University of Macedonia
Department of Balkan, Slavic & Oriental Studies

"Three essays on the behavior of Real Estate and Stock Markets: With applications to Hong-Kong Markets using Nonlinear based approaches".

Argyroudis Georgios

A thesis submitted for the degree of

Doctor of Philosophy in the Department of Balkan, Slavic & Oriental Studies of the University of Macedonia

May 2019

Thessaloniki, Greece

Supervisory Committee

Prof. Siokis Fotios, (supervisor)

Prof. Noulas Athanasios, (member)

Prof. Papapanagos Harry, (member)

Abstract

The main goal of this research is to detect the behavior of the stock exchange and housing market indices of Hong Kong and to assess the complex changes during times of rapid economical mutations, through 3 fundamentally different approaches. We present the multifractal nature of the financial markets based on scaling exponents and the singularity spectrum analysis, and we scrutinize in Hong Kong's stock exchange indices. Furthermore, we will apply both Multifractal Detrended Fluctuation Analysis (MFDFA) and Multifractal Detrended Moving Average (MFDMA) methods to assess both multifractality of each period chosen, and also the connection of the foreshock and aftershock periods. In the case of MFDFA, the HSI for the 2007 and 1994 crisis show the highest level of multifractality, while in the MFDMA method, the findings support that the 1997 crisis and the property market crisis present the highest multifractality level. We also report in our findings that the results from MFDFA and MFDMA could not be used interchangeably. What is more, this research aims to study Hong Kong Real Estate market from two more perspectives. We will provide entropic measurements and efficiency tests of the housing market of Hong Kong, by analyzing the housing market in terms of size and region in order to understand the effects of the sub prime loan crisis in the country. One of the most important findings is the Kowloon area that seemed not impacted by the crisis. Lastly, with the temporal evolution of the indices, we identify periods where the underlying dynamical structure of the market was impacted by certain events like the SARS epidemic and the imposition of Special Stamp Duty on housing. As far as the last perspective is concerned, we would like to show causality patterns between local housing markets and, how causality changes over time, especially during 2007 Great Recession. We used a part of Thermal Optimal Path (TOP), and Granger causality. Our findings confirm the above conclusions and suggest that Kowloon still provides a different reaction to the financial crisis.

Acknowledgements

First of all, I want to express my gratitude to my supervisor, Professor Siokis Fotios, for having offered me a position in his group, not only as a PhD student, but as a master student back in 2013 as well. When I asked him to become part of his team, I told him that he would not regret it. Hope he still believes the same! His enthusiasm, optimism and creativity were contagious. I learned so much valuable skills from him, and without doubt if it was not for him, I would still be trying to write down a proper scientific paper. And trust me, I would not. His valuable feedback on a daily basis has contributed like anything else on this Thesis. I am also grateful for the freedom he gave me, and for always having had an open door for discussions and comments, despite his busy schedule. To be more precise, I had a full access to his office anytime I wanted, showing the highest level of trust and respect to me. I really hope to find someday a way to repay what he has done for me.

I am also grateful to any member of the university for all these fruitful discussions we had about political, economical and philosophical issues. I would like to thank everyone for all those wonderful years.

Lastly, I thank my family. To put it simply, nothing could have been done without them. I would not be where I am today were it not for them. Their unconditional support is during my whole life. And vice versa.

Contents

Table of Contents

List of Figures

List of Tables

CHAPTER 1

1.1 Motivation

The study of the financial sector and the Real Estate markets has become a very challenging issue, especially after the sub prime loan bubble in 2007. The crisis of that period named the Great Recession has caused a number of problems in the US economy and other countries around the globe as well, causing global financial distress. The relationship between the Real Estate market and the real economy encourages researchers to track and produce numerous studies over that subject mainly because of the investment opportunities. Both financial and Real Estate markets attract a lot of investors and institutions to support a portfolio strategy, and also for the scientists to understand how the market works and why a crisis prediction remains such a mystery. From the Efficient Market Hypothesis to multifractal theory, and from permutation entropy and statistical complexity to thermal optimal path model, every research has contributed of making clearer how those markets behave and relate to each other, especially during anxious periods. Still the findings and the results remain ambiguous. This research is divided in three main parts, and each one aims to unveil some respective patterns in the datasets, in order to assess Hong Kong's financial and Real Estate prospects.

1.2 Hypotheses

We apply three hypotheses that define the scope of the research:

- MFDFA and MFDMA cannot be used interchangeably
- Permutation Entropy does not change before and after the Great Recession
- Thermal Optimal Path does not follow Granger causality findings

1.3 Methodology

As far as the first part of the research is concerned, there are many methods in order to find multifractal characteristics in financial time series. Fractals and multifractals are ubiquitous in natural and social sciences (Mandelbrot B. , 1983), (Mandelbrot B. , 1997), (Sornette, 2004). The time series of various data sets have been investigated in many ways, such as rescaled range analysis, spectral analysis

(Hurst H. , 1951), (Mandelbrot & Wallis, 1969), (Mandelbrot & Wallis, 1969), (Mandelbrot & Wallis, 1969), (Mandelbrot B.B., 1968), (Mandelbrot & Wallis, 1969), R/S analysis (Vanderwalle & Ausloos, 1997), partition function method (Sun, Chen, Yuan, & Wu, 2001) (Jiang & Zhou, 2008), fluctuation analysis (Peng C. K., et al., 1992), detrended moving average (Arianos S., 2008) (Arianos S., 2007), detrended fluctuation analysis (Peng C. K., et al., 1992), (Hu, Ivanov, Chen, Carpena, & Stanley, 2001), (Kantelhardt, Zschiegner, Koscielny-Bunde, Havlin, Bunde, & Stanley, 2002), wavelet transform module maxima (Holschneider, 1988), (Muzy J.-F., 1991), (Muzy J.-F., 1993), (Muzy J.-F., 1993), (Muzy J.-F., 1994) (Struzik, 2002), (Turiel, 2003), and detrended moving average (Alessio E., 2002), (Carbone, Castelli, & Stanley, 2004), (2004), (Alvarez-Ramirez J., 2005), (Xu, Ivanov, Hu, Chen, Carbone, & Stanley, 2005). In the first part of this research our aim is to use Multifractal Deterended Moving Average (MFDMA) and Multifractal Detrended Fluctuation Analysis (MFDFA), which are presented later in detail, and apply them on the periods and the time series chosen.

Moving to the next part of the Thesis, we focus to the permutation entropy and informational efficiency, where initially various methods have been suggested for the estimation of the probability distribution. Firstly, the Fourier analysis, introduced by (Powell & Percival, 1979), the wavelet transform by (Rosso & Mairal, 2002), the symbolic analysis by (Daw, Finney, & Tracy, 2003), amplitude statistics (Micco, González, Larrondo, Martin, Plastino, & Rosso, 2008) and permutation entropy, introduced by (Bandt & Pompe, 2002), (Bandt, 2003). Permutation entropy is the only method that derives information from the temporal structure of time series given (Zunino L. , Zanin, Tabak, & Pérez, 2009). There is a vast spectrum of sciences that reveal its simplicity-yet its important role. In the medical area or other applications about EEG, see (Li, Yan, Liu, & Ouyang, 2014) (Keller & Lauffer, 2003), (Li & Richards, 2007), (Keller, Mangold, Stolz, & Werner, 2017), (Olofsen, Sleigh, & Dahan, 2008), (Ouyang, Li, Liu, & Li, 2013), (Cao Y. , Tung, Protopopescu, & Hively, 2004), (Bandt, 2017), (Jordan, Stockmanns, Kochs, Pilge, & Schneider, 2008), (Li, Cui, & Voss, 2008), (Bruzzo, Gesierich, Santi, Tassinari, Birbaumer, & Rubboli, 2008), (Taherkhani, Rahmani, Taherkhani, Akbarzadehd, & Abroshan, 2013), (Bian, Qin, Ma, & Shen, 2012), (Yao, Yang, Bai, & Cheng, 2016), (Weck, Schaffner, Brown, & Wicks, 2015), (Dionisio, Menezes, & Mendes, 2006), (Zunino L. , Zanin, Tabak, & Pérez, 2009). Our goal in this part is to use Shannon and Tsallis

entropy with the complexity entropy causality plane to assess the 2007 crisis.

The third part of the research follows the steps of (Zhou & Sornette, 2006) in order to find the 'optimal thermal causal path', a method that determines the causality relationship between two time series. The investigation of the relationship between time series has already been processed with methods like the Dynamic Time Warping (DTW), see (Shinde & Pawar, 2014), (Keogh & Pazzani, 2001), (Silva & Batista, 2016). Furthermore, for the optimization of the DTW, when using large datasets, various methods have been proposed. Two examples for the implementation of global constraints are the Sakoe-Chiba band and the Itakura parallelogram (Itakura, 1974), (Sakoe & Chiba, 1978). For more details, please refer to (Asuncion & Newman, 2007), (Bratko, 2000), (Keogh & Pazzani, 1999), (Salvador & Chan, 2007), (Keogh & Smyth, 1997).

Another method used for assistance and increase of the performance of the DTW was memoization (Cormen, Leiserson, Rivest, & Stein, 2009). As far as TOP is concerned, this novel non-parametric methodology has been applied to economic data (Zhou & Sornette, 2006), the US inflation rate (Zhou & Sornette, 2007) as well as US stock market indices and US bonds (Guo, Zhou, Cheng, & Sornette, 2011). The TOP (Thermal Optimal Path) method is also found to support the lead-lag relationship of offshore and onshore Renminbi exchange rates (Xu, Zhou, & Sornette, 2017), and for the analysis of the 'one belt-one road' strategy in China (Lai & Guo, 2017). More information about the TOP method is found in (Wang, Tu, Chang, & Li, 2017), (Gong, Ji, Su, Li, & Ren, 2016), (Andersen, 2006), and we should also mention a modified TOP method named TOPS, the 'symmetric thermal optimal path' in (Meng, Xu, Zhou, & Sornette, 2017).

Most of the time series used in this Thesis was found for free at Google and Yahoo, and for some tests we used time series from Bloomberg as well. In some cases we had to shuffle the data in order to compare them with the non-shuffled series. Each index was shuffled more than 1000 times to ensure that it is valid enough to represent randomness. All algorithms and data used for the analysis of this research were conducted in MATLAB.

1.4 Expected contribution

The main goal of this Thesis is to detect the behavior of the stock market and the housing market of Hong Kong, and to see the complex changes during times of

rapid economical mutations, through 3 fundamentally different approaches. We want to understand the multifractal nature of the financial markets, and to delve deeper in Hong Kong's stock exchange indices. Furthermore, we will apply both MFDFA and MFDMA methods to assess both multifractality of each period chosen, but also the connection of the foreshock and aftershock periods. We also want to report whether the results from MFDFA and MFDMA could be used interchangeably.

What is more, this research aims to study the same Real Estate market from two more perspectives. Their successful application to different scientific problems increases our belief that the results could be persuasive. We will provide entropic measurements and efficiency tests of the housing market of Hong Kong, understanding the effects of the sub prime loan crisis in the real estate market. Finally, we would like to show causality patterns between local housing markets and if possible, how causality changes over time, especially during 2007 events.

1.5 Bubbles, and their importance

The main event under analysis in this Thesis is the sub prime loan crisis of 2007. We provide some more information about the bubbles in general, and their importance for the economy. Bubbles have become an interesting subject under analysis that academics started tentatively using it since the 1980s, and they are now part of the modern economic discourse. This is shown from the rapid increase of search phrases like "housing bubble" in Google's Ngram Viewer too. We present in this part information about various bubbles. We also show the importance of the Real Estate housing bubbles, the effects in society, institutions, households, investors and other participants, the need for their investigation, and specifically the moments of turbulence where the markets under pressure reveal their complex structure.

Real estate markets are complex systems, where different economic participants decide to enter or exit the market, and also try to speculate from either a rapid acceleration of the prices or a potential bubble burst. Even though the investors theoretically act according to a strategy chosen, or personal preferences and experience, their actions could affect the whole housing market system, for example through the channel of herding and/or imitation. Therefore, unpredictability and statistical stationarity should not be taken for granted. Real estate housing market bubbles and crashes could affect the lives of many households and investors. The speed of information, and the technological development made the bubble's creation

and burst an important event in the economy. It can weaken or strengthen the trust over corporations, institutions and the economy. An unfortunate trigger event could infect other economies as well.

The Real Estate market of the US has been under analysis from many authors. In the paper of Zhou & Sornette (2006), the real estate market is analyzed as a whole, as separate northeast, south, midwest and west regions, and as each state in order to determine the market dynamics. Supporting the mechanisms that lead to positive feedbacks such as herding, imitating behavior, and investors over-confidence, they found that there are signs of an occurring bubble in mid-2006, because housing prices have accelerated mostly to the West and northeast regions of the country. In a previous paper about the US real estate market of 2000-2003, Zhou & Sornette (2003) supported the view of Alan Greenspan that despite the reduced short-term yields, high mortgage debt, negative savings and the fact that the Real Estate market was part of a huge credit bubble in the country that was developed over the last decades, there were no signs of a significant deflationary risk. At the same time, the analysis supports the idea of either a change of direction or a crash in the UK housing market near the end of 2003.

The aforementioned arguments support the importance of understanding the market bubbles, and why so much analysis is already done. This has started in the 1990s with the time behavior of the S&P500 (Sornette, Johansen, & Bouchaud, 1996) and some other papers that attempted to predict potential crashes (Sornette, 1999), (Johansen, Ledoit, & Sornette, 1998). Later on, agent-based models were also included in the prediction process (Harras & Sornette, 2011). Furthermore, modifications over the existing methods were added in the arsenal of bubble prediction (Filimonov & Sornette, 2013). Papers like the ratio approach (McCarthy & Peach, 2004), (McCarthy & Peach, 2005), the user cost approach (asset-market approach) model (Levin & Wright, 1997a), (Levin & Wright, 1997b), and the Vector error correction model (VECM) (Case & Shiller, 1989), (Quigley, 1999), (Sing, Tsai, & Chen, 2006) should be noted as well.

A very important model is the Log Periodic Power Law (LPPL), an approach proposed by (Johansen, Ledoit, & Sornette, 2000), and (Sornette, 2003), (Sornette, 2004). The paper of Kurz-Kim (2012) is a typical LPPL example that shows how the LPPL works. In his empirical application on daily prices of German stock index (DAX index), the LPPL model is used to reveal the financial crash of 2007. The

author considers time series used for the LPPL calculation from 12 March 2003 until 14 September 2007, taking into account that one day later (15 September 2007), the collapse of Lehman Brothers became public. Firstly he sets the end of the sample in the beginning of January 2007 (02/01/07) until 14 September, to check which date contains the lowest RMSE. Nevertheless, he questions the validity of his finding because the sequence of RMSE of the LPPL model shows several local minimums and non-monotonicity as well. Therefore, his final LPPL model choice is based on the date where the model presented its lowest RMSE, which is 20 June 2007. The actual fitting of the LPPL model over the stock price concludes that LPPL could indeed provide a validate prediction of a potential bubble, while we should take into consideration that the critical time and the real time of the crash had a time distance of more than half a year. Prediction models were also found in various paper using machine learning techniques, see (Lam, Yu, & Lam, 2008) and (McCluskey, 1996).

In the Hong Kong market, the three Hong Kong crashes in a time span of a decade have showed in the best way possible the rise and fall of speculative bubbles and their respective crashes (Sornette, 2004). The first bubble is part of the 1987 crash, where on October 19, 1987, the HSI index peaked at 3,362.4 and closed a few days after on October 26 at 2,241.7, with a cumulative loss of 33.3%. In the United States major indexes lost almost 30% of their value in six days, from October 14, to October 19 1987. At the same time, many countries around the world faced a similar situation, and although the US stock prices advanced back again around 30% after the end of the crash, this was not the case for other countries like Japan. An interesting detail is that we do not clearly know the cause of the crash; nevertheless, the author tries to backfire with his argument some possible reasons like computer trading, derivative securities, illiquidity, trade and budget deficits, and overvaluation.

The second bubble belongs to the group of the “slow crash”. On February 4, 1994, the HSI index closed at 12,157.6, and during the next month on March 3, it closed at 9,802, with a cumulative loss of 19.4%. Two months later, on May 9 it reached another total cumulative loss of 30.7%. The third bubble of this period that ended on August 1997 started a decay until October 17, when an abrupt crash until October 28 led the HSI index from 13.601 to 9,059.9 with a total loss of 33.4%.

In a typical situation of a bubble creation, the bubble starts on an initial smooth pace, supported by a relatively optimistic market. In the next level, the already inflated bubble attracts new investments because the leverage from novel

sources gives a potential for gains to the new entrants. Afterwards, less sophisticated investors participate in the market which results in the leverage to alternate as there is a demand for the market price to rise in a faster rate than the real money is put in the market. After this point, there is no coupling of the market with the actual real wealth, and finally as the price of the market still increases, less investors start to enter the market, a new phase of nervousness beleaguers the prices, and at some point the instability is revealed and the markets collapses (Sornette, 2004).

Many real estate bubbles come first in our mind when we think of the bubble phenomenon. Nevertheless, not only financial and real estate markets faced a bubble creation and burst. Bubbles are everywhere: from the tulip mania, the Mississippi bubble, the south-sea bubble and the British Railway mania, to today's Japanese real estate collapse and the dot.com bubble, until the well-known Great Recession and so on. In some cases, its devastating effects to an economy or through contagion effect to other economies outside the bubble's geographical boundaries are still ongoing. Alan Greenspan has already noted the difficulty of understanding the bubble and its burst (Sornette, 2004), (Greenspan, 2007).

A famous and historical bubble is the tulip mania, the tulip speculation in the republic of the Netherlands. In this case, people started buying and selling tulip bulbs, and we are not sure if the built-up attracted new investment, or the opposite, or even both. The fact was that some trigger event made people to think that the tulip bulb was some sort of "guaranteed investment". This lead to an unprecedented phase and people sold their business, and mortgaged their houses to trade tulips. The tulip mania phenomenon lasted from the mid-1500s until 1636. The crisis started on February 4, 1637, when the bulbs became unsalable, tulip bulbs that valued thousands of US dollars lost their value and the market crashed.

Another very interesting example is the South Sea bubble back in 1720, which lasted for six months until August 1720. In that case, even the famous scientist Isaac Newton had an investment loss of nearly 2000 UK pounds after the crash. After the crash he said: "*I can calculate the motion of the heavenly bodies, but not the madness of people*". The South Sea Company had the monopoly of all trade to the South Sea ports. In return South Sea would assume a portion of the national debt of England, a debt that came out during the War of the Spanish Succession. In 1719 the company asked to assume the whole public debt of the British government. After that, the company used artificial means in order to increase the price of its stock. More

investors (even Dutch Investors) bought the company's stock and as a result there was higher inflationary pressure. At the same period some other imitators emerged hoping to cash from the potential speculation mania. Two examples that define the trend and the expectations of the trader were firstly a project to improve the Greenland fishery and the second project was to import walnut trees from Virginia. At the end the crash was dramatic, and the investors lost their fortunes.

We have also the famous Black Tuesday, on October 29, 1929. That crash put 13 million Americans out of work and it was the reason for a new era of government regulations. Galbraith (1997) writes that "an increasing number of persons were coming to the conclusion - the conclusion that is the common denominator of all speculative episodes - that they were predestined by luck, an unbeatable system, divine favor, access to inside information, or exceptional financial acumen to become rich without work". But even economists like Irving Fisher, professor at Yale University could not foresee. It was just 14 days before the Great Depression when he said "In a few months, I expect to see the stock market much higher than today". His statement clarified how difficult is to make a prognosis of a financial bubble.

Other examples of wrong prediction of a trend follow. We quote in exact words what was told back in the day before the Hong Kong crisis in 1997, to present the difficulty of forecasting. Even the famous economist Irvine Fisher has misunderstood the bubble of 1929. Three days before the crash he said that stock prices have reached a level which looks like a permanently high plateau.

(Credit Suisse First Boston – 15 July 1997): *We believe the supply demand imbalance will continue to drive residential prices during 4Q 1997 and into 1Q 1998. The two extremes will perform best; luxury residential prices should fare particularly well as we see flows of capital into the top end of this sector. Luxury property is also much less sensitive to affordability levels and much more geared to supply and price expectations. The smallest units (below 40 square meters) will also fare well as there is very limited supply at this level, yet they are the most affordable private sector flats.*

(Deutsche Bank – September 1997): *More supply does not mean oversupply when one considers demolition and population growth, which have been higher than forecast. Mass residential property prices are likely to rise at a more subdued rate than before, possibly in line with the nominal GDP growth rate and dominated by end-users, as*

would-be speculators are now aware of the government's seriousness in targeting speculation. Trading up is likely to be a more important source of demand as living standards and aspirations rise, supplemented by the wealth effect of a 419% appreciation in capital values during 1987-1996.

(Goldman Sachs – 22 October 1997: Overweight Major Property Developers): *Without details of the new policies, our initial assessment is that supply should increase and property price growth should slow. Nonetheless, the rising trend of end-user demand due to population growth and increasing net immigration is irreversible… This cushions property prices from undue correction, since they have already fallen by some 15% from the peak in April- May 1997.*

What is more Sornette et al. (2013) support the fact that is difficult to find the presence of a bubble for two main reasons. Firstly, because it is not crystal clear how we define the fundamental value. Secondly, it is difficult to distinguish between an exponentially growing fundamental price, and an exponentially growing bubble price. For example the JLS model is suitable only for endogenous crashes. To be more accurate, the JLS model is for bubbles, not for crashes. Endogenous crashes are preceded by the bubbles that are generated by positive feedback mechanisms of which imitation and herding of the noise traders are probably the dominant ones among the many positive feedback mechanisms inherent to financial system.

The uncertainty behind bubble creation and bubble burst explains why we focus and try to understand the Hong Kong's turbulent period of 1997 and 2007. In the following chapters we report our findings that could shed some light to the financial and housing market of the country.

CHAPTER 2

2.1 Introduction

Financial markets are complex systems that contribute to the society's prosperity and wealth. Participants of a financial system are the numerous individuals, institutions, governments and any other factor that want to generate profit by buying and selling those types of assets. In the housing market the participants are more or less the same, but it is still a unique category of its own; this is because a house can be used either for personal use (a roof over one's head) (Muellbauer & Murphy, 1997) or for investment purposes. What is more, the housing market is a risky and profitable environment, because the Real Estate market is immobile, which means that any property cannot be placed on another region, and indivisible, meaning that no real estate housing property can for instance be split in half. Furthermore, the transaction costs are relatively high, in a sense that households or investors consider carefully each transaction because their future income, portfolio, prosperity and economic status could be radically affected. With the transaction costs we should take into consideration some other extra costs like for example taxes that depend on the Real Estate market of the country under analysis. Furthermore, depending on the country, the level of government intervention over the Real Estate market differs, as well as the market imperfection. Last but not least, the special features derive from the fact that the market is heterogeneous, the Real Estate assets are supposed to be durable, and with an inelastic supply.

We believe that Hong Kong SAR market is a special case study ideal for further exploration. With its glorious history, and a prosperous economy, one of the leading Asian economies, it is the perfect research subject and an ideal laboratory for a Real Estate investigation, based on the fact that it combines a dense population with a high demand for properties, a developed economy with a relatively low interest rates environment and frequent increases in Real Estate prices, where for instance during 2003-2015 the housing prices have tripled. Being part of China since 1997 after the transfer of sovereignty over Hong Kong as a special administrative region, Hong Kong is a western style economic environment in the heart of Asia and China. The island provokes us to shed some light in its volatile periods, and it definitely deserves our attention.

The 2007-2008 financial crisis in the U.S.A. caused an unprecedented international turmoil. The substantial growth in mortgage credit and consequently in housing prices triggered an increase in mortgage delinquencies, as interest rates continued their upward trend. The burst of the housing bubble and the meltdown of the associated prices created a chain of reaction to other financial markets, both domestically and internationally. Asset prices across the globe fell drastically, while financial volatility rose substantially. Among the markets, where prices plummeted considerably, were the housing markets and particularly those that historically exhibited higher volatility.

On the one side it is argued that financial time series are independent without long-range correlation following a Gaussian random process. On the other side, Mandelbrot (1963) claimed that financial time series demonstrate volatility clustering, meaning that large price changes are followed again by large changes in the series. In addition to that, the distribution of the empirical data shows a tendency for fatter tails than the Gaussian distribution. Mandelbrot et al. (1997) referred that all ARCH-type models assume that the data are scale-consistent. Also Di Matteo et al. (2005) supported that scale analysis can be applied in financial data, and we can assess financial time series in different time scales in order to check the effect of the heterogeneous market participants.

