Article

The stock market model with decayed information impact from a socioeconomic view

Zhiting Wang [1], Guiyuan Shi [2], Mingsheng Shang [3] and Yuxia Zhang [1,*]

[1] Physics and Photoelectricity School, South China University of Technology, Guangzhou, 510640, China; 201820127594@mail.scut.edu.cn(Z.W.); zhangyux@scut.edu.cn(Y.Z.)
[2] International Academic Center of Complex Systems, Beijing Normal University at Zhuhai, 519087, China; sgy@bnu.edu.cn
[3] Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China; msshang@cigit.ac.cn
* Correspondence: zhangyux@scut.edu.cn

Abstract: Finding the key factor and possible "Newton's laws" in financial markets has remained the central issue in this area. However, with the development of information and communication technologies, financial models are becoming more and more realistic but complex which is contradictory to the objective law "Greatest truths are the simplest". Therefore, this paper attempts to discover the most critical parameter and establishes an evolutionary model which is independent of micro features. In the model, information is the only key factor and stock price is the emergence of the collective behavior. The statistical properties of the model are significantly similar to the real market. It also explains the correlations of stocks within an industry, which provide a new idea for the study of key factors and core structures in the financial market.

Keywords: econophysics; financial complexity; collective intelligence; emergent property; stock correlation; detrended cross-correlation analysis

1. Introduction

With the massive use of information and communication technologies, we can collect traceable data from almost anyone, and the rise of network science [1] and computational social science [2] has provided opportunities for innovative research in econophysics and sociophysics. In particular, econophysics regards the financial market as a complex system and attempts to depict it more realistically, such as the interactions between traders by network dynamic evolution. Econophysics describes the economic system with many interacting heterogeneous entities (people, firms, institutions, etc.) in the similar way as a physical system, and expect to find similar laws. However, humans are not ideal gas molecules, it is unclear how many and which quantities would be needed for determining and anticipating a given macroscopic, in the sense of collective, observable [3]. Moreover, human beings are adaptable, the study about economic system is bound to be a difficult problem.

Now, researchers have proposed numerous different mechanisms models to depict the microstructure of financial markets. They pursued the most detailed descriptions, such as creating diverse agents and setting rules for interactions between agents and trading rules. To deal with the variables of different individuals, researchers collected the data about traders' behavior through information technology. But individuals rely on different risk preferences and reference points, even if we can reasonably describe the behavior of a single individual, we cannot directly generalize to a group. Since the behavior of individuals is complex and mutable, not all people are the same. For example, in the case of controlling the spread of COVID-19 disease, the spread of the epidemic can be controlled to a certain extent when a strict movement control order restricted their right to independent activities [4]. However, individuals are highly variable in a financial system, microstructure models are not enough to consider the variable adaptability of individuals.

Although traders are unique and unpredictable, the research has exhibited that statistical evidences remain stable relatively accordant to the stability of the statistical properties of particle motion in physics models [5,6]. So in the studies of financial markets, statistical results of different micro models should conform to the general rule. The classical percolation model [7–9] simulates herd behavior. For any pair of agents *i* and *j*, they are linked together with a probability, then agent *i* make buying or selling decision with another probability. The model explains the power-law distribution of stock price returns properly. The two-dimensional Ising model [10] takes into account trader imitation of neighbors, the influence of public information and personal traits, here the influence of public information is a Gaussian distribution. The trader's decision function also has a probability component, and the returns of the model are "fat-tailed" [11,12]. The financial models with network topology [13] also produce the stylized fact of real stock markets by setting the link probability of nodes and performing decision functions. These models share common features. First, they generate a stock trading environment in the form of probability. Second, traders make buy-sell decisions with probability or decision functions. Based on these basic models and their common features, more and more details are introduced to depict a more realistic financial market. Over the past century or so, the stock trading information flow has changed from slow to intensive, traders' literacy from low to high, relationship from simple social relationships to complex social networks. Individual characteristics of traders and the market environment have dramatically changed. Stock trading rules also varied in different countries, for example, there is 10% price limit in China [14]. Nevertheless, no matter what changed the environment or rules, it is observed that stylized facts are robust on different timescales and in different stock markets. Therefore, in the study of the macro laws, statistical properties of the stock market, the key factor should not be the relationship network of traders, the speed of information flow, or the level of literacy of traders which researchers want to introduce. On the other hand, Woolley et al. [15] studied "collective intelligence" and demonstrated that the key factor characterizing "collective intelligence" is not the average or maximum individual intelligence of group members. Collective intelligence appears to be the emergence of collective behavior. In this paper, we consider political, economy and climate as information which is a factor affecting investment decisions. In given information, the behavior of traders emerging with probabilities results in the evolution of stock markets. Here, contrary to the agents model which pursues a realistic and detailed structure, we do not focus on micro features including individual intelligence, interactions between individuals. We established a stock price evolution model with emergence properties in given information and verified its rationality by real market data. In a word, we aim to find the key factor and capture stable macroscopic law in the ever-changing stock market.

