fractal and fractional

Article

Global Collective Dynamics of Financial Market Efficiency Using Attention Entropy with Hierarchical Clustering

Poongjin Cho [1] and Kyungwon Kim [2,*]

[1] Department of Industrial Engineering, Hanyang University, Seoul 04763, Korea
[2] School of International Trade and Business, Incheon National University, Incheon 22012, Korea
* Correspondence: thekimk@inu.ac.kr

Abstract: The efficient market hypothesis (EMH) assumes that all available information in an efficient financial market is ideally fully reflected in the price of an asset. However, whether the reality that asset prices are not informational efficient is an opportunity for profit or a systemic risk of the financial system that needs to be corrected is still a ubiquitous concept, so many economic participants and research scholars have conducted related studies in order to understand the phenomenon of the financial market. This research employed attention entropy of the log-returns of 27 global assets to analyze the time-varying informational efficiency. International markets could be classified hierarchically into groups with similar long-term efficiency trends; however, at the same time, the ranks and clusters were found to remain stable only for a short period of time in terms of short-term efficiency. Therefore, a complex network representation analysis was performed to express whether the short-term efficiency patterns have interacted with each other over time as a coherent picture. It was confirmed that the network of 27 international markets was fully connected, strongly globalized and entangled. In addition, the complex network was composed of two modular structures grouped together with similar efficiency dynamics. As a result, although the informational efficiency of financial markets may be globalized to a high-efficiency state, it shows a collective dynamics pattern in which the global system may fall into risk due to the spread of systemic risk.

Keywords: attention entropy; international financial markets; market efficiency; adaptive market hypothesis; clustering in machine learning; collective dynamics

check for updates

Citation: Cho, P.; Kim, K. Global Collective Dynamics of Financial Market Efficiency Using Attention Entropy with Hierarchical Clustering. *Fractal Fract.* **2022**, *6*, 562. https://doi.org/10.3390/fractalfract6100562

Academic Editor: Carlo Cattani

Received: 2 September 2022
Accepted: 28 September 2022
Published: 3 October 2022

1. Introduction

The dynamics of global financial markets are of interest to numerous researchers and economists. One of the most studied topics is that informational efficiency due to fluctuations in market prices has a large impact on the economy through various channels; therefore, the market efficiency test is related to the appropriate price. The efficient market hypothesis (EMH) is one of the major paradigms of financial economics and is widely known to stock market participants in financial markets [1,2]. According to the hypothesis, in an efficient financial market, all information is ideally fully reflected in the asset price [3]. In other words, when new information about an individual asset or market appears, traders anticipate the new asset value and immediately price it [4]. Although the value can change according to information, the price of an asset is immediately reflected in the price, so it is difficult to predict the future price. This idea is also frequently associated with the concept of a random walk, in which future asset price movements move randomly away from previous prices; therefore, an attempt to predict the future asset price with high informational efficiency is not better than a random guess. These approaches place significant constraints on financial trading, making it impossible to profit beyond efficiency markets through arbitrage trading or certain trading strategies [4].

EMH is one of the very important concepts in financial economics; however, actual financial market movements are not ideally efficient [5]. In fact, most financial assets have an increased autocorrelation in a short period of time and also have a higher correlation with other assets [6]; therefore, this phenomenon is the basis for the idea that the short-term forecasting or arbitrage of an asset is possible. Another piece of evidence showing that financial assets are not ideally informational efficient is that the distribution of the log-return of asset prices is not a Gaussian distribution but a fat-tailed distribution [7,8]. In addition, it is very impractical to predict with random walk models in real short-term forecasts [9], and there is no trading strategy that always yields profits [10]. In theory, it is believed that economic bubbles do not exist and collapses do not occur in informational efficient markets [11]; however, in reality, they are becoming more predictable because of the long-range correlation that these events generate [12]. In the end, how to quantify market efficiency by reflecting these actual market dynamics has been an important challenge for a long time.

The reality that asset prices are not informational efficient is still a ubiquitous concept for many economic players and research scholars. In order to understand the phenomenon of the financial market, it is necessary to analyze whether this reality is an opportunity for profit or a systemic risk of the financial system that needs to be corrected. Despite the growing number of participants that better understand the EMH of the market, it is assumed that most assets remain constant and unchanged over the analysis period; therefore, the realistic topic of changing the informational efficiency of assets over time has been relatively less studied. In addition, in reality, the assumption is very improbable, as many factors influence and change the level of market efficiency over time [13]. Lo and MacKinlay [14] proposed a methodology used to test whether a market is efficient; meanwhile, Lo [15] introduced the adaptive market hypothesis (AMH) based on the assumption that market efficiency evolves over time in order to quantitatively estimate market efficiency. The time-varying methodologies associated with AMH have led to further studies to quantify market effectiveness [16–18]. These concerns and interests can be supported by many studies that try to quantify the efficiency degree of financial assets [19–24]. In this research, in a similar vein, through time-dependent informational efficiency, it is possible to not only capture the collective dynamics in the evolution of market efficiency but also to quantify the stability of various efficiency ranks or clusterings in financial markets.

This study analyzed the collective dynamics of 27 global financial markets that have responded to various global events over the past 15 years. Attention entropy [25] was used to analyze the time-varying informational efficiency of the asset's log-returns. The model was an easily comparable efficiency measure, quantifying the market efficiency across multiple assets, time and data frequencies. In particular, in order to quantify the time-dependent informational efficiency of financial assets, the time evolution of efficiency was defined using attention entropy calculated by sliding time-windows to log-returns. When the value of attention entropy is high, it implies that the randomness of peak points in the data is high and the market is efficient; on the other hand, when the entropy is low, the randomness of peak points decreases and the movement of the market deviates from a random walk. Next, the hierarchical clustering procedure, one of the machine learning methodologies, was applied using attention entropy to identify the collective dynamics of international markets [26,27]. Using the machine learning methodology, a further understanding of how markets are grouped and move similarly over time and how the groups are organically connected to each other and evolve over time was expected. As a result, the results of this research enable market policy makers and researchers to fully understand the characteristics of time-varying dynamics and facilitate the realistic decision making.

The novelty of this research is that it is the first research to analyze the time-varying informational efficiency using attention entropy as an input to hierarchical clustering, confirming the modular structures of the complex network. This study analyzed the time-varying structure of market efficiency in a long sample of approximately 15 years in order to consider the effects of several global events. First, it was confirmed that the major international financial markets can be classified hierarchically into 15 clusters according to the similarity of long-term time-varying evolution in the degree of informational efficiency. In the time-varying patterns of each cluster, the market efficiency is stable only for a short period of time, and most of the time-varying patterns are unstable; however, the clustering based on the long-term behavior may not include important characteristics of short-term interactions between markets. Therefore, to overcome the problem, a dynamic clustering approach was applied to confirm the short-term similar efficiency patterns using a sliding time window. By using these time-varying clusters, secondly, the markets with similar short-term efficiency trends were represented by similar ranks and classified clusters. According to global events, these markets have repeatedly observed unstable movements for 1 to 2.5 years. Finally, the weighted network structure was constructed using the stochastic block model (SBM) approach to identify the collective dynamics based on the complex interactions between markets in time-varying clusters. Through the network, we discovered a modular structure composed of three distinct market clusters, which include the most influential efficiency trends. Since the networks are profoundly dense and entangled, we demonstrated that the efficiency dynamics of global financial markets are collective dynamics that can cause the entire financial system to operate with undoubtedly high informational efficiencies or continue to operate under systemic risk.

This paper is organized as follows: Section 3 describes the attention entropy for measuring the informational efficiency, and the hierarchical clustering and SBM, which reflect the efficiency interactions of international markets; Section 4 describes the experiments and the statistical explanation of the index data; Section 5 discusses the empirical results of this paper; and Section 6 concludes.

2. Literature Review

There have been many studies carried out to determine whether the market becomes inefficient when certain events occur in the market. For example, Gaio et al. [28] confirmed that the Russia–Ukraine conflict had an impact on market efficiency with a multifractal structure, and Hkiri et al. [29] demonstrated that country-specific political shocks in emerging stock markets influenced market efficiency. Since entropy quantifies the certainty of information, various entropy-based models have been developed to measure the market efficiency. For instance, Espinosa-Paredes et al. [30] exhibited large deviations from the efficiency of the WTI market during the COVID-19 pandemic period through singular value decomposition (SVD) entropy, and Wang [31] combined fuzzy entropy and transfer entropy to identify the market inefficiency of energy, metal and financial markets during extreme events such as COVID-19. Shternshis et al. [32] confirmed that market efficiency can also be measured through the high-frequency data with Shannon entropy in the ETF market. Dinga et al. [33] argued that behavioral entropy better explains the behavioral efficiency of the market.

Alternatively, as an effort to determine market efficiency, a study on a multifractal detrended fluctuation analysis (MF-DFA) approach was also conducted. For example, Choi [34] compared market efficiency during the global financial crisis and the COVID-19 pandemic through MF-DFA, and Mensi et al. [35] confirmed through MF-DFA that GCC stock markets are less efficient than global markets, demonstrating the time-varying persistence. Han et al. [36] compared whether the 2015 stock market crash had an effect on efficiency for three sub-markets of the Chinese stock market through MF-DFA. Mensi et al. [37] compared the market efficiency of the European stock market using the MF-DFA approach, where Greece was the most inefficient market. The asymmetric multifractal detrended fluctuation analysis (A-MFDFA) model considers the asymmetric nature of the market efficiency, which is a measure used to better understand the market Cho and Lee [38], Mensi et al. [39] measured the market efficiency of the metals' futures markets during financial and oil crises through the A-MFDFA. Zhuang and Wei [40] ranked the informational inefficiency levels of green finance markets using A-MFDFA.