Scale analysis has been used extensively in various techniques. See for instance (Ausloos, 2000), (Hurst H. , 1951), (Hurst, Black, & Sinaika, 1965), (Peng, Buldyrev, Havlin, Simons, Stanley, & Goldberger, 1994), (Lo, 1991), (Zipf, 1949), (Geweke & Porter-Hudak, 1983), (Percival & Walden, 2000), (Ellinger, 1971), (Ivanova & Ausloos, 1999), and (Sowell, 1992). As a result, the field of econophysics was created, which is a field combining theories and concepts from applied mathematics, statistical physics and non-linear science.

We consider for the research the Hong Kong Stock exchange market (HKEX) as an ideal laboratory for such an investigation, since is one of the largest stock exchanges in terms of capitalization and had exhibited large volatility fluctuations[1]. Stock market exchanges follow a complex architecture, having non-linear properties

[1] In July 2016 the HKEX introduced the so-called Volatility Control Mechanism (VCM) and Closing Auction in an attempt to curb substantial increases or decreases in volatility and prices. According to various reports, volatility in HSI during the closing period can be as much as six times greater than in other developed market indices. Also the HKSE in Asia's second largest stock exchange in terms of market capitalization and the fifth largest in the world.

and the stylized facts call for long-memory, fat tails and multifractality (Mandelbrot B. B., 1963). The importance of the stock exchange general index is well known, since it is considered as a forward looking index, indicating the future direction of the economy. Many publications have already used the stock exchange index to understand the signs of an economy. Oswiecimka et al.(2008) used the DAX index to show the multifractality of bull and bear phase for positive and negative fluctuations in 2 different periods. The slump period shows stronger correlations, irrespective of the sign of fluctuations, while the multifractal spectra are narrower in bear phase than the bull phase. Drozdz et al. (2008) analyzed in detail 30 DAX companies during an 11 year period. They argue that the drawdowns are constantly accompanied by a large separation of one solid collective eigenstate of the correlation matrix, which, concurrently, decreases the variance of the noise states.

On the other side, the draw-ups were found more competitive. In this case the dynamics distribute more uniformly over the eigenstates, which results in a rise of the total information entropy. Hence, increases are more competitive, less collective, and as a result more nonlinear correlated than decreases. Schmitt et al. (2011) analyzed the Chinese currency against euro and the USD. Firstly, they report that multifractality is present in both exchange rates. What is more, they divide the time series into several portions in order to uncover any statistical differences. They also found that during the pegged period there is a change in the power spectra, and when the pegged period ended and the exchange rate was still decreasing, the fluctuations were scaling with multifractal exponents. Wang et al. (2012) focused on the Chinese yuan exchange rate index, and more specifically in the exchange rate system reform period on July 2005. They separated the CIB-CNY Composite Index return series into 2 periods to compare their statistical properties. Results show that both periods exhibit different multifractal results, which is caused by the change in the yuan exchange rate regime in 2008.

We aim to explore and compare the multifractality in the context of time series of the Hong Kong stock market index, and Hong Kong HSI financial and property sub indexes as well. The scope of this chapter is to test the multifractal theory in periods of high stress, before and after the climax of a crisis. The division of data before and after the crisis period describes the concept of bubble creation and burst. The bubble creation starts with a boom phase where stock prices are increasing and as a result the system accumulates energy, and the burst is the phase where prices, after a strong

shock start behaving violently resembling to a certain degree, an aftershock seismic sequence. Multifractality has already been found in long time periods as in (Wang, Wu, & Pan, 2011). We are interested in the forecasting improvement of the housing market, and we think that anxious and turbulent periods can have similar characteristics, thus their analysis should provide useful results. We begin with the crisis of 1994, and later we also focus in 1997 and 2007 crisis. More specifically, in 1997 two major events took place; the transfer of sovereignty that made Hong Kong part of China as a Special Administrative Region and the currency depreciation of Baht that initiated an Asian currency collapse.

In the case of 2007, the 2007 sub prime loan crisis in the US caused a chain reaction in the financial systems around the world. Financial shocks were transmitted very fast, and some of them caused a "contagion" effect to foreign markets as well. We use three indices, the Hong Kong stock market index and two constituencies, named Hong Kong HSI financial and Hong Kong HSI housing index. We want to search for the complexity in various sectors of the economy, and to scrutinize whether the real estate market follows the same trend with the rest, as well as its long-range correlation. For more robust results we will use two multifractal methods in our data, the Multifractal Deterended Moving Average (MFDMA) and Multifractal Detrended Fluctuation Analysis (MFDFA).

Firstly, we would like to review the theory of the methods, including a brief introduction of MFDFA and MFDMA, and after the application of the one-dimensional time series experiments we will assess and compare the empirical results. Both MFDFA and MFDMA followed methods like the Detrended Fluctuation Analysis (DFA), the Detrended Moving Average (DMA), the Rescaled Range Analysis, and the Multifractal Model of Asset Returns methodology, therefore we first provide an overview of those methods before we proceed to the MFDFA and MFDMA.

2.2 Methodology

2.2.1 R/S analysis (Rescaled Range Analysis)

Hurst used for the first time the R/S analysis in order to describe the long-term dependence of water levels in rivers and reservoirs (Hurst H. , 1951). However,

according to Lo (1991), Teverovsky et al. (1999), the method has problems when multiple scale behaviors, heteroskedasticity and short memory were present. The method is also prone to error since the range relies on maxima and minima. In 1991, Lo (1991) proposed a modified version of R/S analysis in which the main difference with the initial R/S was the denominator (Newey & West, 1987).

Even though the modified version could identify long memory when short-term dependence is present (Moody & Wu, 1996), it still had the issue of choosing the truncation lag q. Furthermore, Vanderwalle & Ausloos (1997) showed that the markets will not react efficiently, maybe because the actions of either a government or an institution could be emotionally motivated. This could be the reason why the R/S analysis possibly misses some trend, sequence, or the large and low scale investment horizons.

2.2.2 MMAR (Multifractal Model of Asset Returns)

In Fisher et al. (1997) the Multifracal Model of Asset Returns (MMAR) is presented. When the concept of trading time is combined with multifractal measures, it creates continuous-time stochastic processes that have long tails, long memory in volatility, and they are sufficiently flexible to generate long memory or martingale behavior in log prices. They applied the method to the exchange returns of DM/USD, and they found evidence of multifractality in the series. Also, the scaling properties of DM/USD are robust to seasonal adjustment methods.

What is more, the scaling laws are very similar within subsamples of the data. As a result, a single stationary model can offer an adequate approximation to the data generating process, covering either a long and highly variable sample span, or a broad range of sampling frequencies. The same authors also note that long dependence is displayed by the multifractal model in the absolute value of price increments.

In contrast, price increments could remain uncorrelated. And even though MMAR incorporates long tails, it is not implied that there is "*an infinite variance of returns over discrete sampling intervals*". Furthermore, the MMAR contains the concept of trading time as well as long-dependence. Finally, the main advantages of MMAR are multiscaling, compatibility with the martingale property of returns, long memory in volatility, and scale consistency. But the main disadvantage of the multifractal method is the lack of applicable statistical methods. In their paper, Calvet et al. (1997) concluded that the unique scale contained in earlier financial models

including fractional Brownian motion and Ito processes contrasts with the multiscaling properties of the MMAR method.

2.2.3 DFA (Detrended Fluctuation Analysis)

Multifractal Detrended Fluctuation Analysis owes credit to the Detrended Fluctuation Analysis, which was presented by (Peng, Buldyrev, Havlin, Simons, Stanley, & Goldberger, 1994). In their paper, a modified root mean square analysis was presented in order to analyze physiological data. The goal was to detect long-range correlations in non-stationary time series data. In that way it would be possible to detect healthy subjects and patients with severe heart disease. This could be done by the computation of the detrended fluctuation analysis.

What is more, Grech and Pamula (2008) applied the DFA method in order to calculate the local time-dependent Hurst exponent H_{loc}. H_{loc} is applied as a predictor for the behavior of time series returns. In their case, the WIG (Warsaw Stock Exchange index) time series is used to check if the method works for the Polish time series dataset. They found that a decreasing trend of H_{loc} appears many sessions before the time of the crash occurs. By applying an external influence, they conclude that the WIG behavior changes, thus showing how financial complex dynamical system could be affected. This is a sign, which shows that making predictions for the WIG index is not scientifically justified yet. In Vanderwalle & Ausloos (1997) paper, the DFA method is applied in order to find a series of local long-range correlations. The authors apply from 1 January 1980 until 15 October 1996 an observation box of 2 years length; it starts from the beginning of the data and slides until the end, finding every 4 weeks another a exponent.

It is shown that no matter the trend of the currency dataset, there is a long-range power law correlation of $\alpha = 0.56 \pm 0.01$ in currency fluctuations. On another study similar to (Karlin & Brendel, 1993) about patchiness and long-range power law correlations in DNA, Peng et al. (1994) used again the DFA method in DNA sequences. By applying the DFA to selected DNA sequences, they argue that patchiness is insufficient to be a proof of long-range correlation properties. What DFA method offers is the detection of long-range correlations included in a patchy landscape, and also the ability to recognize long-range correlations that are an artifact of patchiness. Also the DFA "*can identify the length scale of the biased sub regions of the uncorrelated control sequence*".

The computation steps are 3: Firstly, we integrate the interval time series $y(k) = \sum_{i=1}^{k}[B(i) - B_{ave}]$, where $B(i)$ is the ith interbeat interval and B_{ave} is the average interbeat interval. Secondly, the integrated time series is divided into boxes of equal length, n. We fit a least squares line in each box. The y coordinate of the straight line segments is denoted by $y_n(k)$. Thirdly, we detrend the integrated time series, $y(k)$, by subtracting the local trend, $y_n(k)$, in each box. Thus we calculate the root mean fluctuation analysis

$$\boldsymbol{F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2}} \quad (1)$$

This is computed for all time scales, and then we are able to show a relationship between the fluctuation of different box sizes, and the box sizes. When the fluctuation increases, the number of box sizes follows as well. It is therefore a monotonically increasing function. In Figure 1 it is presented the interbeat interval time series $B(i)$ of 1000 beats, and the integrated time series. The vertical dotted lines actually show a box of size n=100, and the solid lines are the fitting line of the series for each box.

Based on the paper of Peng et al. (1995) for the DFA, Leite et al. (2010) used the same methodology to successfully discriminate healthy subjects from those with Chagas disease, and the group with mild to moderate hypertension. Findings revealed that DFA is useful for the Heart Rate Variability, since DFA does not make any assumptions about signal stationarity.

2.2.4 DMA (Detrended Moving Average)

One method similar to the DFA is the Detrended Moving Average (Arianos S., 2007), (Arianos S., 2008). As the authors mentioned, the challenge '*is to get the Hurst exponent H, that is related to the fractal dimension D=2-H, by means of more and more accurate and fast algorithms. The methods of extraction of the scaling exponents from a random signal exploit suitable statistical functions of the series itself*'. The advantage over the DFA is that there is no need for division of the series

in various boxes, and only the use of moving average is necessary. The necessary steps for the dynamic averaging process are:

$$\sigma^2_{DMA} = \frac{1}{N-n} \sum_{i=n(1-\theta)}^{N-n\theta} [y(i) - \tilde{y}_n(i)]^2 \quad (2)$$

$$\tilde{y}_n(i) = \frac{1}{n} \sum_{k=-n\theta}^{n(1-\theta)} y(i-k) \quad (3)$$

The parameter theta is used in the range [0,1] within the moving window n. For theta=0, the $\tilde{y}_n(i)$ is calculated from the past data, for theta=0.5, $\tilde{y}_n(i)$ is calculated with half the past and half the future points within the window n, and finally when theta=1, $\tilde{y}_n(i)$ is calculated with only the future points of the window n.

The applicability of those methods can be assessed only by comparison in time series data; In the report of Shao (2012), four different estimators of long-range correlation time series (FA, DFA, BDMA, CDMA)[2] and three time series data generators (FGN-DH, FBM-RMD, WFBM)[3] were used to compare and rank their performance. According to the report, CDMA and DFA performed the best, and the FA had the worst performance. More specifically, the authors claimed that the FA is more sensitive to the choice of the scaling range than the CDMA and DFA.

2.2.5 MFDFA (Multifractal Detrended Fluctuation Analysis)

In this part we present a more advanced DFA method, which is called the Multifractal Detrended Fluctuation Analysis (Kwapien, 2005). In the pioneering paper of Kantelhardt et al. (2002), MFDFA shows that it can determine the multifractal scaling behavior of time series. The reason why MFDFA is based on DFA is that the latter could avoid correlations that are artifacts of nonstationarities in the time series data.

---

[2] FA is an abbreviation for Fluctuation Analysis, BDMA for Backward Detrending Moving Average, CDMA for Centered Detrending Moving Average

[3] FGN_DH is an abbreviation for Fractional Gaussian Noises-Davies Harte, FBM_RMD for Fractional Brownian Motion-Random Midpoint Displacement, WFBM for Weighted Fractional Brownian Motion

According to the theory, multifractality can be found as a result of broad probability density function, in which case shuffling the data does not remove multifractality, and also as a result of long-range correlations between small and large fluctuations in time, in which case the shuffled series should exhibit unifractality because the long-range correlations are destroyed by shuffling (Barunik J., 2012). In his paper, Zhou (2009) compares the return series with shuffled and surrogate data, and concludes that fat tailed distribution has a key role compared to the temporal structure.

Furthermore, Sun et al. (2001), Sun et al. (2001), and Wei and Huang (2005) point that multifractal analysis can be used for prediction of the increase or decrease of the market index, especially in the case of large fluctuations of the returns. This degree of correlation diminishes and gets weaker as time passes. Also, in the paper of Wang et al. (2011), gold market prices were used to show multifractality in the gold market before and after 1999. For all time scales, the gold market shows multifractality. What is more, they showed persistent fluctuations in the short-term, while large fluctuations are anti-persistent and small fluctuations are persistent in the long-term. The same method applies in the paper of Telesca and Lapenna (2006). Using MFDFA, they support that the seismic phenomenon under analysis present a change from heterogeneity to homogeneity in the aftershock period. This could mean a loss of multifractality compared to the main event.

A very important feature of the study is the number of total observations, 7520, enough to give reliable estimated exponents. One of those exponents is the width of the multifractal spectrum, which shows the wideness of the range of the exponents; the wider the range, the richer multifractality. Also, the $\alpha_0 max$ which equals the value of α when $f(\alpha)$ has its maximum value, shows whether the process has a regular appearance. More regular process comes from larger α_0.

Furthermore, Hasan et al. (2015) investigate the influence of Great Recession in the US on the multifractality of 2 developed, and 5 emerging Asian stock markets, including China and Hong Kong. A very interesting finding for the Chinese stock market is that the $\tau(q)$ vs q plot for the return series of the crisis period deviates marginally from the shuffled series only for negative values of q. This implies the presence of weak correlations of small fluctuations. Also, during the 2007-2008 crisis period, 5 out of 7 cases, Hong Kong, US, S. Korea, Indonesian and Japanese market

exhibit strong multifractality for $q > 0$. These markets were affected by the Great Recession, while the Chinese and Indian markets were not affected as much by the US crisis.

But still we do not have a plethora of papers about the efficiency and long range correlations in the Asian markets. In the paper of Yalamova (2006), multifractality of 5 stock indices is under examination with the use of Multifractal Model of Asset Returns (MMAR). From all five cases and around the drawdown of October 1997, the HSI index has the same foreshock and aftershock Hurst exponent, meaning that the multifractality in the aftershock period stays the same, showing volatility persistence. Another finding is that all indices present decreased α_{min} for the aftershock period, giving a sign of high risk for the investors.

For the same extreme event in 1997, Lim et al. (2008) used 8 Asian stock markets to assess whether the crisis affected their efficiency. Rolling bicorrelation test statistic was used. Results showed that Hong Kong's efficiency was hit the most over the 14 years full sample period. On the other hand, Jiang (2008) notes that the multifractality of the stock market indices is an illusion, and he supports that the multifractal nature in both real and shuffled indices is not significant. In the same way, Bouchaud (2000) claims that it is not clear if financial time series can be distinguished between those with monofractal and multifractal behavior.

In Siokis (2014), extreme events of 3 European countries were under investigation. The analysis of the stock market indices reveals how much the financial assistance (MoU) helped the economies to come back from the crisis. Siokis (2013) analyzed two extreme economic events of the DJIA index. Both crises affected the economy, and it was shown by an increase in multifractality. The crisis of 1929 affected more the economy than the 1987 crisis, especially in the aftershock period. Both crises gave credit to the existence of long-range correlations and the failure of EMH. In Siokis (2012) they investigated the exchange rates of 4 Asian economies. Using a threshold value as a measurement of how volatile the currency price changes, they conclude that all sequences follow a power law. Still, not all shocks behave the same. Siokis (2017) analyzed the multifractality of 3 financial meltdowns in 4 stock market indices.

The results revealed the existence of multifractality during every crisis, with Black Monday being the most complex event. Furthermore, they argued that for Black Monday and Great Recession, multifractality is weaker if we decrease the

threshold level of the higher price changes. Also all extreme events are followed by a period of smaller complexity. This is shown by the transition from a heterogeneous to a homogeneous pattern.

In their paper, Kantelhardt et al. (2002) show the mathematical description of the method: The multifractal generalization of the MF-DFA procedure can be briefly provided as follows. The MF-DFA operates on the time series $x(k)$, where $k = 1, 2, \ldots, N$ and N is the length of the series. We assume that $x(k)$ are increments of a random walk process around the average $\langle \chi \rangle$ and the profile is given by the integration of the signal

$$\boldsymbol{Y(i) = \sum_{k=1}^{N} [\chi(k) - \langle \chi \rangle] \, , i = 1, \ldots, N} \tag{4}$$

Next, the time series $Y(i)$ is divided into $N_s \equiv int(N/s)$ non-overlapping segments of equal length s, starting from both beginning and the end of the time series. Each segment υ has its own local trend that can be approximated by the least-squares fitting of the series. Then we determine the variance

$$\boldsymbol{F^2(v,s) \equiv \frac{1}{s} \sum_{i=1}^{s} \{Y[(v-1)s+1] - y_v(i)\}^2} \tag{5}$$

for each segment ν, $\nu = 1, \ldots, N_s$ and

$$\boldsymbol{F^2(v,s) \equiv \frac{1}{s} \sum_{i=1}^{s} \{Y[N-(v-N_s)s+1] - y_v(i)\}^2} \tag{6}$$

for $\nu = N_s + 1, \ldots, 2N_s$. Here, $y_\nu(i)$ is the fitting line in segment ν. Then, we detrend the series and average over all segments to obtain the qth order fluctuation function

$$\boldsymbol{F_q(s) \equiv \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(v,s)]^{q/2} \right\}^{1/q}} \tag{7}$$

The property of $F_q(s)$ is that for a signal with fractal properties, it reveals power-law scaling within a significant range of s

$$\boldsymbol{F_q(s) \propto s^{h(q)}} \tag{8}$$

and the variable q can take any real value other than zero. In general the exponent $h(q)$ will depend on q. For stationary time series, $h(2)$ is the well-defined Hurst exponent H and thus, $h(2)$ is the generalized Hurst exponent. Multifractal (MF) scaling exponent $\tau(q)$ is related to $h(q)$ through

$$\boldsymbol{\tau(q) = qh(q) - Df} \tag{9}$$

where D_f is the fractal dimension of a geometric support of the multifractal measure and $Df = 1$. The exponent $\tau(q)$ represents the temporal structure of the time series as a function of the various moments q, or τ reflects the scale- dependence of smaller fluctuations for negative values of q, and larger fluctuations for positive values of q. If $\tau(q)$ increases nonlinear with q, then the series is multifractal.

2.2.6 MFDMA (Multifractal Detrended Moving Average)

The MF-DMA can be described as follows (Gu & Zhou, 2010) (Wang, Wu, & Pan, 2011).

Step 1. Consider one time series $a(t)$, $t = 1,2,\dots,N$, where N is the length of the series. We construct the new sequence of cumulative sums

$$\boldsymbol{x(t) = \sum_{i=1}^{t} a(i)} \tag{10}$$

Step 2. Calculate the moving average function of the series of cumulative sums in a moving window (Arianos S., 2007),

$$\tilde{\boldsymbol{x}}(\boldsymbol{t}) = \frac{\mathbf{1}}{\boldsymbol{n}} \sum_{\boldsymbol{\kappa}=-\lfloor(\boldsymbol{n}-\mathbf{1})\boldsymbol{\theta}\rfloor}^{\lceil(\boldsymbol{n}-\mathbf{1})(\mathbf{1}-\boldsymbol{\theta})\rceil} \boldsymbol{x}(\boldsymbol{t}-\boldsymbol{k}) \tag{11}$$

where, n is the window size, $\lfloor x \rfloor$ is the largest integer not larger than x, $\lceil x \rceil$ is the smallest integer not smaller than x, and θ is the position parameter varying from 0 to 1. Hence, the moving average functions of two series consider the past $\lceil (n-1)(1-\theta) \rceil$ data points and the future $\lceil (n-1)\theta \rceil$ data points. In Gu and Zhou (2010), they consider three special cases. The first case $\theta = 0$ refers to the backward moving average (Xu, Ivanov, Hu, Chen, Carbone, & Stanley, 2005), in which the moving average functions are calculated from the past $n-1$ data points.

The second case $\theta = 0.5$ refers to the centered moving average, in which the moving average functions are calculated from half past data points and half future data points. The third case $\theta = 1$ corresponds to the forward moving average, where functions are calculated from future $n-1$ data points. In this paper, we only consider the first case, the backward moving average, $\theta = 0$, since it is evident that it gives the most accurate estimation of the exponents (Gu & Zhou, 2010).

Step 3. Detrend the series by subtracting the moving average functions, and obtain the residual sequence $\varepsilon(i)$ using the equations as,

$$\boldsymbol{\varepsilon}(\boldsymbol{i}) = \boldsymbol{x}(\boldsymbol{i}) - \tilde{\boldsymbol{x}}(\boldsymbol{i}) \tag{12}$$

where, $n - \lfloor (n-1)\theta \rfloor \leq i \leq N - \lfloor (n-1)\theta \rfloor$.

Step 4. Divide the residual sequences into N_n non-overlapped segments with the equal length n, where $N_n = \lfloor N_n - 1 \rfloor$. Each segment can be written as $\varepsilon_\upsilon(i) = \varepsilon(l+i)$ for $1 \leq i \leq n$, respectively, where $l = (\upsilon - 1)n$. Then we can define the fluctuation variance as,

$$\boldsymbol{F_v^2(n) = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_v(i)^2} \tag{13}$$

Step 5. Calculate the qth order fluctuation function using the equation:

$$\boldsymbol{F_q(n) = \left\{\frac{1}{n}\sum_{v=1}^{N_n} F_v^q(n)\right\}^{\frac{1}{q}}} \tag{14}$$

for $q \neq 0$. Where, $F_q(n)$ is the qth order fluctuation function and $F_v^q(n) = \left(F_v^2(n)\right)^{\frac{q}{2}}$. When $q = 0$, according to L'Hôspital's rule, we have

$$\boldsymbol{ln[F_0(n)] = \frac{1}{N_n}\sum_{v=1}^{N_n} ln[F_v(n)]} \tag{15}$$

Step 6. For different values of segment length n, we have the power-law relationship,

$$\boldsymbol{F_q(n) \sim n^{h(q)}} \tag{16}$$

Here, the generalized Hurst exponent $h(q)$ can be obtained by observing the slope of log–log plot of $F_q(n)$ versus n through the method of ordinary least squares (OLS). If $h(2) > 0.5$, the correlations are persistent (positive). An increase is likely to be followed by another increase. If $h(2) < 0.5$, the correlations are anti-persistent (negative). An increase is likely to be followed by a decrease. If $h(2) = 0.5$, the time series display a random walk behavior. The conclusion of multifractality can be obtained from the dependence of $h(q)$ on the values of fluctuation orders q.

The analytical relationship between generalized Hurst exponents based on MF-DMA and the Renyi exponent $\tau(q)$ is,

$$\boldsymbol{\tau(q) = qh(q) - 1} \tag{17}$$

Via a Legendre transform, another important variable set $\alpha - f(a)$ is defined by

$$\boldsymbol{\alpha = h(q) + qh'(q), f(\alpha) = q[\alpha - h(q)] + 1} \quad (18)$$

Here, α is the Holder exponent or singularity strength, which characterizes the singularities in a time series. The singularity spectrum $f(\alpha)$ describes the singularity content of the time series.

