{ARC} is the sign of ARC and can take values of 0, −1 or +1

5. Empirical results and discussion

The performance of each method for pair trading strategy and benchmark are presented in Table 2. Looking at the result, we can see that pair trading strategy in general does not outperform benchmark strategy. In particular, the correlation method generates the highest Information Ratio (IR^*) among 3 methods with 0.531, the second one is the Hurst method with IR^* of 0.374 and the last one is the cointegration method with IR^* of 0.075, while the IR^* of benchmark is 0.609, which is higher than all of 3 methods. Although benchmark strategy gives the highest Annualized Return Compounded (ARC) with 13.36%, its Annualized Standard Deviation (ASD) is also the highest with 21.96%. For the 3 methods applied in pair trading strategy, Hurst method generates the highest ARC with 3.14% while correlation and cointegration method have lower ARC with 2.97% and 0.06% respectively. Three tested methods generate a much lower ASD compared to the benchmark with 5.60% for correlation, 8.37% for Hurst and 8.44 for cointegration method. In terms of adjusted Information Ratio (IR^{**}), we see a slightly different result. The correlation method with IR^{**} of 0.197 can beat the market with IR^{**} of 0.151. Hurst method and cointegration still cannot outperform the benchmark with IR^{**} of 0.078 and 0.001 respectively. What makes the difference in result is that in IR^{**}, we take into account Maximum Drawdown as a measure of risk beside ASD. We can see that the MD of benchmark (53.71%) is significantly higher than the MD of other 3 methods. Among 3 methods, the correlation generates the lowest MD with 7.99%, following the Hurst method with 15.02% and cointegration method with 36.66%.

In general, this result implies that we cannot build a trading strategy that can beat the market in terms of risk-adjusted returns although correlation method may beat the market if we additionally take into account MD as a measure of risk. This result refers to the first hypothesis of this paper, and we can conclude that the market is efficient according to this result.

On the other hand, we should also consider the fact that our result may be inflated due to survivorship bias. Besides, comparing the performance of 3 methods, we can see that Hurst method cannot outperform correlation but performs better than cointegration method. The cumulative returns of all 3 methods and benchmark are illustrated in Fig. 3.

Fig. 3 gives us more information of how each strategy performs in different market condition. We can see that in the bear market condition (market downturn), benchmark performs worse than Hurst and correlation methods, but still better than cointegration method. In bull market condition, benchmark perform the best, while correlation and Hurst methods performs better than cointegration method though still lower than benchmark. Besides, we can see that during the whole rebalancing period, pair trading strategy in general generates much lower volatility of returns compared to the benchmark.

6. Robustness test

In this section, we would verify how the result we obtained above is robust to varying number of pairs traded, varying rebalancing period and varying degree of financial leverage. In the result above, we trade 10 pairs during the period of 6 months with the initial investment of $10 000. To check the sensitivity of this result, we change the number of pairs traded to 5, 20 and 30 pairs, the rebalancing period to 3 months and 1 year and the financial leverage degree to 200% and 400% in each pair.

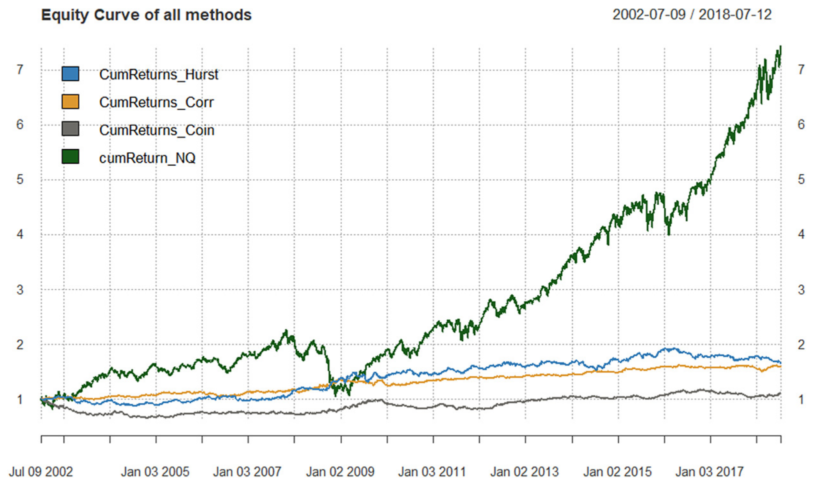

Fig. 3. Equity curve of all methods and benchmark.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 10. The rebalancing period was 6 months. Degree of financial leverage = 100%.

Table 3
Performance of all methods and benchmark — 5, 10, 20 and 30 pairs.

Method	No. of pairs	ARC (%)	ASD (%)	IR*	MD (%)	IR**
Hurst method	5	−1.55	10.30	−0.150	55.06	−0.004
Correlation method	5	5.89	7.25	**0.811**	11.98	**0.399**
Cointegration method	5	−4.84	11.77	−0.411	57.31	−0.035
Hurst method	10	3.14	8.37	0.374	15.02	0.078
Correlation method	10	2.97	5.60	0.531	**7.99**	0.197
Cointegration method	10	0.06	8.44	0.075	36.66	0.001
Hurst method	20	1.11	6.28	0.176	16.61	0.012
Correlation method	20	−1.53	**4.29**	−0.356	26.12	−0.021
Cointegration method	20	−0.08	5.89	−0.130	28.06	−0.004
Hurst method	30	2.41	5.72	0.421	13.00	0.078
Correlation method	30	3.08	4.33	0.712	13.10	0.168
Cointegration method	30	1.00	4.95	0.203	23.12	0.009
Buy & Hold (NASDAQ-100)	_	**13.36**	21.96	0.609	53.71	0.151

Note: ARC: Annualized Return Compounded; ASD: Annualized Standard Deviation; MD: Maximum Drawdown; IR* = ARC/ASD: Information Ratio; IR** = (ARC^2 * sign(ARC))/(ASD * MD): adjusted Information Ratio. Bold font indicates the best results in case of each performance statistics.