The paper is organized as follows: Section 2: A detailed description of the stock market model with decayed information impact. Section 3: Statistical analysis and nonlinear behavior of the proposed model. Section IV: Correlation analysis between stocks in the industry.

2. Stock price model with decayed information impact

Now, the percolation model, the Ising model, and the network topology model have been widely studied. Most of them share a common feature double probability forms. Based on the common feature, we proposed the stock price model of decayed information impact. It includes two components – the generation and decay of market information and the emergence properties of the collective decision-making in given information.

2.1. Information generation and decay:

- Suppose the initial stock price is P_0. The stock market environment is variable daily and is influenced by a series of stochastic events including supply and demand, macroeconomic, political factors, corporate finances and performance, market sentiment, etc. We will coarse-grain all the stochastic events by information just a single influence value. The impact of new information is random variables with a truncated Gaussian distribution $I \sim N(0, \sigma_1^2)$, here $\sigma_1 = \lambda P_0$. Considering the extreme cases (very bad information, very good information), the truncation interval is chosen to be$[-4\sigma_1, +4\sigma_1]$. New information sequence I_t can be obtained by random sampling from the Gaussian distribution.
- Considering that significant events have a sustained impact on the investors, and the impact strength of the information will decay over time, we assume that the information influence I_t decays linearly with time, and the information influence after the i-th day I_i' is expressed as

$$I_i' = \begin{cases} I_t - ai & , I_t > 0 \\ I_t + ai & , I_t < 0 \end{cases} \tag{1}$$

here a is the decay coefficient.

2.2. Stock price evolution process:

The given information determines the theoretical stock price P_t'.

$$P_t' = P_{t-1} + I_t + \sum_{i=0}^{i=t-1} I_i' \tag{2}$$

Traders participate in the game and make decisions based on the given information. Their collective behaviors result in actual stock price. As the investors vary from radicals or conservatives, daredevils or followers, etc, statistical properties of the final actual stock price is stable in the ever-changing stock market. The actual stock price P_t in day t has emergence properties of collective intelligence, which is a random sampling from a truncated Gaussian distribution$P_t \sim N(P_t', \sigma_2^2)$. As the price fluctuation is related to the information, here $\sigma_2 = \frac{1}{3} \times |P_t' - P_{t-1}|$. Considering the extremes, the truncation interval is $[-4\sigma_2, +4\sigma_2]$.

Figure 1 shows the simulated stock price series P_t and the corresponding return series r_t, $P_0 = 3000, \sigma_1 = 20, a = 5$. In Fig. 1, volatility clustering is be observed, large volatility tends to follow large volatility and small volatility tends to follow small volatility.