In order to filter out the economic information from the network structures, a minimum spanning tree (MST) model was widely applied. Bonanno et al. [41] constructed a correlation-based MST network model where the observed topology consists of a star-like structure with one large cluster and a complex multi-cluster structure. In this study, the network was composed of two modular structures and demonstrated a collective dynamic pattern, which differs in that it is vulnerable to the spread of systemic risk. In recent studies, Kuang [42] measured the information flow of the market through transfer entropy, and demonstrated the risk contagion path by constructing an entropy-based network via MST, and Shin et al. [43] discovered a huge cluster during a financial crisis with detrended cross-correlation analysis (DCCA) and the Girvan–Newman method, which is a hierarchical method used to detect communities in complex systems. Balcı et al. [44] confirmed a remarkable change in the coarse graining network in an economic crisis using Voronoi regions. Alves et al. [45] quantified the time-varying efficiency of stock markets with permutation entropy, and demonstrated a modular structure with hierarchical clustering and network representation.

3. Methods

3.1. Attention Entropy

Attention entropy quantifies key patterns through the distribution of intervals, whereas classical entropies are based on the frequency of elements. Attention entropy first defines the key pattern as peak points, which include local maxima and local minima. Then, the intervals between the local values can be computed, and Shannon entropy [46] of the intervals is calculated. For example, given a time series $X = \{x_1, x_2, \ldots, x_n\}$, we can construct the local maxima series of X as $P = \{x_{p_1}, x_{p_2}, \ldots, x_{p_k}\}$ with the index $p = \{p_1, p_2, \ldots, p_k\}$. Similarly, the local minima series of X can be represented as $Q = \{x_{q_1}, x_{q_2}, \ldots, x_{q_l}\}$ with the index $q = \{q_1, q_2, \ldots, q_l\}$. Then, we can compute the interval series from local values to other local values as follows.

$$
\begin{aligned}
I_{max,max} &= \{p_2 - p_1, p_3 - p_2, \ldots, p_k - p_{k-1}\} \\
I_{min,min} &= \{q_2 - q_1, q_3 - q_2, \ldots, q_l - q_{l-1}\} \\
I_{max,min} &= \begin{cases} p_1 - q_2, p_2 - q_3, p_3 - q_4, \ldots & \text{if } p_1 \geq q_1 \\ p_1 - q_1, p_2 - q_2, p_3 - q_3, \ldots & \text{otherwise} \end{cases} \\
I_{min,max} &= \begin{cases} q_2 - p_1, q_3 - p_2, q_4 - p_3, \ldots & \text{if } p_1 \geq q_1 \\ q_1 - p_1, q_2 - p_2, q_3 - p_3, \ldots & \text{otherwise} \end{cases}
\end{aligned} \tag{1}
$$

$I_{max,max}$, $I_{min,min}$, $I_{max,min}$ and $I_{min,max}$ correspond to the interval series from the local maxima to next local maxima, local minima to next local minima, local maxima to next local minima and local minima to next local maxima, respectively. In order to calculate the Shannon entropies of the interval series, such as when composing a histogram, we counted the number of occurrences of unique elements in each set. Then, we can calculate the probability of each element appearing in the set based on the number of appearances. If the probabilities are called as $f = \{f_1, f_2, \ldots, f_m\}$, we can evaluate the Shannon entropy, $H(f)$, as follows.

$$H(f) = -\sum_i f_i \, ln(f_i) \tag{2}$$

Let $f_{max,max}$, $f_{min,min}$, $f_{max,min}$ and $f_{min,max}$ be the probability series of the interval series $I_{max,max}$, $I_{min,min}$, $I_{max,min}$ and $I_{min,max}$, respectively. Then, the attention entropy of the times series X is as follows.

$$H(X) = \frac{f_{max,max} + f_{min,min} + f_{max,min} + f_{min,max}}{4} \tag{3}$$

Yang et al. [25] proposed the attention entropy for classifying time series data, constructing the frequency distribution of interval series. The entropy outperforms state-of-the-art entropy methods, and is expected to show the explanatory power of market efficiency in our study as well. If attention entropy is high, the randomness of peak points of the data increases; therefore, it is expected to achieve a high informational efficiency.

On the other hand, the entropy decreases when decreasing the randomness of peak points, so low informational efficiency is expected.

3.2. Hierarchical Clustering

In order to calculate the similarities of the efficiency degree among 27 international markets with the attention entropy series, we utilized the correlation distance as follows:

$$d(H_i, H_j) = \sqrt{2(1 - \rho(H_i, H_j))}, \tag{4}$$

where $\rho(H_i, H_j)$ is the coefficient of the Pearson correlation between the attention entropy series H_i of the i-th market and the entropy series H_j of the j-th market. For the dynamical clustering analysis, H_i and H_j were obtained by sampling the series of efficiency with a 1-year (252-day) sliding time window. Using the correlation distance in Equation (4), we used four linkage algorithms—"Single", "Average", "Complete" and "Ward"—to determine the hierarchical clustering. The algorithms recursively merged the pair of clusters in the direction of minimally increasing the within-cluster variance. Finally, the number of clusters was determined by maximizing the silhouette score [47]. The higher the value of the score, the better the cluster structure configuration.

3.3. Stochastic Block Model (SBM)

To reflect the complex interactions of markets, the network structure was constructed using the SBM approach. We applied a total of four SBMs to 27 international markets: SBM, degree-corrected SBM (DCSBM), nested SBM (nSBM) and nested DCSBM (nDCSBM). Among them, the one with the smallest minimal description length was considered as the best model. This method directly estimates the probability that each node belongs to a certain cluster, calculating the marginal probabilities that the network will be divided into a certain number of clusters. To evaluate the method, we used 10,000 sweeps of a Metropolis–Hastings acceptance–rejection Markov chain Monte Carlo (MCMC) with several moves, and collected the partitions of the hierarchical network. In addition, the marginal probability that each market belongs to each group was also estimated. In order to identify the centrality of the network, the PageRank centrality [48], a variant of eigenvalue centrality, was also measured, which uncovered the influential nodes of the network.

4. Experiments and Data

4.1. Experiments

For the experiment, from 27 international indexes, we obtained the adjusted daily closing price and took the logarithm to the return of each index. Through the log-returns and the attention entropy, the final input for analysis was calculated. Note that we ensured that the attention entropy could be fundamental data for long-term and short-term dynamics. In order to confirm the long-term efficiency, we constructed a hierarchical clustering model with the correlation distance and silhouette score. For cross-validation, 4 linkages were applied, and we analyzed the correlation distance matrix and dendrogram of the clustering model. Based on the sliding window model for the short-term dynamics, we observed the Kendall rank correlation (Kendall-τ) [49] and adjusted rand index (ARI) [50]. Kendall-τ measures the similarity of informational efficiency ranks between different time points. If high rank correlation is maintained for a certain period of time, it means that a specific event affects the efficiency of the entire system. In order to cross-check the result of Kendall-τ, we also observed ARI, which evaluates the correspondence between two clusters. Furthermore, constructing the SBM network, we drew inferences of the short-term dynamics with the network and PageRank centrality. A step-by-step scenario of the proposed experiment is illustrated in Figure 1.

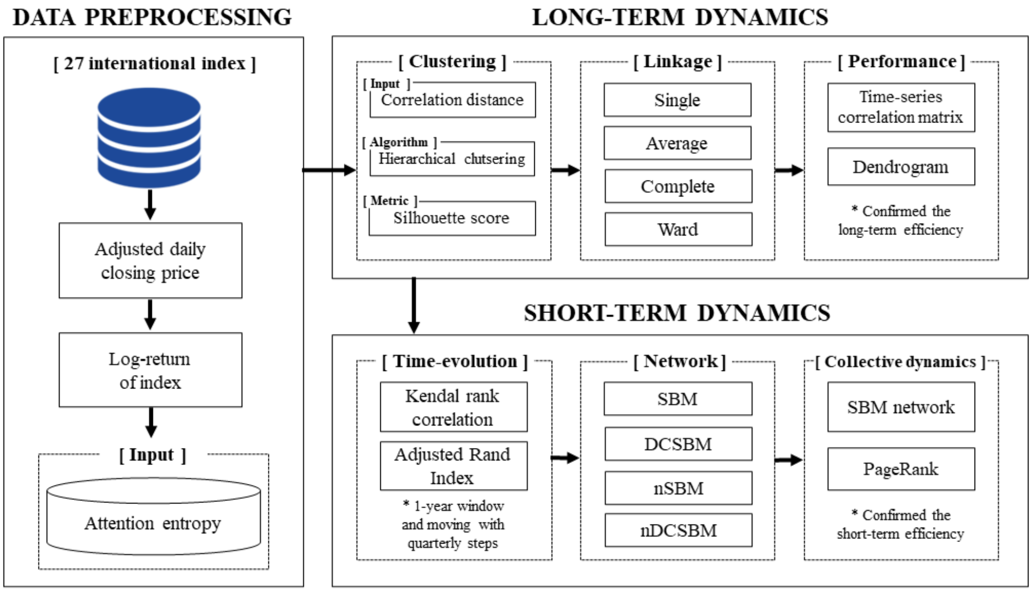

Figure 1. Proposed market efficiency model framework using attention entropy with hierarchical clustering.