6.1. Varying number of pairs traded

Table 3 presents the performance of all three methods with 5 pairs, 10 pairs (the original number of pairs), 20 pairs and 30 pairs. As we can see in Table 3, with 5 pairs, the performance of Hurst and Cointegration method cannot beat the market with ARC of −1.55% and −4.84%; ASD of 10.30% and 11.77%; IR* of −0.150 and −0.411 and IR** of −0.004 and −0.035 respectively, while correlation method performs quite well and can beat the market with IR* of 0.811 and IR** of 0.399 (ARC = 5.89%, ASD = 7.25% and MD = 11.98%). With 20 pairs, no method can beat the market. Especially, only Hurst method can make positive profit with IR* of 0.176 and IR** of 0.012. Both correlation and cointegration methods lead to negative profit with IR* of −0.356 and −0.130 and IR** of −0.021 and −0.004 respectively. With 30 pairs, we can observe that the performance of all 3 methods improve more compared to other number of pairs. In particular, correlation method works well with IR* of 0.712 and IR** of 0.168, which outperforms the market. Hurst and cointegration methods' performance, though better, are still lower than the benchmark with IR* of 0.421 and 0.203 and IR** of 0.078 and 0.009 respectively.

From all of these results, we can conclude that all three methods seem to be sensitive to the number of pairs traded. However, it is not clear whether the performance of all three methods improves with the greater number of pairs. As we can see, Hurst and cointegration method perform better with 10 and 30 pairs and worse with 5 and 20 pairs, while correlation method works especially well with 5 and 30 pairs and quite poor with 20 pairs.

The result we obtained here does not correspond with the result in other papers, such as Alexandra and Dimitriu [41] where they state that cointegration method's performance is superior as compared to the market; or in [27] where

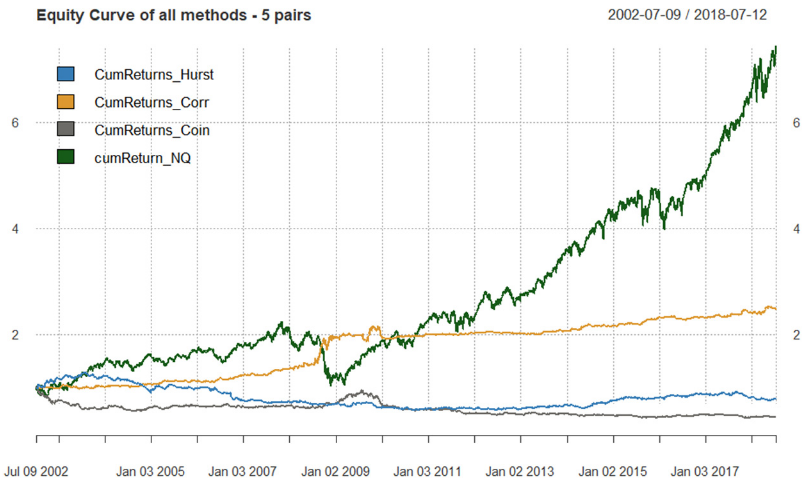

Fig. 4. Equity curve of all methods and benchmark - 5 pairs.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 5. The rebalancing period was 6 months. Degree of financial leverage = 100%.

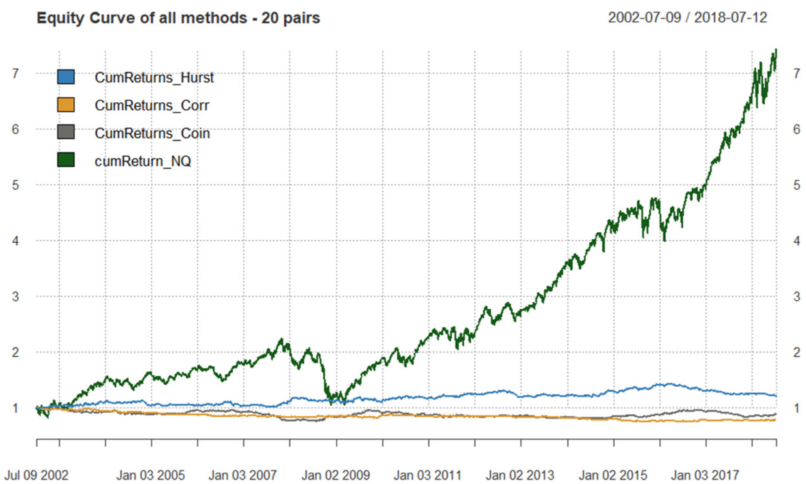

Fig. 5. Equity curve of all methods and benchmark — 20 pairs.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 20. The rebalancing period was 6 months. Degree of financial leverage = 100%.

they claim that the cointegration method is the superior strategy during turbulent market conditions. Similarly, Ramos-Requena et al. [39] found that Hurst method performs better than the correlation method. However, our result, together with the results in other papers, confirms that pair trading strategy performs well and better than the market in the period of market crash, and confirms the market neutral characteristics of pair trading strategy.

