ISSN Online: 2821-1936
Transactions on Data Analysis in Social Science
Journal Homepage: https://transoscience.ir

The Effect of Financial Data Noise on the Long-Term Co-Movement of Stock Markets

S. Mohammadi[1,*], Gh. Mansourfar[2]

[1] Department of Management, Faculty of Social Sciences & Economics, Alzahra University, Iran
[2] Department of Accounting, Faculty of Economic & Management, Urmia University, Iran

ARTICLE INFO

Article History:
Received 25 November 2021
Received in revised form
10 December 2021
Accepted 9 January 2022
Available online 6 March 2022

Keywords:
Financial Data Noise,
Wavelet De-Noising,
Thresholding Function,
Long-Term Co-Movement

ABSTRACT

Due to rapid advances in information technology and the increasing globalization of economies, the interdependence and linkage among international financial markets have become more significant, particularly for investors and portfolio managers who aim to diversify risk and maximize returns. However, one of the main challenges in financial time series analysis is the presence of noise, which often obscures the true underlying patterns of market behavior. This noise component makes it difficult to test financial theories accurately and to detect meaningful long-term relationships among different capital markets. To address this limitation, the wavelet de-noising method was applied to a co-integration model in this study. The objective was to examine the impact of financial time series noise on the long-term dynamics of 16 selected international capital market indices. The weekly closing prices of these indices were analyzed over the study period. The results revealed that after applying the wavelet de-noising technique, the time series exhibited stronger co-integration compared to the raw noisy data. This finding indicates that noise reduction enhances the detection of stable long-term equilibrium relationships among global markets. Furthermore, the use of de-noised time series provides a clearer and more reliable perspective on co-movement analysis, which is crucial for understanding market integration and for making informed investment decisions.

1. INTRODUCTION

One of the essential and challenging issues in economic and financial time series is their behavior, which often differs from what academic theories suggest. Introducing the concept of "Noise," Black explained: "more generally noise makes it difficult to test either practical or academic theories about the way that financial and economical markets work." He also argued that the noise is in contrast with information. "Noise makes financial markets possible, but also makes them imperfect. It keeps us from knowing the expected return on stocks or portfolios." [1]

* Corresponding Author : somayeh_mohammadi20@yahoo.com
Department of Management, Faculty of Social Sciences & Economics, AL Zahra University, Iran

http://dx.doi.org/10.47176/TDASS/2022.9

Therefore, it would be a vital issue to identify the noise component of time series data. Removing noise from data can be perceived as a process of capturing optimal information from noisy signals.

Numerous studies have been documented to find what causes the noise and how one could split it from the observed data. Trueman [2] utilized one riskless and many risky assets and indicated that fund managers are more interested in noise trading via riskier assets. Proposing the two-period lived agent model, De long et al, [3,4] showed that noise traders can affect the price in a way that the price of equity to be different from its fundamental value. They also revealed that noise traders can survive and govern the market in the term of wealth in long run.

Campbell and Kyle [6] studied noise traders' effect on price and its volatility. Employing U.S stock returns during 1871-1986, they showed that the effect of noise on stocks return depends on interest rate. Their findings suggest that when the interest rate is %4 or less, the noise effect can be ignored, but when the interest rate is %5 or above, the noise becomes extremely important in moving stock prices. Siriopoulos and Leontitsis [7] argue that reducing the noise level of data, gives better view of the time series dynamic and smaller prediction error. To this end, they used both BDS (Brock-Dechert-Scheinkman) statistic and LPCA (Local Principal Component Analysis) approaches for nonlinear noise estimation of six western and eastern international capital markets over the period 1988 to 1999. Their study showed that the noise distribution in western markets is mesokurtic and the average NSRs (Noise to Signal Ratio) are about 70%, while eastern markets have NSRs of about 50% in average and the noise distribution is leptokurtic. They also found that DBS statistic results are relatively large for markets with low NSRs and relatively small for those with high NSRs.

Overall, recent studies such as Bandi and Russell [8], Hu [9], Berkman and Koch [10], Kurov [11], Bloomfield et al [12], show that the noise of financial data has a significant effect on stock prices. Therefore, it may change the expected return as well as the risk level in different investment horizons.

In this regard, investors, especially those who have long-term investment perspectives, would prefer to invest in internationally well diversified portfolios rather than in individual or only in local equities [27, 28, 29, 30]. Therefore, international portfolio management should have exact and accurate idea on the possibility of gaining internationally diversified portfolios. Consequently, understanding if the noise of financial time series changes the level of diversification benefit is a crucial question that needs to be answered.

This study contributes to the current literature of risk management by applying a new wavelet based de-noising method (which presents a new algorithm in reducing non-information part of time series) and employing de-noised financial data to check for international portfolio diversification benefits. The combination of wavelet shrinkage method with Johansen and Juselius (1990) co-integration test based on the Vector Autoregressive (VAR) framework would bring a clear outlook toward long run co- movement of international stock markets.

The paper, hence, is organized as follows. Section 2 gives a brief description on data and explains the method. Section 3 reports the result of the wavelet based de noising and co-movement in empirical study. Finally, section 4 comes with a conclusion and some suggestions for further studies.

2. DATA AND METHODOLOGY

2.1. Data

The data used in this study are the weekly closing prices of the sixteen stock market indices during 6/29/2007 to 12/28/2012, which has been acquired from Data-stream database. The data has been divided into two major groups; the first group involves eight developed markets (Australia, Canada, France, Germany, Japan, Netherlands, UK and USA) and the second group includes eight emerging markets (China, Hong Kong, Indonesia, Korea, Malaysia, Singapore, Taiwan and Thailand).