4.2. Data

The dataset used in the study was obtained from historical data of Yahoo Finance. Since we analyzed the time-varying efficiency dynamics of 27 major finance markets over a long period of approximately 15 years, the market efficiency results were accurate only when all values existed at the same time. As a result, missing data or incomplete data were preprocessed through Wall Street Journal data. A total of 27 markets were selected based on Yahoo Finance's World Indices category, and a total of 3912 adjusted daily closing prices were used from March 2007 to July 2022. Therefore, our dataset covered various areas of international financial markets over a long period of 15 years, including the recent global issue of COVID-19. Figure 2 is a scaled value plot for the international financial markets used for the analysis, and Figure 3 is the average plot for the scaled value. For the supplementary explanation of Figure 2, scaled value plots according to geographical region are plotted in Appendix A. Various movement patterns could be confirmed due to the sudden shock in the global financial crisis, the increase due to quantitative easing policy and the recent COVID-19. VIX index, the volatility index of the market, sometimes rises in global events, but not in others, while the dynamic behavior has not clustered recently, but moves similarly and diversely. In the dynamics, we used the log-return, $R_t = \log P(t) - \log P(t-1)$, of international financial market indices as input to analyze the time-varying market efficiency, where $\log P(t)$ and $\log P(t-1)$ were the logarithm closing prices at time t and $t-1$. Table 1 is the descriptive statistics of 27 log-returns. On average, it had a mean (Mean) value of 0.02 and a standard deviation (Stdev) of 1.58. Since the skewness was -0.52, the positive log-return value was widely distributed, and the kurtosis was 11.01, which implies a large difference from the normal distribution. In addition, since the null hypothesis that there is no autoregressive conditional heteroscedasticity (ARCH) effect at lags 10 and 20 was rejected at a significance level of 1%, the conditional heteroscedasticity existed in all market log-returns. In addition, augmented Dickey–Fuller (ADF) also rejected the null hypothesis of the series being stationary at 1% significance level, so the international market log-returns did not move stably in a specific range but moved dynamically.

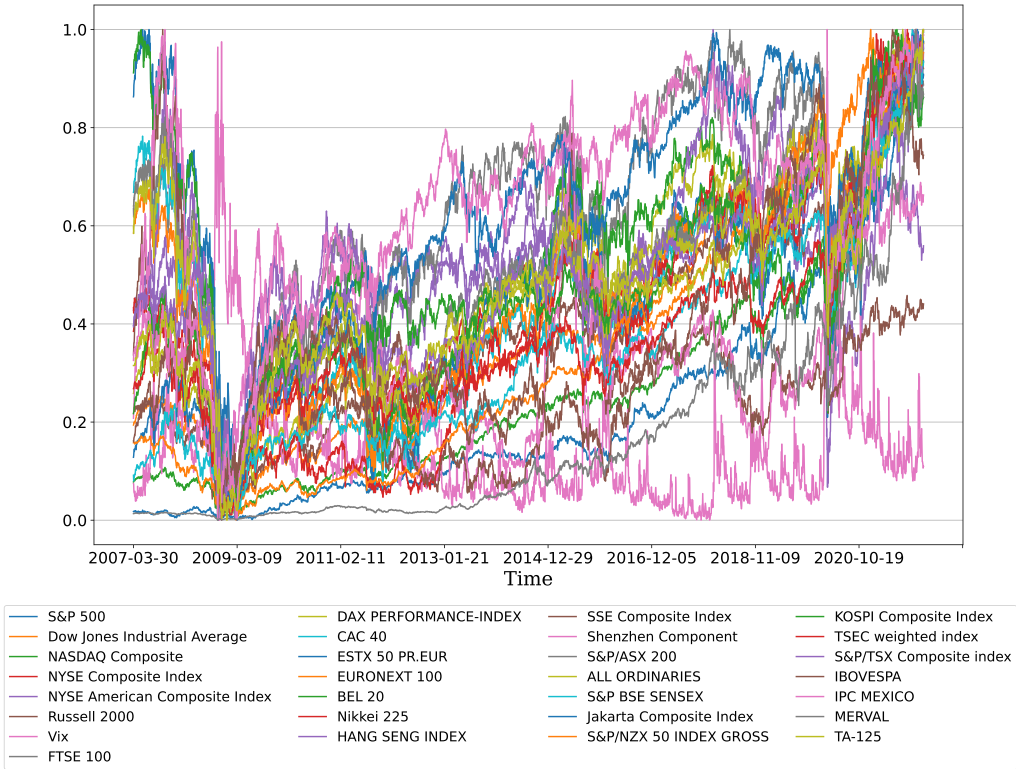

Figure 2. Evolutionary dynamics of the entire samples of international financial market indices between Q1 2007 and Q2 2022 used in the study.

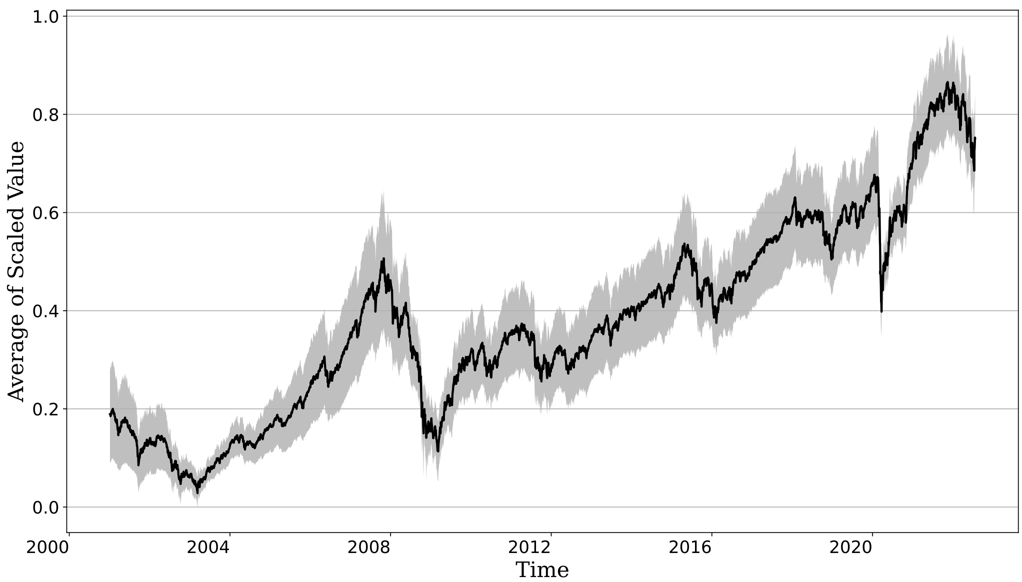

Figure 3. Time evolution of the average scaled value of all 27 market indices between Q1 2007 and Q2 2022 used in the study (gray shaded band means the standard error of the average).

Table 1. Descriptive statistics for the entire sample of international financial log-returns.

	Mean	**Stdev**	**Skewness**	**Kurtosis**	**ARCH(10)**	**ARCH(20)**	**Augmented Dickey–Fuller**
S&P 500	0.07	1.55	0.16	3.53	230.1 ***	250.07 ***	−29.91 ***
Dow Jones Industrial Average	0.03	1.22	−0.49	16.05	1262.23 ***	1311.84 ***	−15.14 ***
NASDAQ Composite	0.05	1.38	−0.50	9.24	998.5 ***	1045.3 ***	−14.49 ***
NYSE Composite Index	0.02	1.32	−0.68	13.41	1225.34 ***	1295.24 ***	−14.79 ***
NYSE American Composite Index	0.01	1.31	−1.00	16.49	984.88 ***	1161.64 ***	−15.2 ***
Russell 2000	0.03	1.62	−0.66	8.60	1146.88 ***	1214.77 ***	−14.73 ***
Vix	0.00	7.70	1.05	6.14	192.18 ***	195.78 ***	−25.58 ***
FTSE 100	0.00	1.20	−0.41	10.03	788.14 ***	834.46 ***	−23.79 ***
DAX PERFORMANCE-INDEX	0.02	1.39	−0.23	8.39	583.26 ***	674.27 ***	−22.63 ***
CAC 40	0.01	1.42	−0.28	8.25	628.1 ***	688.32 ***	−23.59 ***
ESTX 50 PR.EUR	0.00	1.43	−0.32	7.99	589.85 ***	613.74 ***	−29.08 ***
EURONEXT 100	0.01	1.30	−0.38	9.10	692.21 ***	759.31 ***	−23.48 ***
BEL 20	−0.00	1.30	−0.65	10.77	584.09 ***	614.49 ***	−12.24 ***
Nikkei 225	0.01	1.47	−0.45	8.38	1054.16 ***	1098.86 ***	−63.39 ***
HANG SENG INDEX	0.00	1.48	−0.03	8.96	997.94 ***	1106.29 ***	−10.72 ***
SSE Composite Index	0.00	1.53	−0.61	5.44	408.67 ***	472.35 ***	−12.97 ***
Shenzhen Component	0.02	1.78	−0.55	3.45	364.1 ***	426.68 ***	−14.66 ***
S&P/ASX 200	0.01	1.12	−0.69	7.80	1129.81 ***	1175.54 ***	−14.4 ***
ALL ORDINARIES	0.01	1.07	−0.66	9.66	768.79 ***	817.83 ***	−37.55 ***
S&P BSE SENSEX	0.04	1.37	−0.21	13.61	543.65 ***	561.24 ***	−11.96 ***
Jakarta Composite Index	0.03	1.27	−0.58	9.61	508.49 ***	614.77 ***	−18.47 ***
S&P/NZX 50 INDEX GROSS	0.03	0.74	−0.65	8.21	886.56 ***	947.68 ***	−21.0 ***
KOSPI Composite Index	0.02	1.23	−0.55	10.44	966.23 ***	990.46 ***	−12.24 ***
TSEC weighted index	0.02	1.14	−0.42	5.11	488.16 ***	541.14 ***	−14.08 ***
S&P/TSX Composite index	0.01	1.15	−1.08	20.75	1126.95 ***	1262.88 ***	−11.45 ***
IBOVESPA	0.02	1.74	−0.44	10.17	1174.02 ***	1267.65 ***	−26.78 ***
IPC MEXICO	0.02	1.18	−0.02	7.29	724.97 ***	843.78 ***	−27.88 ***
MERVAL	0.10	2.32	−2.67	50.22	37.67 ***	58.42 ***	−61.87 ***
TA-125	0.02	1.09	−1.17	12.07	466.91 ***	527.72 ***	−18.56 ***
Average	0.02	1.58	−0.52	11.01	-	-	-