The equity curves of all methods and the benchmark with 5, 20 and 30 pairs are plotted in Figs. 4–6 respectively.

In general, they all show that pair trading strategy with all methods (Hurst, correlation and cointegration) work roughly the same or poorer compared to the benchmark in market downturn (except for the case with 5 pairs for the correlation method) and always worse in expansion stage. Besides, with 5 pairs, the correlation method's performance seems to be outstanding compared to cointegration and Hurst method in both bull and bear market condition, and it works much better than the market in crisis period. With 20 and 30 pairs, no method can outperform the benchmark in any market conditions. Cointegration method seems to perform poorer than the other two methods for any number of pairs.

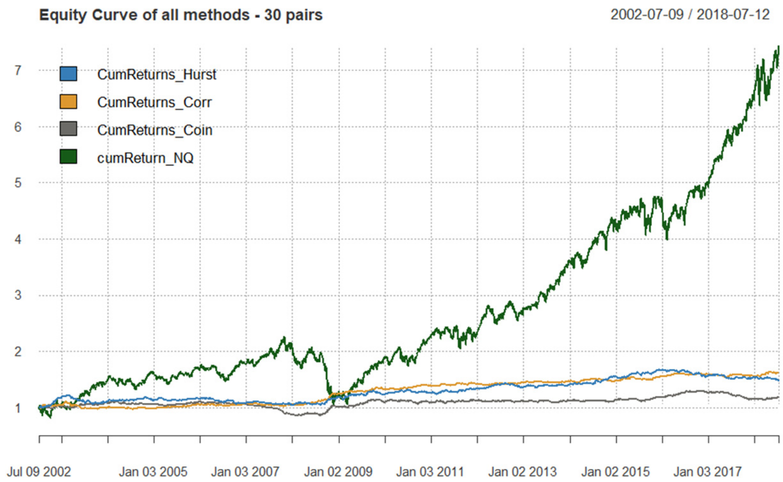

Fig. 6. Equity curve of all methods and benchmark — 30 pairs.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 30. The rebalancing period was 6 months. Degree of financial leverage = 100%.

Table 4
Performance of all methods and benchmark — 3 months, 6 months and 1 year.

Method	Reb.period	ARC (%)	ASD (%)	IR*	MD (%)	IR**
Hurst method	3 months	1.63	9.04	0.180	26.62	0.011
Correlation method	3 months	2.72	5.16	0.527	6.79	**0.212**
Cointegration method	3 months	−0.28	8.11	−0.035	**3.12**	−0.0003
Buy & Hold (NASDAQ-100)	_	10.55	22.45	0.470	53.71	0.092
Hurst method	6 months	3.14	8.37	0.374	15.02	0.078
Correlation method	6 months	2.97	5.60	0.531	7.99	0.197
Cointegration method	6 months	0.06	8.44	0.075	36.66	0.001
Buy & Hold (NASDAQ-100)	_	**13.36**	21.96	0.609	53.71	0.151
Hurst method	1 year	0.40	7.93	0.051	26.06	0.001
Correlation method	1 year	1.77	**4.98**	0.357	8.49	0.075
Cointegration method	1 year	−1.47	9.93	−0.148	3.42	−0.006
Buy & Hold (NASDAQ-100)	_	12.95	20.79	**0.623**	53.71	0.150

Note: Reb.period—rebalancing period. ARC: Annualized Return Compounded; ASD: Annualized Standard Deviation; MD: Maximum Drawdown; IR* = ARC/ASD: Information Ratio; IR** = (ARC^2 * sign(ARC))/(ASD * MD) : adjusted Information Ratio. Bold font indicates the best results in case of each performance statistics.

6.2. Varying rebalancing period

Table 4 demonstrates the performance of all three methods with rebalancing period of 3 months, 6 months (the original rebalancing period) and 1 year, and the number of pairs traded is fixed at 10 pairs for all 3 methods. As we can see in Table 4, with 3 months, only the performance of correlation method (with IR* = 0.527 and IR** = 0.212) outperforms the benchmark (with IR* = 0.470 and IR** = 0.092), while Hurst method, though gain positive result, cannot beat the market with IR* = 0.180 and IR** = 0.011 and cointegration method gains negative result with IR* = −0.035 and IR** = −0.0003. With 1 year rebalancing period, the correlation method has the highest performance with IR* = 0.357 and IR** = 0.075, following the Hurst method with IR* = 0.051 and IR** = 0.001, and cointegration method has lowest performance with IR* = −0.148 and IR** = −0.006. Hence, we can see that no method can beat the market (with IR* = 0.623 and IR** = 0.250) with 1 year rebalancing period.

From all of these results, we can observe that all three methods are quite sensitive to different rebalancing period, but the correlation method seems to be more stable than the other two as it keeps earning positive profit with different rebalancing periods, though it cannot beat the market with longer rebalancing period (6 months and 1 year). The performance of Hurst method drops significantly with 1 year and 3-month rebalancing period as compared to 6-month period. For cointegration method, only 6-month rebalancing period earns positive profit while 3-month and 1 year period earn negative profit. The result we obtained here again does not agree with the result in [41] or in [27] or [39] as we mentioned above.

