Predicting Stock Market Instabilities

A Project Report submitted to the Jawaharlal Nehru University

School of Engineering

2021

Aditya Singla

Registration No.: 14008549

Computer Science Engineering

23-01-2021

Contents

Chapter 1

Introduction

"A mathematician is a device for turning coffee into theorems."

– Paul Erods.

THE stock market data presents minute details of the market and the associated prices of stocks. This large amount of data can be harnessed to understand what really controls the market. Contrary to other fields of science, this large amount of data presents us with the dilemma of choosing just the right set of attributes to predict the next state of market for making the optimum portfolio and maximizing profits. The other set of problems that are areas of interest are predicting an imminent market crash such as the 2008 financial crisis. These crashes are affected by both endogenous and exogenous agents, therefore, adding to the complexity of the whole process.

1.1 Background

What is Stock Market ?

In a stock market stockbrokers and traders buy and sell shares (equity stock), bonds, and other securities. A potential buyer bids a specific price for a stock, and a potential seller asks a specific price for the same stock. When the bid and ask prices match, a sale takes place, on a first-come, first-served basis if there are multiple bidders at a given price. The exchanges provide real-time trading information on the listed securities, facilitating price discovery. [1]

1.2 Aim and Nature of the project

The aim of this project is to develop an indicator that detects emergence of market instabilities by quantifying the intensity of self-organizing processes arising from stock returns' co-movements. In financial markets, phenomena like imitation, herding and positive feedbacks characterize the emergence of endogenous instabilities, which can modify the qualitative and quantitative behavior of the underlying system.

The impossibility to predict the future state of financial systems motivates the use of a synthetic indicator to detect the disruption of an existing equilibrium configuration. The stock market is visualized as an interconnected sub-graph of stock return co-movements. Therby, helping in identification and prediction of up and down market phases.[2]

1.3 Basic algorithms and numerical methods

1.3.1 Efficient Market Hypothesis

The efficient-market hypothesis (EMH) is a hypothesis in financial economics that states that asset prices reflect all available information. A direct implication is that it is impossible to "beat the market" consistently on a risk-adjusted basis since market prices should only react to new information.[3]

1.3.2 Price Time Series

In investing, a time series tracks the movement of the chosen data points, such as a security's price, over a specified period of time with data points recorded at regular intervals.

1.3.3 Log Return

In finance, return is a profit on an investment. It comprises any change in value of the investment, and/or cash flows which the investor receives from the investment, such as interest payments or dividends. It may be measured either in absolute terms (e.g., dollars) or as a percentage of the amount invested. The latter is also called the holding period return.
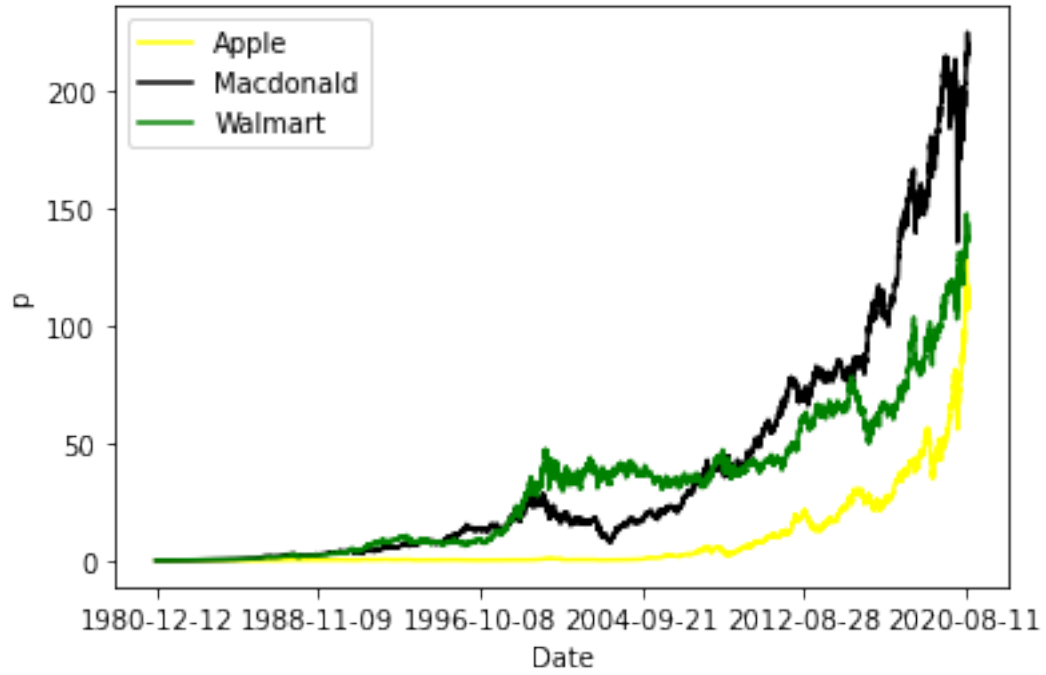
Figure 1.1: The graph shows the price time series of three stocks, that is, Apple, MacDonald and Walmart.

Algorithm 1.1 $x(t)$ represents the return time series and $P(t)$ represents the price time series