*** Significant at the 1% level.

5. Results

5.1. Long-Term Efficiency with Clustering

The results are based on the adjusted daily closing price of 27 international financial markets from March 2007 to July 2022. From these historical time series, we could estimate the logarithm return, R_t, for each market price where t is the closing date. Next, we sampled the R_t with a 252-day sliding time window, roughly corresponding to 1 year of economic trading activity. The time window moves with daily steps, and we can estimate the $H(X)$ for every step. We refer to the attention entropy obtained at time t with a 252-day sliding time window as H_t. Then, a time series for each market was created, as shown in Figure 4 (Upper) for the S&P500, which is an influential indicator of the international equity market. The higher the value of H_t, the more informational efficient at that particular time. Conversely, a decrease in H_t indicates a more inefficient period of the market. We also estimated the average dynamics of the efficiency, H_t, with a daily step over all international markets. Figure 4 (Lower) shows that the average behavior is smoother than the behavior of the S&P500 and appears to reflect major events such as "Global Financial Crisis (2007∼2008)", "Chinese Stock Market Turbulence (2015∼2016)" and "COVID-19 (2020∼)", as H_t shows more lower values around these time periods.

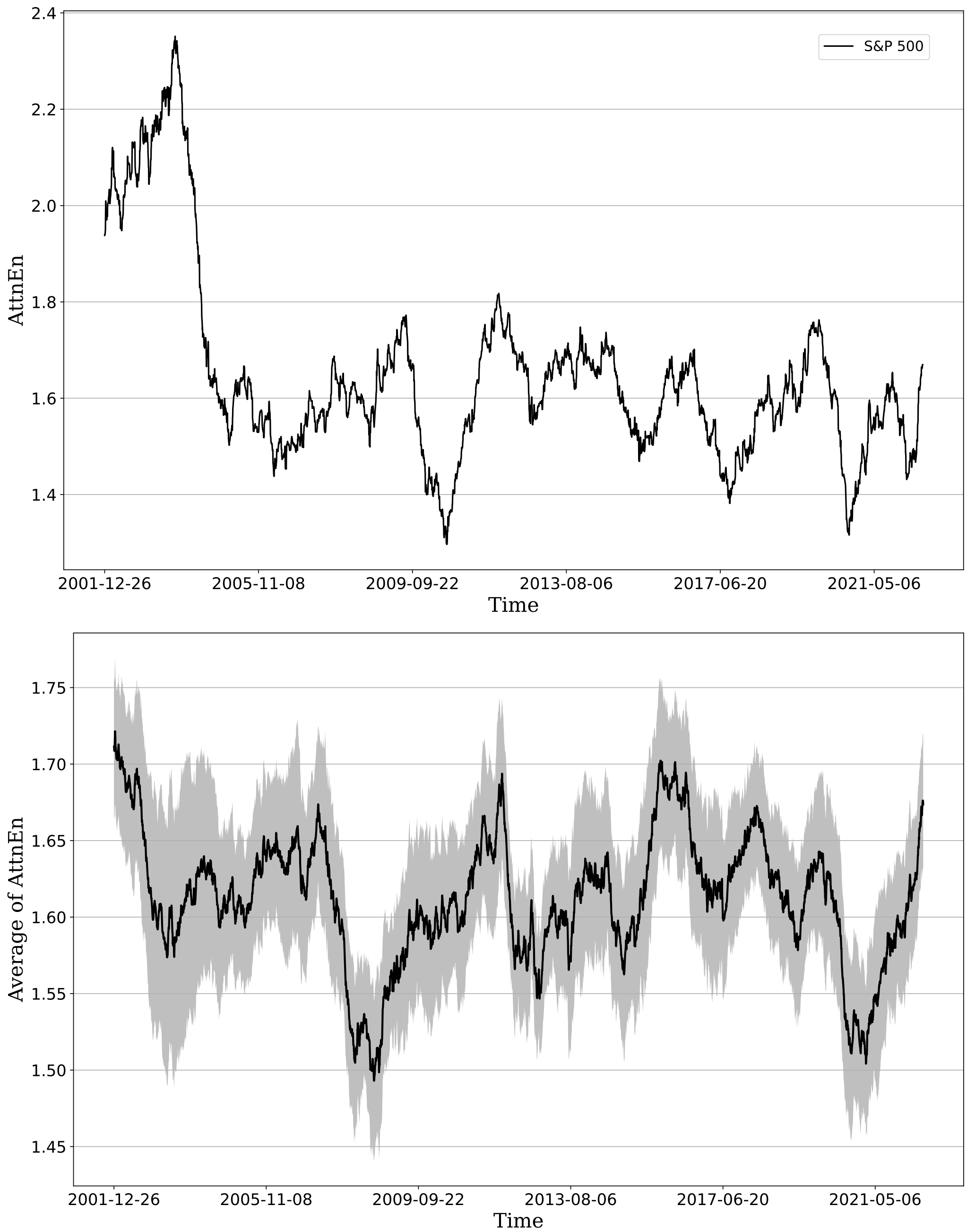

Figure 4. Estimating the informational market efficiency of international markets with AMIM. (**Upper**) Time evolution of the AMIM with daily step of the S&P 500. (**Lower**) Time evolution of the average AMIM of all 27 markets (gray shaded band means the standard error of the average).

When a specific event occurs in the financial market, such as COVID-19, the behavior of investors tends to be synchronized, so the attention entropy decreases with an inefficient period of the market. Therefore, it is possible to identify trading opportunities through forecasting the synchronized behavior; otherwise, their investment strategies propagate shocks to interconnected networks globally, resulting in strong correlations between markets. Similarly, the collective behaviors strongly affect the informational efficiency of markets and naturally cause joint movements of H_t. Accordingly, there are groups with similar informational efficiency trends that can appear as hierarchical structures, so policy makers can manage financial risk and establish effective market policies through the informational efficiency trends. In order to analyze the groups, we calculated the correlation distance of the informational efficiency among all pairs of 27 international financial markets as shown in Figure 5. At the same time, the cluster dendrogram of the distance was estimated using Ward's minimum variance, divided into 15 groups as in Figure 6. In addition, for cross validation, silhouette scores were estimated for all four linkages. As a result, 15 clusters were determined as the optimal number by the Ward's algorithm. The largest cluster consists of three markets and the smallest cluster consists of just one market. Each group of markets shows a similar long-term time evolution of H_t. In addition, most clusters are grouped together by geographically closeness, culturally closeness or countries of a similar economic level as shown in Table 2. In order to evaluate whether the clustering of 15 groups is appropriate, the entropy time series for each group is shown in the right part of Figure 5. There is no overlapping of the time series between different groups, and different entropy series movements are confirmed for each group, so the model is effective in distinguishing between markets with different movements. Therefore, our results indicate that 27 international markets constitute a hierarchical structure according to the evolution of long-term informational efficiency; however, overwhelmingly large clustering is not observed in the correlation distance matrix. In addition, the global long-term analysis does not capture the short-term dynamics, H_t, among the markets.

For example, by understanding the reactions of market traders to global events, investors can manage portfolio risk and identify trading opportunities. In addition, through predicting the inefficiency period, it is possible to provide necessary information to policy makers, thereby contributing to the prevention of risks and the establishment of effective market policies.