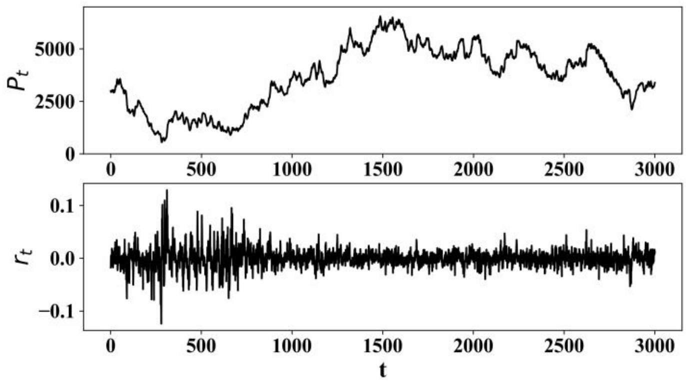

Figure 1. Stock price series of the proposed model and its corresponding return.

3. Descriptive statistics and nonlinear behavior analysis

In this section, we discussed the descriptive statistics and nonlinear behavior of the stock price model with decayed information impact and verified the simulation results with the real stock market. We used real daily closing price data from 2010-01-01 to 2020-

12-03 (T≈2700) including the SSE (Shanghai Composite Index), SZES (Shenzhen Stock Exchange Index), and S&P500 (S&P 500 Index) (https://finance.yahoo.com). Besides, the simulated data length T=3000 matches with the real data (T≈2700).

3.1. Descriptive statistics of returns

"Fat-tailed" stylized fact of returns has been verified in extensive empirical studies [16–18]. It is an important criterion for the reasonableness of price dynamics in the stock model research. Here, the definition of price return is $r_t = lnP_t - lnP_{t-1}$ [19]. The probability density distributions of three simulated and real market returns is shown in Figure 2(a). Simulated and real return distributions are almost identical, they both exhibit distinct "fat-tailed" compared to the Gaussian distribution. Table 1 shows the statistics: mean, standard deviation, maximum, minimum, skew, kurtosis, the results of Kolmogorov-Smirnov test (K-S test) and power-law fit, where the kurtosis of all returns is larger than 3 that is the kurtosis of the Gaussian distribution [20]. In the K-S test, All p-values are very small and all the H-values are 1, so we reject that the distribution of the simulated data and the empirical ones follow the Gaussian distribution at a significance level 5%. Figure 2(b) shows that the cumulative probability distributions of simulated and real market returns follow power-law distribution $P(|r_t| > x) \sim x^{-\alpha}$, α is the power-law exponent. The corresponding power-law exponent values in Table 1 approximately equal to 3, it obeys the "Inverse cubic law" [21].

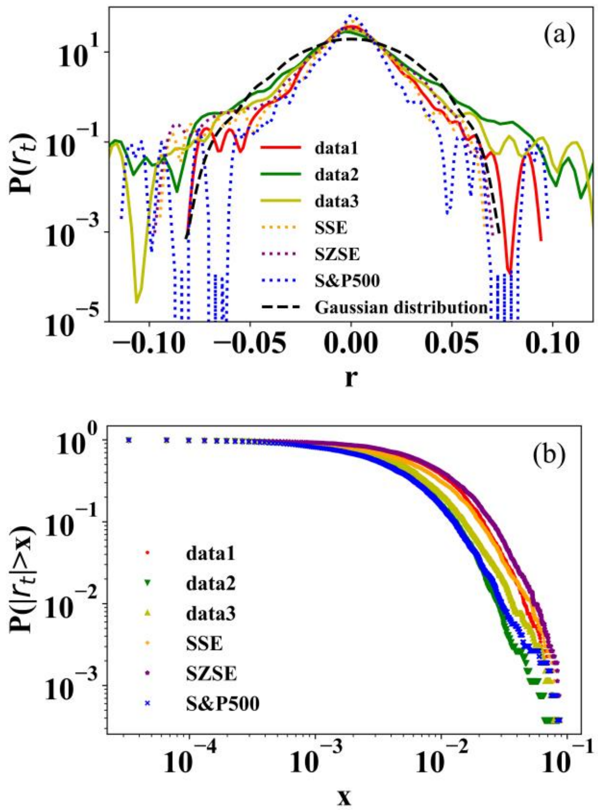

Figure 2. (a) The probability density distributions of simulated and empirical returns (semi-log); (b) The cumulative distributions of simulated and empirical returns (log-log).