$$x(t) = ln\frac{P(t)}{P(t-1)}$$
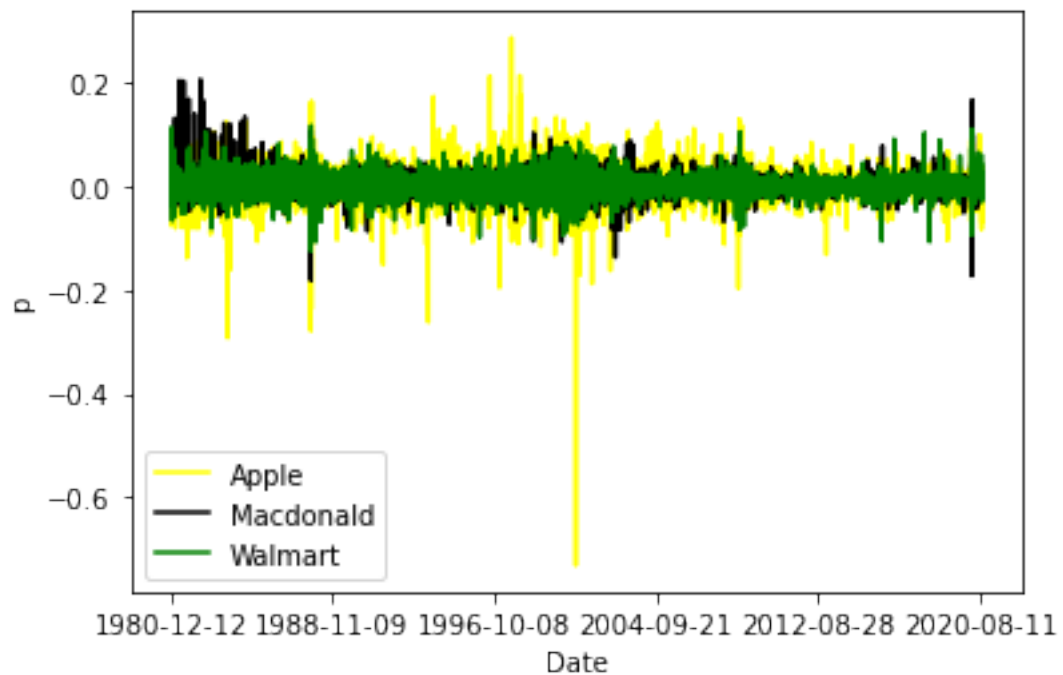
Figure 1.2: The graph shows the log return time series of three stocks, that is, Apple, MacDonald and Walmart.

1.3.4 Arbitrage

Arbitrage is the purchase and sale of an asset in order to profit from a difference in the asset's price between markets. It is a trade that profits by exploiting the price differences of identical or similar financial instruments in different markets or in different forms. Arbitrage exists as a result of market inefficiencies and would therefore not exist if all markets were perfectly efficient.[4]

Algorithm 1.2 Auto-correlation coefficient of a stochastic process

$$\rho_{XX}\left(t_1, t_2\right) = \frac{E\left[\left(X_{t_1} - \mu_{t_1}\right)\left(\overline{X_{t_2} - \mu_{t_2}}\right)\right]}{\sigma_{t_1}\sigma_{t_2}}$$

1.3.5 Autocorrelation

Autocorrelation, also known as serial correlation, is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them. The analysis of autocorrelation is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is often used in signal processing for analyzing functions or series of values, such as time domain signals.[5]

1.3.6 Auto Covariance

In probability theory and statistics, given a stochastic process, the autocovariance is a function that gives the covariance of the process with itself at pairs of time points. Autocovariance is closely related to the autocorrelation of the process in question.

1.3.7 Pearson Correlation Coefficient

In statistics, the Pearson correlation coefficient (PCC, pronounced /ˈpɪərsən/), also referred to as Pearson's r, the Pearson product-moment correlation
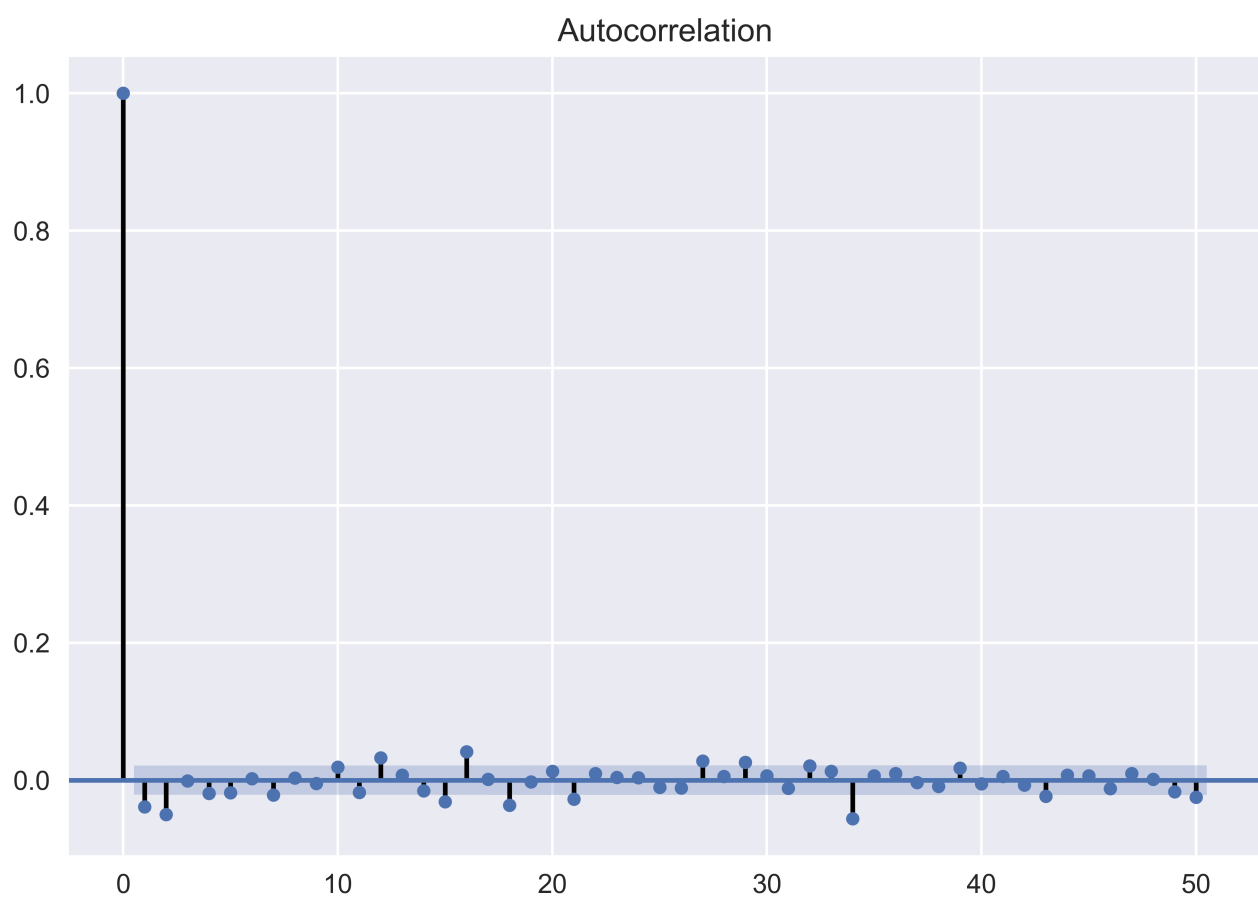
Figure 1.3: The graph shows autocorrelation of microsoft return over lags ranging from 1 to 50.