2.3 Empirical Results

2.3.1 MFDFA results

We examine the nonlinear features of extreme events of HSI, and namely with the 1994, 1997 and 2007 Hong Kong stock index crises. The data consist of daily returns of the stock market index before and after the market crash, and spans from October 1988 to June 1999, June 1992 to February 2003 and August 2002 to February 2013 respectively. We are interested in investigating the whole process of the event, energy accumulation, or bubble rising and the expansion process after the eruption. Therefore the data consist of around nine years of trading days, sufficient enough in revealing the statistical properties and relevant information. The dates of the three crashes were February 4 1994 with a slow accumulative crash of 19.4%. This downfall actually continued for the next two months until May 4, 1994 with a total loss of 30.7%. The next one was on October 17, 1997 with an HSI loss of 33.4%. The third one was on December 3, 2007 according to the NBER announcement that the U.S. economy had entered into a recession, and lasted until June 2009.

For the second crisis of 1997, it should be mentioned that in the paper of Lim et al. (2008) and (www.imf.org, 1998) the authors assume as the starting point of the 1997 crisis the devaluation of the Thai baht in July 1997. Furthermore, datasets do not contain holidays and weekends, although some extreme social and political event could occur at that time as well (Vanderwalle & Ausloos, 1997). It is important to mention that although the duration of a bubble and its burst is still vague, we want to take into consideration the complexity of the market after a tremendous event such the Great Recession, and thus we use that long time series dataset. For the sake of academic comparability, the Financial and Property index, follow the same line of

analysis. The Properties sub-index contains 10 company stocks for the Real Estate area, while the Finance sub-index constitutes of the stocks of 12 financial companies. The daily returns of the stock market indices calculated as $r_t = \ln p_t - \ln p_{t-1}$,where p(t) is the price of the index on day t and r is the rate of return. The multifractal concept is used as a feature of the financial complex systems and we are investigating the multifractal properties of the indices based on the periods of high financial stress as depicted by the HSI (Figure 2). Also descriptive statistics of the data, in terms of mean standard deviation, skewness and kurtosis are presented in Table 1.

We firstly calculate the market complexity of the crises in Hang Seng index. Then we divide the total sample into periods before (foreshocks) and after (aftershocks) the crash in an attempt to identify the changes, if any, in the market dynamics and complexity. Lastly we present the generalized Hurst exponent and we provide some concluding remarks.

In order to analyze multifractality during extreme events, we must check the complexity of all five cases; Even though the crises of 1994 and 1997 are very closed we use both since they represent very important events of the economic and political history of Hong Kong. We would like to scrutinize the complexity and multifractality of the 1987 crash; nevertheless the data were not available. The crash of 1994 was the result of a bubble that started after the 1987 worldwide crash, and ends with a slow crash in 1994. The second crash in 1997 belongs to the 1997 Asian financial crisis, and the wakes of the transfer of sovereignty on July 1st 1997.

Against our belief, Didier Sornette mentioned that "*the simultaneity of the critical times* t_c *of the Hong Kong crash and of the end of the U.S. and European speculative bubble phases at the end of October 1997 are neither a lucky occurrence nor a signature of a causal impact of one market (Hong Kong) onto others, as has often been discussed too naively*". In any case, we want to see whether the exogenous factors can affect more a local economic environment (Sornette, 2004). Lastly, we assume that the last three cases, that is HSI index, HSI property sub-index and HSI financial sub-index have a correlation with the global financial crisis of 2007-2008.

We show first the $f_q(s)$ fluctuation function from all HSI crashes, HSI property and financial index. The $f_q(s)$ functions are straight lines and the slopes change when going from positive moments to negative moments, see Figure 3.

For the MFDFA the scaling parameter ranges from minimum scale s=10 to the maximum scale that is scale=N/10, where N the maximum number of observations. What is more, multifractality can be shown in the multifractal scaling function $\tau(q)$. In Figure 4 we can see the result for the actual and the shuffled data. The reason for the use of shuffled series is that we want to check whether we still have the same level of multifractality by removing the temporal distribution. We use the same scale as for the $f_q(s)$ fluctuation function.

For all five crises, the multifractal scaling function follows a linear trend for negative values of q, and for positive values of q they follow a non-linear pattern, which is a signal of multifractality of the series. Furthermore, if we observe the cases of 1997 and 2007, we can see that both real and shuffled data behave similar for $\tau(q)|_{q<0}$, and when $\tau(q)|_{q>0}$ the series differ, and more specifically the real data follow a non-linear pattern compared to the shuffled data series. In a nutshell, we conclude that the real and shuffled series have a similar trend for small fluctuations, but they differ when the fluctuations are large.

Another way to assess and confirm multifratality is to check the results in **Error! Reference source not found. Error! Reference source not found.** below from the calculation of the q dependence of the Hurst exponent $h(q)$. We use the range of q with $q \in [-10,10]$, and when q increases from -10 to 10, the Hurst exponent decreases, presenting the typical form of multifractality in time series. Furthermore, the multifractal richness is associated with high variability of $h(q)$, which is measured as $\Delta h = h(q_{min}) - h(q_{max})$. We can see that in the 1994 case, the $h(q)$ decreases from 0.71 to 0.31 with a variability $\Delta h = 0.4$. In the 1997 crisis, $h(q)$ falls from 0.7 to 0.34 with a variability of $\Delta h = 0.36$. In 2007 respectively, the decrease measurement is from 0.67 to 0.26 with $\Delta h = 0.41$. For the 2007 property crisis, the hq for the HSI property sub-index falls from 0.66 to 0.34 with $\Delta h = 0.32$.

Lastly, in the HSI financial sub-index, the hq drops from 0.65 to 0.34 and $\Delta h = 0.31$. For every crisis that we measured, it is important to say that none was with a constant slope, and this is an indicator of multifractality. Comparing their Δh, we can conclude that 2007 crisis has the largest Δh of all five cases, meaning that the economic crisis of 2007 in Hong Kong exhibits richer multifractality than the other cases. Very closed to 2007 with $\Delta h = 0.4$, the 1994 crisis presents also rich multifractal results. On the other hand the lowest Δh belongs to the 2007 financial

sub-index and to the 2007 property sub-index. The case of 1997 lies in the middle of all five cases. Taking into consideration the economic and political events of the 90s and the 00s, we see that the rich multifractality of 1994 could be possibly explained from the previous global crisis of 1987 that still caused a "slow crash" and the fact that the country was entering the period before the transfer of sovereignty from the UK to China (Sornette, 2004). In 1997 multifractality is not as high as in 1994 and 2007, but can be explained from the fact that the crisis started a couple of months after the transfer of sovereignty of Hong Kong, on the July 1st 1997.

Finally in the 2007 phase, we observe a very interesting result. We have the richest multifractality and the two poorest multifractality measurements of all five cases. It is probable to say that the 2007 events might be a causal impact after the end of the speculative bubble phases in US market.

Thus the economy of Hong Kong presents signs of dependence with the western economy. On the contrary, the financial and property sub-indices of the country display the lowest richness in multifractality. It can be concluded that both sub sectors of the economy in Hong Kong are strong enough that were not affected as much by exogenous global factors. Therefore, a very first finding is that the housing market of Hong Kong has some kind of protection, since the reaction of the property market was not intense.

Furthermore, to assess multifractality, we convert q and t(q) to a and f(a) by a Legendre transform as $f(\alpha) \equiv \alpha q - \tau(q)$, $\alpha \equiv \frac{d\tau(q)}{dq}$, where $f(\alpha)$ is the fractal dimension of the time series. The degree of multifracality is measured by the $f(\alpha)$ spectra, which is defined as the difference between the α_{max} and α_{min}. It is very crucial to note that the asymmetry in multifractal spectra implies either the dominance of the large or small fluctuations. More specifically, when the left width is larger than the right, $\alpha_0 - \alpha_{min} > \alpha_0 - \alpha_{max}$, this shows the dominance of the large fluctuations in the distribution (Shimizu Y., 2002). In Figure 6, every spectrum is an upside down parabola, and it peaks at f_{max}(at $\alpha(f_{max})$) ranging from min to max.

The results of the α and $f(\alpha)$ lie in different ranges; the 1994 case has $0.21 < \alpha < 0.81$ and $\Delta\alpha = 0.60$. In 1997 we have $0.24 < \alpha < 0.79$ and $\Delta\alpha = 0.55$. In the 2007 crisis, we show that $0.14 < \alpha < 0.77$ and $\Delta\alpha = 0.63$. The property sub-index has $0.26 < \alpha < 0.74$ and $\Delta\alpha = 0.48$. Finally the financial sub-index of the 2007 crisis has $0.25 < \alpha < 0.72$ and $\Delta\alpha = 0.47$. Those numbers show that we have

indeed the richest multifractality in the 2007 and 1994 crisis. The poorest $\Delta\alpha$ and thus the lowest level of multifractality belong to the financial and property sub-indices of 2007. We should note that even though we shuffled the data, there appears multifractality in the shuffled series, meaning that there is long memory in the data.

In the next step of the analysis of multifractality, we use the MFDFA to assess whether those crises affect the market complexity, and more specifically the market complexity before and after of each crisis. Thus, we divide our five cases in equal sized foreshock and aftershock periods, in order to examine how the market participants reacted on each case. From now on, every reference on foreshock will be found in the paper as *fore*, and aftershock as *after* too. Both have the same meaning interchangeably. From our database, we take on every foreshock and aftershock period on average 1320 observations respectively. In the following figures, we present the multifractal scaling function t(q), the q-dependence Hurst exponents, as well as the multifractal spectra $f(\alpha)$.

We compare the foreshock period and aftershock period on every case separately. The richness in multifractality signifies a transition from a homogeneous to heterogeneous pattern. In the multifractal scaling function t(q), we see that the aftershock period is larger than the foreshock period in 1997 and 2007 cases, the third larger difference is on the 1994 case. This is not the case for the two HSI sub-indices of property and financial market, since they have the smallest difference between foreshock and aftershock period.

The richness in multifractality is related to the higher variability of h(q). In Figure 8 we can see that the aftershock period is larger compared to the foreshock period for 1997 and 2007 crises. In the 1994 and 2007 property index the measurements still show some important multifractal difference. On the contrary, 2007 financial crisis presents very low variability between foreshock and aftershock period. Those findings in the financial case and the property one show that the aftershock period was simply a transitory and did not affect the complexity of the market, meaning that they are not easily affected by external factors, always compared to the other cases of our study.

On the other hand, the remaining fore-after relations reveal a higher level of complexity in the market that became more intense after the peak of the event. This means that the market was probably affected and reacted by the crisis events.

We will get the final verdict for the fore-after relation by examining the singularity spectrum of every crisis, and comparing the range of their singularity spectrum. The results are presented in Figure 9. We also present Table 2 with the singularity spectrum of every crisis. It is found that the 2007 crisis displays a wider $\Delta\alpha$, 0.24 compared to 0.22 for the 1997 crisis. The singularity spectrum for the financial index is 0.14, while it is very interesting the fact that $\Delta\alpha$ for 1994 and property cases are negative. This indicates that the main shock during the crash was probably transitory, and thus did not change the complexity of the market in that specific period. On average, there is an average $\Delta\alpha$=0.08 for all cases, which means that the average aftershock singularity spectrum is broader by 0.08 than the foreshock spectrum.

Furthermore, we search the correlation with the generalized Hurst exponent, H(q), where q=2. According to the theory, when H=0.5 there is no correlation of data and thus they follow a random walk. If H>0.5, we have a long-term correlation of data with persistent behavior, and in contrast when H<0.5, we have an anti-persistent behavior. In total the financial crisis show a pure "Markovian" behavior with H(q)=0.5. In Table 3, we see that 1994 and 1997 total crises reveal an persistent behavior with H(q)= 0.51 and H(q)=0.52 respectively, which is nevertheless very closed to 0.5.

On the other hand, 2007 and property crises behave similarly with an H(q)=0.48, presenting an anti-persistent behavior, which is again closed to pure noise. In 1994 a change from a foreshock persistent behavior with H(q)=0.53 to an aftershock anti-persistent behavior of 0.49 can be explained by the fact that the market participants overreact before the beginning of the crash, following some imitation trend and having the same herding behavior. After the crash (a critical point in time), they initiate an anti-herding behavior, because the available information was not well absorbed and there is still some panic in the market. Apart from 1994, it is very interesting to notice that in the rest foreshock periods we have an anti-persistent behavior that is followed by another aftershock anti-persistent behavior. The exception is the 1997 aftershock, which is under the farm of the Efficient Market Hypothesis.

2.3.2 MFDMA results

In a similar way like the MFDFA example, we examine the nonlinear features of extreme events of HSI, and namely with the 1994, 1997 and 2007 Hong Kong stock index crises. The data for each crisis consist of 2640 trading days ≈ almost 10 years of the stock market index. They span from October 1988 to June 1999, June 1992 to February 2003 and August 2002 to February 2013 respectively. The reason for this amount of database usage is that we are interested in investigating the whole process of the event, i.e. energy accumulation, or bubble rising and the expansion process after the eruption.

Therefore the data consist of around ten years of trading days, sufficient enough in revealing the statistical properties and relevant information of those highly volatile and nervous periods. The dates of the three crashes were February 4 1994 with a slow accumulative crash of 19.4%. This downfall actually continued for the next two months until May 4, 1994 with a total loss of 30.7%. The next one was on October 17, 1997 with an HSI loss of 33.4%, and finally December 3, 2007 with the loss amounting to 3%. The use of Multifractal Detrended Moving Average methodology in the same database, assures us for the validity of the multifractal results, and enables for a more persuasive explanation of the extreme events in Hong Kong market. Therefore, in this part we will present the empirical results of the MFDMA.

In a brief review for the MFDMA, Yudong and Chongfeng (2012) conclude that unlike previous studies, in a time period less than a year crude oil future markets are weakly persistent, while in more than a year long-term period crude oil future markets are strong anti-persistent. Furthermore, in the paper of Yudong et al. (2011), multifractality is found by using the MFDMA method in 12 USD exchange rates. Also they support that extreme events contribute in multifractality of the USD/EUR exchange rate series. Another point is that the extreme events might cause the negative fractal dimension of exchange rate series, and finally the multifractal degrees of the shuffled series are smaller compared to the real series. In another paper, Zhou et al. (2013) used the MFDMA and they found multifractality in three time series in CSI 300 spot return series. More specifically, after shuffling, surrogating and removing the extreme values from the data, they concluded that extreme events and fat distribution are the main causes of multifractality.

In Wang et al. (2014), they reported that the best performance has the

backward MFDMA (with θ=0). It is also mentioned that local generalized Hurst exponent using MFDMA that has been proposed could be used for texture image segmentation. Further application of their method could be used to "*target identification and location in remote sensing images, lesion area extraction and analysis in crop leaf images, and tumor recognition in medical images*". Nevertheless, according to Mali et al. (2016) there are some flaws in multifractality. In the case of multiparticle production data and since the signal length is very important, both MFDMA and MFDFA could not discriminate the statistical noise from the signal.

Firstly, we present the log-log plots of $f_q(s)$ fluctuation functions versus the scale s from all five HSI crashes and HSI property and financial index. The $f_q(s)$ functions are straight lines and the slopes change when going from positive moments to negative moments. The differences between the q-order RMS are larger for small segment sizes than for the larger ones. This is caused by the ability of the small segments to separate more effective the periods of large and small fluctuations, while the large segments average out those differences in magnitude. And as we go from small segments to large segments, multifractal time series become similar to monofractal (Ihlen, 2012). The scaling parameter s is $10 \leq s \leq N/10$, where N the total number of observations. In our case we have on average maximum 2640 observations on each crisis. In the Figure 10 below we can see the fluctuation functions $f_q(s)$ with respect to the scale n for various qs.

Multifractality is presented in the multifractal scaling function $\tau(q)$. In Figure 11 we can see the results for the real and the shuffled data in all 5 cases. We want to check whether we still have the same level of multifractality by removing the temporal distribution from our data. We use the same scale as for the $f_q(s)$ fluctuation function. The multifractal scaling function follows a linear trend for the negative values of q. When the values of q are positive we see that all five cases have an inclination, which is a sign of multifractality. One more important thing to notice is the fact that both shuffled and real data have a similar trend when $\tau(q)|_{q<0}$ and when $\tau(q)|_{q>0}$ the two series change direction; shuffled data has a linear trend, and even though this is the rule in the crises under research, the larger difference is revealed in 1997 and 2007 property.

In a similar way with the MFDFA, we convert q and t(q) to a and $f(\alpha)$ using a

Legendre transform as $f(\alpha) \equiv \alpha q - \tau(q)$, $\alpha \equiv \frac{d\tau(q)}{dq}$, where $f(\alpha)$ is the fractal dimension of the time series. The degree of multifractality on each crisis will be quantified by the difference of α_{max} and α_{min}. Also $f(\alpha)$ informs when events with α scaling exponents occur. As we aforementioned in the findings of Shimizu (2002) for the MFDFA method, when the left width is larger than the right one, it shows the dominance of the large fluctuations in the distribution. Furthermore, to see the level of multifractality on each crisis separately, we have to measure the $\Delta\alpha$, which is the difference of $\alpha_{max} - \alpha_{min}$. The results of the α and $f(\alpha)$ lie in different ranges; the 1994 case has $0.26 < \alpha < 0.88$ and $\Delta\alpha = 0.62$. In 1997 we have $0.16 < \alpha < 0.82$ and $\Delta\alpha = 0.66$. In the 2007 crisis, we show that $0.24 < \alpha < 0.8$ and $\Delta\alpha = 0.56$. The property sub-index has $0.25 < \alpha < 0.88$ and $\Delta\alpha = 0.63$. Finally the financial sub-index of the 2007 crisis has $0.21 < \alpha < 0.8$ and $\Delta\alpha = 0.59$.

The crisis of 1997 has the highest $\Delta\alpha$, or the richest multifractality, followed by the property crisis of 2007. Although the differences in multifractality are not that big compared to each other, we do not have the same richness rating as in the MFDFA method. MFDMA and MFDFA are different methods, and thus the results differ. In the MFDMA method, the findings support that the 1997 crisis and the property market crisis present the highest multifractality. They are followed by 1994, financial and 2007 crisis.

On the other hand, MFDFA method presents the crisis list in another sorting; the highest multifractality belongs to the 2007 crisis, followed by 1994, 1997, property, and financial crisis. It is important to note, that the width of the spectra for the 1994, 2007, and financial crisis remains wide, meaning that a contribution of a broad probability distribution in the observed multifractality cannot be excluded. So even though we shuffled the data in order to remove temporal correlations, there is still long memory in the data. For more details, see (Zhou W.-X. , 2009).

In the next step of the analysis of multifractality, we use the MFDMA to assess whether those crises affect the market complexity, and more specifically the market complexity before and after of each crisis. We divide our five cases in equal sized foreshock and aftershock periods, in order to check the reaction of the market participants. From our database, we take again on every foreshock and aftershock period on average 1320 observations respectively. In the following figures, we present the multifractal scaling function $\tau(q)$, and the multifractal spectra $f(\alpha)$.

We compare the foreshock period and aftershock period on every case separately. The richness in multifractality signifies a transition from a homogeneous to heterogeneous pattern. In the multifractal scaling function $\tau(q)$, we can see that except for the 1994 crisis, the aftershock period is larger than the foreshock period, giving a first sign of a multifractal and rich in complexity aftershock period. By looking in Figure 13 below we can say that for positive values of q, 1997 and financial aftershock periods have the larger difference in multifractality than the property and 2007 crises. Lastly, larger foreshock period of 1994 could possibly mean that the market's complexity has not changed after the extreme event.

Figure 14 presents the upside down parabolas of the foreshock and aftershock periods. Just like in $\tau(q)$ analysis, from a first look, aftershock periods show a higher range $\Delta\alpha$, except for 1994 crisis. One very interesting phenomenon is the knotting of the multifractal spectrum that should be explained. It is well developed in both foreshock and aftershock period of the financial crisis. The length of the time series is a very important factor, because finite signals are assumed to be by construction monofractals; nevertheless, they can reveal some artificial multifractal structure (Czarneski & Grech, 2009).

The reason is that extreme events tend to be more rare in finite time series than in infinite time series or very long-term time series. The author searches to find the influence of the series length to the width of multifractality. He supports the notion that in time series with "no visible distortion", we have "good" data that provide strong multifractal properties, and in time series with some extreme events like financial crashes, we have "bad" data that can result in poorer multifractal scaling $F_q(s) \sim s^{h(q)}$. This poor scaling is the actual reason of "twists" and knotting in singularity spectrum $f(\alpha)$. It is also confirmed that $h(q)$ function has non-monotonic character, and there is non-quadratic $f(\alpha)$ dependence because of the knotting presence. Table 2 presents the singularity spectrum of every crisis. It is found that the 2007 financial crisis displays the widest Δα, 0.10. In 1997 there is 0.05 multifractal spectra difference between aftershock and foreshock.

The 2007 and property crisis follow with $\Delta\alpha = 0.02$ and $\Delta\alpha = 0.01$ respectively. Just like in the MFDFA method the Δα for 1994 is negative. This indicates that the main shock during the crash was probably transitory, and thus did not change the complexity of the market in that specific period. Lastly, there is an

average Δα=0.01 for all cases, which means that the average aftershock singularity spectrum is broader by 0.01 than the foreshock spectrum.

2.3.3 Time dependent analysis

In this part we will analyze the temporal evolution of multifractality in time series. The a-exponent measures how persistent a given signal is (Grech & Mazur, 2004). For $\alpha = {}^{1}/_{2}$ we do not pay much attention since it is an evidence of the EMH. For $\alpha > {}^{1}/_{2}$ and $\alpha < {}^{1}/_{2}$ there is some long-range correlation, persistent and anti-persistent time series corresponding to fractional Brownian motion. During time series returns with long-range correlations and before a dramatic breakdown, the difference between α_{max} and α_{min} becomes larger, and the a-exponents increase rapidly. We use the concept of local a-exponent in order to measure the strength of local correlations. Thus we create a time-window, or observation box that is calculated in the period $< i - N + 1, j >$, where N is the length of the time-window. And since the whole time-window has to be covered with τ-sized boxes, the last box used in the DFA method is in the period $< i - N + 1, i - \left[\frac{N}{\tau}\right]\tau + 1 >$, where $[.]$ means the integer part. The choice of N is a matter of debate. There are many different theories about the golden point between number of observations and reliable results. For instance, Grech & Mazur (2004) support the idea of using no more observations than a trading year (N=240), in case of financial series with daily closure signal. Their actual ideal N lies between $190 < N < 230$. They also support that we can predict an extreme event only if the considered time-window contains clear-long lasting trends before the crash.

The same concept of a-exponents is presented also in Siokis (2017). The author presents the dynamic aspect of the stock market indices by using a sliding window of 1320 days, sufficient enough to ensure comparability, and the shift of the window is set to 22 trading days, which is around one business month. The temporal evolution of multifractality is presented by calculating the α_0 and the width between α_{max} and α_{min}. The results show a significant pattern change in the indices before and after crisis period. What is more, the measured width is broader during the extreme event. We apply firstly a separate sliding window algorithm in MATLAB in order to get the window shifts, and later we use an adjusted code, using the original MATLAB code from the paper of Gu and Zhou (2010). In Figure 15 we show the

variation of both α_0 (panel 1), which is the value where the spectrum has $f(\alpha)=1$, and width $W= \alpha_{max} - \alpha_{min}$ (panel 2), in other words the range of the spectrum.

The first thing that we notice is that both parameters under study have strong variability and they are characterized by dynamical changes during time. Also, we find characteristic changes of their pattern during periods under stress. For the 1994 event we do not have enough past data to show a clear pattern of dynamical change. This is not the case for the rest of the data. As far as the 1997 event is concerned, the parameter starts increasing from its lowest point in January 1998 which corresponds to the window of January 1993 to January 1998 and reaches its maximum levels in April 2000, (corresponding to the window of April 1995 to April 2000) and fluctuated around that level until May 2001.

In particular, the highest increase of the α_0 value is documented right after the transfer of sovereignty over Hong Kong from the United Kingdom to the People's Republic of China on July 1st 1997 and the effect of the Asian currency crisis. Panel 2 plots the time variation of the singularity spectrum width showing great variability during the two events. Clearly, based on the width value for the two events there is a gain of multifractality with the 1997 event depicting a greater degree of multifractality and for a longer time period. The width increases after June 1997, which corresponds to the window June 1992 to June 1997 and fluctuates in high levels until May 2001, which corresponds to the window May 1993 to May 2001. As for the 2007 crisis, the α_0 parameter starts to increase in October 2007, reaching its maximum in September 2008 and stays in high levels until October 2010. After October 2010, the trend is reversed and the α_0 value started decreasing.

Surprisingly, around July 2012 the width of the spectrum started increasing again for a short period of time. This increase could be attributed mainly to two events. Firstly, in June 2012, HKEX announced its cash offer to acquire the London Metal Exchange (LME), the world's premier metal exchange and the acquisition was completed in December 2012. Secondly, beginning in June 2012 HKEX started the process of creating an over the counter (OTC) Clearing House, offering OTC derivatives clearing services. This process finalized in November 2013.