Figs. 7 and 8 illustrate the equity curves of all strategies and benchmark with 3-month and 1 year rebalancing period respectively. For 3-month rebalancing period, correlation and cointegration methods outperform the benchmark while

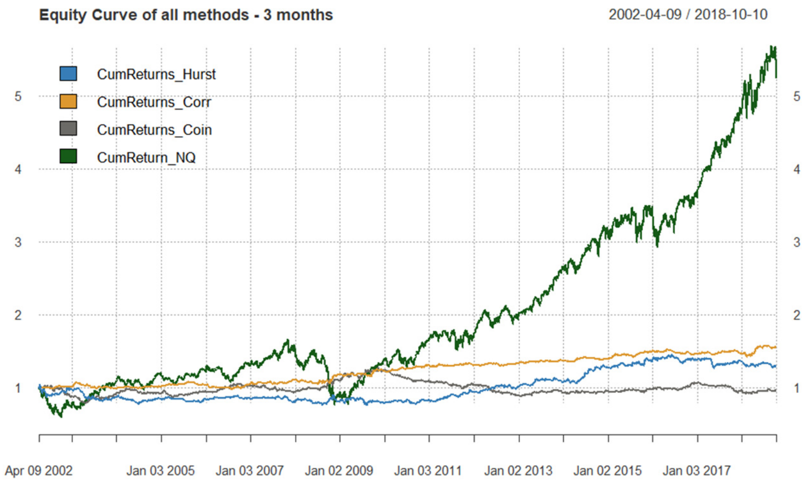

Fig. 7. Equity curve of all methods and benchmark - 3 months.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 10. The rebalancing period was 3 months. Degree of financial leverage = 100%.

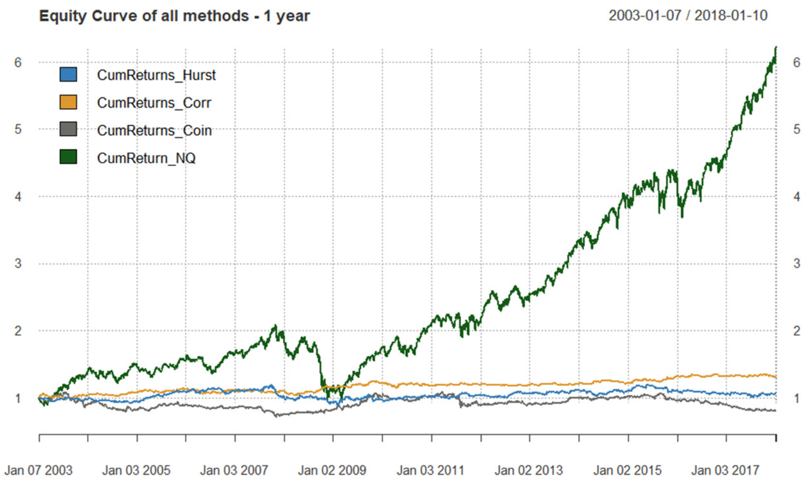

Fig. 8. Equity curve of all methods and benchmark - 1 year.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 10. The rebalancing period was 1 year. Degree of financial leverage = 100%.

Hurst method performs roughly the same as the benchmark in bear market condition. For 1 year rebalancing period, all three methods perform the same or a little bit lower than the benchmark in bear market condition. In the bull market condition, no method can beat the market for both 3-month and 1 year rebalancing period. However, in general, correlation method seems to perform better than the other two methods in all market conditions.

6.3. Varying degree of financial leverage

Table 5 shows the result for varying degree of financial leverage, so instead of invest \$10 000/N in each pair, we try to invest \$10 000/(N/2) and \$10 000/(N/4) in each pair, taking into account that we can earn higher return with reasonable level of volatility. The number of pairs is fixed at 10 and rebalancing period is fixed at 6 months. With \$10 000/(N/2) investment in each pair, the IR* of correlation, Hurst and cointegration methods are 0.327, 0.513 and 0.069 respectively, which cannot beat the market with IR* = 0.609. We can see that with double degree of financial leverage, we gain almost double ARC for each method (5.47% for Hurst, 5.69% for Correlation and 1.17% for Cointegration), but we also get almost double ASD (16.71% for Hurst, 11.10% for Correlation and 17.13% for Cointegration), in total making our IR* not improve at all. In terms of IR**, only Correlation method can beat the market with IR** = 0.188, both Hurst and cointegration method

Table 5
Performance of all methods and benchmark - N/2, N/4 and N.

Method	DFL	ARC (%)	ASD (%)	IR*	MD (%)	IR**
Hurst method	$10000/(N/4)	8.66	33.35	0.260	52.24	0.043
Correlation method	$10000/(N/4)	**13.76**	21.79	**0.631**	28.40	**0.305**
Cointegration method	$10000/(N/4)	−3.14	38.43	−0.082	91.59	−0.003
Hurst method	$10000/(N/2)	5.47	16.71	0.327	28.74	0.062
Correlation method	$10000/(N/2)	5.69	11.10	0.513	15.52	0.188
Cointegration method	$10000/(N/2)	1.17	17.13	0.069	62.73	0.001
Hurst method	$10000/N	3.14	8.37	0.374	15.02	0.078
Correlation method	$10000/N	2.97	**5.60**	0.531	**7.99**	0.197
Cointegration method	$10000/N	0.06	8.44	0.075	36.66	0.001
Buy & Hold (NASDAQ-100)	_	13.36	21.96	0.609	53.71	0.151

Note: ARC: Annualized Return Compounded; ASD: Annualized Standard Deviation; MD: Maximum Drawdown; IR* = ARC/ASD: Information Ratio; IR** = (ARC^2 * sign(ARC))/(ASD * MD) : adjusted Information Ratio. Bold font indicates the best results in case of each performance statistics.