coefficient (PPMCC), or the bivariate correlation, is a statistic that measures linear correlation between two variables X and Y. It has a value between +1 and −1. A value of +1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation.[6]

1.4 Interesting and relevant questions

Imitation, herding behaviors, and positive feedbacks among market participants have been recognized as phenomena leading to endogenous instabilities. Herding behaviors spread when the knowledge about other investors' allocation decisions influences personal strategies, meaning that investors tend to use similar investment practices to those applied by other market participants even when this is not justified by their own information set, while positive feedbacks can induce the underlying system to accumulate instabilities that lead to new configurations as a self-fulfilling mechanism. Hence, a strengthening of interactions among asset prices may emerge due to market euphoria, which drives prices to a sharp rise or, by contrast, to phenomena of financial turmoil, which induce fire sales and stock market crashes.[2]

Chapter 2

METHOD

THE experience of the natural sciences in coping with complex systems suggests a parsimonious stochastic approach. Agents in large economic systems will display heterogeneity in terms of their different micro motives, degrees of deliberation and information-processing capabilities. The variability of human behaviour can be quantified in a tractable way using statistical laws. Statistical physics shows that relatively simple models with plausible behavioural rules have the potential to replicate key empirical regularities of stock markets. The complexity of globalized financial markets has brought to the fore systemic vulnerabilities.[9][10]

Algorithm 2.1 x(t) denotes the logarithmic return time series

$$y\left(k\right) = \sum_{t=1}^{k} \left[x\left(t\right) - <x>\right]$$

Algorithm 2.2 k_{nm} denotes the starting time fo m^{th} box of size n.

$$F\left(n\right) = \frac{1}{M} \sum_{m=1}^{M} \sqrt{\frac{1}{N} \sum_{k=k_{nm}}^{k_{nm}+n} \left[y\left(k\right) - y_{nm}\left(k\right)\right]^2}$$

2.1   Detrended Fluctuation Analysis

DFA of y(k) enables us to reveal long range correlations in x(t) by getting rid of trends. The integrated y(k) is divided into M=$\left\lceil \frac{N}{n} \right\rceil$ overlapping boxes (subintervals).

The slope of F(n) to ln(n) determines "scaling exponent", which can be used to characterize the fluctuations. If α = 0.5 , the time series x(t) is uncorrelated (white noise) — this condition will be considered as the null hypothesis satisfying the geometric Brownian motion model of the efficient markets hypothesis. The case α =1, x(t) corresponds to the 1/f noise or flicker noise. If α = 1.5, x(t) can be represented as a 1/f 2 noise — a random walk series (the best known non stationary series) is exactly 1/f2 noise.[7]
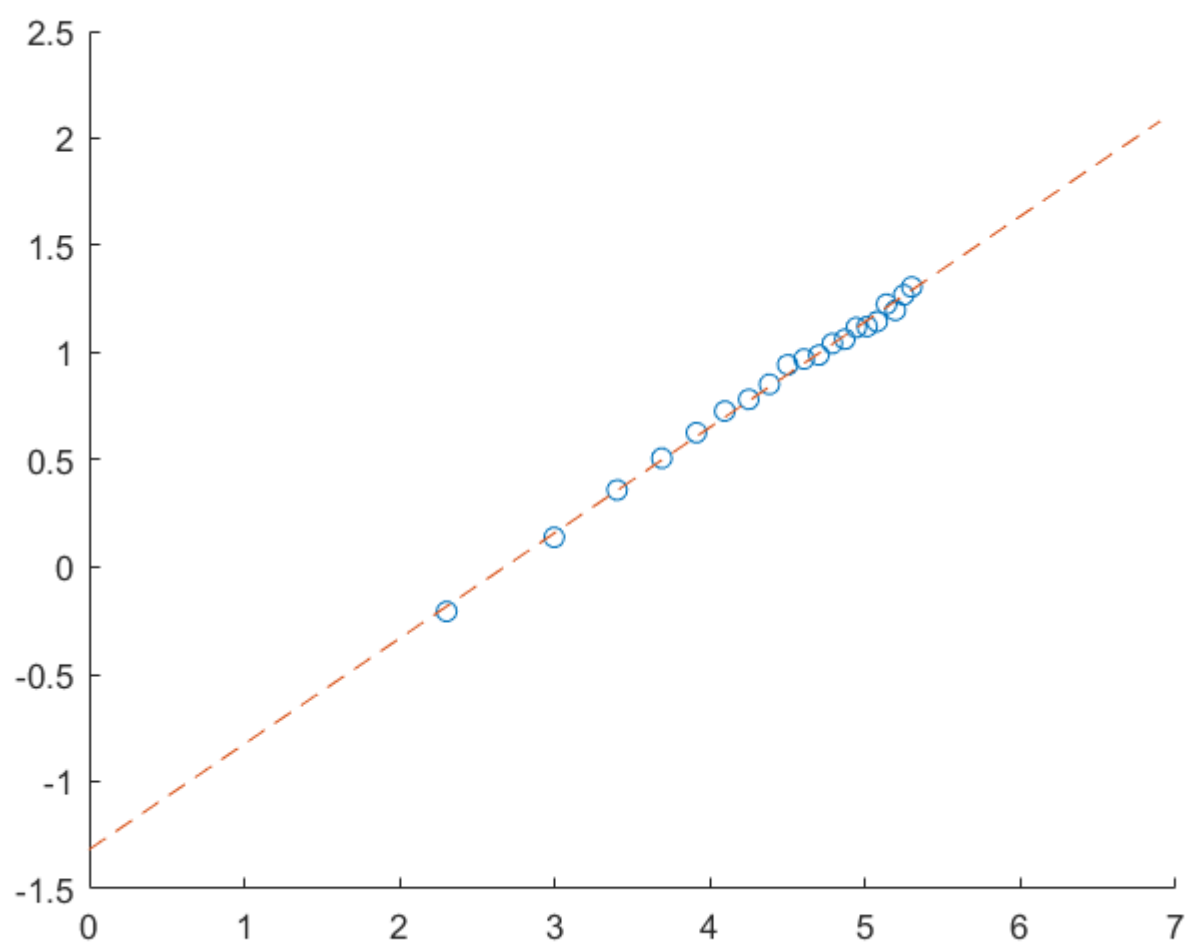
Figure 2.1: DFA of randomm number generator