2.2. Methodology

2.2.1. Wavelet transform

Wavelet is a powerful mathematical tool, which its base lays in Fourier transform. The wavelet transform converts a time series to scale (frequency) domain; therefore, certain characteristics of time series that are not visible can be highlighted. The wavelet transforms splits data into low pass (approximate) and high pass (detail) portions using father and mother wavelets. In Discrete Wavelet Transform (DWT), Mother wavelets ψ (known as wavelet function), which are basic functions of wavelet transform, capture the high frequency component of a signal, while father wavelets φ (known as scaling function) are good to describe the smooth and trend part of a signal. In addition, mother wavelet integrates to zero, whereas father wavelet integrates to one. Mathematical definition of ψ and φ is as follows:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}}\psi(2^{-j}t - k) = 2^{-\frac{j}{2}}\psi\left(\frac{t-2^{j}k}{2^{j}}\right) \tag{1}$$

$$\varphi_{j,k} = 2^{-\frac{j}{2}}\varphi(2^{-j}t - k) = 2^{-\frac{j}{2}}\varphi\left(\frac{t-2^{j}k}{2^{j}}\right) \tag{2}$$

J, k ϵ Z;

Where k is the index for translation (location) of the wavelet and j is the index for dilation (size) of the wavelet.

The wavelet version of signal $y(t)$ can be written as:

$$y(t) = \sum_k v_{J,k}\,\varphi_{J,k}(t) + \sum_k w_{J,k}\psi_{J,k}(t) + \sum_k w_{j-1,k}\psi_{j-1,k}(t) + \cdots + \sum_k w_{1,k}\psi_{1,k}(t) \tag{3}$$

Where J is the number of multi-resolution components, $v_{J,k}$ are called the smooth coefficients, and $w_{J,k}$ are called the detailed coefficients. These coefficients are defined by:

$$v_{J,k} = \int y(t)\varphi_{J,k}(t)\,dt \tag{4}$$

$$w_{J,k} = \int y(t)\psi_{J,k}(t)\,dt \tag{5}$$

2.2.2. Wavelet de-nosing

Unlike classic Fourier analysis, which can only eliminate noise from certain patterns over the full-time horizon, wavelet analysis can cope with multi-scales and more complex data. As a result, it is better suited for financial time series. [13] Wavelet de-noising is a powerful and simple nonlinear technique used to reduce the level of noise in time series, and to obtain an actual signal from the noisy observations. In 2002, Ramsey identified the concept of time series smoothing versus de-noising. He confirmed that in an observed signal when both the signal and noise part are non-smooth, or the signal includes discontinuities or regime shifts, such situation requires de-noising not smoothing because by smoothing the important part of signal, which you want to capture, it will be obliterated or lost [14]. Hence, in this article the wavelet de-noising method is used.

Assuming a signal given by $Y = X + \varepsilon$, (Y is observed data, X states noiseless or clear data and ε is error term),

The following are the three fundamental steps of de-noising utilizing the wavelet coefficient shrinking technique [15]:

Apply a wavelet transform W to the data to compute the wavelet coefficient matrix $\boldsymbol{\omega}$:

$$\omega = WY = WX + W\varepsilon \tag{6}$$

1. Modify (e.g. threshold or shrinkage) to detail coefficients of $\boldsymbol{\omega}$ to obtain the estimate $\widehat{\omega}$ of wavelet coefficient of X:

$$\omega \to \widehat{\omega} \tag{7}$$

$$\widehat{\omega} = \widehat{\sigma}\delta_{\lambda}(\widehat{\omega}/\widehat{\sigma}) \tag{8}$$

2. To obtain the de-noised estimate, reverse convert the changed coefficients:

$$\hat{X} = W^{-1}\widehat{\omega} \tag{9}$$

In using wavelet de-noising algorithm, three important factors should be chosen correctly.

1. The shrinkage function that determines how the thresholds are applied to the data.
2. Noise estimate.
3. Shrinkage rules to determine the thresholdλ.

2.2.2.1. Thresholding function

Wavelet thresholding is the signal estimation technique that develops the capabilities of signal denoising [16]. Threshold methods are classified into two main types including hard thresholding and soft thresholding. Figure1 shows two common theresholding functions: hard and soft functions intervals [-1, 1].

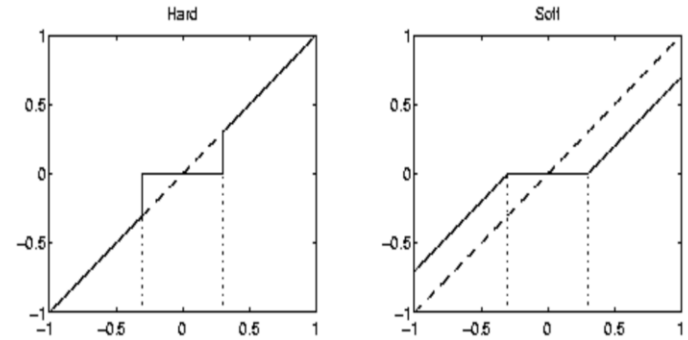

Fig. 1. Shrinkage functions. The $\boldsymbol{x}$ axis implies detail wavelet coefficient "$\boldsymbol{\omega}$" and the $\boldsymbol{y}$ axis represent the corresponding thresholding function "$\boldsymbol{\delta_\lambda(\omega)}$". [17]

Mathematical definitions of these two shrinkage functions are:

Hard tresholding function:

$$\delta_\lambda^H(\omega) = \omega I_{\{|\omega|>\lambda\}} \tag{10}$$

Soft tresholding function:

$$\delta_\lambda^S(\omega) = sgn(\omega)(|\omega| - \lambda)\omega I_{\{|\omega|>\lambda\}} \tag{11}$$

By assuming λ as threshold, δ_λ indicates the shrinkage function, which determines how threshold is applied to the data and $\hat{\sigma}$ as estimator of standard deviation of the noise.