Lastly in panel 3, the α_{max} and α_{min} are depicted along with the associated shuffled data. For the 1997 event, great variability is recorded both for the α_{max} and α_{min} values, indicating a change from homogeneous to heterogeneous dynamics. During that period the spectrum is ''richer'' in structure, with wider range of fractal

dimensions. As for the 2007 event, the α_{min} value drops sharply at the beginning of the crisis, prior to any movement made by the α_{max} value. A major point should be made about the shuffled series, both in the first and in the third panel, where we see that the shuffled series behave like a random walk with, and they are more or less close to 0.5-0.55.

In Figure 16 we present the time evolution of the two sub-indices. The analysis starts from February 1999 extending it to June 2016, in an attempt of capturing the dynamics of the indices. Same pattern is depicted for the 2007 event for both Properties and Finance indices. The α_0 parameter fluctuates between 0.48 and 0.55 until April 2008 where the parameter records the lowest value. From that point on, an increase of the α_0 and W parameters is observed in association with the main event of the crisis, indicating an enhancement of the multifractality degree. The maximum value is achieved just around August 2010.

As for the degree of multifractality (Panel 2), the width of the singularity spectrum appears to be time-varying and intense. In particular, the degree of multifractality is characterized by a clear increase just prior to the main shock, increasing thus the degree of market complexity. The most striking feature in the plot of width for the Finance Index is the significant and more pronounced change of the time pattern just prior to the outbreak of the crisis. In particular, the width shows a sudden increase just around August 2007, where the credit crunch in the US financial markets took place. It seems that the credit crunch had global implications, since the new financial products and the packaging of asset backed securities, composed of risky mortgages were sold to banks, international investors and to pension funds worldwide. Thus, the results support the view that a shock in the US financial markets is transmitted immediately to other developed financial markets.

Lastly, Panel 3 depicts the temporal variation of the α_{max} and α_{min} values, for both sub-indices. The value of the α_{min} for the Properties Index decreases around the end of 2007, while for the Finance there is a drop in the value again in August 2007. Therefore, the degree of multifractality exhibits strong time variations, which are associated with the dynamic evolution of the major shocks during the crisis periods.

2.3.4 MFDMA-MFDFA: A comparison

In the next step of our analysis, we compare the results from both MFDFA and MFDMA. The reason is to provide evidence of multifractality in the time series. In their paper, Ruan and Zhou (2011) show that both MFDFA and MFDMA unveil that their database, that is the 1 min intertrade durations possess multifractal nature. Even though they both support the existence of multifractality, it is shown that the two methods do not have similar multifractal spectra. In another study, Xi et al. (2016) compare the 2D-MFDFA and 2D-MFDMA methods. They search using 2D-MFDFA and 2D-MFDMA for the differences between those two methods, as well as how the parameter settings are influenced.

Apart from describing the differences in the calculation methodology, they also check how the adjustment of parameters influences the final result of the two methods. Their first finding was that all three ($\theta = 0, \theta = 0.5, \theta = 1$) 2D-MFDMA algorithms perform better than the 2D-MFDFA, having smaller root mean square values of the differences $\Delta\tau(q)$ and $\Delta H(q)$. The only exception was the 2D-MFDMA ($\theta = 1$). The second finding was that 2D-MFDMA($\theta = 0, \theta = 0.5$) has better performance since they are insensitive of the sample size. For the case of the same sample size and the smaller scale value, the shorter the runtime of the critical step of 2D-MFDMA. In 2D-MFDFA we have the exactly opposite results as far as runtime is concerned. But with the change of the scaling range, the runtime of 2D-MFDMA becomes longer and the runtime of 2D-MFDMA becomes shorter (Xi, Zhang, Xiong, & Zhao, 2016).

Finally, in their paper about multifractality, Gu and Zhou (2010) compared all three cases of MFDMA with different parameter theta ($\theta = 0, \theta = 0.5, \theta = 1$), and the MFDFA too. The results showed that the MFDMA with theta=0 outperformed all the other three, exhibits the best performance with more accurate estimates for multifractal analysis.

Figure 17 presents MFDFA and MFDMA spectra results of every index in the same plot. Both MFDMA and MFDFA display a monotonic upside parabola showing in all cases the existence of multifractality. Apart from 2007 crisis where MFDFA shows a larger $\Delta\alpha$ than the MFDMA, all the other cases result the opposite. The exact multifractal spectral differences are shown in **Error! Reference source not found.** below. Since we compare the MFDFA with the MFDMA($\theta = 0$), according to the

comparative studies it is expected to find more multifractality in the latter case. That said, even though we can find similarities between the findings, we are still comparing two different methods, and we expect the MFDMA to present more reliable results.

2.4 Conclusion

We investigated the behavior the HSI stock market index and its constituencies during the crises of 1994, 1997 and 2007, and also before and after the 1997, and 2007 U.S. financial crisis. Based on statistical mechanics theory we calculate the multifractality using MFDFA and MFDMA algorithms for a wide time range of the anxious times, as well as before and after the outbreak periods of the crisis. In the case of MFDFA, the HSI for the 2007 and 1994 crisis show the highest level of multifractality. And when we divide our five cases in equal sized foreshock and aftershock periods to assess whether an extreme event has actually affected multifractality, we see that 1997 and 2007 have the highest difference in multifractality between before and after crisis. In addition we searched for the correlation with the generalized Hurst exponent. Apart from the financial sub-index that was found to follow a pure "Markovian" behavior with H(q)=0.5, the rest were found to follow a persistent or anti-persistent path, yet closed to 0.5.

In the MFDMA method, the findings support that the 1997 crisis and the property market crisis present the highest multifractality level. In the foreshock and aftershock division, we found that 1997 and financial aftershock periods have the larger difference in multifractality. Finally, using time dependent analysis we show the variation of both α_0, and width $W = \alpha_{max} - \alpha_{min}$. For the HSI index, both α_0 and width W fluctuate during 1997 and 2007, while the width increases too. As for the constituencies, they fluctuate during the 2007 crisis, and the most noticeable feature in the plot of width for the Finance Index is the significant and more pronounced change of the time pattern just prior to the outbreak of the crisis.

In a nutshell, the findings support that even though multifractality was found in different levels in the presented time series and the indices reacted to the extreme events, we are not able to reject the first hypothesis that both MFDFA and MFDMA could be used interchangeably. The findings are not the same because the methods differ in the process, meaning that the creation steps of MFDFA and MFDMA might affect the verdict of the indices and so the outcome of each case.

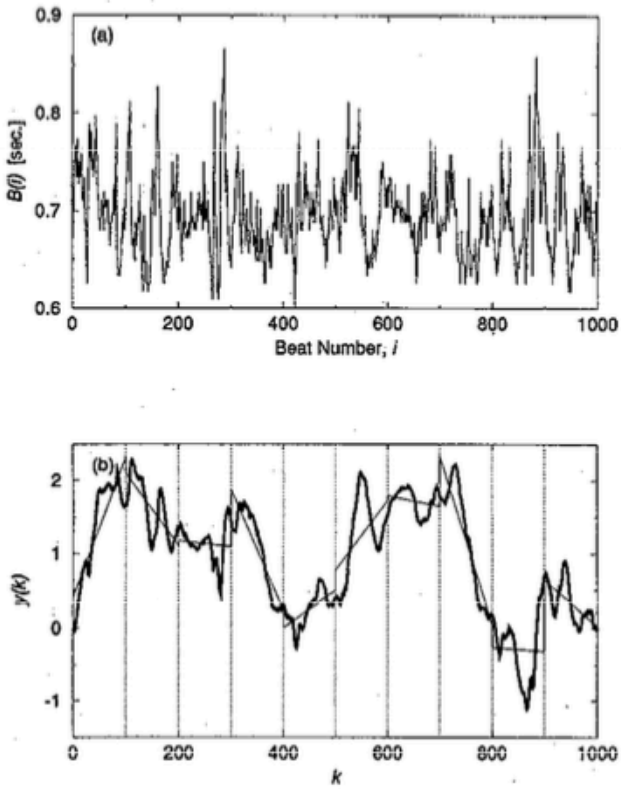


Figure 1: (a) The interbeat interval time series B(i) of 1000 beats. (b) The integrated time series: $y(k) = \sum_{i=1}^{k}[B(i) - B_{ave}]$, where B(i) is the interbeat interval shown in (a). The vertical dotted lines indicate box of size n=100, the solid straight line segments represent the "trend" estimated in each box by a least-squares fit.

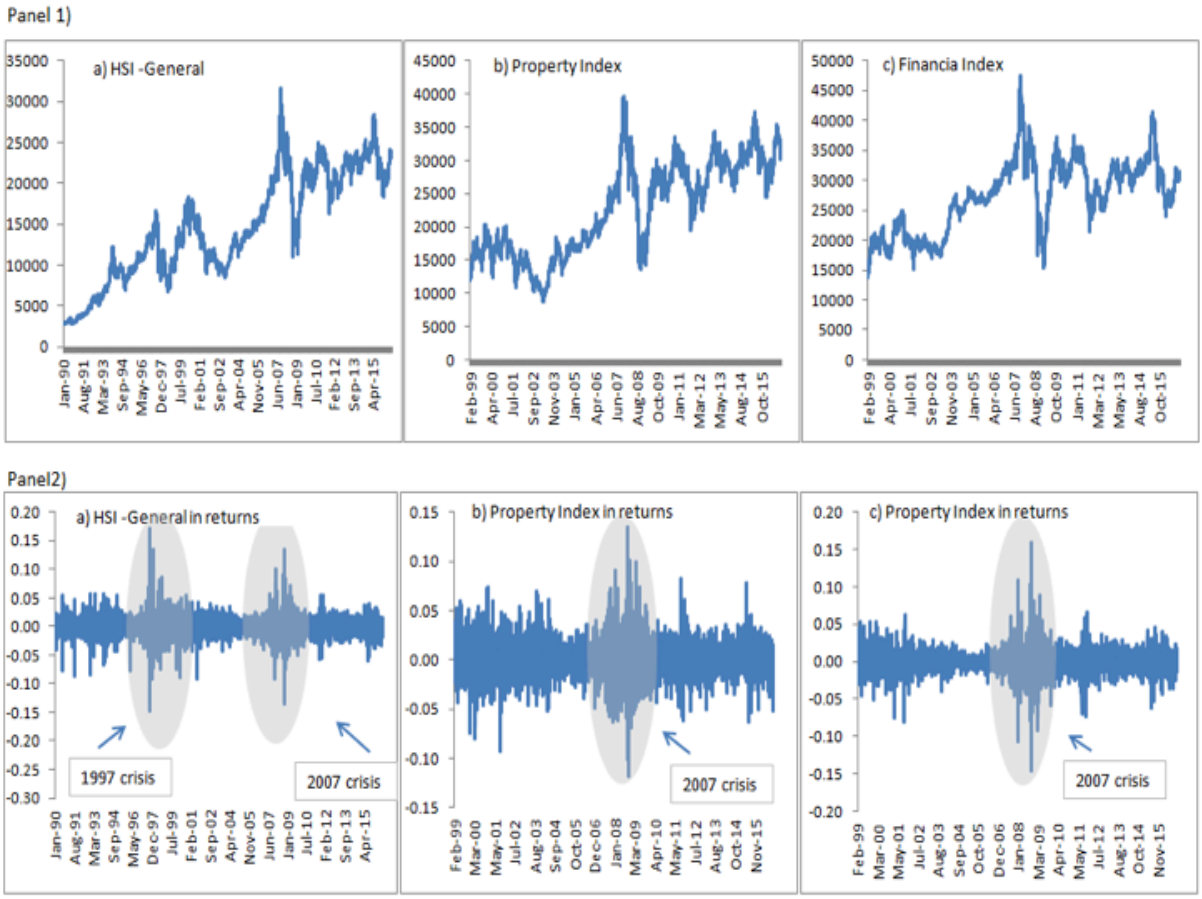


Figure 2: The Hang Seng Index and two constituents, the Property and Financial Indices. In Panel 1, the indices are in levels and in Panel 2, in returns. Shaded areas depict the crisis.