2.2 Separating stocks on the basis of sector

Sector breakdowns help portfolio managers and investors determine the allocation of funds within a portfolio. If an investor wants to create a diversified portfolio, the portfolio should include stocks from a variety of sectors. For smaller investors seeking to create a diversified portfolio, they can easily do by investing in an index exchange-traded fund (ETF). However, if an investor is only interested in investing in, for example, technology or energy-based businesses, they can, of course, confine their investing to only the sectors they are interested in.

Overview of the S&P sectors:

2.2.1 Information Technology

The information technology – IT – sector consists of companies that develop or distribute technological items or services, and includes internet companies. Technology products include computers, microprocessors, and operating systems. Example of companies in this sector includes big names like Microsoft Corporation, Oracle Corp., and Mastercard Inc. This sector has seen a lot of change in recent years because of the rapid rise in technology-based companies.

2.2.2 Health Care

Health care consists of medical supply companies, pharmaceutical companies, and scientific-based operations or services that aim to improve the human body or mind. Familiar names include Johnson & Johnson, a medical device and pharmaceutical company that owns Tylenol, and Abiomed, which manufactures medical implant devices.

Cannabis companies are a new, but rapidly growing, part of the health care sector. Currently, the more well-known ones include Canopy Growth Corp. and Aurora Cannabis, with market caps of $23 billion and $12 billion, respectively.

2.2.3 Financials

The financial sector includes all companies involved in finance, investing, and the movement or storage of money. It includes banks, credit card issuers, credit unions, insurance companies, and mortgage real estate investment trusts (REITs). Companies within this sector are usually relatively stable, as many are mature, well-established firms. Banks in this sector include Bank of America Corp, JPMorgan Chase & Co., and Goldman Sachs. Other notable sector names include Berkshire Hathaway, American Express, and Aon plc.

2.2.4 Consumer Discretionary

Discretionary consumer products are luxury items or services that are not necessary for survival. The demand for these items depends on economic conditions and the wealth of individuals. Products include cars, jewelry, sporting goods, and electronic devices. Luxury experiences include trips, stays at hotels, or dining in a posh restaurant. Most companies in this sector are easily recognized. Some examples include Starbucks, Best Buy, and Amazon.

2.2.5 Communication Services

The communication services sector consists of companies that keep people connected. This includes internet providers and phone plan providers. The more exciting part of the sector includes media, entertainment, and interactive media & services companies. Netflix Inc. and Walt Disney Co. are considered part of the communication services sector. Other companies within this sector include AT&T, CBS Corp., and Facebook.

2.2.6 Industrials

Industrials include a wide range of companies, from airlines and railroad companies to military weapons manufacturers. Since the range of companies is so large, the sector has 14 different industries. Two of the largest industries are Aerospace & Defense and Construction & Engi-

neering. The best known names within this sector are Delta Air Lines and Southwest Airlines, FedEx Corporation, and Boeing Company.

2.2.7 Consumer Staples

Consumer staples companies provide all the necessities of life. This includes food and beverage companies, household product providers, and personal product providers. Consumer staple companies are well known, since people see their products in stores regularly. For example, Procter & Gamble is a famous company within this sector, which produces bleach and laundry detergent under brand names such as Dawn and Tide. Another example is Kroger, which is the largest supermarket chain in the U.S.

2.2.8 Energy

The energy sector consists of all companies that play a part in the oil, gas, and consumable fuels business. This includes companies that find, drill, and extract the commodity. It also includes the companies that refine the material and companies that provide or manufacturer the equipment used in the refinement process. Companies such as Exxon Mobil and Chevron extract and refine gas, while companies like Kinder Morgan transport fuel to gas stations.

2.2.9 Utilities

Utility companies provide or generate electricity, water, and gas to buildings and households. For example, Duke Energy generates and distributes electricity, and Southern Company provides gas and electricity. Many utility companies are developing more renewable energy sources.

2.2.10 Real Estate

As the name suggests, the newest addition to the S&P sectors includes Real Estate Investment Trusts (REITs), as well as realtors and other companies. The real estate sector makes up 2.9% of the S&P 500. Companies in the sector include American Tower Corp., Boston Properties, and Equinix.

2.2.11 Materials

Companies within the materials sector provide the raw material needed for other sectors to function. This includes the mining companies that provide gold, zinc, and copper, and forestry companies that provide wood. Companies that are not typically associated with materials but are in the sector include container and packaging companies such as the Intertape Polymer Group, a company that produces tape.[8]

2.3 Finding stocks and associated sectors which have a trend with high DFA

We will use a rolling window technique. For each window find the stocks which had highest memory. Categorise these stocks with respect to the sector they belong to. The window time is 400 days. Therefore, for such a large time interval if the stock is exhibiting memory then it means that it corresponds to a bubble in the stock market. Analysing how trend in one sector relates to the trend of another sector.

Chapter 3

Data

The data of S&P 100 stocks is obtained from yahoo finance for the time interval 31-12-2004 00:00:00 to 30-11-2012 00:00:00. The data was accessed throguh yahoo finance website. The database developed had 101 attributes containing date and associated adjusted close values for all the S&P 100 stocks. The stocks were further classified into sectors. The stock prices of all the stocks in a particular sector were then added to get sectoral data of stock market for further analysis.

Chapter 4

Analysis

4.1 Sector wise impact of 2008 stock market crash

The sector wise impact of the 2008 stock market crash can be analysed through the Figure 4.1. The correlation matrix of the different sectors can be seen in Figure 4.2.