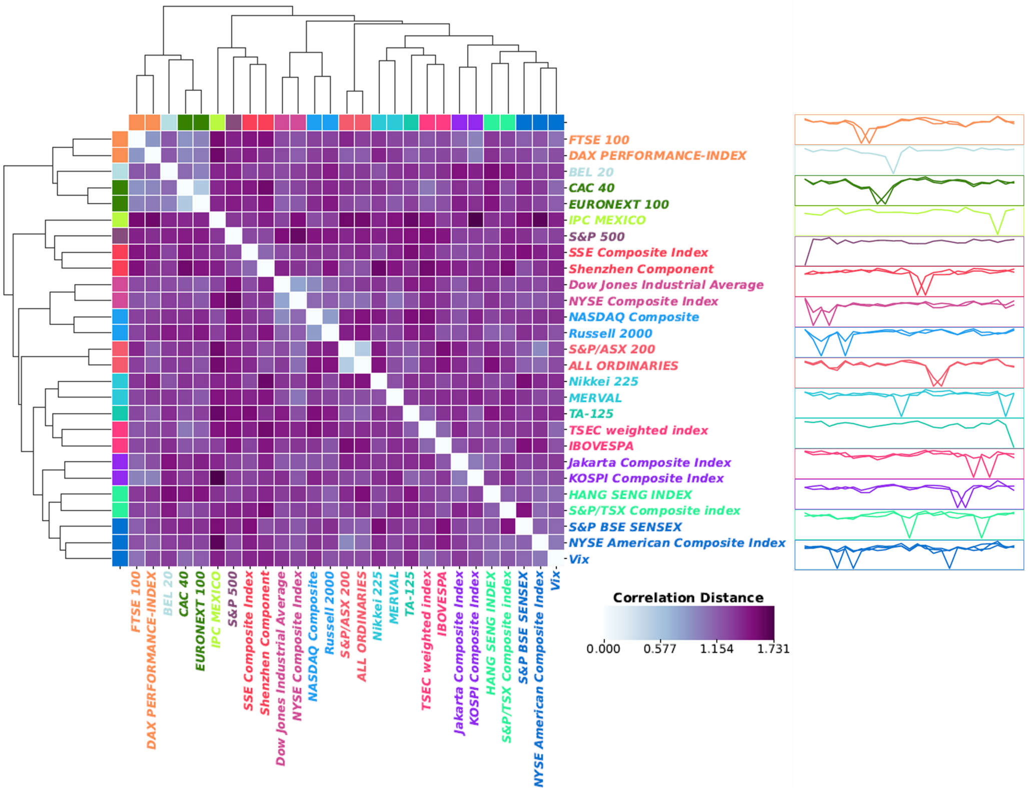

Figure 5. Long-term hierarchical clustering of informational efficiency patterns of 27 international markets. The colored squares below the dendrogram branches indicate the 15 clusters obtained by maximizing the silhouette score.

Table 2. The collection of country lists according to the long-term hierarchical clustering of efficiency patterns.

Cluster	Market	Country
1	FTSE 100 DAX PERFORMANCE-INDEX	United Kingdom Germany
2	BEL 20	Belgium
3	CAC 40 EURONEXT 100	France France
4	IPC MEXICO	Mexico
5	S&P 500	United States
6	SSE Composite Index Shenzhen Component	China China
7	Dow Jones Industrial Average NYSE Composite Index	United States United States
8	NASDAQ Composite Russell 2000	United States United States
9	S&P/ASX 200 ALL ORDINARIES	Australia Australia
10	Nikkei 225 MERVAL	Japan Argentina
11	TA-125	Israel
12	TSEC weighted index IBOVESPA	Taiwan Brazil
13	Jakarta Composite Index KOSPI Composite Index	Indonesia Republic of Korea
14	HANG SENG INDEX S&P/TSX Composite index	Hong Kong Canada
15	S&P BSE SENSEX NYSE American Composite Index Vix	India United States United States

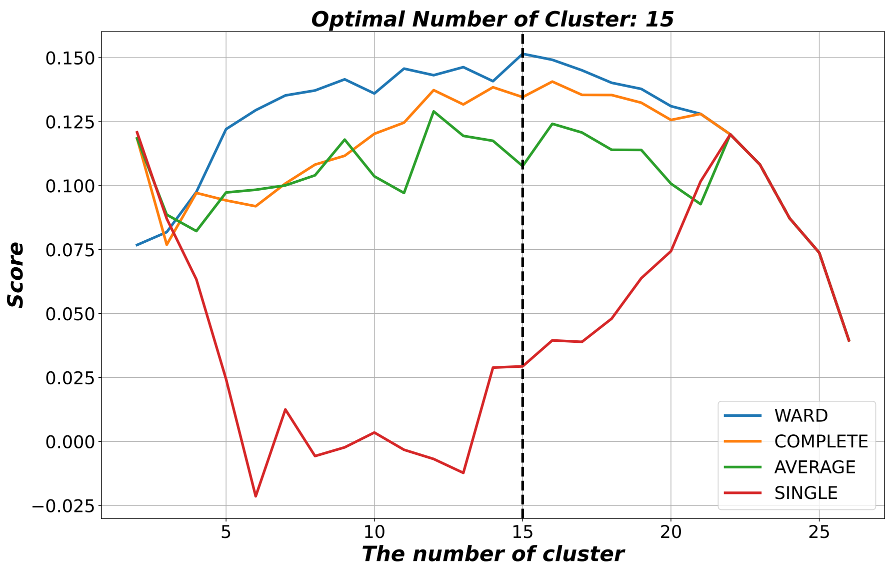

Figure 6. Determining the optimal number of clusters using silhouette scores of four linkages.

5.2. Short-Term Efficiency with Sliding Window

To estimate the short-term collective dynamics of the efficiency degree, we sampled the H_t with a 1-year sliding window, and the window moves with quarterly steps so that the periods of the extracted samples do not overlap. Next, we created a ranking of the efficiency degree within each time window and investigated the stability of rankings by calculating the Kendall-τ for all pairs of time windows as shown in Figure 7 (Left). The names of rows and columns represent the last date for each time window, and the coefficient values of Kendall-τ indicate how similar the rank of informational efficiency is from the past to next quarters. In addition, the large diagonal blocks with high values mean the stable efficiency ranks over long periods. Similar to the results of Figure 4 (Lower), we can observe the large diagonal blocks with approximately 1$\sim$2.5 years of the periods by major finance events, whereas most efficiency ranks are stable for the many short periods by the small blocks in the other periods.

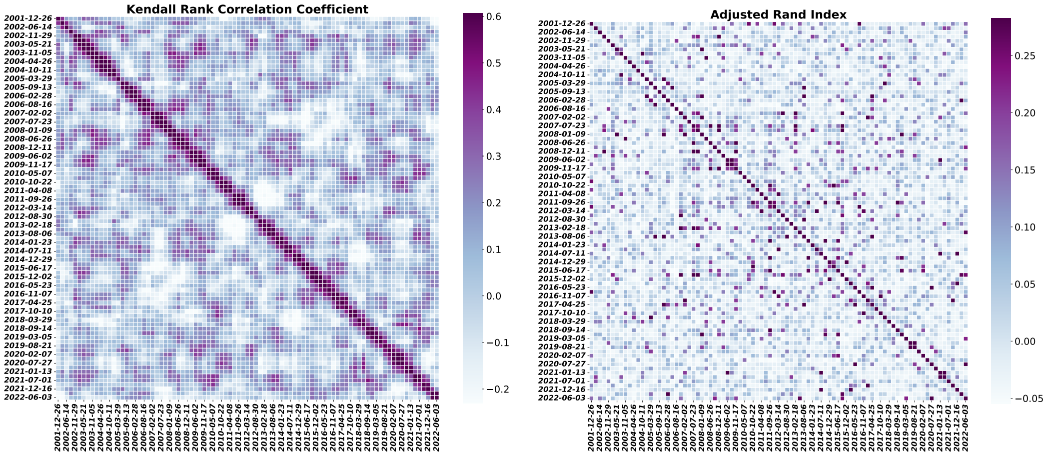

Figure 7. Short-term stability analysis of informational efficiency rankings and clusterings of 27 international markets moving with quarterly steps. (**Left**) Time evolution plot of the Kendall rank correlation coefficient (kendal-τ) for all pairs of efficiency ranks at different time windows. (**Right**) Time evolution plot of the adjusted rank index (ARI) for all pairs of similar efficiency clusters at different time windows.

In order to analyze the short-term dynamics, we also calculated the efficiency degree of the clusters for each time window using the same clustering approach of H_t. For all clusters, the ARI was estimated to investigate the stability of the same group. ARI measures the concordance between two clusters by counting the number of values placed in the same cluster while controlling situations that may overlap by chance. If ARI is close to 0, the two clusters are inferior to random partitions, and if ARI is close to 1, they are perfectly identical. For each time window, therefore, the degree of similarity between markets was determined based on the informational efficiency. Figure 7 (Right) shows the matrix plot by ARI of all cluster pairs. Since the diagonal blocks of up to approximately 1 to 1.5 years are observed, the similar efficiency collections do not remain stable over long periods of time affected by certain events, which is similar to the result of Kendall-τ, but seems to be more conservative. In addition, the large blocks have been observed in other events, except the Global Financial Crisis in 2008, and the same grouping continues in the recent COVID-19 events.