Table 1. Descriptive statistics, power-law fit and K-S test of returns

Data	Mean	Std	Max	Min	Skew	Kurtosis	K-S test		α
							P-value	H	
x_1	0.00004	0.0172	0.1294	-0.1240	0.2987	6.6543	8.1208*10^ (-9)	1	3.5784
x_2	-0.00002	0.0217	0.1601	-0.2125	-0.3848	9.3611	1.8554*10^ (-10)	1	4.0968
x_3	0.00005	0.0182	0.1520	-0.1367	-0.0234	8.3799	4.2418*10^ (-10)	1	3.8109
S&P500	0.00004	0.0111	0.0934	-0.1066	-0.9710	15.2922	4.0739*10^ (-18)	1	3.4624
SSE	0.00002	0.0136	0.0060	-0.0887	-0.8969	6.1958	1.6704*10^ (-10)	1	3.5277
SZSE	0.00001	0.0164	0.0625	-0.0895	-0.7368	3.7987	5.8053*10^ (-7)	1	3.4777

3.2. Nonlinear statistical analysis of returns

The analysis of nonlinear statistical behavior can characterize the chaotic behavior of a dynamic system. It is usually applied to financial time series analysis. In empirical economics and economic physics, a number of studies have investigated the nonlinear properties of financial markets [22–24]. Here we also apply the chaotic approach to analyze the nonlinear behavior of simulated return series and compare them with real market return series.

3.2.1. Correlation dimension analysis

The correlation dimension method is a measure of the complexity of dynamical systems that distinguishes deterministic systems (including low-dimensional chaos) and stochastic systems [25]. According to the method of Grassberger et al. [26], the correlation dimension can be calculated when the appropriate embedding dimension m and time lag τ are selected for the phase space reconstruction. For an m-dimensional phase space, the correlation integral *C(r)* is calculated by

$$C(r) = \lim_{N\to\infty} \frac{2}{N(N-1)} \sum_{i,j=1,i\neq j}^{N} \Theta\left(r - \left|X_i - X_j\right|\right) \tag{3}$$

where Θ is the step function. The appropriate choice of r enables the correlation dimension of the system D to describe as

$$D = \lim_{r\to 0} \frac{log_2 C(r)}{log_2 r} \tag{4}$$

A common method is to fit the $\log_2 C(r)$ and $\log_2 r$ by least squares, and the slope is the correlation dimension D. For random sequences, D increases linearly with the embedding dimension m with no saturation, while for deterministic chaotic sequences, D increases with m to a certain position to reach saturation, and the saturation m is the correlation dimension D of the time series attractor. Figure 3 shows the correlation integral $\log_2 C(r)$ and $\log_2 r$ in different embedding dimensions m. Figure 4 shows the correlation dimension. It is observed that all correlation dimensions increase with m, and reach saturation at a certain position. It can be seen that all the returns have deterministic noise which means the systems are chaotic. The simulated data from the proposed model well match with the real market data.