Analysis:

1. As is evident from the plot (Fig. 4.1), all the sectors saw a drastic fall in their stock prices in 2008, a result of the 2008 financial crisis.

2. The fall in the prices for financials and healthcare was the most drastic. This significant drop in the stock prices for financials can be attributed to the fact that the 2008 financial crisis affected some of the

largest investment banks of the time, mutual funds, etc. Also, bearing a direct impact on the savings of people, it may be expected that this drastically impacted their ability to spend in healthcare services.

3. It may thus be concluded that the performance of all sectors on a stock market are impacted even if one of the sectors experiences a shock, as was experienced in 2008.

4.2 Sectors with long term memory during the period of 2006-07

According to the Efficient Market Hypotheisis (EMH), stocks always trade at their fair value on exchanges, making it impossible for investors to purchase undervalued stocks or sell stocks for inflated prices. Therefore, it should be impossible to outperform the overall market through expert stock selection or market timing, and the only way an investor can obtain higher returns is by purchasing riskier investments.

The onset of long term trends coincide in case of real estate and materials. Similarly such relations in trends can be seen for other pairs of sectors as well. The relation in trends can be found for: Real Estate and Materials , Industrials and Utilities , Consumer discretionary and consumer staples , Information technology and communication services , health care and financials. These trends can be utilised to make an investment strategy.

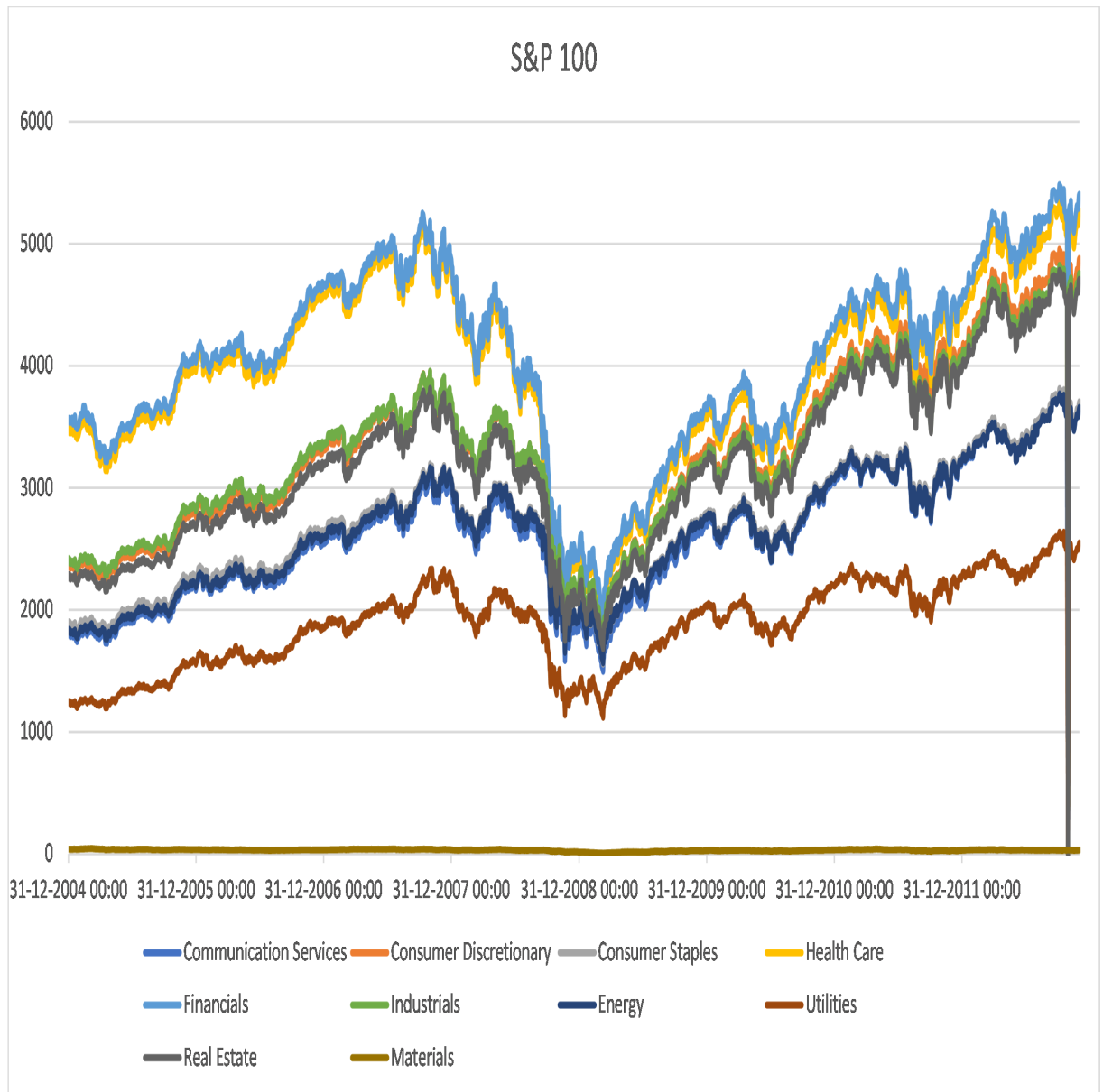Analysis on the basis of data plotted in Figures 4.3-4.7:

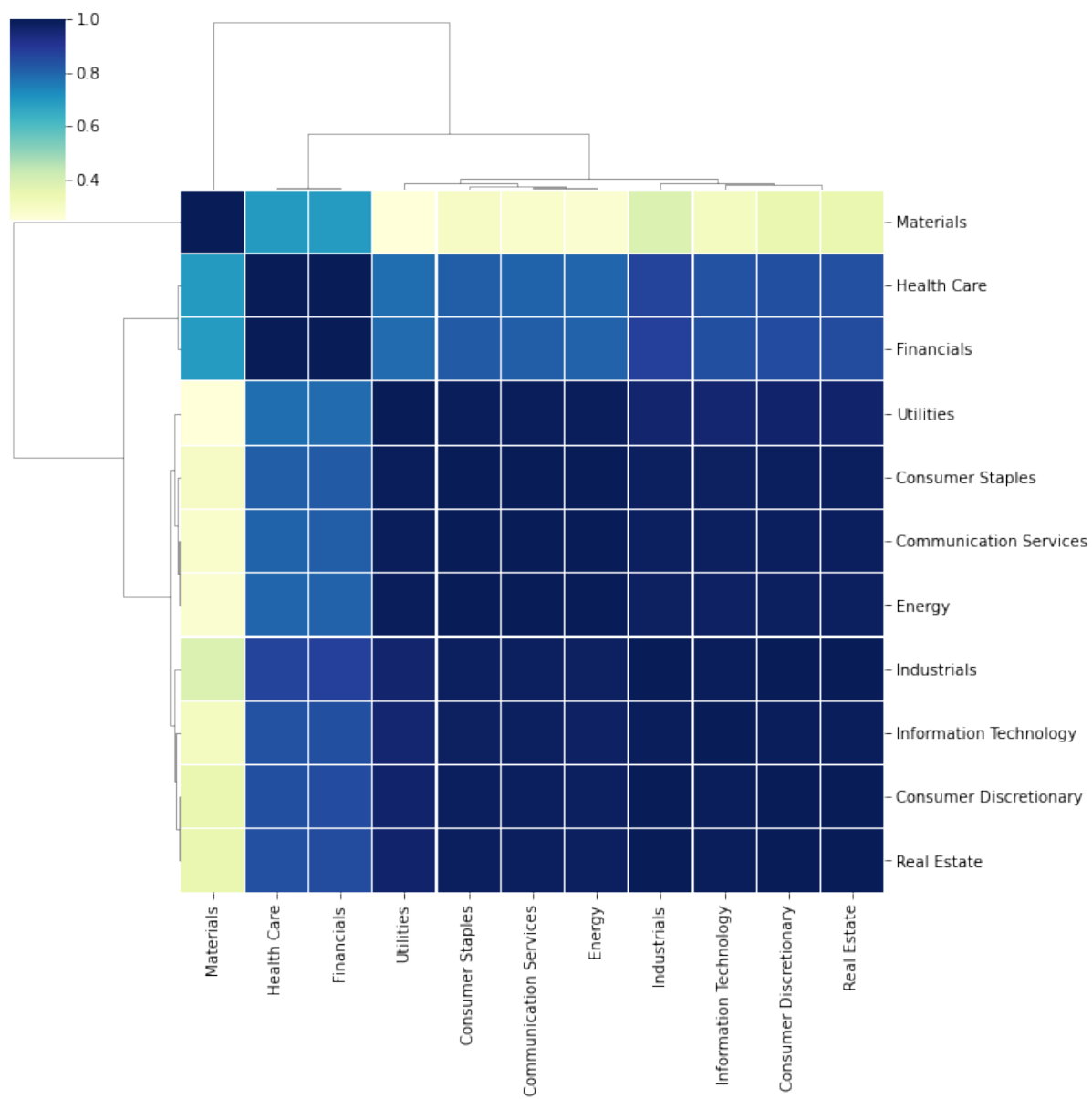Figure 4.1: Sector wise impact of 2008 stock market crash
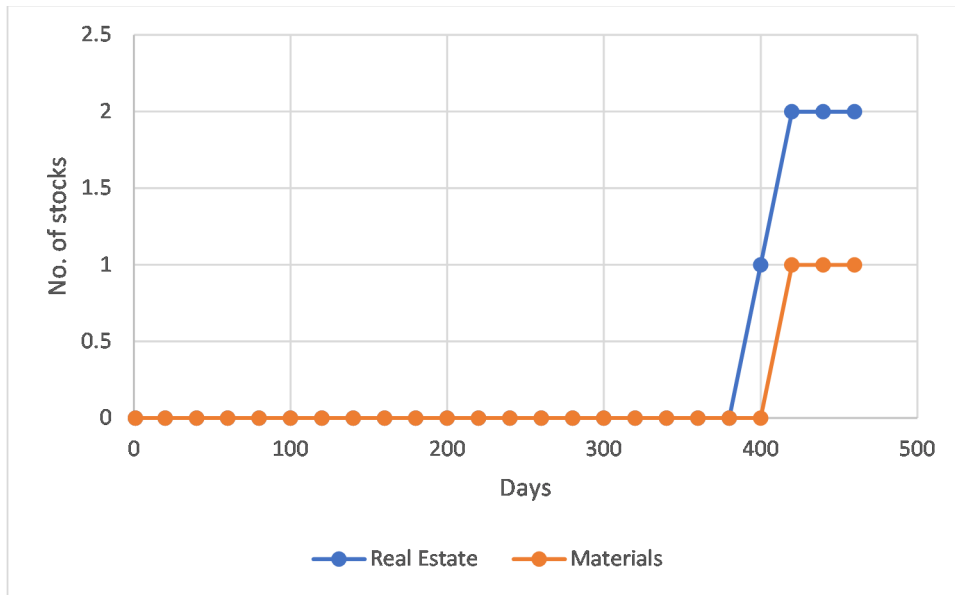
Figure 4.2: Correlation matrix sector wise