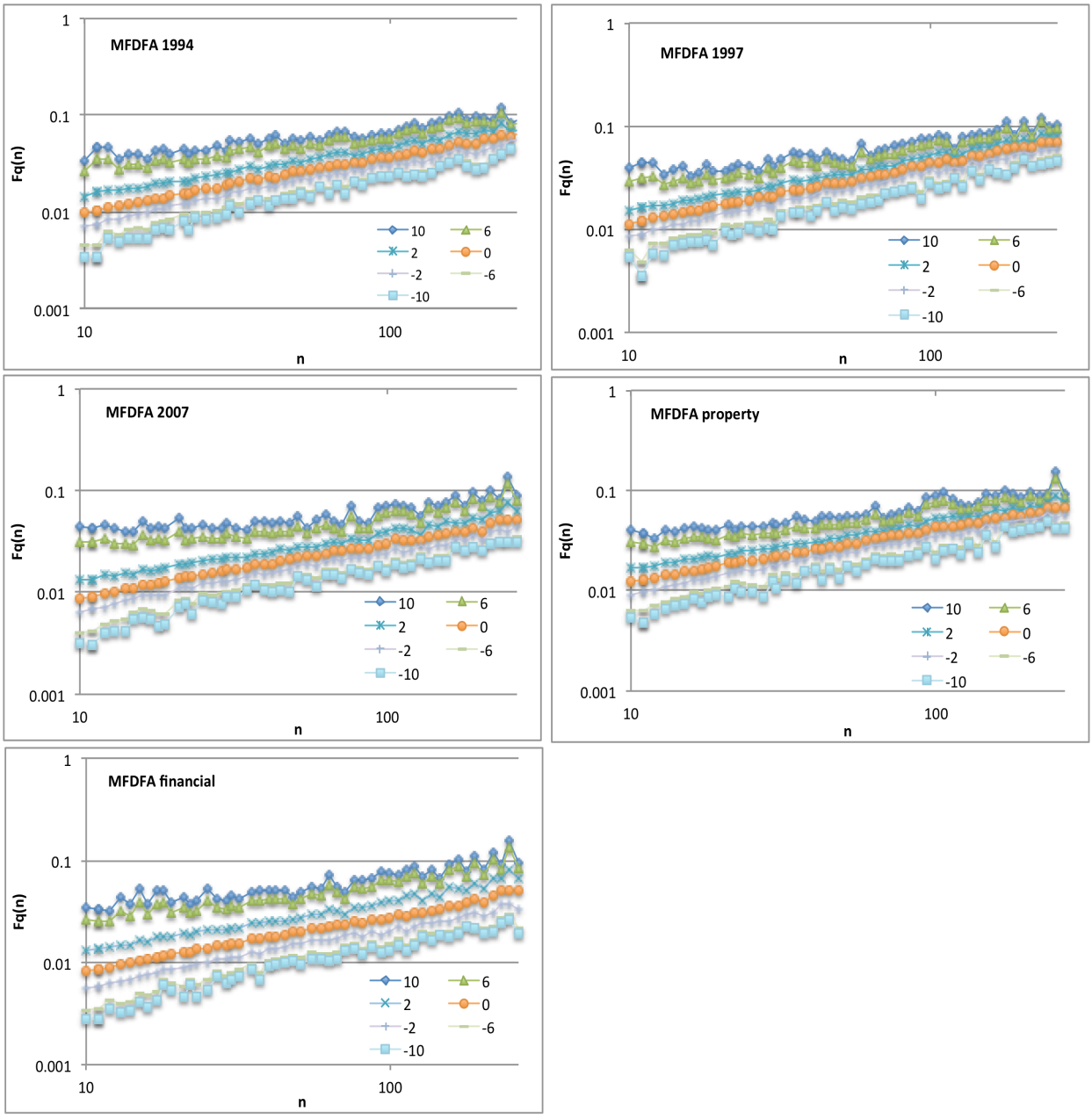


Figure 3: The MFDFA functions Fq(n) versus the time scale n in log-log plot

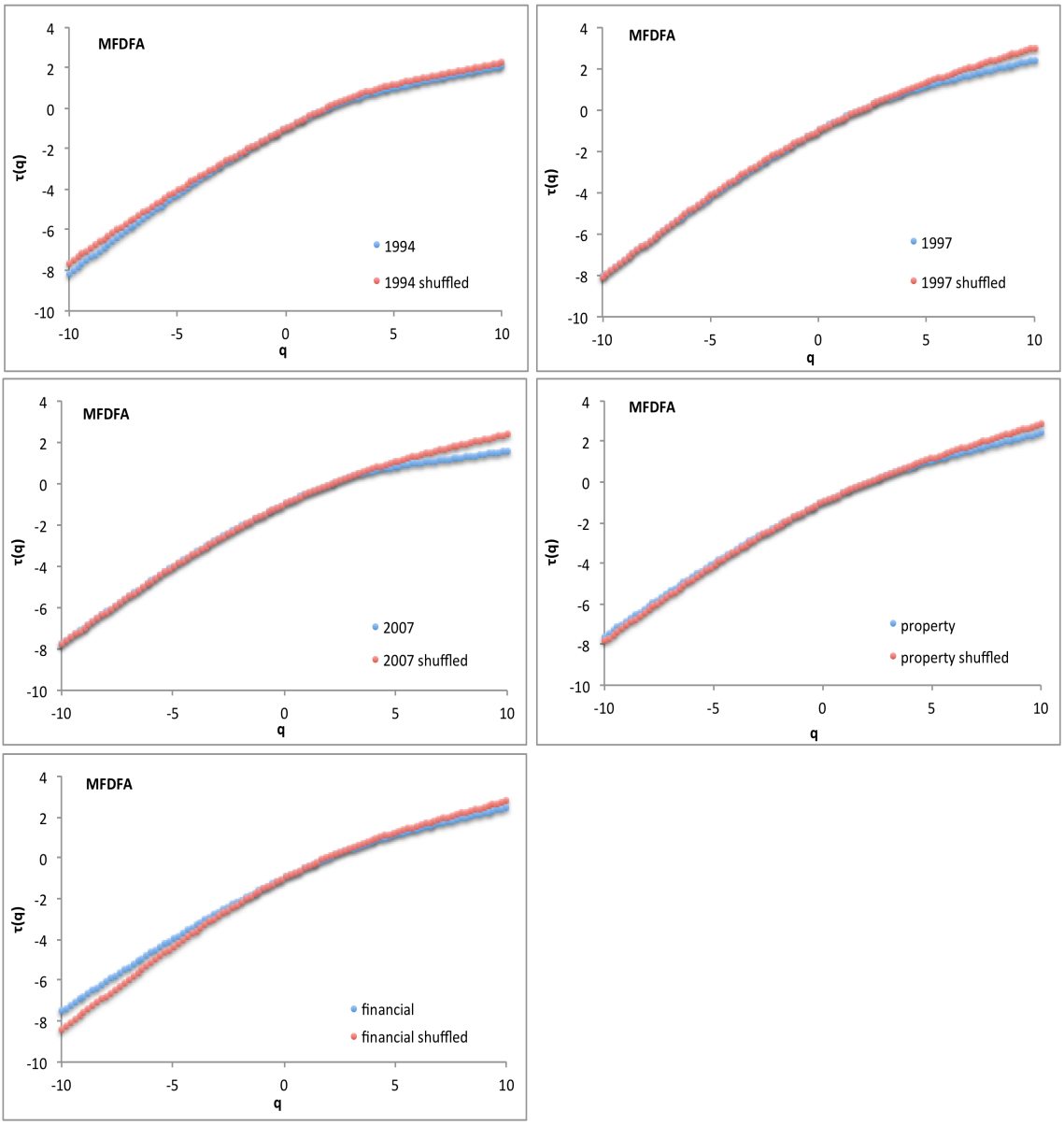


Figure 4: Multifractal scaling exponents for both actual and shuffled data

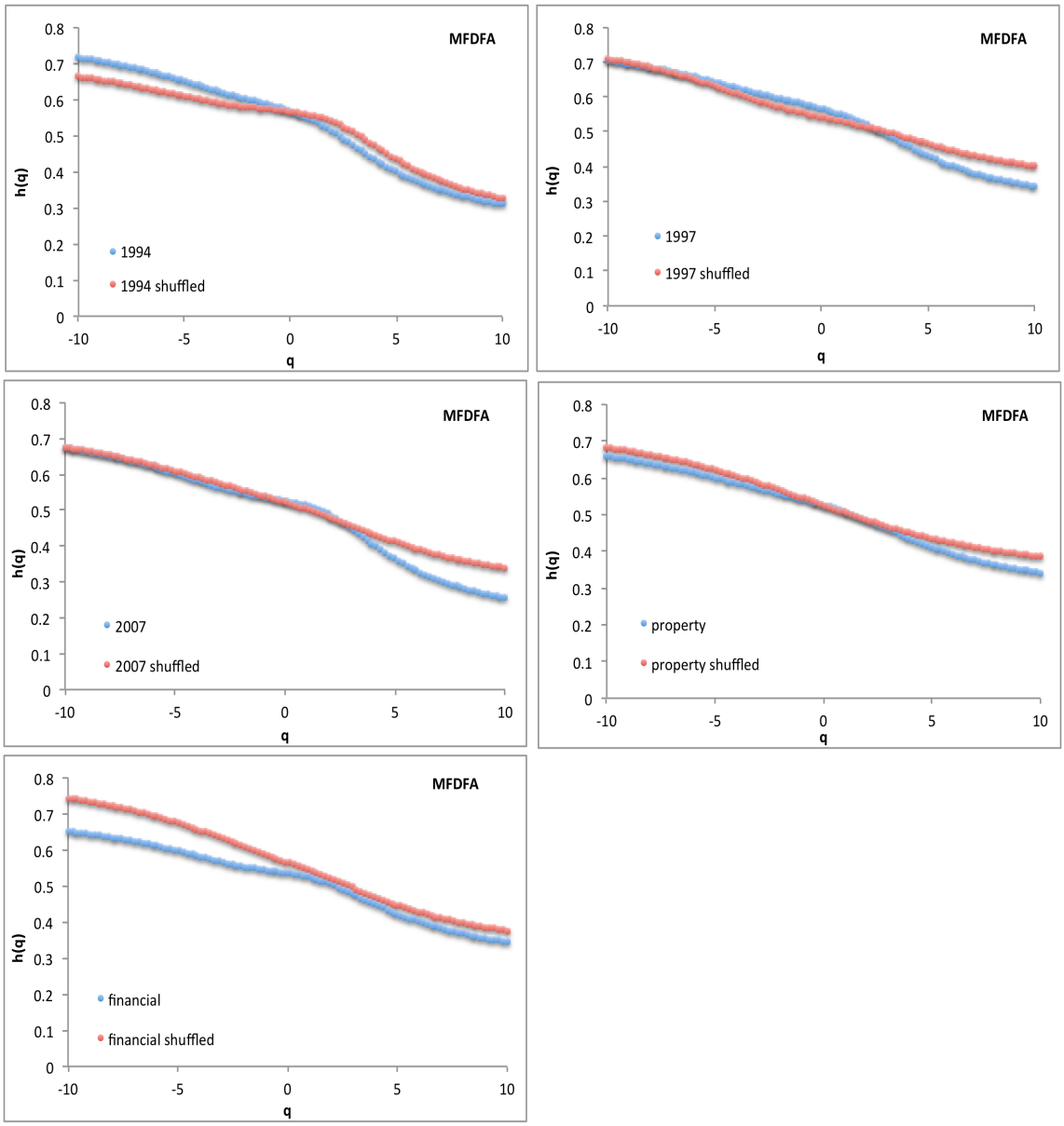


Figure 5: the Hurst exponent of every crisis and their shuffled series

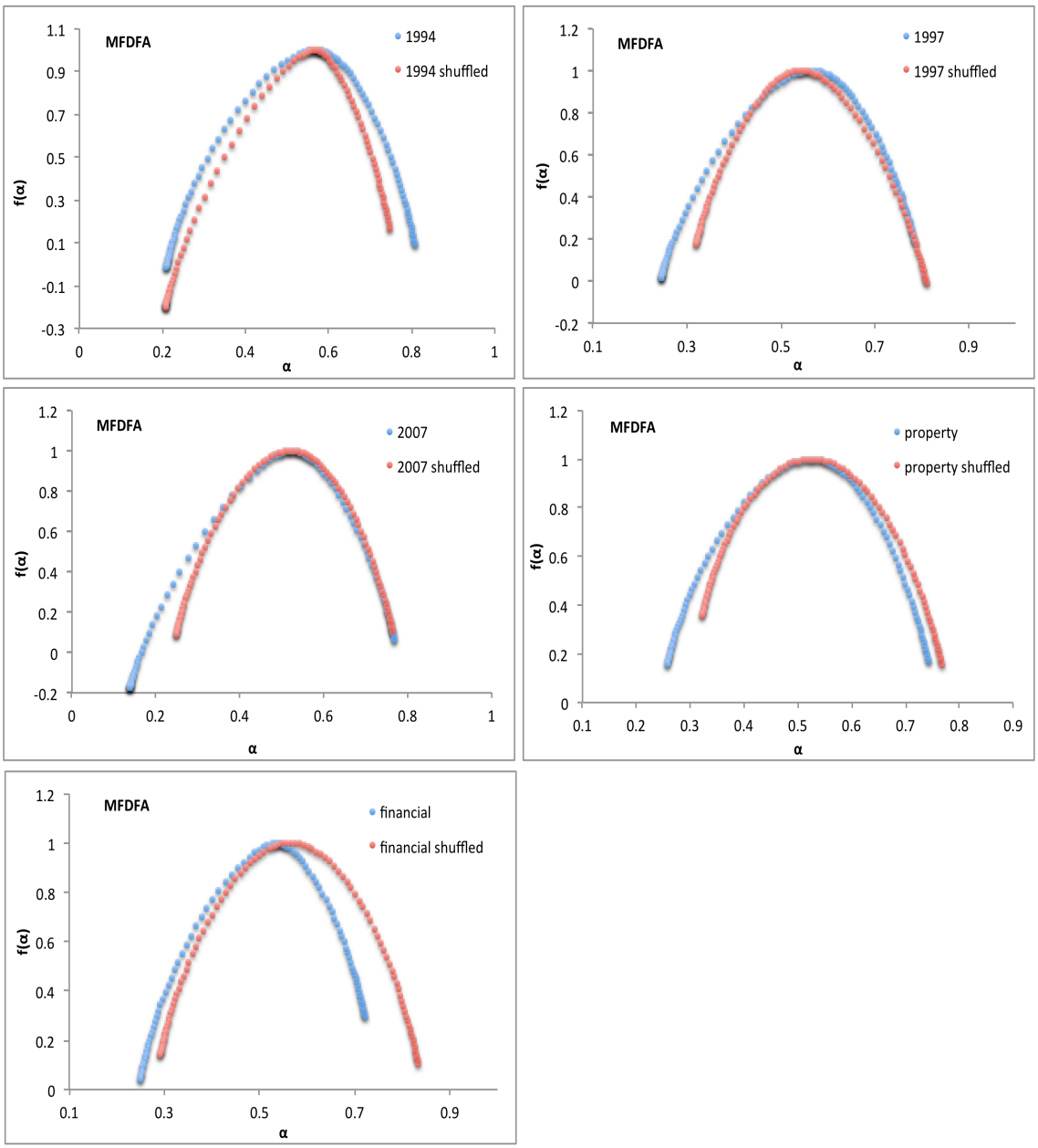


Figure 6: Fractal dimensions for both actual and shuffled data

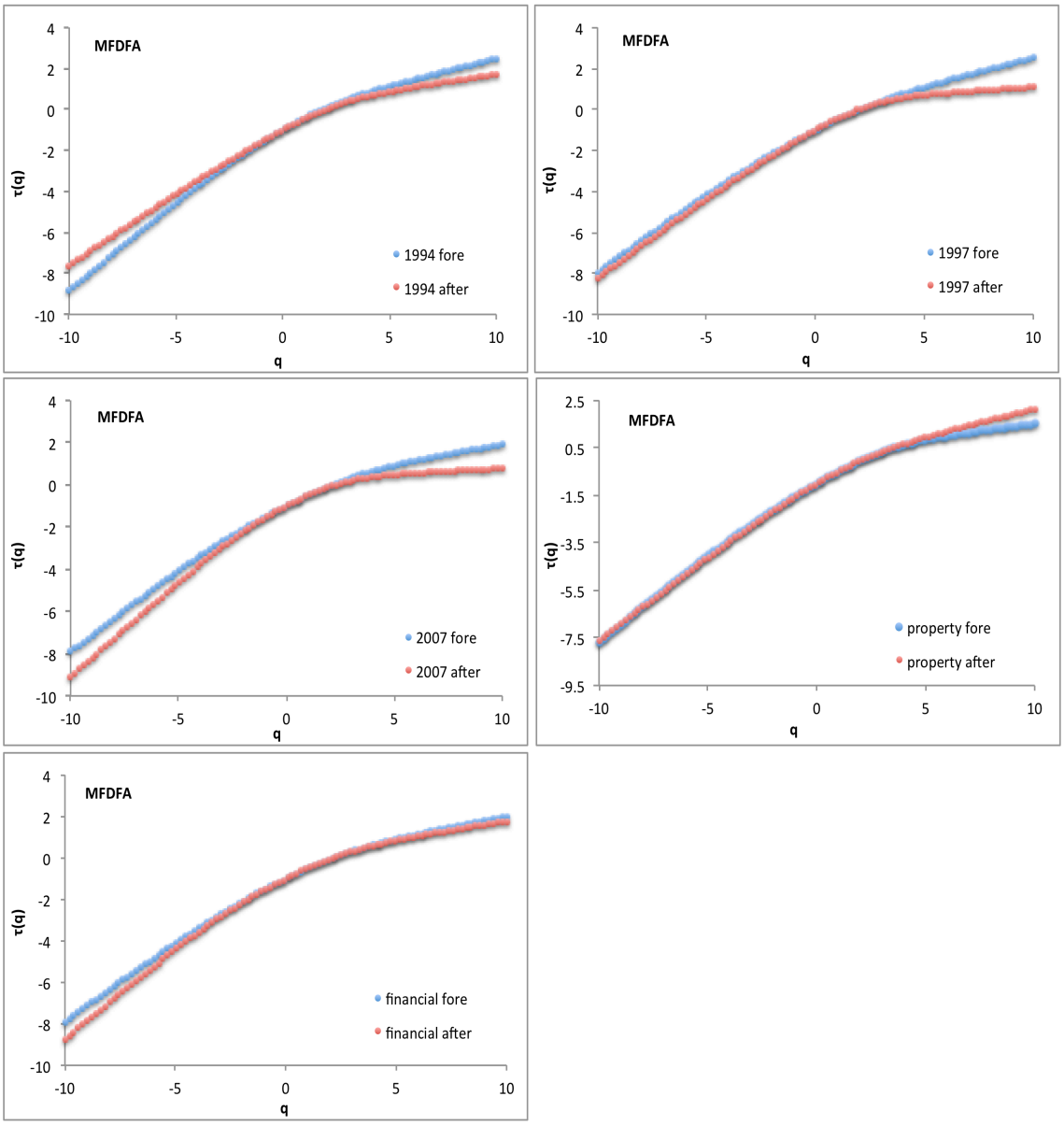


Figure 7: Multifractal scaling exponents prior and after every crisis

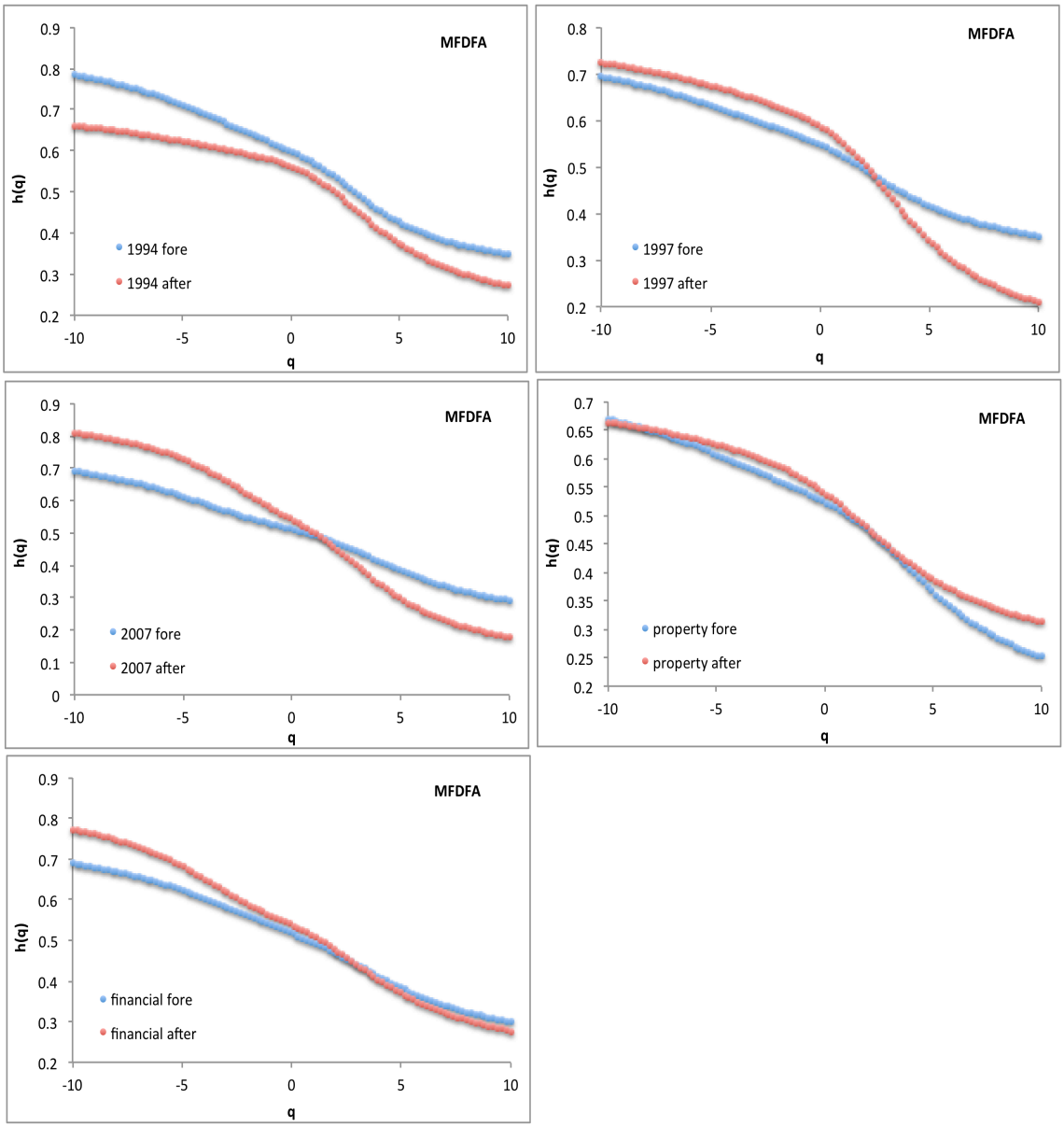


Figure 8: the Hurst exponent prior and after every crisis

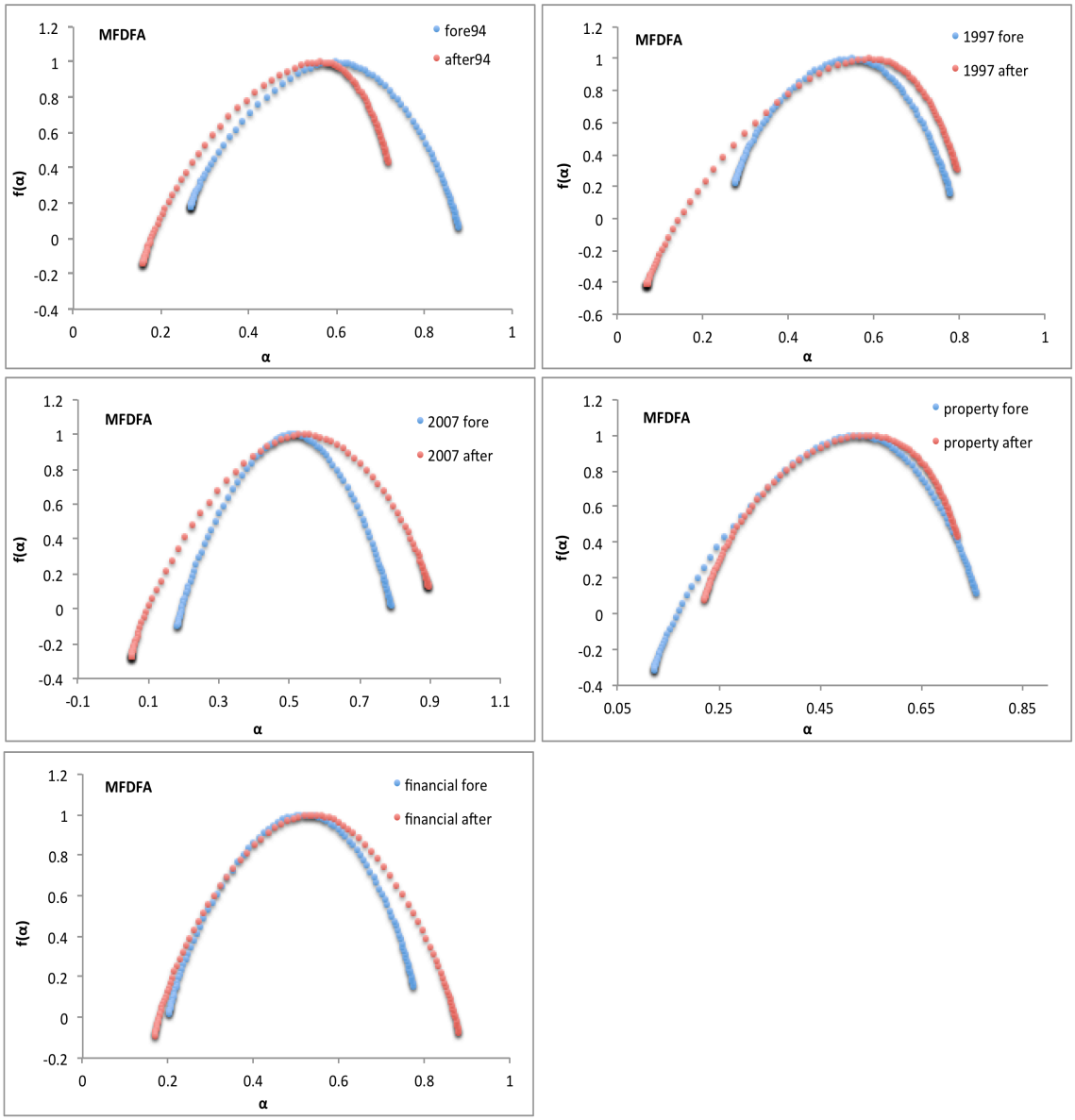


Figure 9: Fractal dimensions prior and after every crisis

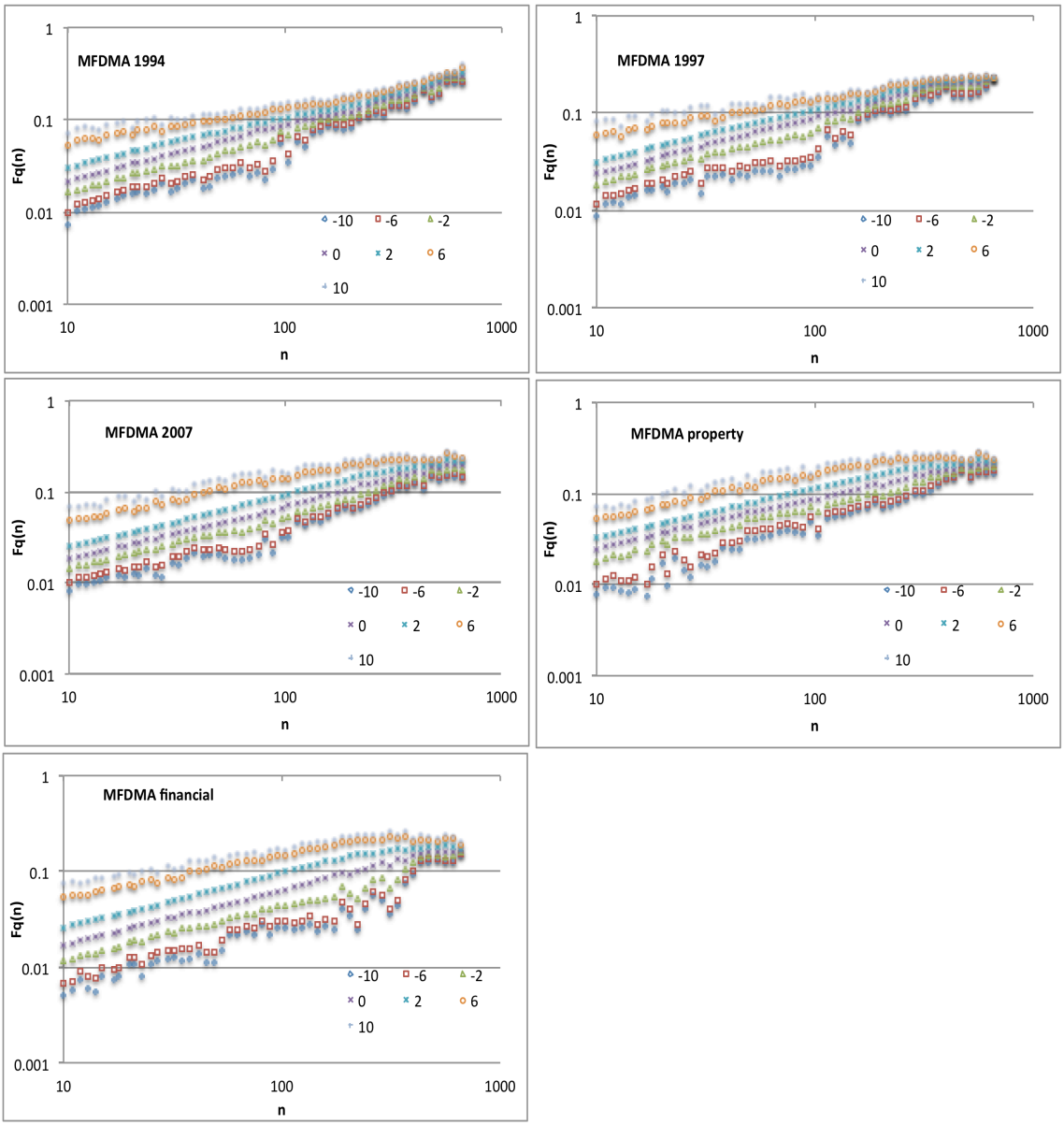


Figure 10: The MF-DMA functions Fq(n) versus the time scale n in log-log plot

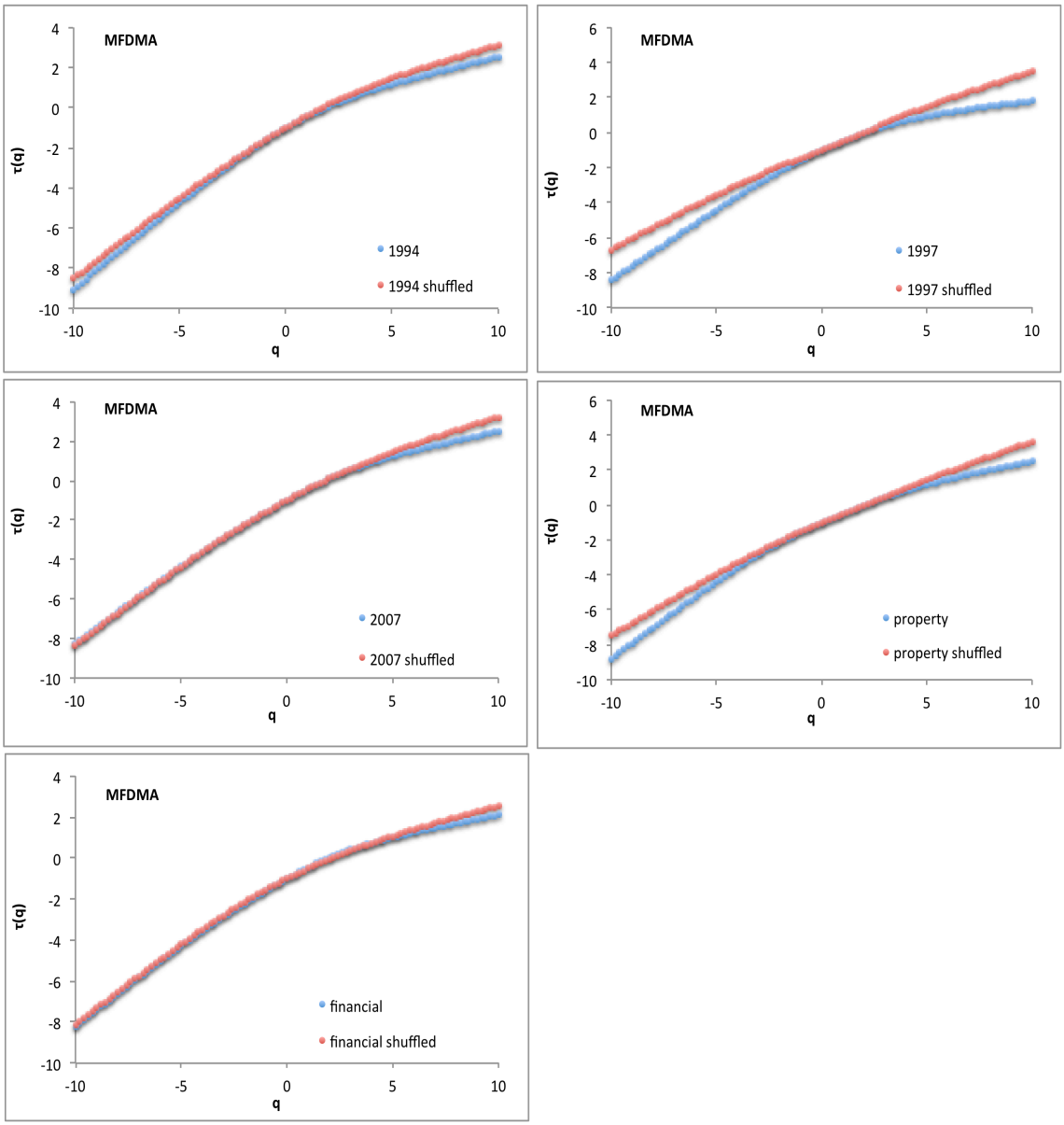


Figure 11: Multifractal scaling exponents for both actual and shuffled data

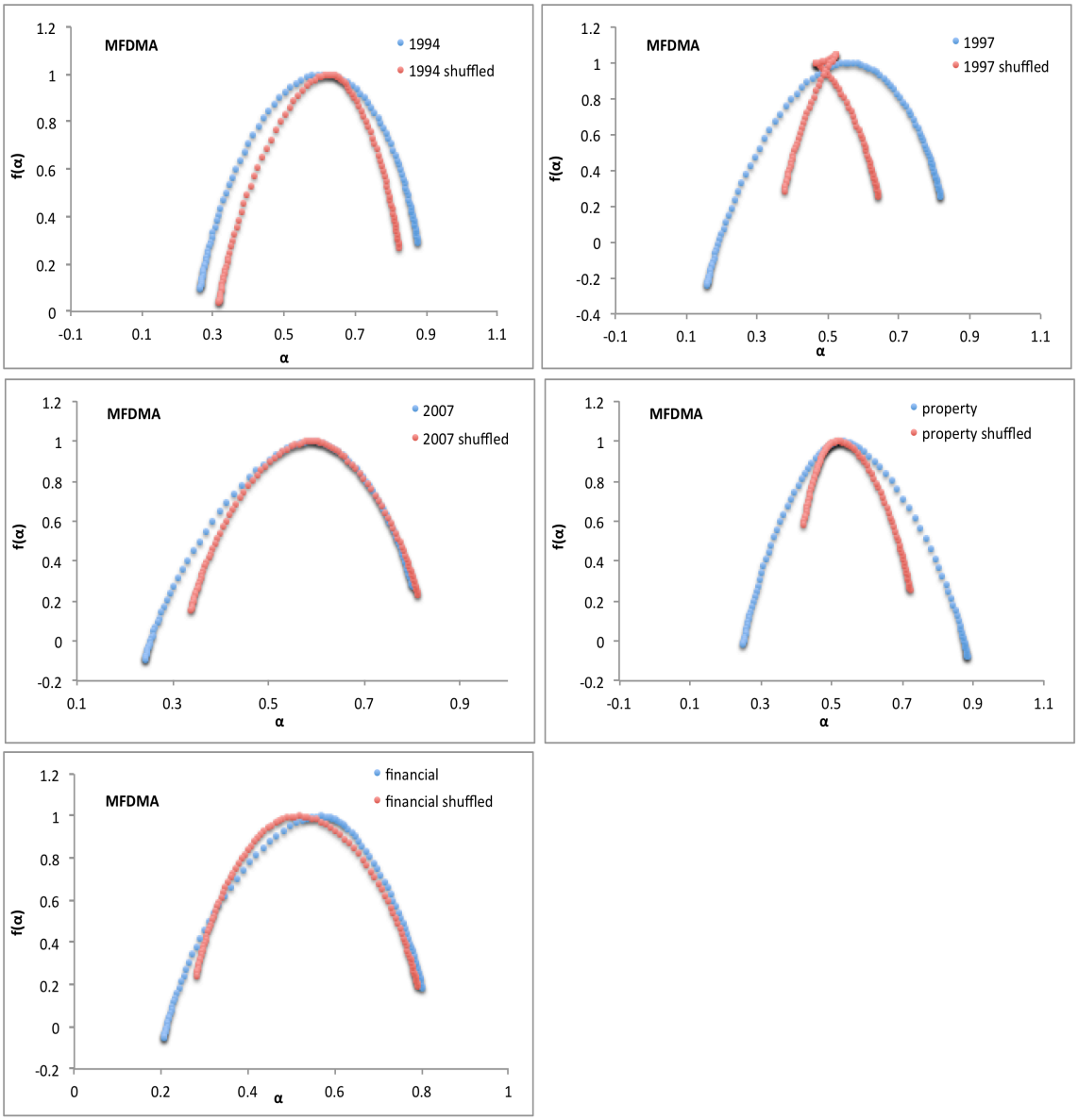


Figure 12: Fractal dimensions for both actual and shuffled data

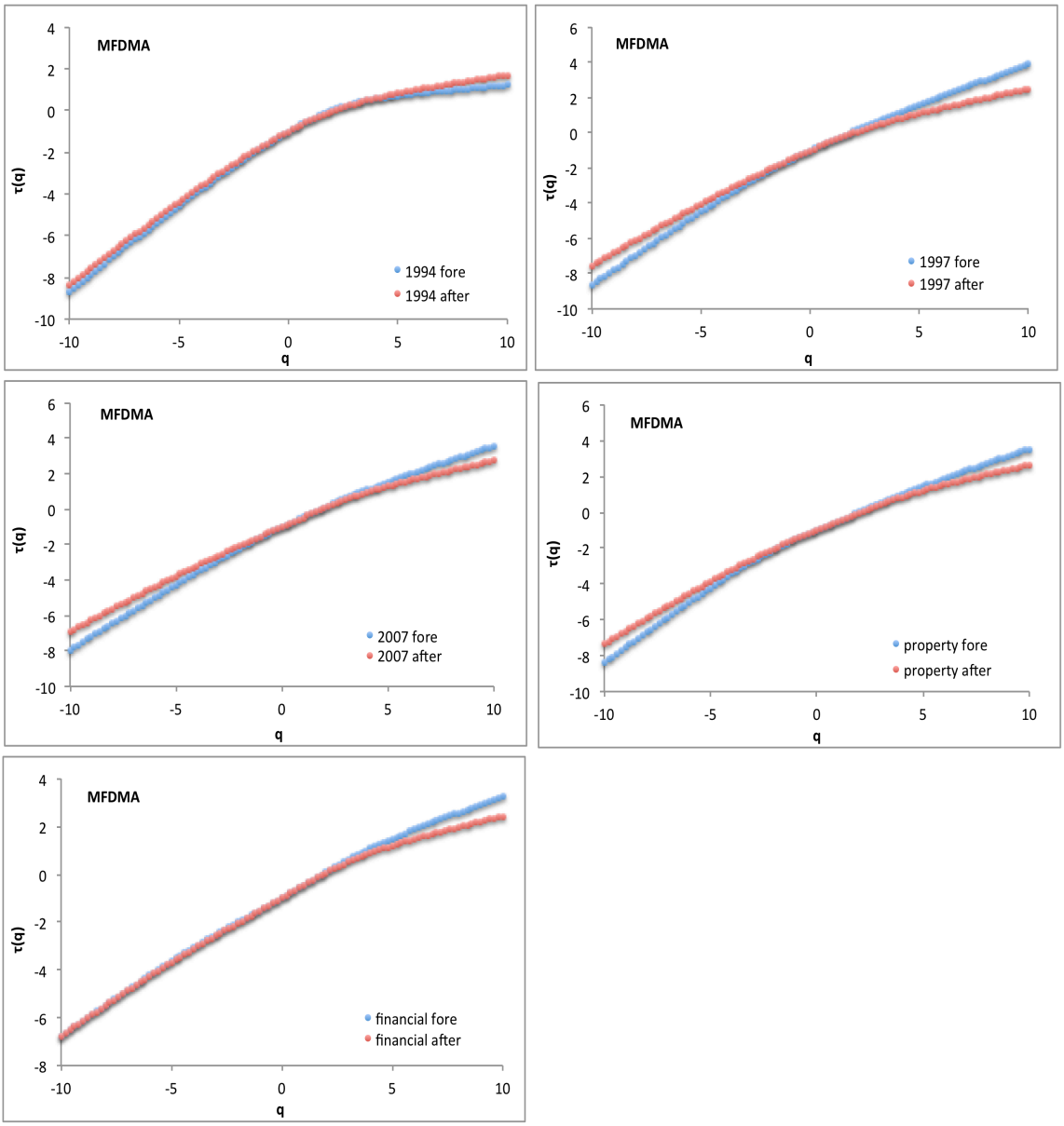


Figure 13: Multifractal scaling exponents prior and after every crisis

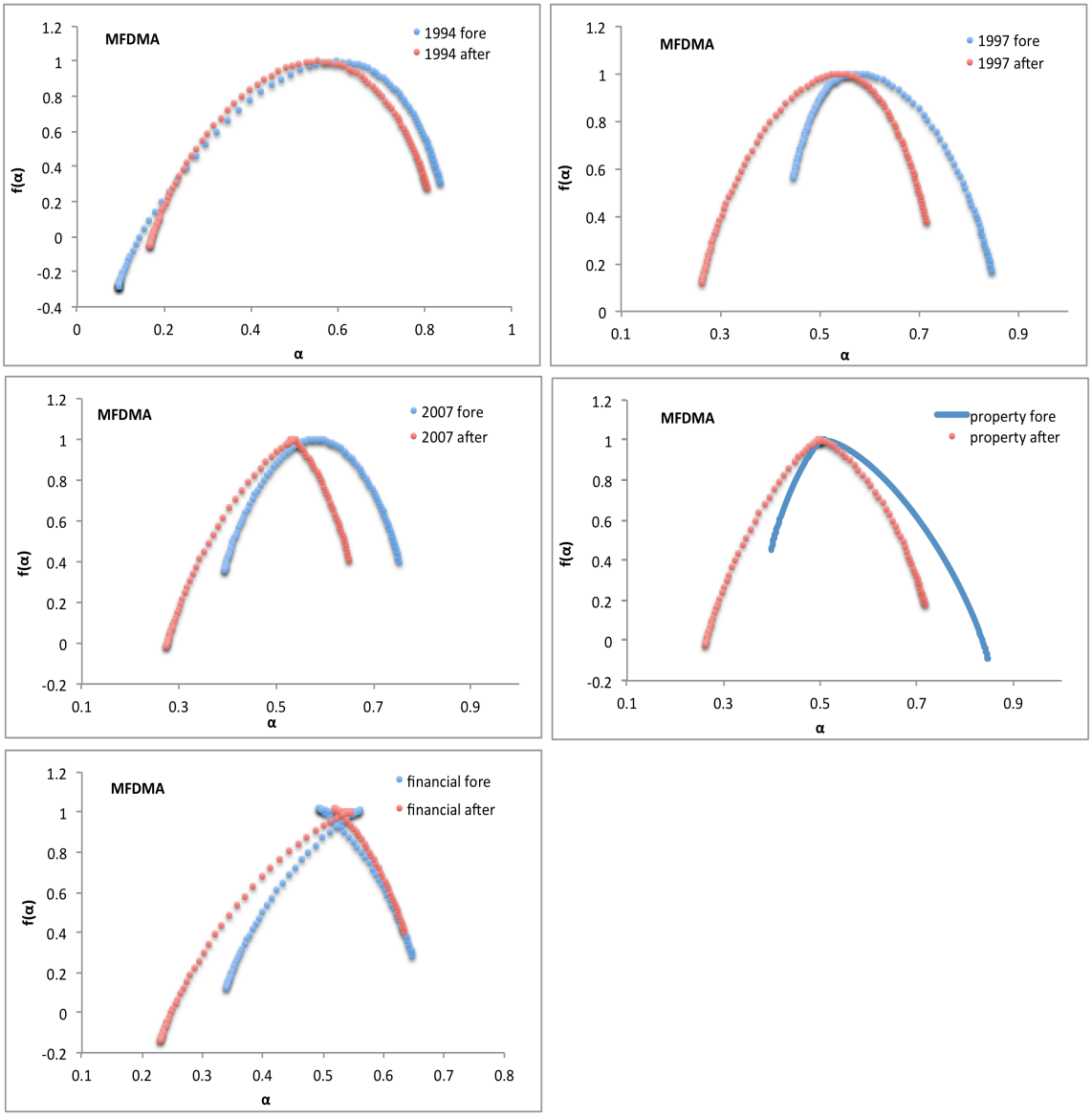
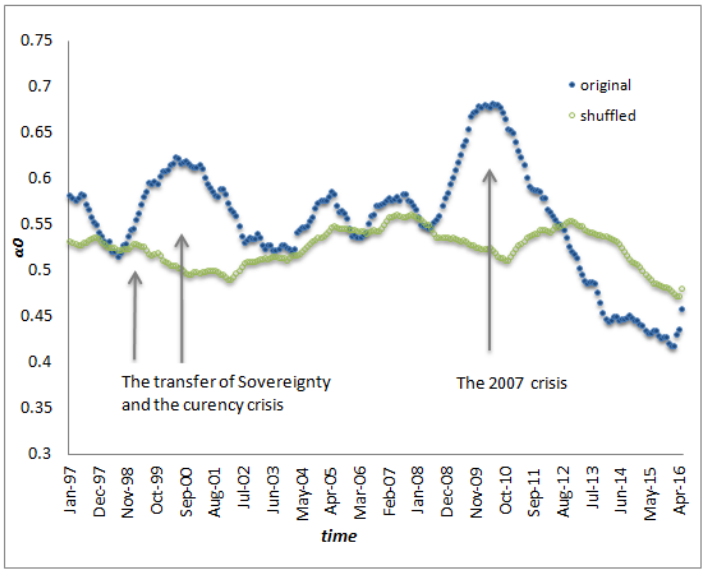
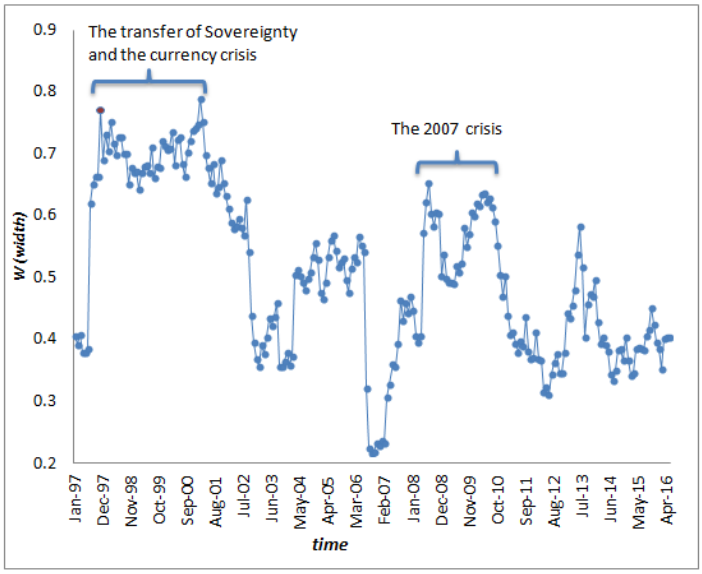


Figure 14: Fractal dimensions prior and after every crisis

I)


original
shuffled
α0
The transfer of Sovereignty and the curency crisis
The 2007 crisis
time

II)


The transfer of Sovereignty and the currency crisis
The 2007 crisis
W (width)
time

III)

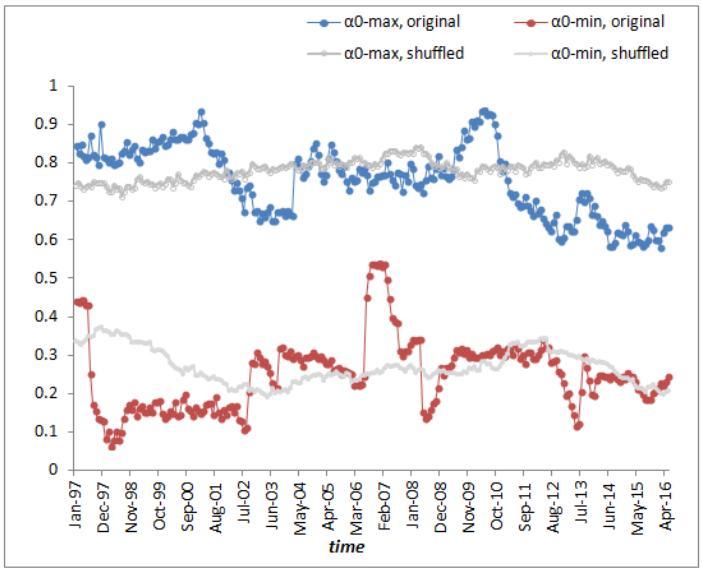

α0-max, original
α0-min, original
α0-max, shuffled
α0-min, shuffled
time

Figure 15: Time variation of the Hang Seng index for the parameter α_0, the width W and the values of α_{max} and α_{min}. The data span is from Jan. 1992 to Jun. 2016.

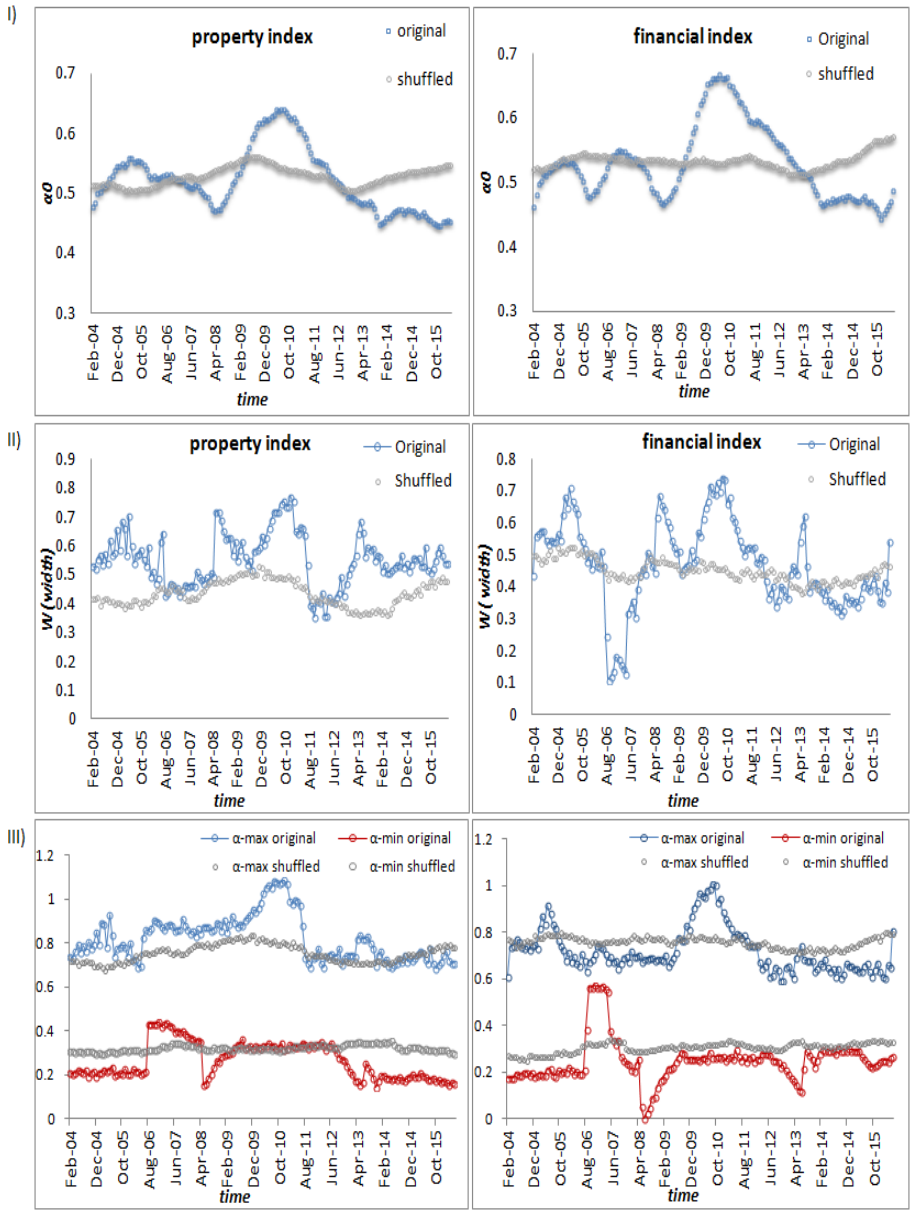


Figure 16: Time variation of the Property and Financial indices for the parameter α_0 (singularity spectrum maximum), the width W and the values of α_{max} and α_{min}. The available data span is from Feb. 1999 to Jn. 2016

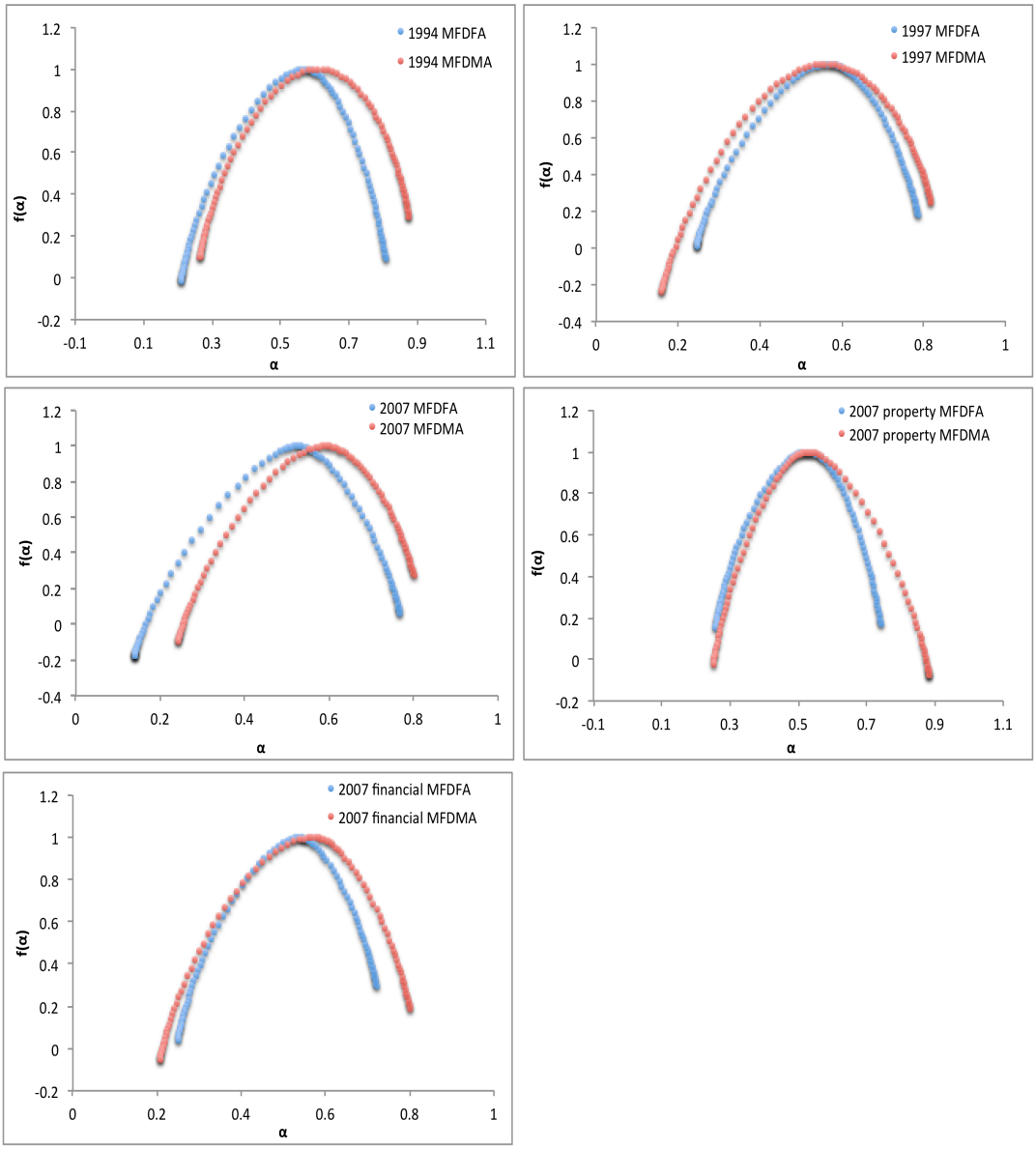


Figure 17: Fractal dimensions for the comparative study of MFDFA and MFDMA

Period	Mean	Std deviation	Skewness	Kurtosis
1997 crisis	0.00017	0.0182	0.048	11.1134
2007 crisis	0.000301	0.0158	0.0438	12.332
property crisis	0.000325	0.0194	0.2089	7.0617
finanial crisis	0.000181	0.0162	0.0838	15.5569
1997 prior	0.000633	0.0149	-0.3689	6.1309
1997 after	-0.000293	0.021	0.2276	11.0598
2007 prior	0.000767	0.0111	-0.0491	5.4757
2007 after	-0.000166	0.0194	0.1052	10.3218
property prior	0.00083	0.0159	0.2883	5.1792
property after	-0.00018	0.0224	0.2029	6.6998
financial prior	0.00058	0.0093	0.0349	7.0785
financial after	-0.000219	0.0209	0.1207	10.9191

Table 1: Mean, standard deviation, skewness and kurtosis, presented for 1997 general, and 2007 general, property and financial crises. All events are divided in pre crisis and post crisis period

Event	(Δα)after - (Δα)fore
1994	-0.05
1997	0.22
2007	0.24
2007(property)	-0.14
2007(financial)	0.14