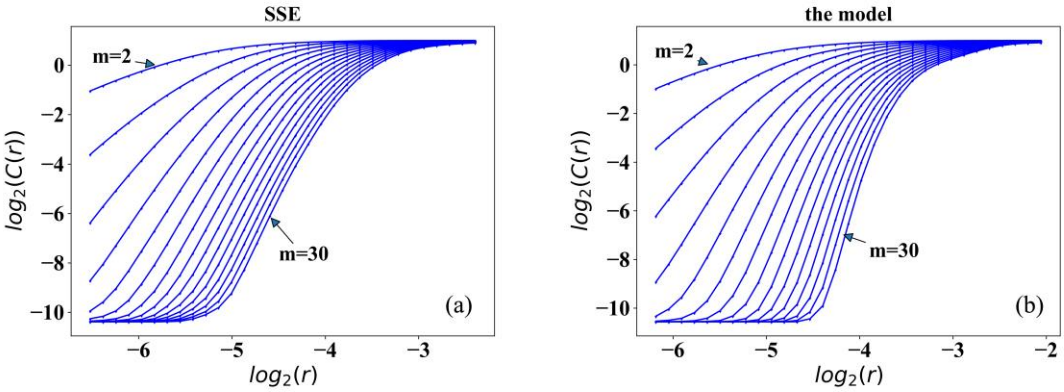

Figure 3. Correlation integral results of return series from SSE (a), the model (b).

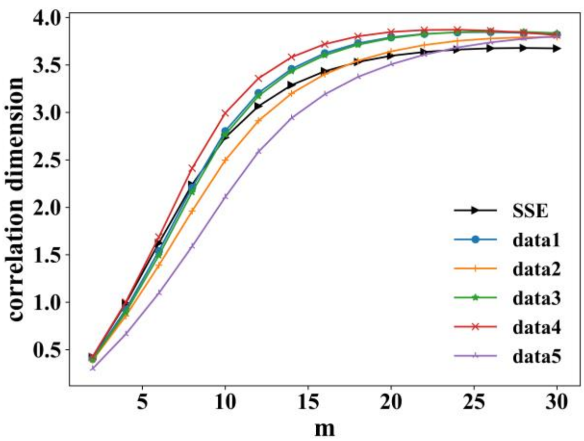

Figure 4. Correlation dimension of returns from SSE and 5 simulated data.

3.2.2. Lyapunov exponent analysis, sample entropy analysis and Hurst exponent

To further compare the chaotic behavior of simulated and empirical returns, we calculated the maximum Lyapunov exponent for each stock price series by the algorithm of Rosenstein et al. [27], the Hurst exponent by the rescaled range analysis[28], and the sample entropy method by Richman et al. [29]. In Table 2, the maximum Lyapunov exponents of both simulated and empirical returns are larger than 0 which indicates that they have similar chaotic properties. Second, similar sample entropy values indicate similar complexity. Hurst exponent is slightly larger than 0.5, which means that the simulated and real returns have similar long memory.

Table 2. The maximum Lyapunov exponent (m=10), Sample Entropy (m=2) and Hurst exponent of returns from the model and empirical market.

Data	MLE	Sample Entropy	Hurst Exponent
Data1	0.0778	1.7497	0.6281
Data2	0.0762	1.6832	0.6364
Data3	0.0773	1.7033	0.6478
Data4	0.0757	1.7401	0.6152
Data5	0.0575	1.4901	0.5840
SSE	0.0628	1.7889	0.5238
SZSE	0.0842	1.8750	0.5176
S&P500	0.0639	1.4902	0.5022

4. Correlation analysis of stocks

The correlation between stocks is an important criterion to weigh the correlation of stock market risk level and portfolio rationality. Studies on the properties of stock correlation show that the stronger correlations between stocks are, the higher risk in the corresponding asset portfolio is [30]. Usually, stocks belonging to the same industry are more correlated because they are influenced by the same external information, including natural climate, macro policies, raw materials, and other factors [31]. The stocks in an industry have strong correlations and risky portfolios, so the rational investments usually cover different industries. In our model, stocks rise or fall are affected by external information, so the model can be considered to study the correlation between stocks.

This section investigated the correlation of stock returns within per industry in China by the detrended cross-correlation analysis (DCCA) [32,33] and calculated their distributions. ρ is the detrended cross-correlation coefficient, $-1 \leq \rho \leq 1$. There are 28 industries in Shenwan Industry Classification Standard, we selected 16 industries from 2016-01-01 to 2020-12-10 ($T \approx 1200$), which contain a sufficient number of stocks (the number of stocks $N > 30$). Simulated stock data in an industry: the initial stock price is same, to avoid the sensitivity to initial conditions, we selected the data from 6000 to 7500 steps in the simulation (T=1500), and got 100 stocks under the same historical information series.