Table 6
Performance of all methods and benchmark in additional OOS period.

Method	ARC (%)	ASD (%)	IR*	MD (%)	IR**
Hurst method	2.04	8.87	0.230	34.00	0.014
Correlation method	2.65	5.84	0.454	11.19	0.108
Cointegration method	2.09	8.55	0.244	36.66	0.014
Buy & Hold (NASDAQ-100)	15.34	22.74	0.675	53.71	0.192

Note: ARC: Annualized Return Compounded; ASD: Annualized Standard Deviation; MD: Maximum Drawdown; IR* = ARC/ASD: Information Ratio; IR** = (ARC^2 * sign(ARC))/(ASD * MD): adjusted Information Ratio. Bold font indicates the best results in case of each performance statistics.

with IR** of 0.062 and 0.001 cannot beat the market. Similar to IR*, we get almost double ARC but also double ASD and MD for each method, overall making IR** not better.

With $10 000/(N/4) investment in each pair, the Information Ratio (IR*) of correlation method can outperform the market with 0.631, which is mainly because correlation method can earn higher ARC with lower ASD compared to the market. For Hurst method, the IR* of 0.260 cannot beat the market and is also lower than the original performance (without financial leverage) as it earns 4 times ARC (8.66%) but also has more than 4 times ASD (33.35%). For cointegration, it has negative profit with ASD of −3.14%, which is the cost of employing financial leverage as our loss is amplified as compared to not using financial leverage, together with very high ASD (38.43%), in total leading to IR* of −0.082. In terms of adjusted Information Ratio (IR**), correlation method can 2 times outperform the benchmark with IR** of 0.305 as its MD (28.40%) is as half of the MD of benchmark (53.71%), while Hurst and cointegration method still perform poorer with IR** of 0.043 and −0.003 respectively as their MD (53.34% for Hurst and 91.59% for cointegration) are almost the same or even higher than that of benchmark.

From the analysis of the result in Table 5, we can see that the return of all methods improves significantly with higher degree of financial leverage, but it also results in higher standard deviation and maximum drawdown. Overall, only correlation method can outperform the market while Hurst and cointegration methods cannot. Hence, we can say that the result we gain is less sensitive to the degree of financial leverage than to the number of pairs or rebalancing period as we discussed above. This result again does not correspond with other papers as we mentioned above where they state that cointegration method is superior as compared to the market [41] or Hurst method can outperform correlation method [39].

Figs. 9 and 10 plots the equity curves of all three methods and the benchmark for double degree of financial leverage and quadruple degree of financial leverage respectively. Both of the graphs show that in recession period, Hurst and correlation methods perform better than the benchmark while cointegration performs poorer. In expansion period, all three methods perform lower than the benchmark for double degree of financial leverage, but Hurst and correlation method performs significantly better for quadruple degree of financial leverage. However, as we indicated above, their equity curves are characterized by much higher volatility, which is shown clearly in Fig. 10.

6.4. Additional out-of-time results

As an additional part of sensitivity analysis we decided to add the results based on the newest data for the period between the end of the first version of our research and the moment of publication of our paper, i.e. 31/12/2018 – 01/07/2021. Table 6 presents the results of all tests methods for th whole data period from 01/01/2002 to 01/07/2021 while the fluctuations of equity curves and benchmark are presented in Fig. 11.

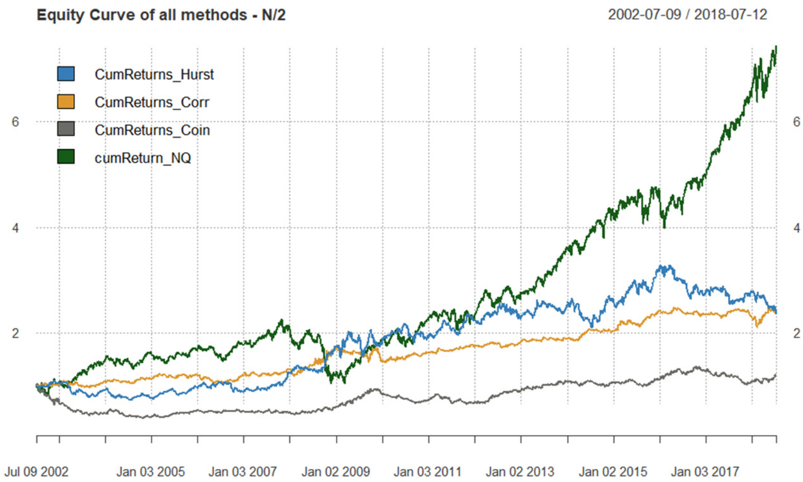

Fig. 9. Equity curve of all methods and benchmark - N/2.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 10. The rebalancing period was 6 months. Degree of financial leverage = 200%.