The hard threshold function tends to have bigger variance and it is unstable [16]. However, soft theresholding produces a smoother estimation, because all coefficients are being pushed toward zero [17]; therefore, in this study the soft thresholding method is applied in order to obtain the optimal part of time series from noisy observations.

2.2.2.2. Noise estimate

Because noise structure must be estimated from observable data in many cases, the options for estimating noise scale include selecting the functional form of the estimator and selecting the detail coefficients to include in the calculation [15]. In addition, Peters introduced the Fractal Market Hypothesis (FMH) in 1991, which posits that market price dynamics may be caused by the interaction of actors with distinct time horizons and varied interpretations of information. [18]. Because of the two above-mentioned reasons, at the first step, both un-scaled white noise and scaled with noise estimator will be applied to data. Afterward, the standard deviation of each method will be calculated. On the other hand, Estimator, which would lead less amount of standard deviation, will be chosen as the main noise estimator of the study.

2.2.2.3. Shrinkage rules to determine the threshold λ

Threshold determination is a critical issue. A small threshold may produce a noisy outcome, whereas a big threshold may cut a major portion of the signal, resulting in the loss of key signal features [19]. Wavelets provide four basic rules to establish the threshold λ, which are described below.

2.2.2.3.1. Universal

A Fixed form of threshold proposed by Donoho and Johnstone [5] can be explained as $\lambda = \hat{\sigma}\sqrt{2\log N}$, where N is the data size and $\hat{\sigma}$ is the noise standard deviation which is estimated as $\hat{\sigma} = \frac{MAD_j}{0.6745}$. MAD is the Median Absolute Deviation of wavelet coefficients at scale j and 0.6745 is the normalization factor [16, 20, 13 15, 19].

2.2.2.3.2. SURE

Donoho and Johnstone's [5] Steins Unbiased Risk Estimator is an adaptive approach that is dependent on shrinkage function and multiresolution level [14].

Assume $W = [\omega_1, \omega_2, \dots \dots \omega_N]$ is a vector consisting of square wavelet coefficients ranging from tiny to large. Choose the smallest value $r_b(b^{th}r)$ from the risk vector, which is provided as,

$$R = \{r_i\}_{i=1,2\dots\dots N} = \frac{[N - 2i + (N - i)\omega_i + \sum_{k=1}^{l} \omega_k}{N} \tag{12}$$

as well as the risk value the chosen threshold is $\lambda = \sigma\sqrt{\omega_b}$, where ω_b is the b^{th} squared wavelet coefficient, (coefficient at lowest risk) chosen from the vector W, and σ is the noisy signal's standard deviation [19].

2.2.2.3.3. Heuristic SURE

When the wavelet coefficient decomposition is sparse, a hybrid technique that combines the Universal and SURE thresholds outperforms SURE [17]. This rule aims to circumvent SURE's limitation [20]. When the signal-to-noise ratio is exceedingly low, the SURE technique provides poor estimation. In this scenario, the fixed form threshold technique provides more accurate threshold estimation. Given that the universal method's threshold is λ_1 and Rigrsure's threshold is λ_2, Heuristic SURE calculates the threshold as follows:

$$\lambda = \begin{cases} \lambda_1 & A > B \\ \min(\lambda_1, \lambda_2) & A \geq B \end{cases} \tag{13}$$

in which $A = \frac{S-N}{N}$ and $B = (logN)^{3/2}\sqrt{N}$. N denotes the length of the wavelet coefficient vector, and S is the sum of squared wavelet coefficients, which is defined as $S = \sum_{i=1}^{N} \omega_i^2$ [19]. SureShrink is the name given to this hybrid approach when paired with a soft shrinkage function in the literature [15].

$S = \sum_{i=1}^{N} \omega_i^2$ [19] is the sum of squared wavelet coefficients and N is the length of the wavelet coefficient vector. SureShrink is the name given to this hybrid approach when paired with a soft shrinkage function in the literature [15].

2.2.2.3.4. Minimax

Using the Minimax principle, this approach determines the threshold λ. It employs a defined threshold to achieve Minimax performance for mean square error when compared to an ideal process. In statistics, the Minimax concept is used to develop estimators. Because the de-noised signal can be assimilated to the estimator of the unknown regression function, the Minimax estimator is the choice that achieves the smallest maximum Mean Square Error over a given collection of functions (MSE) [19]. The threshold is obtained by

$$\lambda = \begin{cases} \hat{\sigma}(0.3936 + 0.1829log_2^N) & N > 32 \\ 0 & N < 32 \end{cases} \tag{14}$$

in which $\hat{\sigma} = median\,\frac{|\omega|}{0.6745}$ and ω is the wavelet coefficient vector at unit scale and N is the length of signal vector [19].

In this study in order to select the most favorable thresholding rule, two steps would be taken.

One: Four thresholding methods will be applied to the data

Two: Signal to Noise Ratio (SNR) will be computed for all principles. In addition, the threshlding methods with highest SNR would be chosen as the shrinkage rule determination.

SNR is defined as the rate of mean to standard deviation ($^{\mu}/_{\sigma}$). As noise part generally has a zero mean, its reduction does not affect the mean of the series [21]. Thus, the mean of the series can be used as a proxy of actual part of time series.