5.3. Short-Term Efficiency with SBM Network

In international financial markets, Figure 7 shows that the short-term patterns of informational efficiency change with time and differ from the patterns obtained from the long-term window data. The simple partitions of international markets found in the study may not be sufficient to capture the complex interactions that exist in real financial markets; therefore, we analyzed the market using a complex network to account for the entangled complex interactions. The nodes of network represent the 27 international markets, and the links indicate that the two markets are clustered together over time at least once. In addition, the connection between two markets is weighted according to the number of times that those markets are grouped into the same cluster. In addition, there may be a modular structure in which certain markets are frequently grouped into the same cluster and exhibit high edge weights. In order to analyze the possibility, SBM was used, and a total of four types of SBM were applied to the international financial markets, where the nested SBM without degree correction was the best model to explain our network structure. Through the process, the aggregated short-term information was easily transformed into a global coherent picture in which the informational efficiency dynamics of 27 markets over some periods were highly correlated. The strength of the connections also indicates the degree of interactions between the markets.

Figure 8 shows the complex network of the 27 international markets. The results suggest that the majority are almost always allocated to the same cluster. The network consists of three network modules, and is indicated by different node colors. In addition, through 10,000 sweeps of a Metropolis–Hastings acceptance–rejection MCMC, the probabilities of belonging to certain clusters were measured, and they were displayed together in each partition. The largest network module includes 15 markets and the second partition contains the remaining 11 markets. Since the NYSE Composite Index is highly related to the second cluster, it is better to view the entire system as two modules. Figure 9 demonstrates the centrality ranking of the network using PageRank. The NYSE Composite index (United States), Vix (United States) and NASDAQ Composite (United States) emerge as the most influential international markets, whereas the TSEC weighted index (Taiwan), MERVAL (Argentina) and SSE Composite Index (China) are the least influential for the informational efficiency dynamics. Similar to the result of Figure 8, we can observe that the NYSE Composite Index plays an important role in the network.

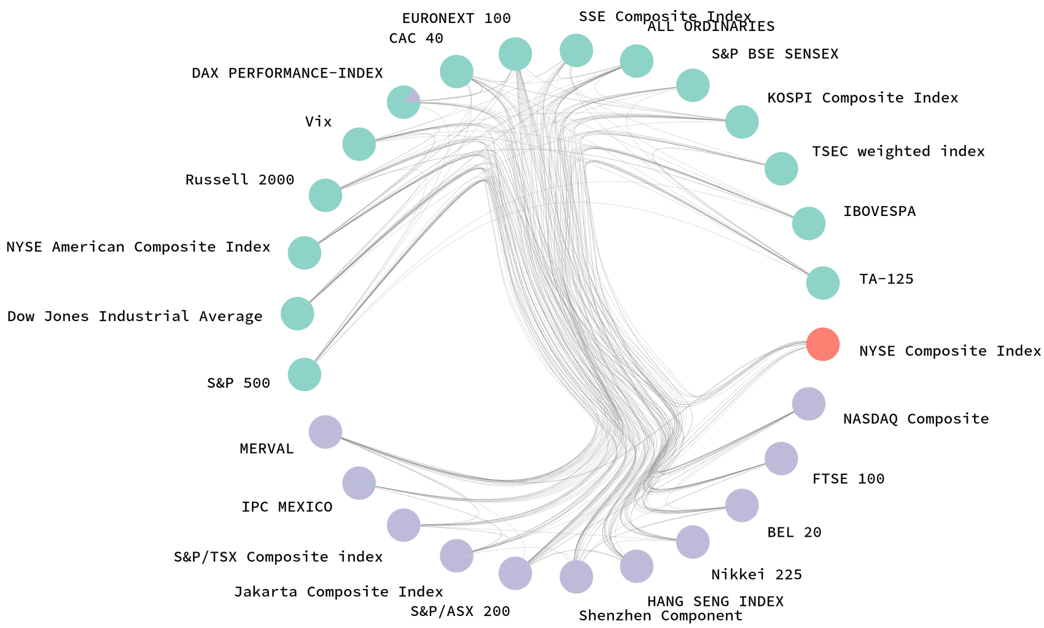

Figure 8. The network structure of international markets exhibiting similar short-tern efficiency trends. By estimating the nested SBM without degree-corrected model, there are two network partitions indicated by the different colors, and edge widths are proportional to the weights of the connections.

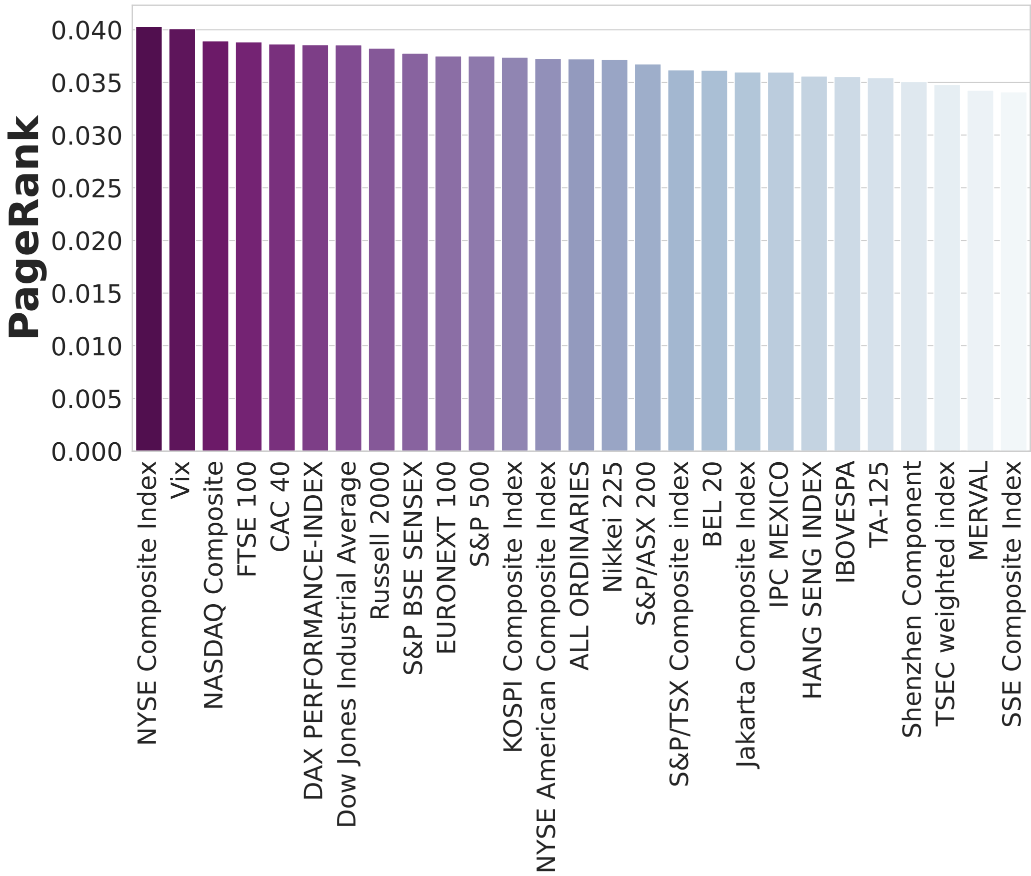

Figure 9. Centrality ranking by the PageRank suggesting the most influential international finance markets for the informational efficiency evolution.

Although there are many exceptions, the geographical proximity or similar cultures seems to play an important role in distinguishing the partitions. Furthermore, the network is a complete graph since it has all possible connections for all markets. The international finance markets, therefore, are strongly globalized with respect to their informational efficiency, meaning that a large number of markets may have the same efficiency level in simultaneous periods. In other words, it implies that there may be a systemic risk that low-efficiency states propagate, and, at the same time, high-efficiency states may also appear globally. In addition, the inequality of edge weights means that international markets can have a modular structure. Specifically, unlike other markets, the market pairs with high edge weights are much more grouped into clusters. Although the association of international markets is quite entangled, the meaning of the structure suggests that some similar market groups may become more similar to each other, and the degree of efficiency may maintain the partitioned structures.

6. Discussion and Conclusions

The problem of quantifying market efficiency by reflecting the inefficient dynamics of financial markets has long been a major challenge. In this paper, using attention entropy, we examined the dynamic market efficiency of 27 global financial markets during the past 15 years. It was found that the international financial markets can be hierarchically clustered into 15 groups with a similar long-term efficiency trend; however, the long-term clusters may be insufficient for understanding the short-term collective dynamics of informational efficiency. In fact, the short-term patterns of efficiency have changed over time, so there is a difference from the long-term results. The short-term efficiency ranks are stable only for a relatively short period, and, similarly, clusters with similar efficiency trends do not maintain stable movement for a long time. Therefore, based on the facts, the complex network representation was analyzed to understand how the collective dynamics of informational efficiency in global financial markets interact with each other via various instabilities over time. The network of 27 international markets is fully connected, confirming that it is strongly globalized and entangled. In addition, the complex network demonstrates that two modular structures are grouped together with similar efficiency dynamics, and a high or low efficiency appear in both modules at the same time. In other words, global markets are highly intertwined with each other in terms of efficiency; however, the degree of efficiency can be particularly pronounced for markets belonging to each group. Therefore, the complex global financial network revealed in our study suggests that it is not only a systemic risk that may occur due to the spread of a low efficiency, but a high efficiency may also appear at the global level as well.