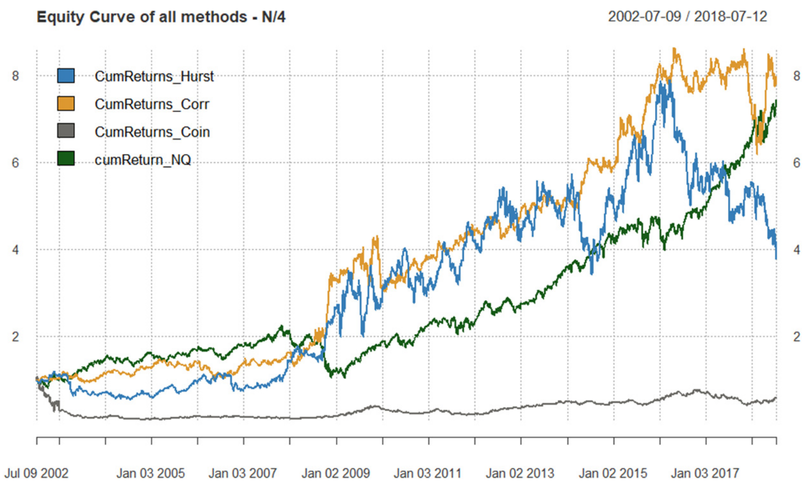

Fig. 10. Equity curve of all methods and benchmark - N/4.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 31/12/2018. The number of pairs traded was 10. The rebalancing period was 6 months. Degree of financial leverage = 400%.

Based on the results presented in Table 6 and Fig. 11 we can say that the additional 2.5 year did not change them a lot. Correlation method is still the best one among all three tested pair trading strategies (IR*=0.454 in comparison to 0.23 and 0.244) and is a little bit worse than the benchmark (IR*=0.675 for Nasdaq100). These differences between pair trading strategies and Nasdaq100 are confirmed when we take into account IR** (0.192 for Nasdaq100 is still better then 0.108 for Correlation method which is much better then for two other pair trading strategies).

These results confirm that the methodology used was not dependent on the selected data and even on new data period the initial results still hold.

7. Conclusion

This paper aims to explore an alternative method of pair selection for a pair trading strategy. The main hypotheses of this paper are (H1) *whether the Efficient Market Hypothesis work i.e. whether the Hurst exponent can generate a trading strategy that outperforms the market*, and (H2) whether the Hurst exponent is a superior pair selection model for pair trading strategy as compared to correlation and cointegration method. Based on these two hypotheses, the research

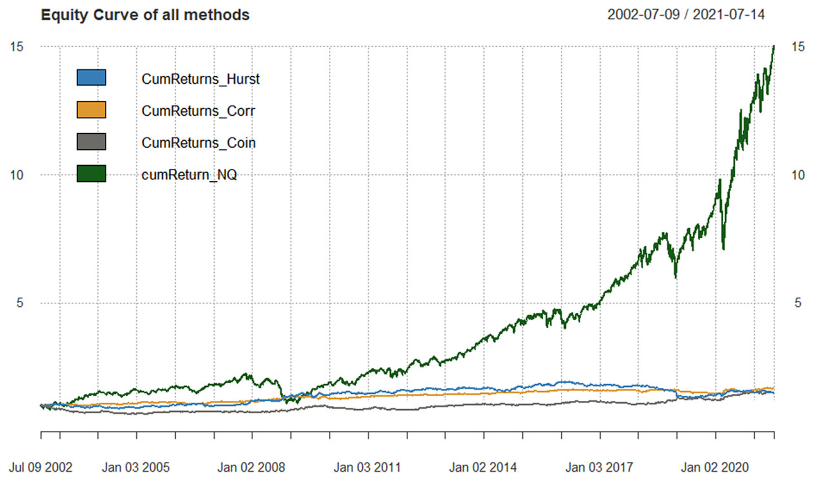

Fig. 11. Equity curve of all methods and benchmark in additional OOS period.
Note: The result of pair trading strategy with signals based on Hurst exponent, correlation and cointegration methods compared with the results of buy & hold strategy for NASDAD-100. The trading period started on 01/01/2002 and ended on 01/07/2021. The number of pairs traded was 10. The rebalancing period was 6 months. Degree of financial leverage = 100%.

questions are constructed as follows: *whether our results are robust to* (RQ1) *varying number of pairs selected*, (RQ2) *varying rebalancing period*, and (RQ3) *varying degree of financial leverage.*

The dataset used for this empirical research is 103 stocks listed in NASDAQ-100 (updated until 31/12/2018) and the NASDAQ-100 index itself. The data is collected on a daily basis over the period from 01/01/2000 to 31/12/2018. The pairs are established from the pool of these 103 stocks. We take advantage of the Hurst exponent, which can be used to classify the pattern of time series (mean-reverting, momentum, or random walk). Based on that, we select the pairs with the lowest Hurst exponent, indicating a mean-reverting pattern, which is a key condition to be profitable in pair trading strategy. Generalized Hurst exponent methodology with first moment (q = 1) is applied to calibrate Hurst exponent for each pair. For the correlation method, we apply the Pearson correlation coefficient to estimate the correlation between 2 series of log returns, and for the cointegration method, we employ Engle–Granger's two-step approach to test for cointegration between 2 time series of prices. To generate buy and sell signals for the pair trading strategy, a volatility breakout model is applied with upper and lower thresholds calculated based on the exponential moving average and rolling standard deviation of the time series. The window size of the exponential moving average, the window size of rolling standard deviation, and multiplier *m* are parameters to be optimized i.e. the set of parameters that generate the best performance on in-sample data will be selected. The performance of the Hurst method is compared with the benchmark (buy & hold NASDAQ-100 index) as well as with the correlation and the cointegration methods in terms of annualized return compounded, annualized standard deviation, maximum drawdown, Information Ratio, and adjusted Information Ratio.