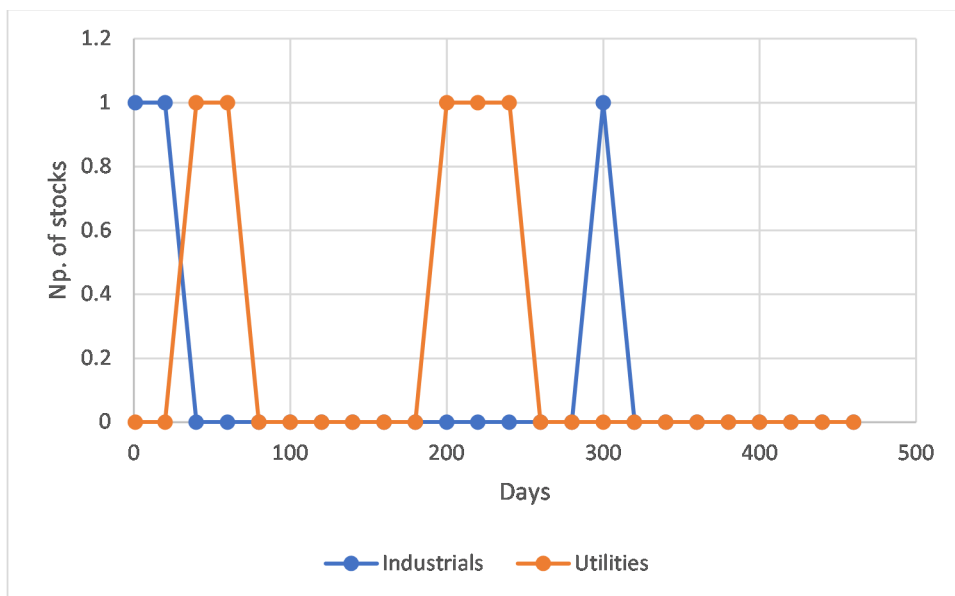Figure 4.3: No. of stocks belonging to the real estate and materials sector with long term memory

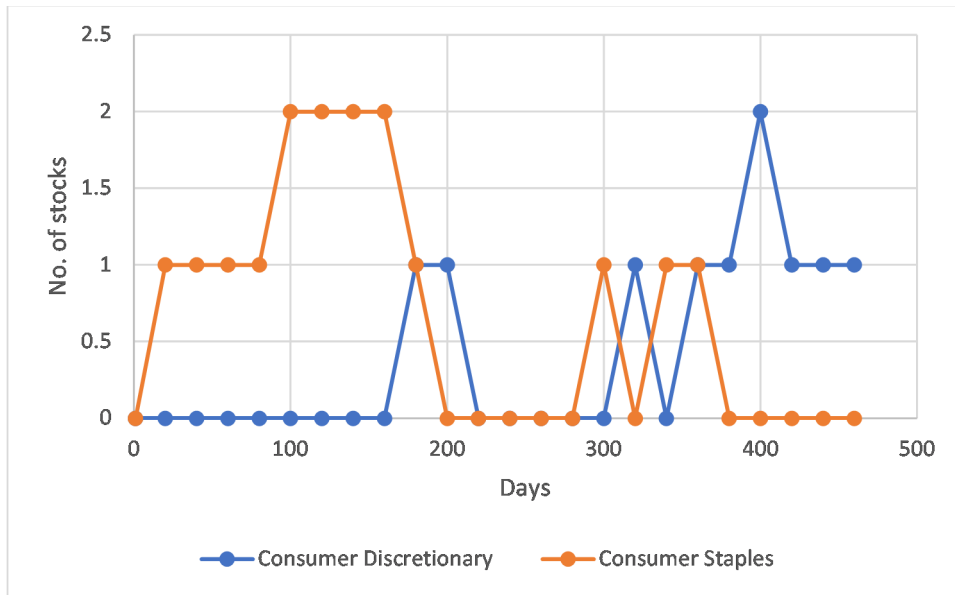Figure 4.4: No. of stocks belonging to Industrials and Utilities sector with lomg term memory

Figure 4.5: No. of stocks in Consumer Discretionary and Consumer staples sector with long term memory
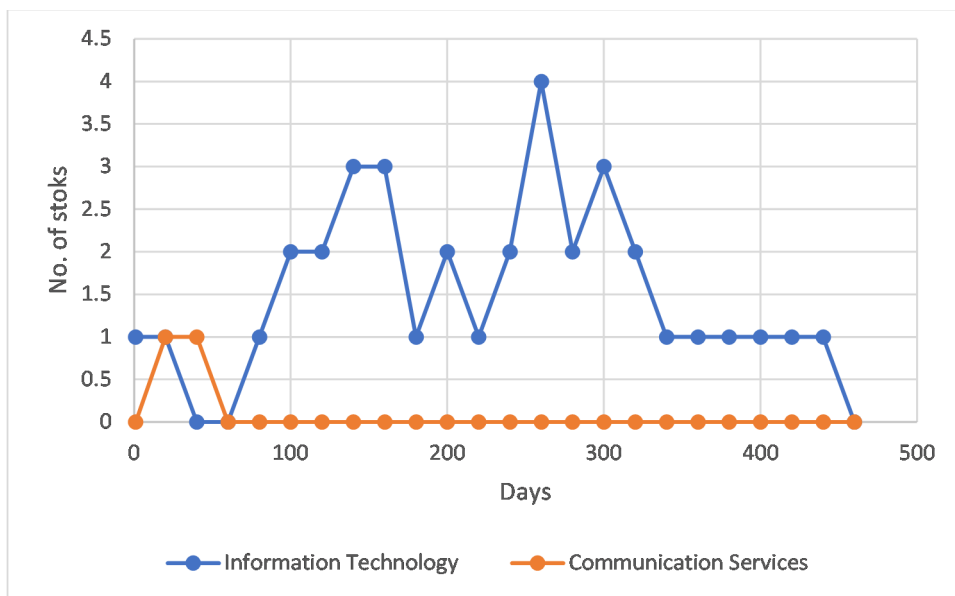