2.2.3. Co-movement model

Stock market co-movements and interrelationships are typically explored using estimate to determine the presence of co-integration vectors. Co-integration studies reveal whether share prices from multiple markets are linked in the long run. As a result, co-integration restricts the potential benefits of long-run diversity. In contrast, the lack of co-integration supports the benefits of foreign diversity for long-term investors [22]. However, before performing a co-integration test, the data series' non-stationarity must be proved.

The Johansen [23, 24] and Johansen and Juselius [25] approaches based on the Vector Autoregressive (VAR) framework are used to test the long term correlations between equity prices. In this case, a VAR model of order z is considered:

$$X_t = A_1 X_{t-1} + \cdots + A_z X_{t-z} + \mu + +\varepsilon_t \tag{15}$$

where X_t is a $(n \times 1)$ matrix of the non-stationary I(1) weekly index price series, each of A_i

is a $(n \times n)$ matrix of coefficients, μ is a vector of constants, δ_t is a vector of trend coefficients and ε_t is a vector of innovations.

VAR model can be rephrased in VECM framework as:

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{z-1} \lambda_t \, \Delta X_{t-1} + \delta_t + \varepsilon_t \tag{16}$$

$$\theta = \sum_{i=1}^{z} A_i - I_i \qquad \lambda_i = -\sum_{j=i+1}^{z} A_i$$

Co-integration test uses two test statistics, the λ-trace statistics and the maximum eigenvalue statistics (λ-max statistics) which are defined as:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{p} \ln\,(1 - \hat{\lambda}_i) \tag{17}$$

$$\lambda_{max}(r, r+1) = -T ln(1 - \hat{\lambda}_{r+1}) \tag{18}$$

When there is a dispute between these two test statistics, Johansen and Juselius [25] suggest that the emphasis should be on -trace statistics rather than -max statistics. Furthermore, Johansen's recommendation is employed in this study to identify the appropriate lag duration, so that the VAR residuals are Gaussian or serially uncorrelated [26].

2.2.4. Steps of data analysis

To explore the effect of data noise on the long-term co movement of stock markets, Johansen co-integration test was applied to the data twice, before and after de noising. In order to de-noise the market price time series, Daubechies8 (db8) in Discrete Wavelet Transform (DWT) is applied to data, which were decomposed to five level (since the data in DWT should be dyadic, the maximum level in this study is five). Afterward, Heuristic SURE method through soft thresholding function (SureShrink) was applied to the data to reduce the noise component, and to modify noise free wavelet coefficient. Figure 2 shows wavelet de-noising procedure.

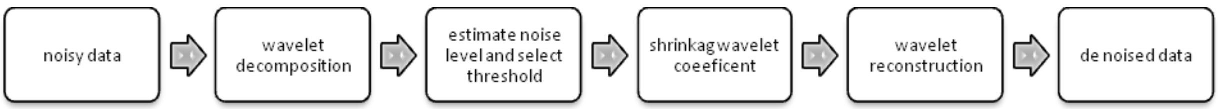

Fig. 2. Wavelet de-noising procedure

3. RESULTS

3.1. Choosing appropriate noise estimation and thresholding function

3.1.1. Noise estimate

As discussed in methodology, in order to discover the most favorable noise estimator, both un-scaled white noise and scaled with noise estimator were applied to data. Afterward, the standard deviation of each method was calculated. Results are reported in table1.

Results show that using scaled white noise estimator leads to a smaller standard deviation (SD). For example, the SD of de noised price series of Australia with the assuming noise type as an un-scaled white noise is 70.44. While, this statistic for de-noised series, considering scaled white noise structure, is 68.39. Thus, the scaled white noise is chosen to be removed from the data. Therefore, in the rest of study, the word noise means scaled white noise.

Table 1. standard odeviation of denoise time series via two diffrent noise type

countries	standard deviation of de noised data by different noise structure	
	un-scaled white noise	scaled white noise
developed		
Australia	70.44	68.39
Canada	99.87	96.26
France	103.60	101.23
Germany	73.86	71.85
Japan	19.91	19.34
Netherlands	117.51	115.43
UK	77.20	75.7
USA	80.34	78.52
emerging		
China	551.94	529.49
Hong-Kong	121.31	117.63
Indonesia	107.01	105.77
Korea	207.13	202.34
Malaysia	231.00	228.63
Singapore	90.12	88.22
Taiwan	41.14	39.89
Thailand	22.85	22.7

3.1.2. Thresholding function

Each four thresholding methods were applied to the data. Subsequently, the Signal to Noise Ratio (SNR), as introduced in method, was computed for all the principles. Results are reported in table2. The highest SNRs are captured from Universal and Heuristic SURE methods. Therefore, these techniques generate a better smooth series as compared to the other two methods in sample time series. Since Heuristic SURE is flexible and has the ability of both Universal and SURE procedure at the same time, this method is supposed to be the Shrinkage rules for determining the threshold λ.