Overall, the result shows that the Hurst method cannot generate a strategy that can outperform the market and its performance is superior as compared to the cointegration method but it is not as compared to the correlation method. This result does not correspond with the result of some other papers in the past, which states that the cointegration method is superior compared to other methods. On the other hand, the result implies that the market is efficient, so we cannot build a trading strategy and earn an excess return on the market, which corresponds with the efficient market hypothesis, though the question of whether the market is efficient or not still remains debated in the literature. Moreover, pair trading strategy in case of all methods can sometimes outperform the market in recession stage but cannot in the expansion stage, which confirms the market neutrality of pair trading strategy. However, our result is quite sensitive to the different number of pairs traded and rebalancing period and less sensitive to the financial leverage degree. In specific, it is not clear whether the Hurst method's performance improves with the greater number of pairs or longer rebalancing period. On the other hand, it is clear that all methods' returns and volatility are greater with a higher degree of financial leverage, hence the information ratio does not improve, except for the correlation method with the quadruple degree of financial leverage.

To conclude, from the obtained result, we can state that the market is efficient in general (first hypothesis) and Hurst's method of pair selection is superior as compared to the cointegration method but is not as compared to the correlation method (second hypothesis). Furthermore, the obtained result is sensitive to varying number of pairs (first research question) and varying rebalancing period (second research question) and less sensitive to varying degree of financial leverage (third research question). The results are summarized in Table 7.

There are some limitations of this paper that can be improved in future work. First, we should mention the impact of survivorship bias due to the lack of full historical data. It may inflate the performance of all methods, especially when

Table 7
The reference to research hypotheses and questions.

	Verification
RH1	No reason to reject
RH2	Hurst's method of pair selection is better than the cointegration method but is worse than the correlation method
RQ1	The obtained result is sensitive to varying number of pairs
RQ2	The obtained result is sensitive to varying rebalancing period
RQ3	The obtained result is less sensitive to varying degree of financial leverage

Note: The detailed content of each research hypotheses and questions are presented in the Introduction section.

we employ a mean-reverting pattern in a pair trading strategy. This issue may be addressed if we get a full constituent of the NASDAQ 100 index over the period 2000–2018. Second, we use the Pearson correlation coefficient to calculate the correlation between 2 series of log returns, which may suffer some limitations. In specific, Pearson correlation only measures linear relationships, but in practice, most stock price movements have non-linear dependencies. Besides, the correlation coefficient varies with different time frames and hence leads to different estimates of correlation parameters. Regression is also sensitive to outliers, so spurious correlations can occur [19]. To solve this issue, we can try to use other conditional correlation models such as the dynamic conditional correlation (DCC) model of Engle [42] or the time-varying correlation (TVC) model of Tse and Tsui [43]. Third, we apply the Engle–Granger two-step approach to test for cointegration of two time series. However, the major issue in the Engle–Granger method stays in the choice of the dependent and independent variable when we run the regression to extract residuals as it may lead to inconsistent conclusions [24]. This issue may be addressed by alternative tests such as Phillips–Ouliaris and Johansen's.