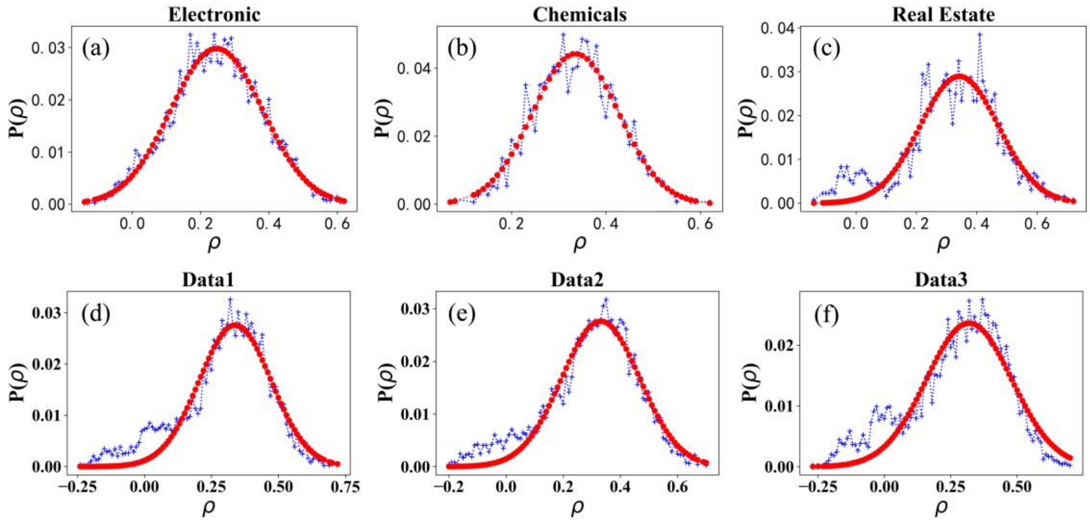

Figure 5. The distribution of ρ from Chemicals(a), Real Estate(b), Electronic(c),and 3 simulated data(d)(e)(f).

Figure 5 shows the distribution of the correlation of stock returns within an industry. Figure 5 (a)(b)(c) are three empirical data examples and (d)(e)(f) are three simulated ones that are generated in different historical information series. It can be seen in Figure 5 that ρ distributions within each of the 16 industries show a regular single-peaked distribution. The most probable correlation coefficients ρ_m are around 0.3, which indicates that the model is consistent with the real market, and most stocks have weak positive correlations within an industry. Figure 6 shows the most probable correlation coefficients ρ_m which of within the 16 industries and the three simulated data. The three simulated data peaks are 0.34, 0.33, and 0.32, which all lying within the peak range from 0.21 to 0.43 in the real market. Moreover, since each set of simulated data is generated in given the same historical information series, there is probable that the stock market evolution will recur when there is similar information series. In our proposed model, the correlation of the simulated stock with the same historical information can be analogized statistically to the correlation of the stocks within China's industry. It is a supplement method of stock correlation research which helps get IPOs' prices accurately and investors obtain a better portfolio strategy.

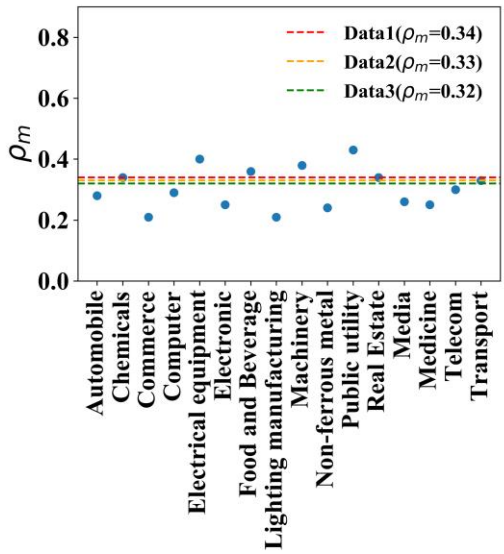

Figure 6. The ρ_m of stocks in 16 industries and 3 simulated data.