Table 2. Signal to Noise ratio of time series with four different thresholding methods

SNR ($^{\mu}/_{\sigma}$)					
		de-noised prices			
countries	Noisy prices	Universal	SURE	Heuristic SURE	Minimax
developed					
Australia	5.03	5.21	5.06	5.19	5.16
Canada	6.05	6.31	6.10	6.29	6.24
France	4.30	4.41	4.32	4.40	4.38
Germany	4.69	4.85	4.72	4.83	4.81
Japan	6.48	6.72	6.54	6.69	6.65
Netherlands	3.76	3.85	3.74	3.84	3.82
UK	4.86	4.98	4.89	4.96	4.95
USA	6.36	6.54	6.40	6.51	6.49
emerging					
China	5.24	5.50	5.29	5.46	5.41
Hong-Kong	5.75	5.95	5.79	5.93	5.81
Indonesia	3.15	3.19	3.16	3.19	3.18
Korea	4.49	4.63	4.52	4.60	4.59
Malaysia	4.51	4.57	4.53	4.56	4.56
Singapore	5.24	5.37	5.27	5.36	5.33
Taiwan	5.60	5.81	5.64	5.78	5.75
Thailand	3.08	3.25	3.23	3.25	3.24

3.2. Descriptive Analysis

Descriptive analyses of weekly prices of eight developed and eight emerging market indices are performed to find out the properties of the data; results are presented in tables 3 and 4. Statistic illustrate that, the sample mean remained almost unchanged by de-noising process, while the standard deviations of de-noised time series are reduced. This means that, all information except noise is kept [21]. The null hypotheses of normal distribution in both noisy and de-noised weekly prices are rejected for all developed and emerging markets.

Table 3. Descriptive analysis of developed markets price indices, before and after de-noising

	Australia	Canada	France	Germany	Japan	Netherlands	UK	USA
	Noisy data							
Mean	355.15	605.32	445.80	346.94	129.35	442.76	375.56	511.86
Std. Dev	70.59	100.07	103.67	73.95	19.97	117.65	77.33	80.46
Skewness	-0.75	-0.83	0.82	0.503	0.99	0.84	0.46	-0.60
Kurtosis	3.40	3.54	2.65	2.51	3.15	2.81	3.06	2.54
Jargue-Bera	29.60	36.95	33.62	15.07	47.33	34.65	10.26	19.97
Probability	0.0001	0.0001	0.0001	0.0005	0.0001	0.0001	0.0006	0.0001
	De noised data							
Mean	355.23	605.33	445.84	346.99	129.36	442.83	375.59	511.84
Std. Dev	68.39	96.26	101.23	71.85	19.34	115.43	75.70	78.52
Skewness	-0.85	-0.92	0.86	0.54	1.07	0.87	0.48	-0.62
Kurtosis	3.47	3.59	2.64	2.50	3.11	2.80	3.03	2.53
Jargue-Bera	37.17	44.95	37.01	17.22	55.14	37.03	11.12	21.50
Probability	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0039	0.0001

Table 4. Descriptive analysis of emerging market price indices, before and after de-noising

	China	HongKong	Indonesia	Korea	Malaysia	Singapore	Taiwan	Thailand
	Noisy data							
Mean	2892.28	698.34	337.23	930.27	1042.93	472.67	230.72	73.69
Std. Dev	552.29	121.52	107.13	207.34	231.13	90.23	41.20	22.86
Skewness	-0.0008	-0.85	-0.48	-0.86	-0.43	-.1.26	-0.88	-0.09
Kurtosis	4.60	3.47	2.25	3.07	2.16	3.81	3.57	2.10
Jargue-Bera	30.91	37.70	17.66	35.52	18.41	83.60	41.20	10.18
Probability	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0062
	De noised data							
Mean	2892.28	698.27	337.28	930.57	1043.04	472.80	230.70	73.72
Std. Dev	529.49	117.63	105.77	202.34	228.63	88.22	39.89	22.70
Skewness	-0.07	-0.95	-0.49	-0.92	-0.46	-1.32	-0.98	-0.11
Kurtosis	4.43	3.56	2.21	3.01	2.13	3.82	3.64	2.09
Jargue-Bera	23.76	46.79	19.16	40.75	19.13	91.32	51.45	10.57
Probability	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0051

Figures 3 and 4 show both noisy and de-noised price time series of Australia and China. It is noticeable that the primary behavior of prices (which is the most imperative factor in the long-term co-movement test) remains unaffected during wavelet based de-noising. However, short-term deviations of prices from the main behavior line are reduced. In addition, this procedure continues in time series. Therefore, de-noising through wavelet shrinkage gives clear attitude toward long- term behavior of markets price.