Market efficiency is not only an opportunity for profit, but it can also be an early signal of an impending crisis in the market. Our findings will help regulators to assess the current state of the market and predict the future in order to measure the financial risk. For example, by understanding the reactions of market traders to global events, investors can manage portfolio risk and identify trading opportunities. In addition, through predicting the inefficiency period, it is possible to provide necessary information to policy makers, thereby contributing to the prevention of risks and the establishment of effective market policies. Therefore, the quantification of the time-varying informational efficiency will help us to identify the transaction risk of numerous economic actors in advance. In this study, it was confirmed that the modular structures were constructed with similar efficiency dynamics through the SBM network, which leads to a different result from the study of Bonanno et al. [41], which composes one large cluster around a large hub. Since the data and methodology are different from our model, it seems difficult to make a simple comparison; therefore, in future studies, a comparison with existing methods, such as an MST-based network, can be attempted. In addition, a directed graph can be used to identify the efficiency direction between markets; moreover, attention entropy obtained through other time frequency data can be compared.

Author Contributions: Conceptualization, K.K.; methodology, K.K.; software, K.K.; validation, K.K.; formal analysis, K.K.; investigation, K.K.; resources, K.K.; data curation, P.C.; writing—original draft preparation, K.K. and P.C.; writing—review and editing, K.K. and P.C.; visualization, K.K. and P.C.; supervision, K.K.; project administration, K.K. and P.C.; funding acquisition, P.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2022-00166875).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

For the supplementary explanation of Figure 2, scaled value plots for attention entropy according to geographical region are plotted as follows.

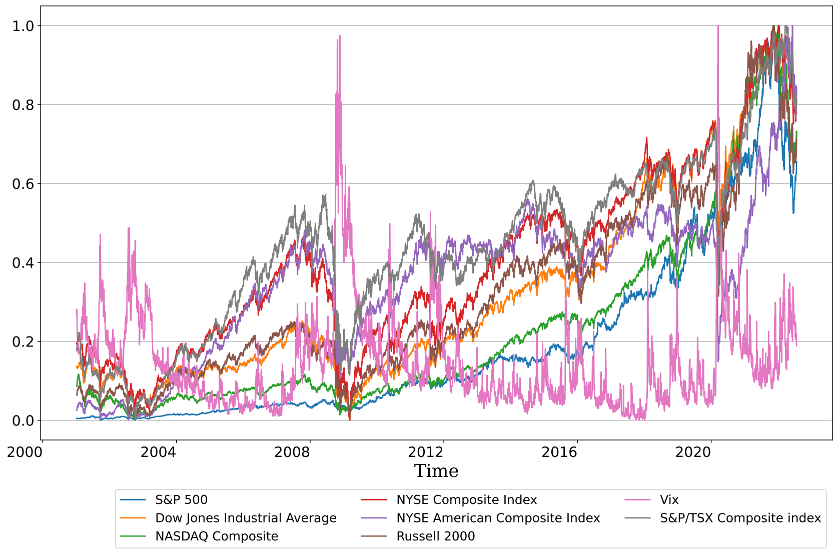

Figure A1. Evolutionary dynamics of North American financial market indices.

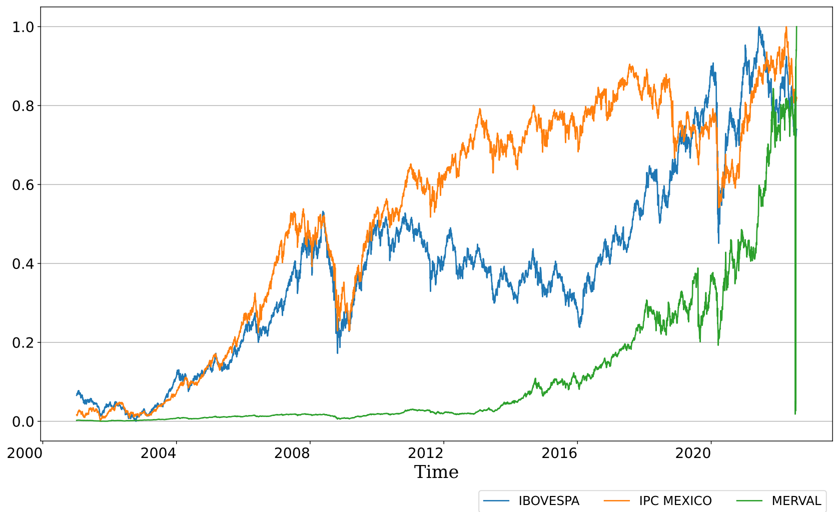

Figure A2. Evolutionary dynamics of South American financial market indices.

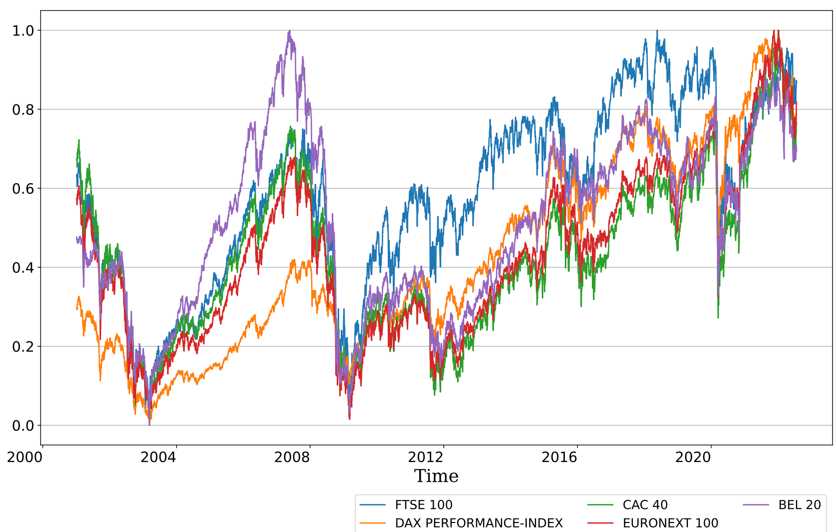

Figure A3. Evolutionary dynamics of European financial market indices.

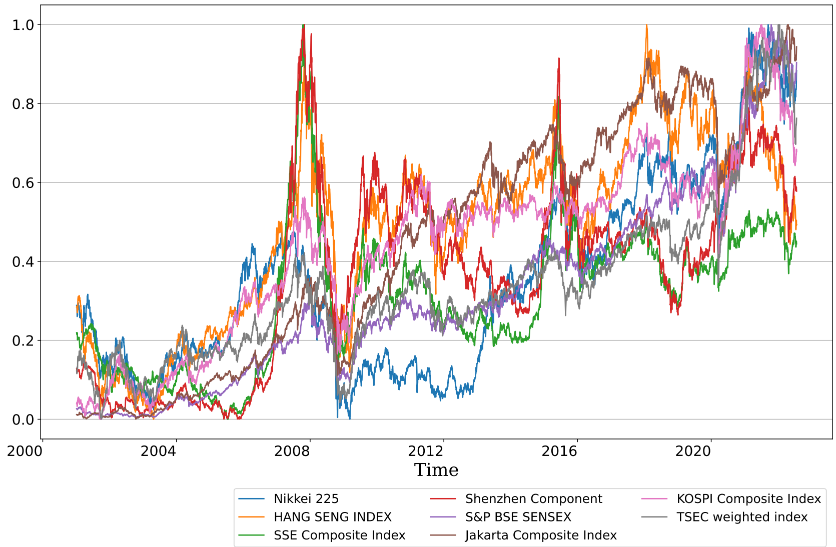

Figure A4. Evolutionary dynamics of Asian financial market indices.

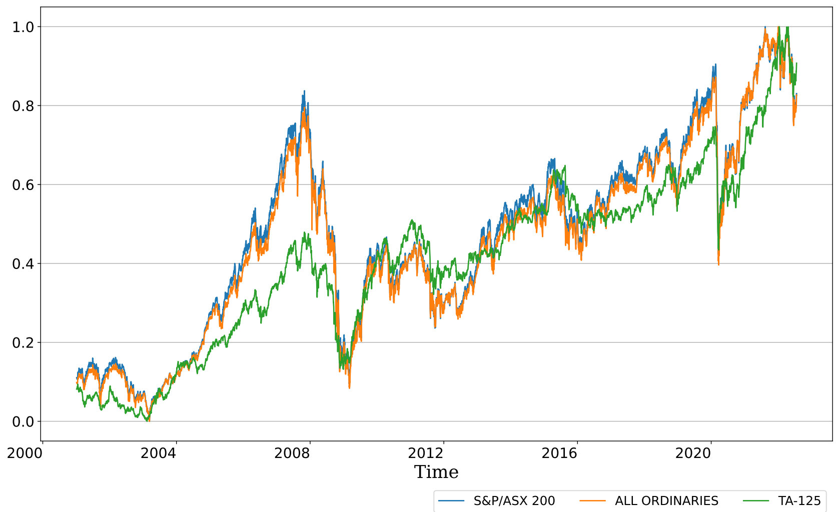

Figure A5. Evolutionary dynamics of other financial market indices.