avg.crash	0.082

Table 2: MFDFA singularity spectrum of every crisis

The generalized Hurst exponent (q=2)			
Event	Total sample	Foreshock	Aftershock
1994	0.51	0.53	0.49
1997	0.52	0.49	0.5
2007	0.48	0.47	0.45
Property	0.48	0.47	0.47
Financial	0.5	0.46	0.47
Average	0.5	0.49	0.48

Table 3: Generalized Hurst exponent for all crises, including the period before and after the shock

Event	(Δα)after - (Δα)fore
1994	-0.1
1997	0.05
2007	0.02
2007(property)	0.01
2007(financial)	0.1
avg.crash	0.01

Table 4: MFDMA singularity spectrum of every crisis

Event	Δα(MFDFA-MFDMA)
1994	-0.013
1997	-0.118
2007	0.348
2007(property)	-0.149
2007(financial)	-0.123

Table 5: Singularity differences between MFDFA and MFDMA

CHAPTER 3

3.1 Introduction

In the second essay we focus in the efficiency of Real Estate market and the turbulence period of 2007, which might reveal useful information about the housing market indices of Hong Kong. We scrutinize the special characteristics and the changes that took place during the crisis, so we could be able to react better during the next potential crisis, and might also give the policymakers and the regulators some extra knowledge to control or deflate easier such a life-changing event.

The second part contains the presentation of permutation entropy where Shannon, Tsallis entropy and the complexity entropy causality plane (CECP) are under consideration. We look forward to answer on the hypothesis that Permutation Entropy does not change before and after the Great Recession. Any other discussion addressing some other issue will be peripheral at most.

The 2007-2008 Great Recession was a turbulent period that caused an international turmoil. The substantial growth in mortgage credit and consequently in housing prices triggered an increase in mortgage delinquencies, as interest rates continued their upward trend. The burst of the housing bubble and the meltdown of the associated prices, across the country created a chain of reaction to other financial markets, both domestically and internationally. According to NBER it started in the US on December 2007 until June 2009 and was triggered by the subprime mortgage market crisis in the country, causing a global financial contagion. Asset prices across the globe fell drastically, while financial volatility rose substantially. Among the markets, where prices plummeted considerably, were the housing markets and particularly those that historically exhibit high volatility. The housing market of Hong Kong attracts our interest for further analysis because of the reasons above.

Once again, the celebrated Efficient Market Hypothesis (EMH) is under the shadow of doubt by many participants and analysts. Fact of the matter is that the Real Estate market could reveal its complexity during some anxious time, in another global economy too. The linked exchange rate regime, and the economic correlation dynamics, could lead the US happenstances to permute the economy of Hong Kong SAR as well. In this paper, Hong Kong's Real Estate market is under analysis to

assess any change in efficiency in the housing market of Hong Kong as a result of the 2007 financial crisis.