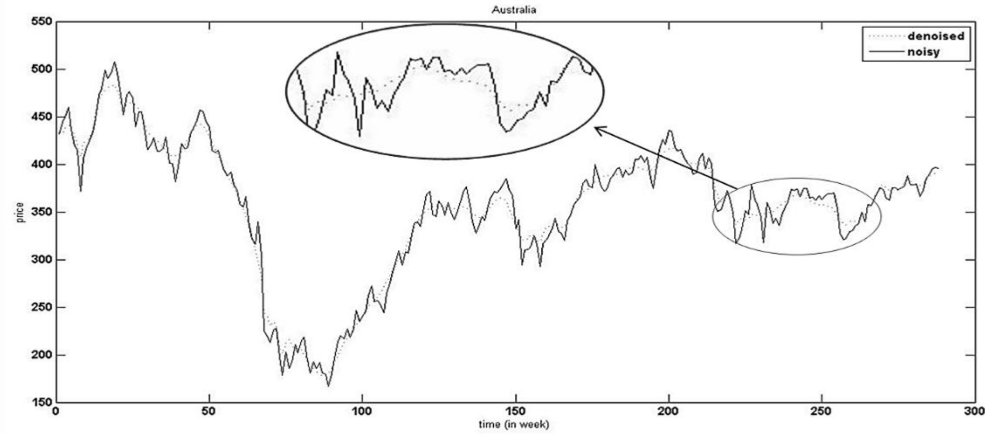

Fig. 3. Noisy and de-noised series of Australia stock market price

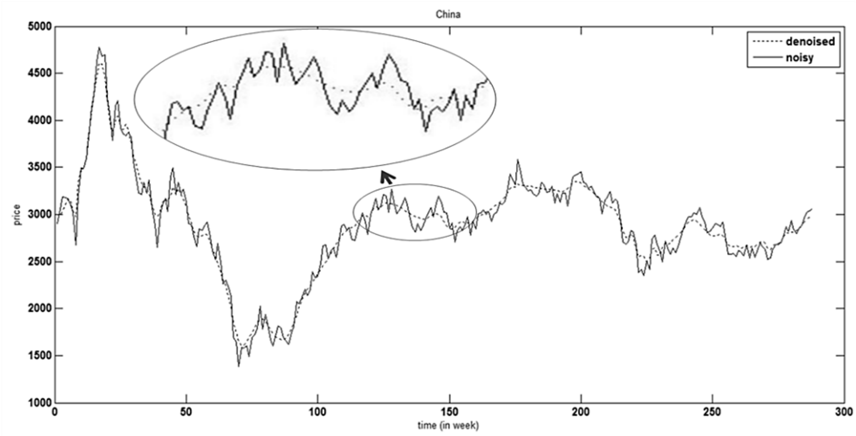

Fig. 4. Noisy and de-noised series of China stock market price

3.3. Unit root test

Non-stationarity of data is checked through unit root tests. Augmented Dickey and Fuller (ADF) and the non-parametric Phillips and Perron (PP) tests are applied to investigate the presence of stochastic non-stationarity in the data. For the ADF test, the number of lags is optimized by minimizing the Schwarz Information Criteria (SIC). If the test statistic differs significantly from zero, the null hypothesis is rejected in support of the notion that the series is stationary. Tables 5 and 6 report the results of the unit root tests for weekly index prices and returns.

ADF and PP tests suggest that the levels of both noisy and de-noised variables across the sample countries contain unit roots. Consequently, all variables follow stochastic trends in their levels. In addition, tests indicate that non-stationarity is effectively removed using first difference of variables. Therefore, all variables are integrated for the first order. As the null hypothesis of non-stationarity among variables could not be rejected, the long-run linkage among them can later be investigated using the multivariate co-integration models.

Table 5. Unit root test of Noisy and De-noised weekly index price and returns; developed markets

	Noisy data				**De-noised data**			
	ADF		Phillips- Perron		ADF		Phillips- Perron	
country	level P-value	1st.D P-value	level P-value	1st.D P-value	level P-value	1st.D P-value	level P-value	1st.D P-value
Australia	(0.6985)	(0.0001)	(0.6850)	(0.0001)	(0.4785)	(0.0015)	(0.7402)	(0.0001)
Canada	(0.6204)	(0.0001)	(0.6307)	(0.0001)	(0.4564)	(0.0004)	(0.6903)	(0.0001)
France	(0.6043)	(0.0001)	(0.5942)	(0.0001)	(0.5385)	(0.0001)	(0.6928)	(0.0001)
Germany	(0.7096)	(0.0001)	(0.6510)	(0.0001)	(0.6441)	(0.0001)	(0.7583)	(0.0001)
Japan	(0.4185)	(0.0001)	(0.4694)	(0.0001)	(0.5123)	(0.0001)	(0.5849)	(0.0001)
Netherlands	(0.7449)	(0.0001)	(0.6692)	(0.0001)	(0.4894)	(0.0001)	(0.7540)	(0.0001)
UK	(0.6630)	(0.0001)	(0.6854)	(0.0001)	(0.7312)	(0.0001)	(0.7305)	(0.0001)
USA	(0.6287)	(0.0001)	(0.6416)	(0.0001)	(0.6413)	(0.0001)	(0.7476)	(0.0001)

Table 7. Unit root test of Noisy and De-noised weekly index price and returns; emerging markets

	Noisy data				De-noised data			
	ADF		Phillips- Perron		ADF		Phillips- Perron	
country	level P-value	1st.D P-value	level P-value	1st.D P-value	level P-value	1st.D P-value	level P-value	1st.D P-value
China	(0.5714)	(0.0001)	(0.4849)	(0.0001)	(0.3657)	(0.0001)	(0.6233)	(0.0001)
Hong-Kong	(0.7838)	(0.0001)	(0.6767)	(0.0001)	(0.4737)	(0.0006)	(0.7423)	(0.0001)
Indonesia	(0.6789)	(0.0001)	(0.7166)	(0.0001)	(0.7521)	(0.0001)	(0.8018)	(0.0001)
Korea	(0.6793)	(0.0001)	(0.6921)	(0.0001)	(0.5434)	(0.0001)	(0.7786)	(0.0001)
Malaysia	(0.7225)	(0.0001)	(0.7439)	(0.0001)	(0.6034)	(0.0327)	(0.7606)	(0.0001)
Singapore	(0.7361)	(0.0001)	(0.6515)	(0.0001)	(0.6000)	(0.0001)	(0.7293)	(0.0001)
Taiwan	(0.6938)	(0.0001)	(0.5855)	(0.0001)	(0.4860)	(0.0059)	(0.6208)	(0.0002)
Thailand	(0.8322)	(0.0001)	(0.8030)	(0.0001)	(0.6162)	(0.0001)	(0.8057)	(0.0001)

ADF and PP tests suggest that the levels of both noisy and de-noised variables across the sample countries contain unit roots. Consequently, all variables follow stochastic trends in their levels. In addition, tests indicate that non-stationarity is effectively removed using the first difference of the variables. Therefore, all variables are integrated in the first order. As the null hypothesis of non-stationarity among variables could not be rejected, long-run linkage among them can be investigated using the multivariate co-integration models.