5. Conclusions

With the development of Internet and communication technologies, various financial micro models are devoted to introducing plenty of individual traits and relationships into financial market models. However, empirical studies show that no matter how the market environment changes, the stylized facts remain stable which are independent of the speed of information, the literacy of trader, the wealth of society, or the closeness of the relationship. That is to say there must be the so-called "Newtons laws" in the stock market which has no relation with the micro-characteristics and individual characteristics. So we studied the once stock market models, and try to find the nature of the systems. From the percolation model, Ising model, and network topology financial model, we got a common feature that trading environment and group decision are generated with probability. We ignore the micro items that do not affect macro statistical properties and establish the stock price model with decayed information impact macroscopically. It recaptures the stylized facts and the chaotic characteristics of the real stock market, confirming that the key factor that affects stock price fluctuations is information, not the individual characteristics of investors or the sparseness of the relationship network. Stock prices are the emergence of collective behavior in given information (current and decayed). Besides, the model generated different stock price series in the same historical information, which be analogous to the stocks in the same industry. Similar single-peaked distribution proving that the model can be effectively used in stock correlation research and history recur rules. It opens a new way to selecting rational portfolios, complements current industry correlation research methods and providing theoretical support. The paper provides a useful framework for understanding stock price evolution through the emergence of collective intelligence. And we finding the possible key factor of stock price fluctuation and the essence of the financial market at a macro level.

Author Contributions: writing—original draft preparation, Z.W.; writing—review and editing, Y.Z., G.S and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant No. 61503140)