The Real Estate sector, and particularly the housing market have been much less studied compared to other economic assets, and little is known as far as the efficiency of Real Estate markets is concerned, especially after the financial contagion between two global economies. It is important to fill this gap in the literature. The reasons behind the excessive speed of the market contamination during the Great Recession could be the technology progress and the intertwined relations among world economies.

Bond et al. (2006) found existence of contagion among Australia, Hong Kong, Singapore, Japan and the US. Semi-positive signs of contagion were found in Wilson and Zurbruegg (2004), where little evidence of contagion from the 1997 Asian crisis was showed from the Thailand Real Estate market to Australia, Malaysia, Hong Kong and Singapore. Nevertheless, this is not always the case. In the paper of Hatemi and Roca (2011) it is shown the lack of contagion between the Real Estate market of Australia, Japan and UK. In the same line of thought, Fry et al. (2010) showed no contagion between the Real Estate markets of US, the UK, Australia, Japan, Germany and Hong Kong.

Moreover, an important discourse about the Real Estate market is the theory of Efficient Market Hypothesis. Under the farm of EMH, efficient prices reflect all available information, and follow a rather "Random walk" where it is impossible to make a profit systematically, or above average return (Fama, 1970) (Fama, 1991). The main dogma of the proponents of EMH is the ability of the stock market to be memory-free. According to the EMH discourse, the market is able to absorb the information of yesterday's price change, thus no influence in today's price change is present. The famous example in Malkiel (1973) of a chimpanzee throwing darts to select a portfolio that would do as well as the portfolio of an expert displays the notion about the EMH.

Nevertheless, the assumption is made that prices are random walks was set into question. According to the empirical evidence, levels of inefficiency can be found even to the most competitive markets. Shiller (2015) questioned why an efficient stock market should be that volatile since the actual price should have been the same as the optimal forecast. After the dot-com bubble at the end of the twentieth century, a new trend that invalidates the axioms of the EMH especially during

nervous moments has raised. In this case, many reports found evidence of pricing irregularities during short term. After all, as Grossman and Stiglitz (1980) noted, prices cannot reflect all available information, because if they did, there would not have been any incentives for the informed traders to spent their resources to receive compensation.

The Efficient Market Hypothesis theory was unconfirmed in the findings of numerous articles, and its validity is kept into controversy. For example, Barkoulas and Baum (1996) found some evidence of long memory in their analysis. Volatility clustering and long memory was also a study of Bentes et al. (2008). Ito and Sugiyama (2009) showed that the market efficiency in US stock market fluctuates through time; it was more inefficient during the late 80s, and became efficient around 2000.

Moreover Di Matteo et al. (2005) found in the results of their study an inclination from pure Brownian motion, and more specifically the 3 months maturity that is strongly influenced from the decisions of central bank. In Eom et al. (2008) the authors found a positive relationship between the predictability of the hit rate and the degree of efficiency, concluding that the prediction of future price changes is feasible by using the Hurst exponent as a measurement. Bariviera et al. (2012) studied the long varying behavior of sovereign and corporate bond markets of seven EU countries. Using the DFA method and the sliding window of four years, they found evidence of different memory dynamics in both bond indices after the Great Recession.

Against the theory of randomness by (Bachelier, 1964), Zhang (1999) shows a new approach of the conditional entropy, which measures possible profit pockets that can be revealed from the inefficiency margin in market dynamics. Zunino et al. (2009) in an attempt to discriminate market dynamics resulted that developed stock markets have lower number of forbidden patterns and higher normalized permutation entropy than the emerging ones. In Zunino et al. (2010), an innovative statistical tool, the complexity-entropy causality plane is used to classify the stage of stock market development. Developed stock markets show higher entropy, and lower permutation complexity, while the emergent stock markets behave in the opposite way, showing signs of time correlation and inefficiency.

The commodity markets are under analysis in Zunino et al. (2011) in order to find periods of time where signs of long range dependency are present and the

underlying dynamics is predictable. The complexity-entropy plane is also found in Rosso et al. (2012), Rosso et al. (2007). Furthermore, Zunino et al. (2012) presents a complexity-entropy causality plane for 30 bond indices, a non-exhaustively studied sector. The findings confirm that permutation entropy is higher for the developed countries, and market size is correlated with permutation entropy, showing that developed and emerging bond markets differ. In Bariviera et al. (2013) the informational efficiency of European sovereign bond markets is under analysis.

The establishment of the common currency and the 2008 financial crisis are investigated. Results showed that the former had a smaller standard deviation in unpredictability degree, while the latter clearly presented much higher standard deviation of the permutation entropy, meaning that the common currency has hidden the differences between the European markets, and in contrast the crisis uncovered them. In Serinaldi et al. (2014) the position of chaotic deterministic systems are located in the upper left region of the plane, while the stochastic and noisy signals are closed to the limit point (1,0). Bariviera et al. (2015) use the concepts of permutation entropy and permutation statistical complexity to reveal the temporal correlation of interest rates, and more specifically the Libor rate. Using sliding window to show the evolution of these quantifiers, a complex behavior is revealed, and temporal correlations in Libor rates are shown. Zunino et al. (2016) analyzed the efficiency of the European corporate bond sectorial indices before and during the financial crisis of 2007.

The sectors that are connected to the financial economy like financial banks, insurance, basic resources and financial services have presented lower informational efficiency. On the other hand, sectors related to the real economy like healthcare, food & beverage and utilities presented less change in the informational efficiency. As for the construction sector, which is closely related to the Real Estate market, it lost much of its informational efficiency during the crisis. In Bariviera et al. (2017), the efficiency of crude oil market is questioned. The authors concluded that during important economic events the efficiency of the oil has changed.

Lastly, several works applied the entropy concept to quantify the efficiency level in various financial markets, for example see (Oh, Kim, & Eom, 2007) where the investigation of 17 countries with the use of approximate entropy (ApEn) for two periods, namely 1984-1998 and 1999-2004, have shown that the Asian and African markets (excluding Japan) have a lower efficiency level in contrast with the North

American and European markets where the efficiency level is higher. In Risso (2008), the authors support that the decrease of informational efficiency of an economy increases the probability of a stock market crash. The most important finding was that the Russian recently established stock market is the most inefficient, and on the other side the US stock market is considered as the most efficient. Alvarez-Ramirez et al. (2012) studied a long 80 year period of the US stock market. By using entropy as a measure of relative market efficiency, they claimed that the Great Recession lowered the US stock market efficiency, while the least volatile period was found between 1973 and 2003.

In Siokis (2018), the informational efficiency is observed for instruments of the US bond, money and stock exchange markets prior and after the 2007 collapse of Lehman Brothers. The results have shown that after the credit crunch the efficiency level of money and stock exchange market falls considerably, while the bond market keep the same efficiency level. The paper of Martina et al. (2011) aims to shed some light on the crude oil price movements, as well as its connection to extreme events like the Iraq War, and various macroeconomic conditions. The conclusions state that, in the last 25 years, periods of recessions in the US overlap with periods of reduced entropy levels. This shows that an economic downturn affects the long-term market complexity, in contrast with some extreme events like the Iraq War, which affected only the short-term market complexity.

In the same sector, Ortiz-Cruz et al. (2012) try to figure the crude oil price changes. By taking into consideration the WTI crude oil daily data during 1986-2011, they confirmed that the crude oil market is found efficient, apart from the early 1990s and late 2000s, two periods that overlap with US recessions. Furthermore, they stated that as the informational efficiency decreases, the probability for another US recession increases.

The aim of this part is to examine the evolution of the informational efficiency of the Real Estate market in Hong Kong on a size-based and on a region-based perspective during the 1998-2017 period. We would like to investigate the effects of the U.S Real Estate collapse on other housing markets and particularly on Hong Kong area. The rationale for testing this particular region is that the Hong Kong housing market could be considered as an ideal laboratory, based on the robust demand for properties, the relatively low interest rates environment and the large frequent swings in property prices. More precisely, house prices have tripled in Hong Kong from 2003

to 2015 and construction surged sharply.

We also want to understand the changes in the statistical characteristics of the Real Estate market during a period of anxious time. By identifying the statistical properties of the housing market indices under stress, we expect to improve our perception about the mechanisms that determine the dynamics. Taking the present research as an experience, we also believe to develop in the future suggestive diagnostic models in predicting Real Estate downturns. We are studying meticulously the informational efficiency before and after the financial meltdown of 2007. The Real Estate housing price indices of Hong Kong are region-based and size-based, and namely: Hong Kong Mass, Hong Kong Island, Kowloon, New Territories East, New Territories West, the General Housing Price index for overall flat sizes, Large Housing Price index and Medium/Small Housing Price index. Flats with net floor area over 1076 sq. ft. belong to the large housing price index[4]. In the following section we describe the used technique to estimate the degree of informational efficiency.

The remainder of this second part of the research is organized as follows. In the section 2 we describe the information-permutation theory. In Section 3 we describe the data used followed by the presentation and discussion of the empirical results obtained from the different instruments of the financial markets. In section 4 we summarize the findings of the analysis and conclude.

3.2 Methodology

3.2.1 Permutation entropy

A classic topic in the field of financial economics is the informational efficiency. Also, a very challenging issue about time series is to deduce information from its evolution through time. Entropy is a general concept of information theory, and it is expressed in terms of a discrete set of probabilities. The probability distribution function (PDF) is mainly used to quantify the information content of a system. This is done using various measures of entropy, for instance Shannon entropy, Tsallis entropy, Kolmogorov Sinai entropy, or Approximation entropy.

The calculation of the associated probabilistic distribution comprises as the first step of calculating the so-called entropy quantifier, for a given time series. The

[4] For more information, see (www1.centadata.com).

information content can be quantified via a probability distribution function (PDF) P that describes the information sharing of the observable quantity. Various methods have been suggested for the estimation of the probability distribution. Firstly, the Fourier analysis, introduced by Powell & Percival (1979), 2) the wavelet transform by Rosso and Mairal (2002), 3) the symbolic analysis by Daw et al. (2003), 4) amplitude statistics Micco et al. (2008) and, 5) permutation entropy, a procedure introduced by Bandt & Pompe (2002) (B&P).

Apart from permutation entropy, these classical methods present significant limitations in their application; they are computationally expensive, they require long time series data, they fail to capture chronological relationships, their performance is not satisfactory with non-linear chaotic regimes, and fundamental knowledge about the system is a prerequisite. The point is to use a simple, fast and effective method without using all the structure from the time series (Bandt, 2003). In 2002, Bandt and Pompe introduced a new complexity measure, which was a combination of entropy and symbolic dynamics, capable of measuring the time series complexity, called permutation entropy (Bandt & Pompe, 2002).

Compared to the drawbacks mentioned above, in the case of permutation entropy we have no need for prior knowledge, and the method seeks only for order relations of the neighboring values without the notion of magnitudes. It is the only method that derives information from the temporal structure of time series given (Zunino L. , Zanin, Tabak, & Pérez, 2009). The B&P approach is based on the Shannon entropy, and several scientific applications have already used and applied this technique; there is a vast spectrum of sciences that reveal its simplicity-yet its important role. For example Ruiz et al. (2012) measure the volatility in electricity power markets with the use of topological entropy, permutation entropy, and modified permutation entropy. Simulated data report that their results are superior to those based on other measures of dispersion. Also the relation of Kolmogorov-Sinai entropy with permutation entropy is extensively analyzed in Amigo and Keller (2013).

In the medical area, Li et al. (2014) used PE to measure the changes in electroencephalogram (EEG) and to discriminate among three seizure phases. Other uses refer to sleep, epileptic absences and anesthesia as well. In similar applications about EEG, Keller and Lauffer (2003) used symbolic dynamics in order to visualize and extract qualitative changes and information about data related to epileptic activity. In Li and Richards (2007), permutation entropy is under investigation and

comparison with the sample entropy for the prediction of the absence seizures.

Results showed that permutation entropy could be used as a tracker of the dynamical change of EEG data, meaning that transient dynamics prior to a seizure can be found. More specifically, more than 50% of the experiments could detect the pre-seizure state with an average anticipation time of 4.9 sec. Keller et al. (2017) review variants of permutation entropy, being skeptical about the validity of the of the information that the new implementations offer. They also state that the complexity of the permutation entropy could be dealt with the use of a more sophisticated idea, like machine learning. Furthermore, Olofsen et al. (2008) support the permutation entropy and its ability to be used in the electroencephalography and pharmacodynamics. With minimal pre-processing, the composite permutation entropy index (CPEI) could track the EEG changes because of anesthesia, presenting delta waves, spindle-like waves and loss of high frequencies, concluding that CPEI can be considered as a promising measure of GABAergic anaesthetic drug effect. The change of EEG prior to an absence seizure is considered as detectable, and the multiscale permutation entropy (MPE) is used in Ouyang et al. (2013).

The experiments presented an obvious permutation entropy reduction when the patient goes from seizure-free to seizure state, and what is more the characteristics of EEG data can detect the differences between seizure-free state, pre-seizure and lastly seizure state. Also, Cao et al. (2004) confirm the validity of permutation entropy, by analyzing a transient Lorenz system, a transient logistic map, and clinical EEG data. They claim in their findings that permutation entropy can detect bifurcations, and the commencement of epileptic seizures. Bandt (2017) presents a new version of permutation entropy. This new method can be considered as distance to white noise. The results showed that it could be applied to diagnose the sleep disorders by classifying sleep stages using EEG data. Jordan et al. (2008) used in their study electroencephalographic data from patients undergoing general anesthesia, and the findings confirm that they can distinguish unconsciousness from consciousness. Li et al. (2008) compare the results of permutation entropy and approximate entropy that focus on measuring the sevoflurane effect in electroencephalographic data. Results conclude that permutation entropy estimate more effectively the drug effect.

In Bruzzo et al. (2008) the authors found that although permutation entropy could be used to separate preictal and interictal phases, a useful finding for epileptic events, it is not fully reliable for seizure predictions. Another research about the effect

of isoflurane anesthetic drug effect can be found in Li et al. (2013). Moreover, Taherkhani et al. (2013) reveal the difference between healthy people and patients with heart failure, using PE and Detrended Fluctuation Analysis (DFA) with cardiac heart interbeat signals. A modification of permutation entropy, called modified permutation entropy is also used in this area with the goal of distinguishing heart rate variability (HRV) signals under various physiological and pathological conditions (Bian, Qin, Ma, & Shen, 2012). In Traversaro et al. (2017) permutation entropy is compared with a Bayesian Missing Data Imputation in heart rate variability (HRV) data, and the latter outperforms the existing methodologies. The paper of Unakafova and Keller (2013) propose efficient methods to compute time series complexity, in order to be able to use very large time series data in real-time.

Celebrating the tenth anniversary of the original work of permutation entropy, Zanin et al. (2012) review the main applications of the method on the economical and biomedical sector. Permutation entropy is also applied in mechanical engineering topics; see (Yao, Yang, Bai, & Cheng, 2016) a research that presents permutation entropy and fault diagnosis in railway rolling bearings, and some other various uses refer to astrophysical plasmas (Weck, Schaffner, Brown, & Wicks, 2015).

Another field of research that permutation entropy supports is the financial area. The paper of Zhou et al. (2013) deals with the concept of entropy in the field of asset pricing and portfolio selection, and compares it with traditional methods. Dionisio et al. (2006) analyzed the Portuguese stock market. They support the effectiveness of entropy over the variance, and note that irrespective of the location of concentration, the entropy can measure diffuseness of the density, while the variance can measure concentration only around the mean. What is more, they used in their study 1856 daily closing prices of 23 Portuguese stock market stocks from 28/06/1995 to 30/12/2002. The statistical analysis showed that the entropy is sensitive in the portfolio diversification, since it decreases when more assets are included in the portfolio. This makes the number of possible states of the portfolio to decline, together with the uncertainty.

In Hou et al. (2017) the Shanghai Stock Exchange, and Shenzhen Stock Exchange are under investigation, to define whether the complexity of a market increase or decrease. Using two significant time windows, one from 2006 until 2011, and the other for the period 2014-2016, they showed that permutation entropy falls drastically in both periods, meaning that the economic rise was followed by gigantic

stock crashes. A non-parametric independence test was conducted by Matilla-Garcia and Marin (2008) with the use of permutation entropy. The authors test the random walk hypothesis of the DJIA, S&P500, and three exchange rate returns. Similar analysis about entropy, symbolic dynamics, and non-parametric test could be found in Matilla-Garcia (2007). Also, Zanin (2008) analyzed the appearance of forbidden patterns in various economical indicators like NYSE, Dow Jones and 10-year Bond interest rate. For more information about forbidden patterns, see (Amigo, Zambrano, & Sanjuan, 2007), (Rosso, Carpi, Saco, Ravetti, Larrondo, & Plastino, 2012).

The complexity of arbitrary time series is calculated using an ordinal pattern, and then time causality is assessed from the comparison of the adjacent values. Ordinal time series analysis can be found in (Keller, Sinn, & Emonds, 2007). The main advantages of permutation entropy as set forth by Zunino et al. (2009) are the robustness, simplicity, and its invariance considering nonlinear monotonous transformations. Furthermore, its applicability to any type of time series, a large sample size is not required, like for instance in fractal analysis, and last but not least it provides a fast computational algorithm. For more information, see (Soriano, Zunino, Rosso, Fischer, & Mirasso, 2011).

The degree of structure in a process is not quantified by randomness measures, and therefore it is necessary to obtain measures of statistical complexity. Those measures were firstly introduced by Lopez-Ruiz et al. (1995), which manage to identify essential dynamics, and distinguish different degrees of periodicity and chaos. The embedding dimension D defines the number of symbols that formulae the ordinal pattern, and guides for the necessary length of time series in a way that the condition $M >> D!$ is necessary for reliable statistics. Also τ, which shows the time delay, is a measurement of successive points in the symbol sequences. These cause the problem of multiple testing.

Moreover, Riedl et al. (2013) presented an exhaustive literature research in their paper, to show how experts apply PE on different scientific problems. In this research, the table for economics and environmental studies is taken into consideration. Despite their effort, the chosen parameters presented later remain an important issue that remained to be analyzed, and a new method of choosing the right parameters would probably help the optimization of permutation entropy, and become an idea for another research. For more analysis and information, see (Zunino, Soriano, Fischer, Rosso, & Mirasso, 2010), (Soriano, Zunino, Larger, Fischer, & Mirasso,

2011), (Li, Li, Liang, Voss, & Sleigh, 2010), and (Tenreiro Machado, Costa, & Quelhas, 2011).

3.2.2 The calculation of permutation entropy

The calculation of permutation entropy essentially contains an algorithmic sequence, which measures information based on the absence or the occurrence of specific permutation patterns of the ranks of values. Seven steps should be followed in order to compute the permutation entropy of $\{x_i\}$ time series of length N. Despite the value of the return time series, only price time series are used (Zunino L. , Zanin, Tabak, & Pérez, 2009).

Given a time series $\{x_t: t = 1,..., M\}$, an embedding dimension D > 1, and a time delay τ, we consider the ordinal patterns of order D which are generated by a segment $s\rightarrow(x_{s-(D-1)\tau}, x_{s-(D-2)\tau},...,x_{s-\tau}, x_s)$. The ordinal pattern of the time series is the permutation $\pi= (r_0,r_1,\ldots,r_{D-1})$ of the index set (0,1,...,D–1) corresponding to the ranking of the x in ascending order, namely $x_{s-r_{D-1}\tau} \leq x_{s-r_{D-2}\tau} \leq \cdots \leq x_{s-r_1\tau} \leq x_{s-r_0\tau}$. In order to get a unique result we consider that if $x_{s-r_i\tau} = x_{s-r_{i-1}\tau}$ then $r_i < r_{i-1}$. For all the D! possible permutations π_i of order D, their associated relative frequencies can be computed by

$$\boldsymbol{p(\pi_i) = \frac{\#\{s|1 + (D-1)\tau \leq s \leq M \text{ has ordinal pattern } \pi_i\}}{M-(D-1)\tau}} \qquad (19)$$

where # stands for frequency of occurrence of π. Also $P = \{p(\pi_i), i=1,\ldots,D!\}$. The permutation entropy is defined as the Shannon entropy of this probability distribution, i.e.,

$$\boldsymbol{S[P] = -\sum_{i=1}^{D!} p(\pi_i) \ln p(\pi_i)} \qquad (20)$$

A typical example to make things clear is the following. Consider the time series $x = 5, 8, 10, 11, 7, 12, 4$ with N=7.

Step 1: The permutation order is set to n=3. We have n! = 6 possible permutations,

with π_1 = 1,2,3, π_2 = 1,3,2, π_3 = 2,1,3, π_4 = 2,3,1, π_5 = 3,1,2, and π_6 =3,2,1.

Step 2: Initialize $i = 1$ and $z_{j=1,\dots,6} = 0$.

Step 3: The rank sequence of the selected values 5, 8, 10 is 1, 2, 3.

Step 4: It is equal to π_1, therefore z_1 is increased to 1.

Step 5 and Step 3: Because i = 1 < 7 − 3, the next values 8, 10, 11 are selected which have the rank sequence 1, 2, 3.

Step 4: It is again equal to π_1, therefore z_1 is increased to 2. The loop between Step 3 and Step 5 is then passed through three more times, which leads in the end to z_1 =2, z_2 =0, z_3 =1, z_4 =2, z_5 =0, and z_6 =0.

Step 6: The values of the counters are divided by sum = 5 which leads to $p'_1 = {}^{2}/_{5}$, $p'_2 = 0$, $p'_3 = {}^{1}/_{5}$, $p'_4 = {}^{2}/_{5}$, $p'_5 = 0$, and $p'_6 = 0$.

Step 7: On the basis of the non-zero p'_j, the permutation entropy of order 3 is

$H_3 = -\left({}^{2}/_{5}\, log_2\left({}^{2}/_{5}\right) + {}^{1}/_{5}\, log_2\left({}^{1}/_{5}\right) + {}^{2}/_{5}\, log_2\left({}^{2}/_{5}\right)\right) \approx 1.5$ and the permutation entropy per symbol is $h_3 = {H_3}/_{2} \approx 0.76$.

From the same paper of Riedl et al. (2013), the existence of equal values in the time series data should be solved with one of those strategies:

1. The ranks of the values are determined in accordance to their order in the sequence.
2. The identity is eliminated by adding white noise with the strength of the stochastic term being smaller than the smallest distance between values.
3. The values get the same rank number within the regarded sequence, for example 3, 6, 2, 3, 8 leads to the rank sequence 2, 4, 1, 2, 5.

When we apply the permutation entropy, what we expect to find is some skewed probability distribution, meaning that those pattern occurrences reveal the workings of the underlying system. The ones that do not appear are the forbidden patterns.

3.2.3 Complexity entropy causaslity plane

We get the Normalized permutation entropy by dividing the Shannon entropy by lnD!

$$\boldsymbol{H_S[P]} = \left[-\sum_{\boldsymbol{i=1}}^{\boldsymbol{D!}} \boldsymbol{p(\pi_i)} \ln\left(\boldsymbol{p(\pi_i)}\right) \right] / \ln \boldsymbol{D!} \tag{21}$$

The Normalized permutation entropy lies between 0 and 1. $H_s[P]=1$ when there is pure random sequence and our ignorance is maximal, while $H_s[P]=0$ when we can predict the outcome with certainty.

The results of permutation entropy are linked to the choice of the embedding dimension. The latter depends heavily on the length M of the time series in a way that M>>D! must be satisfied. B&P proposed $3 \leq D \leq 7$, arguing that a D of 2 would not work properly because there are few distinct states, while a D>7 would cause memory restrictions, and as Staniek and Lehnertz (2007) also state, one cannot have reliable statistics about the ordering dynamics. The embedding dimension is set to D=4, but also D=3 and D=5 are used for comparison. The time delay τ, which is the time separation between symbols is set in our paper as τ=1.

We also search for the degree of correlational structure in the time series with the use of statistical complexity (Lopez-Ruiz, Mancini, & Calbet, 1995), (Rosso, Larrondo, Martin, Plastino, & Fuentes, 2007). In Lopez-Ruiz et al. (1995), the measure of complexity is proposed as a distance from the equiprobable distribution of the accessible states of the system. The reason why we care about it is given in (Tsallis, 2014).

The definition of complexity is not universal because it's difficult to define it in an exact way. In the literature, various measures of complexity have been proposed. For instance, in Sprott (2004) as a measure of a system's self-organization "capacity", in Ott et al. (1994), and in Sprott (2004) where the concomitant's attractor dimension is a measure of its complexity, and the Kolmogorov-, information-, algorithmic-, or Chaitin-complexity in Solomonoff (1964), Kolmogorov (1965), Chaitin (1966) where it is measured according to the smallest computer program possible that could produce an observed pattern, and also the measure of the specific amount needed about the past in order to predict the future (Crutchfield & Young, 1989). To make things clear a good line of thought is to understand and exclude the

processes that are not complex, for example the white noise random process, or those that exhibit periodic motion. These two circumstances can be regarded as "trivial".