References

1. Cootner, P.H. *The Random Character of Stock Market Prices*; MIT Press: Cambridge, MA, USA, 1964.
2. Fama, E.F. Efficient capital markets: A review of theory and empirical work. *J. Financ.* **1970**, *25*, 383–417. [CrossRef]
3. Fama, E.F. The behavior of stock-market prices. *J. Bus.* **1965**, *38*, 34–105. [CrossRef]
4. Malkiel, B.G. The efficient market hypothesis and its critics. *J. Econ. Perspect.* **2003**, *17*, 59–82. [CrossRef]
5. Sornette, D. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*; Princeton University Press: Princeton, NJ, USA, 2003.
6. Stanley, H.E.; Amaral, L.A.; Gabaix, X.; Gopikrishnan, P.; Plerou, V. Similarities and differences between physics and economics. *Phys. Stat. Mech. Appl.* **2001**, *299*, 1–15. [CrossRef]
7. Gopikrishnan, P.; Plerou, V.; Amaral, L.A.N.; Meyer, M.; Stanley, H.E. Scaling of the distribution of fluctuations of financial market indices. *Phys. Rev. E* **1999**, *60*, 5305. [CrossRef]
8. Mantegna, R.N.; Stanley, H.E. Scaling behaviour in the dynamics of an economic index. *Nature* **1995**, *376*, 46–49. [CrossRef]
9. Johansen, A.; Ledoit, O.; Sornette, D. Crashes as critical points. *Int. J. Theor. Appl. Financ.* **2000**, *3*, 219–255. [CrossRef]
10. Preis, T.; Moat, H.S.; Stanley, H.E. Quantifying trading behavior in financial markets using Google Trends. *Sci. Rep.* **2013**, *3*, 1–6. [CrossRef]
11. Preis, T.; Stanley, H.E. Bubble trouble. *Phys. World* **2011**, *24*, 29–32. [CrossRef]
12. Sornette, D. Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2522–2529. [CrossRef]
13. Lim, K.P.; Brooks, R. The evolution of stock market efficiency over time: A survey of the empirical literature. *J. Econ. Surv.* **2011**, *25*, 69–108. [CrossRef]

14. Lo, A.W.; MacKinlay, A.C. The size and power of the variance ratio test in finite samples: A Monte Carlo investigation. *J. Econom.* **1989**, *40*, 203–238. [CrossRef]
15. Lo, A.W. The adaptive markets hypothesis. *J. Portf. Manag.* **2004**, *30*, 15–29. [CrossRef]
16. Urquhart, A.; McGroarty, F. Are stock markets really efficient? Evidence of the adaptive market hypothesis. *Int. Rev. Financ. Anal.* **2016**, *47*, 39–49. [CrossRef]
17. Ito, M.; Noda, A.; Wada, T. The evolution of stock market efficiency in the US: A non-Bayesian time-varying model approach. *Appl. Econ.* **2016**, *48*, 621–635. [CrossRef]
18. Noda, A. A test of the adaptive market hypothesis using a time-varying AR model in Japan. *Financ. Res. Lett.* **2016**, *17*, 66–71. [CrossRef]
19. Zunino, L.; Tabak, B.M.; Pérez, D.G.; Garavaglia, M.; Rosso, O.A. Inefficiency in Latin-American market indices. *Eur. Phys. J. B* **2007**, *60*, 111–121. [CrossRef]
20. Zunino, L.; Zanin, M.; Tabak, B.M.; Pérez, D.G.; Rosso, O.A. Forbidden patterns, permutation entropy and stock market inefficiency. *Phys. Stat. Mech. Appl.* **2009**, *388*, 2854–2864. [CrossRef]
21. Zunino, L.; Bariviera, A.F.; Guercio, M.B.; Martinez, L.B.; Rosso, O.A. On the efficiency of sovereign bond markets. *Phys. Stat. Mech. Appl.* **2012**, *391*, 4342–4349. [CrossRef]
22. Rocha Filho, T.M.; Rocha, P.M. Evidence of inefficiency of the Brazilian stock market: The IBOVESPA future contracts. *Phys. Stat. Mech. Appl.* **2020**, *543*, 123200. [CrossRef]
23. Sánchez-Granero, M.; Balladares, K.; Ramos-Requena, J.; Trinidad-Segovia, J. Testing the efficient market hypothesis in Latin American stock markets. *Phys. Stat. Mech. Appl.* **2020**, *540*, 123082. [CrossRef]
24. Szarek, D.; Bielak, Ł.; Wyłomańska, A. Long-term prediction of the metals' prices using non-Gaussian time-inhomogeneous stochastic process. *Phys. Stat. Mech. Appl.* **2020**, *555*, 124659. [CrossRef]
25. Yang, J.; Choudhary, G.I.; Rahardja, S.; Franti, P. Classification of interbeat interval time-series using attention entropy. *IEEE Trans. Affect. Comput.* **2020**. [CrossRef]
26. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
27. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]
28. Gaio, L.E.; Stefanelli, N.O.; Pimenta, T.; Bonacim, C.A.G.; Gatsios, R.C. The impact of the Russia-Ukraine conflict on market efficiency: Evidence for the developed stock market. *Financ. Res. Lett.* **2022**, *50*, 103302. [CrossRef]
29. Hkiri, B.; Béjaoui, A.; Gharib, C.; AlNemer, H.A. Revisiting efficiency in MENA stock markets during political shocks: Evidence from a multi-step approach. *Heliyon* **2021**, *7*, e08028. [CrossRef]
30. Espinosa-Paredes, G.; Rodriguez, E.; Alvarez-Ramirez, J. A singular value decomposition entropy approach to assess the impact of COVID-19 on the informational efficiency of the WTI crude oil market. *Chaos Solitons Fractals* **2022**, *160*, 112238. [CrossRef]
31. Wang, X. Efficient markets are more connected: An entropy-based analysis of the energy, industrial metal and financial markets. *Energy Econ.* **2022**, *111*, 106067. [CrossRef]
32. Shternshis, A.; Mazzarisi, P.; Marmi, S. Measuring market efficiency: The Shannon entropy of high-frequency financial time series. *Chaos Solitons Fractals* **2022**, *162*, 112403. [CrossRef]
33. Dinga, E.; Oprean-Stan, C.; Tănăsescu, C.R.; Brătian, V.; Ionescu, G.M. Entropy-Based Behavioural Efficiency of the Financial Market. *Entropy* **2021**, *23*, 1396. [CrossRef] [PubMed]
34. Choi, S.Y. Analysis of stock market efficiency during crisis periods in the US stock market: Differences between the global financial crisis and COVID-19 pandemic. *Phys. Stat. Mech. Appl.* **2021**, *574*, 125988. [CrossRef]
35. Mensi, W.; Hamdi, A.; Yoon, S.M. Modelling multifractality and efficiency of GCC stock markets using the MF-DFA approach: A comparative analysis of global, regional and Islamic markets. *Phys. A Stat. Mech. Appl.* **2018**, *503*, 1107–1116. [CrossRef]
36. Han, C.; Wang, Y.; Ning, Y. Analysis and comparison of the multifractality and efficiency of Chinese stock market: Evidence from dynamics of major indexes in different boards. *Phys. A Stat. Mech. Appl.* **2019**, *528*, 121305. [CrossRef]
37. Mensi, W.; Tiwari, A.K.; Al-Yahyaee, K.H. An analysis of the weak form efficiency, multifractality and long memory of global, regional and European stock markets. *Q. Rev. Econ. Financ.* **2019**, *72*, 168–177. [CrossRef]
38. Cho, P.; Lee, M. Forecasting the Volatility of the Stock Index with Deep Learning Using Asymmetric Hurst Exponents. *Fractal Fract.* **2022**, *6*, 394. [CrossRef]
39. Mensi, W.; Vo, X.V.; Kang, S.H. Upward/downward multifractality and efficiency in metals futures markets: The impacts of financial and oil crises. *Resour. Policy* **2022**, *76*, 102645. [CrossRef]
40. Zhuang, X.; Wei, D. Asymmetric multifractality, comparative efficiency analysis of green finance markets: A dynamic study by index-based model. *Phys. A Stat. Mech. Appl.* **2022**, *604*, 127949. [CrossRef]
41. Bonanno, G.; Caldarelli, G.; Lillo, F.; Micciche, S.; Vandewalle, N.; Mantegna, R.N. Networks of equities in financial markets. *Eur. Phys. J. B* **2004**, *38*, 363–371. [CrossRef]
42. Kuang, P.C. Measuring information flow among international stock markets: An approach of entropy-based networks on multi time-scales. *Phys. A Stat. Mech. Appl.* **2021**, *577*, 126068. [CrossRef]
43. Shin, K.H.; Lim, G.; Min, S. Dynamics of the Global Stock Market Networks Generated by DCCA Methodology. *Appl. Sci.* **2020**, *10*, 2171. [CrossRef]

44. Balcı, M.A.; Batrancea, L.M.; Akgüller, O.; Nichita, A. Coarse Graining on Financial Correlation Networks. *Mathematics* **2022**, *10*, 2118. [CrossRef]
45. Alves, L.G.; Sigaki, H.Y.; Perc, M.; Ribeiro, H.V. Collective dynamics of stock market efficiency. *Sci. Rep.* **2020**, *10*, 1–10. [CrossRef]
46. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
47. Aranganayagi, S.; Thangavel, K. Clustering Categorical Data Using Silhouette Coefficient as a Relocating Measure. In Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), Sivakasi, India, 13–15 December 2007; Volume 2; pp. 13–17.
48. Page, L.; Brin, S.; Motwani, R.; Winograd, T. *The PageRank Citation Ranking: Bringing Order to the Web*; Technical Report 1999-66; Stanford InfoLab: Stanford, CA, USA, 1999.
49. Kendall, M.G. A new measure of rank correlation. *Biometrika* **1938**, *30*, 81–93. [CrossRef]
50. Rand, W.M. Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* **1971**, *66*, 846–850. [CrossRef]