Acknowledgments: Z.W. is obliged to Jiawei Zhang for several enlightening discussions and suggestions for writing.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Barabási, A.-L. The network takeover. Nat. Phys. 2012, 8, 14–16.
2. Lazer, D.; Pentland, A.; Adamic, L.; Aral, S.; Barabasi, A.-L.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; et al. Social science. Computational social science. Science 2009, 323, 721–723.
3. Caldarelli, G.; Wolf, S.; Moreno, Y. Physics of humans, physics for society. Nat. Phys. 2018, 14, 870.
4. Kraemer, M.U.G.; Yang, C.-H.; Gutierrez, B.; Wu, C.-H.; Klein, B.; Pigott, D.M.; du Plessis, L.; Faria, N.R.; Li, R.; Hanage, W.P.; et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science (80-.). 2020, 368, 493–497.
5. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. Rev. Mod. Phys. 2009, 81, 591.
6. Perc, M. The social physics collective. Sci. Rep. 2019, 9, 16549.
7. Cont, R.; Bouchaud, J.-P.; others Herd Behavior And Aggregate Fluctuations In Financial Markets. Macroecon. Dyn. 2000, 4, 170–196.
8. Eguiluz, V.M.; Zimmermann, M.G. Transmission of Information and Herd Behavior: An Application to Financial Markets. Phys. Rev. Lett. 2000, 85, 5659–5662.
9. Ren, F.; Zheng, B. Generalized persistence probability in a dynamic economic index. Phys. Lett. A 2003, 313, 312–315.
10. Zhou, W.X.; Sornette, D. Self-organizing Ising model of financial markets. Eur. Phys. J. B 2007, 55, 175–181.
11. Cont, R. Empirical properties of asset returns: stylized facts and statistical issues. Quant. Financ. 2001, 1, 223–236.
12. Maganini, N.D.; Filho, A.C.D.S.; Lima, F.G. Investigation of multifractality in the Brazilian stock market. Phys. A-statistical Mech. Its Appl. 2018, 497, 258–271.
13. Zhao, H.; Zhou, J.; Zhang, A.; Su, G.; Zhang, Y. Self-organizing Ising model of artificial financial markets with small-world network topology. {EPL} (Europhysics Lett. 2013, 101, 18001.
14. Wan, Y.-L.; Wang, G.-J.; Jiang, Z.-Q.; Xie, W.-J.; Zhou, W.-X. The cooling-off effect of price limits in the Chinese stock markets. Phys. A Stat. Mech. its Appl. 2018, 505, 153–163.
15. Woolley, A.W.; Chabris, C.F.; Pentland, A.; Hashmi, N.; Malone, T.W. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. Science (80-.). 2010, 330, 686 LP – 688.
16. Gopikrishnan, P.; Plerou, V.; Nunes Amaral, L.A.; Meyer, M.; Stanley, H.E. Scaling of the distribution of fluctuations of financial market indices. Phys. Rev. E 1999, 60, 5305–5316.
17. Qiu, T.; Zheng, B.; Ren, F.; Trimper, S. Return-volatility correlation in financial dynamics. Phys. Rev. E 2006, 73, 65103.
18. Zhang, J.W.; Zhang, Y.; Kleinert, H. Power tails of index distributions in chinese stock market. Phys. A Stat. Mech. its Appl. 2007, 377, 166–172.
19. Wang, Y.; Zheng, S.; Zhang, W.; Wang, J. Complex and Entropy of Fluctuations of Agent-Based Interacting Financial Dynamics with Random Jump. Entropy 2017, 19, 10.
20. Balanda, K.P.; Macgillivray, H.L. Kurtosis: A Critical Review. Am. Stat. 1988, 42, 111–119.
21. Gopikrishnan, P.; Meyer, M.; Amaral, L.A.N.; Stanley, H.E. Inverse cubic law for the distribution of stock price variations. Eur. Phys. J. B 1998, 3, 139-140.
22. HSIEH, D.A. Chaos and Nonlinear Dynamics: Application to Financial Markets. J. Finance 1991, 46, 1839–1877.
23. Alves, P.R.L.; Duarte, L.G.S.; da Mota, L.A.C.P. Detecting chaos and predicting in Dow Jones Index. Chaos, Solitons & Fractals 2018, 110, 232–238.
24. Zhu, H.; Zhang, W. Multifractal property of Chinese stock market in the CSI 800 index based on MF-DFA approach. Phys. A Stat. Mech. its Appl. 2018, 490, 497–503.
25. Grassberger, P.; Procaccia, I. Dimensions and entropies of strange attractors from a fluctuating dynamics approach. Phys. D Nonlinear Phenom. 1984, 13, 34–54.
26. Grassberger, P.; Procaccia, I. Measuring the strangeness of strange attractors. Phys. D Nonlinear Phenom. 1983, 9D, 189-208.
27. Rosenstein, M.T.; Collins, J.J.; De Luca, C.J. A practical method for calculating largest Lyapunov exponents from small data sets. Phys. D Nonlinear Phenom. 1993, 65, 117-34.
28. Couillard, M.; Davison, M. A comment on measuring the Hurst exponent of financial time series. Phys. A Stat. Mech. its Appl. 2005, 348, 404–418.
29. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Circ. Physiol. 2000, 278, H2039--H2049.
30. Eom, C.; Park, J.W. Effects of common factors on stock correlation networks and portfolio diversification. Int. Rev. Financ. Anal. 2017, 49, 1–11.
31. Guo, X.; Zhang, H.; Tian, T. Development of stock correlation networks using mutual information and financial big data. PLoS One 2018, 13.
32. Kristoufek, L. Measuring correlations between non-stationary series with DCCA coefficient. Phys. A-statistical Mech. Its Appl. 2014, 402, 291–298.
33. Ferreira, P.; Pereira, É.J. de A.L.; Silva, M.F. da; Pereira, H.B. Detrended correlation coefficients between oil and stock markets: The effect of the 2008 crisis. Phys. A Stat. Mech. its Appl. 2019, 517, 86–96.