3.4. Long-term co movement analysis

Co-integration method of Johansen [23, 24] and Johansen and Juselius [22] is applied to assess the long-run linkage within eight developed and eight emerging equity markets, before and after de-noising. The results of co-integration tests among developed markets are reported in table 7. Panel A reports the existence of one co-integrating vector which is confirmed among the developed markets as the λ-trace statistic (168.17) surpasses the critical value (159.53) of 5% significance level Before de –noising.

Once again, co-integration test is applied for the wavelet constructed de-noised price indices. Among these equity markets, two co-integrating vectors are observed at 5% significant level (second part of panel A).

The result of co integration test using Max-Eigen Statistic is summarized in panel B of table 7. As it is apparent, de-noised data (second part of panel B) reports the existence of one co integrating vector at 5% of significant level.

Table 8. Results of co-integration tests among developed markets are reported using Noisy and De-noised data

test statistics	Countries in the Group		H0: r = 0	H1: r ≤ 1	H2: r ≤ 2	H3: r≤ 3	H4: r ≤ 4	H5: r ≤ 5
Panel A								
λ-trace	The developed markets	**Result of co integration test by noisy data**						
		Trace Statistic	168.17*	118.27	82.89	53.20	35.01	20.44
		Critical Value 5%	159.53	125.62	95.75	69.82	47.86	29.80
		Results of co integration test by de-noised data						
		Trace Statistic	213.32*	133.52*	95.22	64.46	34.98	12.86
		Critical Value 5%	159.53	125.62	95.75	69.82	47.86	29.80
Panel B								
Max-Eigen	The developed markets	**result of co integration test by noisy data**						
		Max-Eigen Statistic	49.90	35.39	29.69	18.19	14.57	11.76
		Critical Value 5%	52.36	46.23	40.07	33.88	27.58	21.13
		results of co integration test by de-noised data						
		Max-Eigen Statistic	79.80*	38.30	30.76	29.48	22.12	7.32
		Critical Value 5%	52.36	46.23	40.07	33.88	27.58	21.13

Table 9. Results of co-integration tests among emerging markets are reported using Noisy and De-noised data

test statistics	Countries in the Group		H0: r = 0	H1: r ≤ 1	H2: r ≤ 2	H3: r≤ 3	H4: r ≤ 4	H5: r ≤ 5
Panel C								
λ-trace	The emerging markets	**result of co integration test by noisy data**						
		Trace Statistic	185.38*	120.76	82.90	51.57	28.27	15.12
		Critical Value 5%	159.53	125.62	95.75	69.82	47.86	29.80
		results of co integration test by de-noised data						
		Trace Statistic	251.79*	181.26*	120.47*	76.28*	36.63	19.15
		Critical Value 5%	159.53	125.61	95.75	69.82	47.86	29.80
Panel D								
λ-max	The emerging markets	**result of co integration test by noisy data**						
		Max-Eigen Statistic	64.62*	37.86	31.33	23.31	13.15	8.00
		Critical Value 5%	52.36	46.23	40.07	33.88	27.58	21.13
		results of co integration test by de-noised data						
		Max-Eigen Statistic	70.53*	60.78*	44.19*	39.65*	17.48	10.64
		Critical Value 5%	52.36	46.23	40.07	33.88	27.58	21.13

This specifies that these stock markets are strongly linked, while using noisy data, this statistic confirms no co integration vectors (first part of panel A).

Therefore, in the long-term none of the price indices in developed stock markets can arbitrarily drift away from other market indices. Both λ-trace and Max-Eigen statistics indicate that, using de-noised price indices clarify that the possibility of gaining international portfolio diversification within developed markets is not noticeable. Meanwhile, noisy data show a reduced amount of co integration vectors. This confirms that using noisy data leads to imperfect conclusion.

Noise effect is even more considerable among emerging markets. Table 8 reports the results of noise impact on co-integration tests among emerging stock markets. Via noisy data, (first part of panel C) one co-integrating vector is observed (λ-trace statistic (185.38) among emerging equity markets, with an improvement (159.53) as the critical value, at 5% significant level.

Results of applying co-integration test to de-noised data using λ-trace statistic, are summarized in the second part of panel C of table 8. Four co-integrating vectors are reported in 5% significant level. Results obtained from Max-Eigen statistic, completely confirm the findings via λ-trace statistic (panel D of table 8). This indicates that benefit of portfolio diversification is very low amongst emerging markets and in the long run these markets are heavily co-integrated. This detection would be hidden when using noisy observations.

4. DISCUSSION

In this paper, the effect of financial data noise on long run linkage of sixteen stock market (developed and emerging) indices was examined by a novel de-noising method (wavelet shrinkage). The key findings contain two main features. First, the Heuristic SURE method through soft thresholding function gives better smooth data in comparison with other wavelet de-noising rules in de-noising price indices. Second, the noise components of time series change the long- run behavior of stock market price indices. Applying co-integration test for noisy and de-noised data, study showed that noisy time series lead to imperfect observation. By analyzing the data, study explained that, hidden noise of time series, decrease co-integration vectors in the long run, which lead to erroneous expectation from long-term diversification benefit. In other words, while the benefit of diversification is low and investors with long-term horizons cannot gain very much, noisy data show diversification opportunity. In addition, findings of the study confirm that noise effect is very significant amongst emerging markets rather than in developed markets.

Further research studies in this direction might include testing robustness of many financial theories and models using de-noised data (i.e. efficient market hypotheses, Capital Asset pricing model, Arbitrage pricing theory, etc.) in order to divulge the behavior of actual parts of the time series.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Black, F. (1986). Noise. *Journal of Finance, 41*(3), 529–543.