CRediT authorship contribution statement

Quynh Bui: Software, Data curation, Methodology, Writing – original draft, Formal analysis, Visualization, Investigation, Conceptualization. **Robert Ślepaczuk:** Conceptualization, Data curation, Methodology, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] H.E. Hurst, Long-term storage capacity of reservoirs, Trans. Am. Soc. Civ. Eng. 116 (1951) 770–808.
[2] L. Bachelier, Theory of speculation (translation of 1900 French edition), in: P.H. Cootner (Ed.), The Random Character of Stock Market Prices, MIT Press, Cambridge, MA, 1900, pp. 17–78.
[3] B.B. Mandelbrot, Fractals and Scaling in Finance, Springer, New York, 1997.
[4] P.K. Clark, A subordinated stochastic process model with finite variance for speculative prices, Econometrica 41 (1973) 135–155.
[5] B.B. Mandelbrot, The variation of some other speculative prices, J. Bus. 40 (1967) 393–413.
[6] E.F. Fama, The behavior of stock-market prices, J. Bus. 38 (1965) 34–105.
[7] E.E. Peters, Fractal Market Analysis – Applying Chaos Theory to Investment and Economics, Wiley, 1994.
[8] H.E. Hurst, R. Black, Y.M. Sinaika, Long-Term Storage in Reservoirs: An Experimental Study, Constable, London, 1965.
[9] A.W. Lo, Long-term memory in stock market prices, Econometrica 59 (1991) 1279–1313.
[10] M. Ausloos, Statistical physics in foreign exchange currency and stock markets, Physica A 285 (2000) 48–65.
[11] T. Di Matteo, T. Aste, M.M. Dacorogna, Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development, J. Bank. Finance 29 (2005) 827–851.
[12] T. Di Matteo, T. Aste, M.M. Dacorogna, Scaling behaviors in differently developed markets, Physica A 324 (2003) 183–188.
[13] B. Qian, K. Rasheed, Hurst exponent and financial market predictability, in: Proceedings of the Second IASTED International Conference on Financial Engineering and Applications, 2004.
[14] Y. Kopeliovich, Can Technical Trading Be Profitable – Evidence from Volatility Markets, University of Connecticut, 2013.
[15] I.P. Antoniades, G. Brandi, L. Magafas, T. Di Matteo, The use of scaling properties to detect relevant changes in financial time series: A new visual warning tool, Physica A 565 (2021) 125561.
[16] S. Lee, Correlation shifts and real estate portfolio management, J. Real Estate Portf. Manag. 1 (9) (2003) 45–57.
[17] R. Spurgin, G. Martin, T.A. Schneeweis, Method of estimating changes in correlation between assets and its application to hedge fund investment, J. Asset Manag. 1 (3) (2002) 217–230.
[18] E. Kalotychou, S.K. Staikouras, G. Zhao, On the dynamics of asset correlations: What matters to portfolio managers, 2009.
[19] G. Meissner, Correlation trading strategies – Opportunities and limitations, J. Trading (2015).
[20] S. Riedinger, Demystifying Pairs Trading: The Role of Volatility and Correlation, SSRN, 2017.
[21] R.F. Engle, C.W.J. Granger, Co-integration and error correction: Representation, estimation, and testing, Econometrica 55 (1987) 251–276.
[22] U. Yule, Why do we sometimes get nonsense-correlations between time series? - A study in sampling and the nature of time series, J. Roy. Statist. Soc. 89 (1) (1926) 11–63.
[23] C. Granger, P. Newbold, Spurious regressions in econometrics, J. Econometrics 2 (2) (1974) 111–120.
[24] J. Armstrong, Principles of Forecasting: A Handbook for Researchers and Practitioners, Springer Science & Business Media, 1, 2001.

[25] P.C.B. Phillips, S. Ouliaris, Asymptotic properties of residual based tests for cointegration, Econometrica 58 (1990) 165–193.
[26] W. Chaovalitwongse, P. Pardalos, P. Xanthopoulos, Computational Neuroscience, Springer Science & Business Media, 2010.
[27] H. Rad, R.K.Y. Low, R. Faff, The profitability of pairs trading strategies: Distance, cointegration and copula methods, Quant. Finance 16 (10) (2016) 1541–1558.
[28] J.F. Caldeira, G.V. Moura, Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy, Braz. Rev. Finance 11 (1) (2013) 49–80.
[29] E.F. Fama, Efficient capital markets: A review of theory and empirical work, J. Finance 25 (2) (1970) 383–417.
[30] A. Cowles, H. Jones, Some a posteriori probabilities in stock market action, Econometrica 5 (3) (1937) 280–294.
[31] M. Latoszek, R. Ślepaczuk, Does the inclusion of exposure to volatility into diversified portfolio improve the investment results? Portfolio construction from the perspective of a polish investor, Econ. Bus. Rev. 6 (20) (2020) 46–81.
[32] J. Jabłecki, R. Kokoszczyński, P. Sakowski, R. Ślepaczuk, P. Wójcik, Volatility as an Asset Class. Obvious Benefits and Hidden Risks, Peter Lang, Frankfurt am Main, 2015.
[33] M. Kijewski, R. Ślepaczuk, Predicting Prices of S & P500 Index using Classical Methods and Recurrent Neural Networks, Working Papers of Faculty of Economic Sciences, University of Warsaw, WP 27/2020 (333), 2020, https://www.wne.uw.edu.pl/index.php/download_file/5769/4919/.
[34] A.W. Lo, Risk management for hedge funds: Introduction and overview, Financ. Anal. J. 57 (6) (2001) 16–33.
[35] S. Brown, W. Goetzmann, R. Ibbotson, Offshore hedge funds: Survival and performance 1989–1995, J. Bus. 72 (1999) 91–118.
[36] C. Ackerman, R. McEnally, D. Revenscraft, The performance of hedge funds: Risk, return and incentives, J. Finance 54 (1999) 833–874.
[37] B. Liang, Hedge funds: The living and the dead, J. Financ. Quant. Anal. 35 (3) (2000) 309–326.
[38] A.L. Barabasi, T. Vicsek, Multifractality of self-affine fractals, Phys. Rev. A 44 (1991) 2730–2733.
[39] J.P. Ramos-Requena, J.E. Trinidad-Segovia, M.A. Sánchez-Granero, Introducing Hurst exponent in pair trading, Physica A 488 (39–45) (2017) 0378–4371.
[40] K. Kość, P. Sakowski, R. Ślepaczuk, Momentum and contrarian effects on the cryptocurrency market, Physica A 523 (2019) 691–701.
[41] C. Alexandra, A. Dimitriu, The Cointegration Alpha: Enchanced Index Tracking and Long-Short Equity Market Neutral Stragies, ICMA Centre Discussion Papers in Finance, Henley Business School, Reading University, 2002, http://dx.doi.org/10.2139/ssrn.315619.
[42] R.F. Engle, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, J. Bus. Econom. Statist. 20 (2002) 339–350.
[43] Y.K. Tse, A.K.C. Tsui, A multivariate generalized autoregressive conditional heteoscedasticity model with time-varying correlations, J. Bus. Econom. Statist. 20 (2002) 351–362.