Figure 4.6: No. of stocks belonging to Information Technology and communication services sectors with long term memory
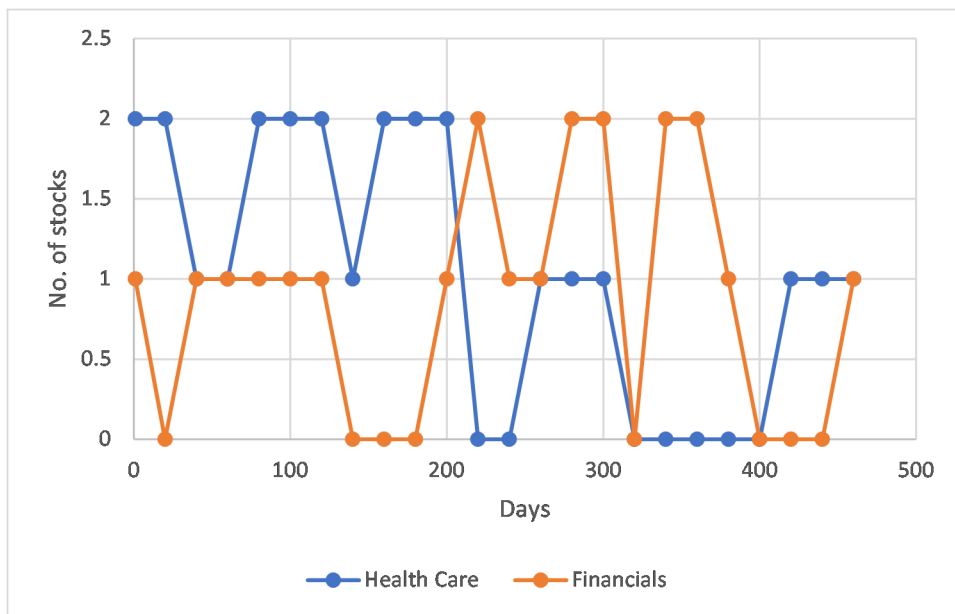
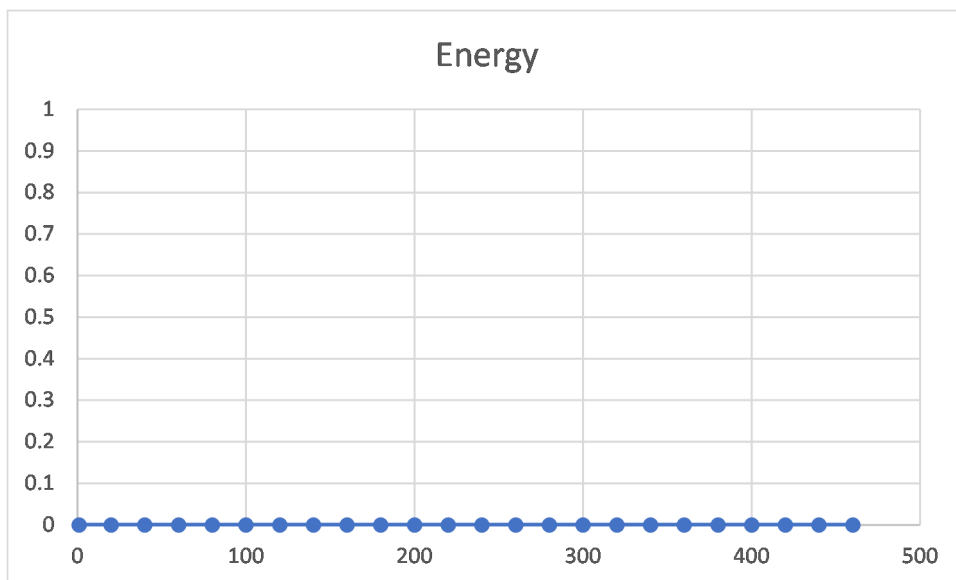Figure 4.7: No. of stocks belonging to health care and financials sectors with long term memory

Figure 4.8: No. of stocks belonging to energy sector with long term memory

1. As can be seen from Fig. 4.3., both real estate and materials show a positive correlation. Increase in stock prices in one sector sees a similar trend in the other. This may be attributed to the fact that an upward trend in real estate sector implies an exapected increase in demand for materials too. Greater confidence in materials will lead to a stronger performance of the stocks in materials sector.

2. In a circular economy, an increase in industrial performance would lead to increase in disposable income, either in the form of increased wages for labour or in the form of increased dividends from increased profits for the shareholders. This would lead to an increase in demand for consumer discretionary/ utilities, resulting in stronger expected performance for these sectors.

3. Income =Expenditure+ Savings=> An increase in disposable income would lead to increase in either expenditure or savings or both. This may explain the rise in performance for consumer discretionary.

4. An almost stable performance for utilities because of price elasticity.

5. Increased performance for financials after regulations after 2008.

6. The strong growth in Information Technology and long term memory in stocks of this sector can be related to the fact that Apple launched its revolutionary iPhone on January 9, 2007.

Chapter 5

Conclusion

The project has been able to identify long term trends in the stock market. The sectoral analysis points out interdependence of stock prices.

During the 2008 stock market crash all sectors were not equally affected. While there was drastic fall in stcok prices for Information Technology, health care and financials, the stock prices for Utility and consumer staple sectors had marginal fall and soon recovered.

Long term trends in a particular and correlation with other sectors can be used to predict market instabilities.

Bibliography

[1] Hayes, Adam. "How Does the Stock Market Work?", https://www.investopedia.com/articles/investing/082614/how-stock-market-works.asp.

[2] Spelta, Alessandro. "A behavioral approach to instability pathways in financial markets." nature communications, https://www.nature.com/articles/s41467-020-15356-z.

[3] DOWNEY, LUCAS. "Efficient Market Hypothesis (EMH)." Investopedia, https://www.investopedia.com/terms/e/efficientmarkethypothesis.asp.

[4] Chen, James. "Arbitrage." Investopedia, https://www.investopedia.com/terms/a/arbitrage.asp#:~:text=Arbitrage%20is%20the

[5] Smith, Tim. "Autocorrelation." Investopedia, https://www.investopedia.com/terms/a/autocorrelation.asp.

[6] Kenton, Will. "Pearson Coefficient." Investopedia, https://www.investopedia.com/terms/p/pearsoncoefficient.asp#:~:text=The%20Pears

[7] SERLETIS, APOSTOLOS. "DETRENDED FLUCTUATION ANALYSIS OF THE US STOCK MARKET." World Scientific, https://www.worldscientific.com/doi/10.1142/S0218127408020525.

[8] "What are the S&P Sectors?" S&P sectors, corporate finance institute, https://corporatefinanceinstitute.com/resources/knowledge/finance/the-sp-sectors/.

[9] Farmer, J. Doyne. "The Economy Needs Agent-Based Modeling." https://www.researchgate.net/publication/51437577_The_Economy_Needs_Agent-Based_Modeling.

[10] Onnela, J-P. "Asset Trees and Asset Graphs in Financial Markets." Physica Scripta, https://iopscience.iop.org/article/10.1238/Physica.Topical.106a00048.