[2] Trueman, B. (1988). A theory of noise trading in securities markets. *Journal of Finance, 43*(1), 83–95. https://doi.org/10.1111/j.1540-6261.1988.tb02590.x

[3] De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy, 98*(4), 703–738. https://doi.org/10.1086/261703

[4] De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1991). The survival of noise traders in financial markets. *Journal of Business*. [Details missing add volume, issue, pages].

[5] Donoho, D. L., & Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association, 90*(432), 1200–1224. https://doi.org/10.1080/01621459.1995.10476626

[6] Campbell, J. Y., & Kyle, A. S. (1993). Smart money, noise trading and stock price behaviour. *Review of Economic Studies, 60*(1), 1–34. https://doi.org/10.2307/2297810

[7] Leontitsis, A., & Siriopoulos, C. (2006). Nonlinear forecast of financial time series through dynamical calendar

correction. *Applied Financial Economics Letters, 2*(5), 337–340. https://doi.org/10.1080/17446540500461786
[8] Bandi, F. M., & Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics, 79*(3), 655–692. https://doi.org/10.1016/j.jfineco.2005.01.005
[9] Hu, S.-Y. (2006). A simple estimate of noise and its determinant in a call auction market. *International Review of Financial Analysis, 15*(4–5), 348–362. https://doi.org/10.1016/j.irfa.2006.02.004
[10] Berkman, H., & Koch, P. D. (2008). Noise trading and the price formation process. *Journal of Empirical Finance, 15*(2), 232–250. https://doi.org/10.1016/j.jempfin.2006.10.005
[11] Kurov, A. (2008). Information and noise in financial markets: Evidence from the e-mini index futures. *Journal of Financial Research, 31*(3), 247–270. https://doi.org/10.1111/j.1475-6803.2008.00239.x
[12] Bloomfield, R., O'Hara, M., & Saar, G. (2009). How noise trading affects markets: An experimental analysis. *Review of Financial Studies, 22*(6), 2275–2302. https://doi.org/10.1093/rfs/hhn102
[13] Chai, J., Du, J., Lai, K. K., & Lee, Y. P. (2015). A hybrid least square support vector machine model with parameters optimization for stock forecasting. *Mathematical Problems in Engineering, 2015*, 1–7. https://doi.org/10.1155/2015/231394
[14] Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics, 6*(3). https://doi.org/10.2202/1558-3708.1090
[15] Fodor, I. K. (2003). Denoising through wavelet shrinkage: An empirical study. *Journal of Electronic Imaging, 12*(1), 151. https://doi.org/10.1117/1.1525793
[16] Karthikeyan, P., Murugappan, M., & Yaacob, S. (2012). ECG signal denoising using wavelet thresholding techniques in human stress assessment. *International Journal on Electrical Engineering and Informatics, 4*(2), 306–319. https://doi.org/10.15676/ijeei.2012.4.2.9
[17] Gençay, R., Selçuk, F., & Whitcher, B. (2002). An introduction to wavelets and other filtering methods in finance and economics. *Waves in Random Media, 12*(3), 399.
[18] Anderson, N., & Noss, J. (2013). The fractal market hypothesis and its implications for the stability of financial markets. [Incomplete reference please provide journal or publisher].
[19] Verma, N., & Verma, A. K. (2012). Performance analysis of wavelet thresholding methods in denoising of audio signals of some Indian musical instruments. *International Journal of Engineering Science and Technology, 4*(5), 2040–2045.
[20] Phinyomark, A., Phukpattaranont, P., & Limsakul, C. (2012). The usefulness of wavelet transform to reduce noise in the SEMG signal. In *EMG methods for evaluating muscle and nerve function*. InTech. https://doi.org/10.5772/25757
[21] Mehr, N. B. (2011). Portfolio allocation using wavelet transform [Doctoral dissertation, ProQuest UMI Dissertation Publishing].
[22] Ibrahim, M. H. (2006). Financial integration and international portfolio diversification: US, Japan and ASEAN equity markets. *Journal of Asia-Pacific Business, 7*(1), 5–23. https://doi.org/10.1300/j098v07n01_02
[23] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics & Control, 12*(2–3), 231–254. https://doi.org/10.1016/0165-1889(88)90041-3
[24] Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica, 59*(6), 1551–1580. https://doi.org/10.2307/2938278
[25] Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxford Bulletin of Economics and Statistics, 52*(2), 169–210.
[26] Johansen, S. (1992). Testing weak exogeneity and the order of cointegration in UK money demand data. *Journal of Policy Modeling, 14*(3), 313–334. https://doi.org/10.1016/0161-8938(92)90003-U
[27] Mansourfar, G., Mohamad, S., & Hassan, T. (2010). The behavior of MENA oil and non-oil producing countries in international portfolio optimization. *Quarterly Review of Economics and Finance, 50*(4), 415–423. https://doi.org/10.1016/j.qref.2010.06.007
[28] Imen, G. M., & Rim, A. (2012). A dynamic analysis of financial contagion: The case of the subprime crisis. *Journal of Business Studies Quarterly, 4*(2).
[29] Lean, H. H., & Teng, K. T. (2013). Integration of world leaders and emerging powers into the Malaysian stock market: A DCC-MGARCH approach. *Economic Modelling, 32*, 333–342. https://doi.org/10.1016/j.econmod.2013.02.013
[30] Chen, M.-P., Chen, P.-F., & Lee, C.-C. (2014). Frontier stock market integration and the global financial crisis. *North American Journal of Economics and Finance, 29*, 84–103. https://doi.org/10.1016/j.najef.2014.05.004