In our research, we are dealing with the case of the statistical complexity measure, a family of entropic non-triviality measures C where we have very small complexity for large amounts of either order or disorder The basic condition claims that complexity should vanish for systems of total order and disorder. Hence we expect to have its maximum value somewhere between order and disorder. In the second definition of statistical complexity measure we confirm that complexity grows with increasing disorder, and in the third definition it grows with increasing order.

Some new definitions of complexity have been released (Calbet & Lopez-Ruiz, 2001), (Lopez-Ruiz, Mancini, & Calbet, 1995), (Anteneodo & Plastino, 1996), (Feldman & Crutchfield, 1998), (Poeschel, Ebeling, & Rose, 1995). The complexity is defined from the distribution function describing the system. This type of complexity is based on the notion of "disequilibrium", and is referred as Lopez-Ruiz-Mancini-Calbet (LMC) complexity. The advantage of disequilibrium-based complexity is that it is easy to calculate and has some interesting properties to show. Nevertheless as the size of a system increases, or equivalently, when the distribution function becomes continuous the complexity does not behave properly. According to the theory of complexity definition, the LMC complexity C is C=DH, with H as the entropy, and D the disequilibrium term. The system can be in one of the total N possible states i. The discrete distribution function f_i defines the probability of the system being in state i.

The system is defined in a way that, if completely isolated, it will reach equilibrium, and all the states will have the same probability f_e. As a result $f_i \geq 0$, after the normalization the I must hold such that $I \equiv \sum_{i=1}^{N} f_i = 1$. The equilibrium distribution function is $f_e = {}^{1}/_{N}$, and the disequilibrium is defined as $D \equiv \sum_{i=1}^{N} (f_i - f_e)^2$.

Thus a complexity measure is developed, such as the so-called Jensen-Shannon complexity C_{JS}. The Jensen-Shannon complexity is a functional of the discrete distribution P of N probabilities associated with the time series (Martin, Plastino, & Rosso, 2006). It is a range of C_{JS}'s values provided in the range of C_{min} and C_{max} given for an entropy value. For that matter a complexity measure is developed, such as the so-called Jensen- Shannon complexity C_{JS}. Once normalized such that $0 \leq C_{JS} \leq 1$ then

$$C_{JS}[P] = -2\frac{S\left[\frac{P+P_e}{2}\right] - \frac{1}{2}S[P] - \frac{1}{2}S[P_e]}{\frac{N+1}{N}\log(N+1) - 2\log(2N) + \log N}H[P] \quad (22)$$

where S is the Shannon entropy, *H[P]* is the normalized Shannon entropy and P_e*={1/N,...1/N}* is the uniform probability. The Jensen-Shannon complexity has been introduced in nonlinear dynamics analysis to detect essential details of the dynamics and discriminate different degrees of periodicity and chaos (Lamberti, Martin, Plastino, & Rosso, 2004). The disequilibrium is defined in terms of the divergence, which quantifies the difference between two probability distributions. So in order to understand how the evolution of complexity works, we cannot use a complexity versus time plot, because the second law of thermodynamics defines that the entropy increases monotonically with time, meaning that $\frac{dH}{dt} \geq 0$. So another way to depict the complexity evolution is the trade of the time axis with the measure of entropy. An important note is that the time evolution specifies the shrinking or the stretching of the entropy horizontal axis. This entropy-time case is not a trivial issue (Latora & Baranger, 1999).

3.2.4 Tsallis-q Entropy

Even though the Jensen-Shannon complexity is widely used and accepted, in this paper we apply also monoparametric generalizations of the statistical complexity and normalized Shannon entropy, based on the Tsallis q entropy. We use multiple q (in Shannon entropy we have q=1), which supply with different weights the underlying probabilities of the system, in order to understand whether they can reveal different meanings of complexity. Therefore, we compare $H_S[P]$ and $C_{JS}[P]$ with $H_q[P]$ and $C_P[P]$.

In the paper of Zunino et al. (2008), Tsallis entropy is used to analyze the fractional Gaussian noise and fractional Brownian motion, and is was shown that Tsallis entropy offers better discrimination of the processes compared to the Shannon entropy, and for a similar research we refer also to Zunino et al. (2008). In the

findings of Ribeiro et al. (2017), and Martin et al. (2006), we get the normalized Tsallis entropy as

$$\boldsymbol{H_q[P] = \frac{\frac{1}{q-1}\sum_{j=1}^{N}\left((p_j) - (p_j)^q\right)}{\frac{1-N^{(1-q)}}{q-1}}} \tag{23}$$

For the complexity measure $C_P[P]$, divergence and Q_0 are also calculated as

$$\boldsymbol{Jensen\ divergence = \frac{1}{2(q-1)}\sum_{\square=1}^{N}(p_j)^q\left\{\left(\frac{p_j+\frac{1}{D!}}{2}\right)^{1-q} - (p_j)^{1-q}\right\} + \frac{1}{2(q-1)}\sum_{j=1}^{N}\left(\frac{1}{D!}\right)^q\left\{\left(\frac{p_j+\frac{1}{D!}}{2}\right)^{1-q} - \left(\frac{1}{D!}\right)^{1-q}\right\},} \tag{24}$$

$$\boldsymbol{Q_0 = (1-q)\left\{1 - \left[\frac{(1+N^q)(1+N)^{(1-q)}+(\square-1)}{2^{(2-q)}N}\right]\right\}^{-1},} \tag{25}$$

The complexity for the Tsallis entropy is calculated as

$$\boldsymbol{C_P[P] = Q * H_q[P]} \tag{26}$$

where the disequilibrium $Q = Jensen\ divergence * Q_0$.

For further analysis about Tsallis entropy, complexity and Jensen divergence, please refer to (Jauregui, Zunino, Lenzi, Mendes, & Ribeiro, 2018), (Tsallis, 2017), (Raj & Wiggins, 2008), (Grosse, Bernaola-Galvan, Carpena, Roman-Roldan, Oliver, & Stanley, 2002), (Lamberti, Majtey, Borras, Casas, & Plastino, 2013), (Lamberti & Majtey, 2003), (Bentes & Menezes, 2012), (Tsallis, Anteneodo, Borland, & Osorio, 2003), (Tsallis, 1988), (Bentes, Menezes, & Mendes, 2008), (Mendes, Evangelista, Thomaz, Agostinho, & Gomes, 2008), (Tsallis, 2016), (Kalimeri, Papadimitriou,

Balasis, & Eftaxias, 2008), (Huang, Yong, & Hong, 2016), (Tsekouras & Tsallis, 2004), and (Kowalski & Plastino, 2012).

3.3 Empirical results

The data used are called CCL, aka "Centa-City Leading Index", which is a weekly index-containing contract prices from transactions taken in Centaline Property Agency Limited. It reflects the overall price trend of Real Estate market of Hong Kong. In this paper we shall refer to CCL as the "General index". The Large and Small/Medium housing indices are both sub-indices of the General index, showing the trend of flats over and under 1076 sq. ft. (100 sq.m.). We also analyze the Mass index, a general trend of all regions of Hong Kong, and its four sub-indices named from the regions chosen: Hong Kong Island, Kowloon, New Territories East and New Territories West. Figure 18 presents the dataset of every CCL index.

Our data comprise of weekly region based and size based measurements before and after the 2007 U.S. sub-prime loan crisis. Data are taken from Centa Data, covering the period from April 1998 to September 2017, divided into two equal non-overlapping segments. The first period named 'prior-crisis' spans from April 1998 to December 2007, while the 'crisis/post crisis' spans from January 2008 to September 2017. Each sub-period consists of 505 weekly observations. The reason behind the division of the weekly time series is to obtain information about Great Recession and its connection with the Real Estate market of Hong Kong. Even though daily measurements might have shown interesting results as well, they are not available in our case.

In this application, permutation entropy and the statistical complexity are calculated for embedding dimension D=4 and embedding delay τ=1. We also performed the analysis for two extra embedding dimensions D=3 and D=5. The results seem to be independent of the embedding dimension selected for the symbolic reconstruction of the original time series. Therefore, apart from some abstract presentation of the other two embedding dimensions the rest part of this analysis will comment mainly findings for the D=4 case even though all results are available. We can clearly discriminate the findings for D=3, 4, 5 and their relative position in Figure 19. We choose this pattern length based on the B&P criterion that satisfies M>>D!.

Figure 20 presents the Shannon Complexity Entropy Causality Plane (CECP) for all indices of Real Estate market in Hong Kong for D=4 and D=5. They are separated firstly as size based and then as location based. Panels 1 and 2 show the size based and region based indices for D=4 and in Panel 3 and 4 of Figure 20, we presented for the sake of comparison the case of D=5 as well. By analyzing the D=4 case, we see that the three quantifiers for the prior period are located in the area of higher permutation entropy, ranging from 0.914 to 0.929 and lower permutation complexity between 0.079 and 0.095.

In contrast, the after crisis period quantifiers are located on a level with lower permutation entropy and higher permutation complexity. In detail, the former spans 0.878 to 0.91, and the latter spans from 0.095 to 0.126. We can see that the pre crisis measurements for the general index, large and small/medium size flats show a higher efficiency level in comparison with the after crisis quantifiers, which means that there is more memory after the 2007 crisis and the price of last week influence next week's price.

A very interesting measurement is the quantifier of the large size flat index for the after crisis period. It is located closer to the core of the pre-crisis group, where the pre-crisis large market measurement belongs as well, giving the intuition that the Great Recession has not affected the large flat Real Estate market efficiency that much. Nevertheless, all pre-crisis indices are more efficient than the ones in the period after the crisis. In Table 6, the values of entropies and their statistical complexity, as well as the average and standard deviation of both sub-periods with D=4 and τ=1 are presented.

Next we analyze the CECP for four sub-regions of Hong Kong, namely Hong Kong Island, Kowloon, New Territories East, and New Territories West. Various clusters of points in different efficiency groups are the major difference of Panel 2 compared to size-based Panel 1 of Figure 20. We firstly note that both mass after and mass prior are the least efficient compared to their respective sub-indices. In the case of Hong Kong Island, New Territories East and New Territories West we found lower efficiency in the post crisis period. Their prior crisis measurements form a small group meaning that the housing price dynamics are very similar.

Nevertheless, Hong Kong Island indices have a shorter distance on the plane, a sign of a comparatively more stable market. A very interesting finding lies in the Kowloon index, which shows lower permutation entropy and higher permutation

complexity during the pre-crisis period than after. One possible explanation would be the fact that Kowloon is part of the decentralized trend because of the Hong Kong Island high prices in the office and retail sector; thus Kowloon's housing market is unavoidably affected and as a result, an inelastic demand for the region from the outsiders as well as the Hong Kongers kept the market protected and more efficient since the 2007 downturn. What is more, similar to Hong Kong after index, the Kowloon prior index is located closed to the cluster of higher efficiency levels.

In Zunino et al. (2010), it is argued that if time series is purely random, then according to Fama's notion the permutation entropy reaches the maximum level of 1, and also complexity bottoms to 0. In a nutshell, this optimal combination (1,0) that a financial market can reach characterizes the strong form of the EMH, where the market is strongly efficient and reflects both private and public information. To enlighten things more, shuffled series are also shown in the second figure, located to the bottom right of each panel, and they are also presented in Figure 21 for both D=4 and D=5 zoomed to their scale.

Furthermore, permutation entropy and permutation complexity depict high correlation. Also, the stochastic processes are characterized by information redundancy. Thus, we perform the analysis of fractional Brownian motions (fBms) using different Hurst exponents. There is a wide use of correlated stochastic processes for financial time series modeling. In our case, we generate 1000 numerical independent realizations of fBms with Hurst exponent $H \in \{0.10, 0.20, 0.30, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.80, 0.90\}$ of length N=1010 data points. We used the command wfbm in MATLAB to complete the task, which follows the algorithm proposed by (Abry & Sellan, 1996), (Bardet, Lang, Oppenheim, Philippe, Stoev, & Taqqu, 2003). The fBms curves are illustrated in Figure 20 too, along with the Real Estate indices. The quantifiers are located closed to the fBms, meaning that they share some dynamical properties. The fBms in the range (0.5-0.7) could be used for modeling purposes. Furthermore, some of our measurements lean close to H = 0.5, giving a sign of uncorrelated dynamics.

Next, we consider Tsallis q entropy and complexity, and present in Figure 22 and Figure 23 the q-complexity-entropy curve. Apart from D=5 case, the D=3, 4 results of our Real Estate indices from panel a to panel h resemble those of the fractional Brownian motion in Figure 24. They all form a loop with different size for

the prior and after crisis period. More specifically, all indices revealed broader loops for the after crisis period, apart from the Kowloon index, where the loop for the period after 2007 is smaller. Except for the latter case, the results confirm the concept that during periods of Real Estate turbulence, housing market indices show larger Hurst exponents. When D=5, the series do not form a loop in any case, meaning that some permutations π_i do not appear. In Carpi et al. (2010), it is reported that for D=4, time series of fractional Brownian motion should be more than 500 in order to avoid forbidden patterns; here we are a notch above 500, so this might be an issue for the incompleteness of the time series loop.

The calculation of the generated time series for the fractional Brownian motion is completed following the procedure of (Hosking, 1984). We apply the Hurst exponents (0.2, 0.3,…,0.9) for the embedding dimension D=4. The plots presented in Figure 24, are the average values from 100 realizations of 1000 length time series of the entropy and complexity causality plane, with the parameter q ranging from 10^{-4} to 1000 with 1000 separated steps. Here we have to declare, that the initial methodology in Ribeiro et al. (2017) uses much broader range of data; the average values from 100 realizations of $2^{\wedge 17}$ length time series of the entropy and complexity causality plane, with the parameter q ranging from 10^{-4} to 1000 with 10^{-4} separated steps, which means 10 million points. Our limited resources in raw computer power and the finite knowledge of programming in order to write and run a possibly more optimized and effective algorithm imposed us to do some alterations in order to run our initial algorithm in the best way possible. Nevertheless, the conclusions do not alter for the Real Estate indices.

Furthermore, as in (Ribeiro, Jauregui, Zunino, & Lenzi, 2017) we present for Hurst exponent h= (0, 0.1, 0.2,…1) the minimum value of the normalized entropy as a function of q at $q = q^*H$, and the maximum value of the complexity as a function of q as well, when $q = q^*C$. The functions in Figure 25 represent the largest contrast between H_q and C_q calculated for the system distribution P and the uniform distribution P_e. In the first panel q^*H is positively correlated with the hurst exponent until a specific point and then starts to fall, while the q^*C is increasing monotonically as h increases.

Next from Figure 26 to Figure 29 we present the $q = q^*H$ and $q = q^*C$ for the housing indices. Dimension D=3 presents the largest q^*H for both prior and after

the crisis indices, followed by D=4 and D=5, respectively. On the other hand, the values of q^*C, are almost the same for all indices, no matter the value of D. All the values of q that optimize H_q and C_q are available in Table 7.

3.3.1 Time evolution

From Figure 30 to Figure 35 we investigate the time evolution of the quantifiers in an attempt to draw dynamically the inefficiency level, and to test if the underlying process changes before and after the Great Recession. Considering a 100 weeks window, which corresponds to about 500 business days (or two years) and shifting through the time series with a step of 1 week, we measure the time variation of both permutation entropy and statistical complexity. This methodology can grasp any change in the underlying stochastic process by observing the way measurements locate progressively through time.

On each instrument there are two time periods, the first spans from 1998 to 2007, and the second from 2008 to 2017. We detect the changing position over time for both Shannon (Figure 30, Figure 31) and Tsallis entropy (Figure 32, Figure 33), and we can definitely conclude that both methods behave similarly. In the case of Tsallis entropy we calculate every index twice with the Kullback-Leibler (Figure 34, Figure 35) and Jensen-Tsallis method. Although both methods behave similarly, clearly the entropy complexity causality plane discriminates between the two periods, as seen by the different locality of the points. The efficiency level in the cases of Hong Kong Island, Mass, Small, Large and General index remains almost to the same locus for both periods. New Territories East, and New Territories West lean to a lower permutation and higher complexity area as we move from the pre-crisis to the after-crisis period. The Kowloon index increases its efficiency to a higher permutation and lower complexity area as we reach the after-crisis period.

We depict in the next figures the normalized Shannon entropy and Tsallis entropy quantifiers. The results of every method for every index are presented in Figure 38, Figure 39 and Figure 36, Figure 37 respectively. Especially for the Tsallis case, we performed the entropy quantifiers for q = 1.4, 1.45, and 1.5, and we find that

all the indices behave similarly. In order to compare their trend we do not present in Figure 40 all indices.

For the sake of parsimonious and aesthetically correct figures, and since the General index gives similar results with the Large, Small/Medium, Mass and NT West indices, we chose to display only the General index. On the other hand, Hong Kong and NT East index give more or less similar results too. In this case we take into consideration the former one. Lastly, Kowloon index is one trend of its own. With an eyeballing of panel 1 and panel 2 in the Figure 40, we notice that the information efficiency for the Real Estate indices is clearly affected by the 2007 crisis. But before the 2007 nosedive, the quantifier had another steep fall too, which could be explained by the severe acute respiratory syndrome (SARS) epidemic on March 2003 (Lam, Yu, & Lam, 2008). We see that after the virus outbreak, the market replied gradually but fully effective, changing the trend of the entropy.

In 2007 there is a drop of the information efficiency level when we are entering the Great Recession period through time. More specifically, General index follows a downward trend after the sub prime crisis, which lasted until the last quarter of 2009, which was the first low level of a triple dip in permutation entropy. The reason for the second low after 2009 could have been the January 2010 resigns of the five Hong Kong MP's to revive the campaign for civil liberties and western style democracy. Probably fear dominated the housing market, which might have reacted nervously in a potential political instability. The last decrease in permutation entropy values takes place in the first quarter of 2012, where a Special Stamp Duty on housing was imposed and also, the announcement of sovereign downgrade of nine European economies (including France, Austria, Spain and Italy).

After the end of the triple dip all indices followed more or less similar upward trend in both entropy quantifiers, until the first quarter of 2017. At that time, all housing price quantifiers apart from the indices of Hong Kong, NT East and NT West fall again. An explanation is the government's project to raise land supply, announcing on February 23 the 2017-2018 Land Sale Programme, to increase by 18900 the supply of flats in Hong Kong. Some information about the specific regions of land supply can be seen in Figure 41 below, which was found in Jones Lang LaSalle http://www.jll.com/.

Although many important events affect the projection of the entropy quantifier, the transition from the pre crisis to the after crisis period in 2007 is still the

main subject of research. The Hong Kong Island index and the two New Territories indices have a stable transition after the 2007 event, with increasing entropy until the end of the time series. This might means that the housing market of the aforementioned indices is an efficient Real Estate market.

Another important finding of our research is the Kowloon index, which had an increasing efficiency level trend after the crisis that remained in high levels too. A good explanation of this trend could be the increasing inflow of investors' capital from Mainland China in Kowloon housing market to hedge against RMB depreciation (Hui, Ng, & Lau, 2011). A second one is the investors' rush-in that want to find a strategic gateway from the Mainland China and could justify this stable trajectory, and finally the Hong Kong's government intention to transform Kowloon East into a second Central Business District (CBD) is also another reason. In CBRE (2012) it is noted that as investors become cost conscious, they prefer cheaper areas from the CBD of Hong Kong, to split their operations in decentralized areas like Kowloon.

3.4 Conclusion

We investigated the behavior of housing market indices of Hong Kong area before and after the 2007 U.S. financial crisis. Based on normalized permutation entropy and statistical complexity we calculate the informational efficiency level before and after the outbreak periods of the crisis. We found that the financial crisis modified the degree of randomness and the complexity of the housing indices. The indices, both in terms of size and region, in most cases exhibit lower permutation entropy and higher complexity for the after crisis period, meaning that the financial crisis has affected the dynamical structure of the indices, increasing their regularity and predictability. But for Kowloon index, the quantifiers of the after-period exhibit the opposite results denoting greater informational efficiency or decreased regularity and loss of complexity. This outcome seems to be robust by utilizing the Tsallis-q entropy methodology.

In addition, by taking a 100-week long window, shifting through the time series with a step of one week, we calculate the time variation of the permutation entropy for representative indices. Both, Shannon and Tsallis-q entropy methods display very similar trends and values. It seems that they capture well the impact of the financial meltdown on Hong Kong's housing market indices, but with a small time delay. Again, based on time variation the Kowloon index depicts a different trend with the quantifiers to increase their values even during the outbreak of the crisis. Lastly, the permutation entropy captures well the other major events such as the SARS epidemic, the Eurozone crisis and the resignations of the house representatives.

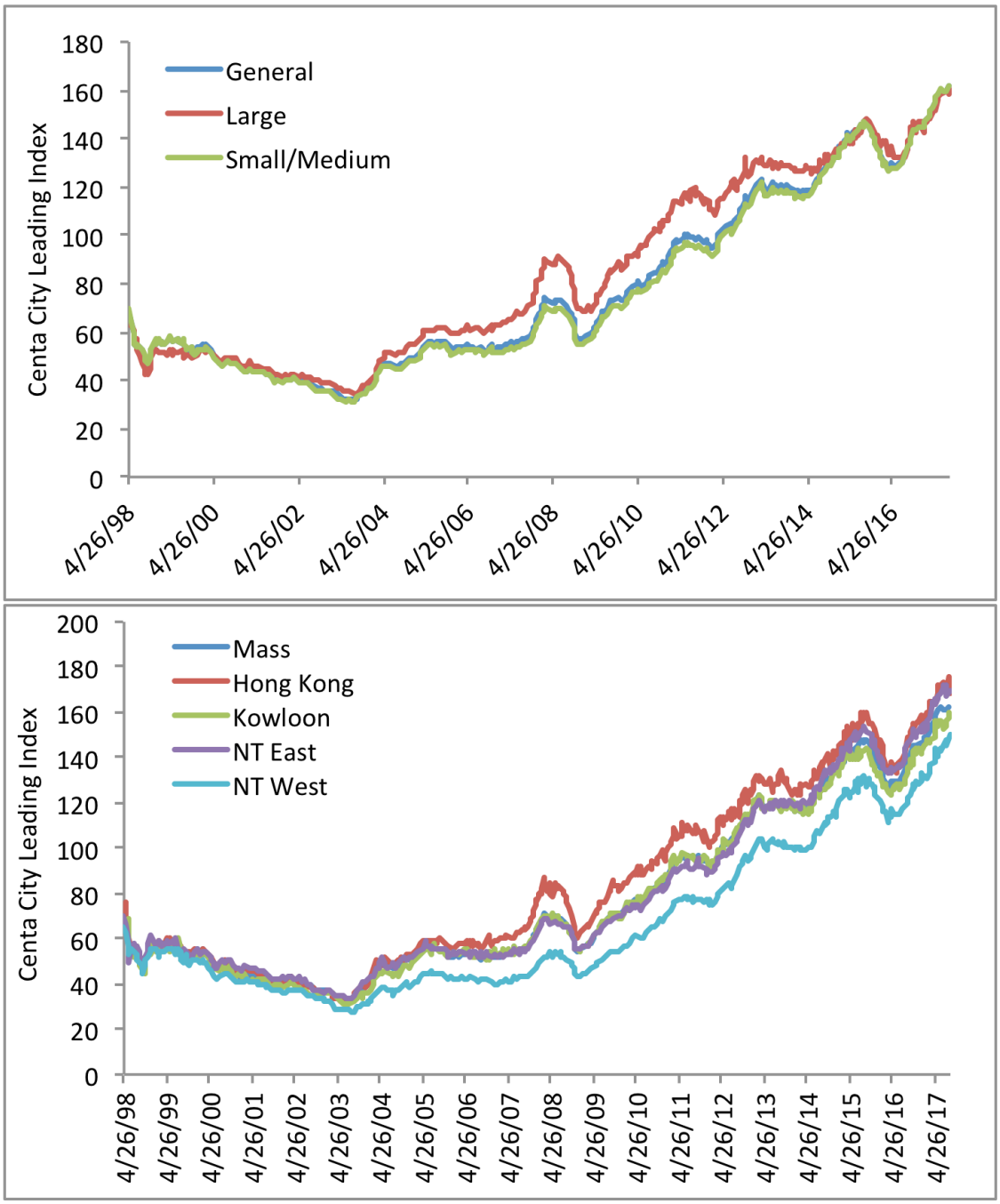


Figure 18: Hong Kong weekly property indices

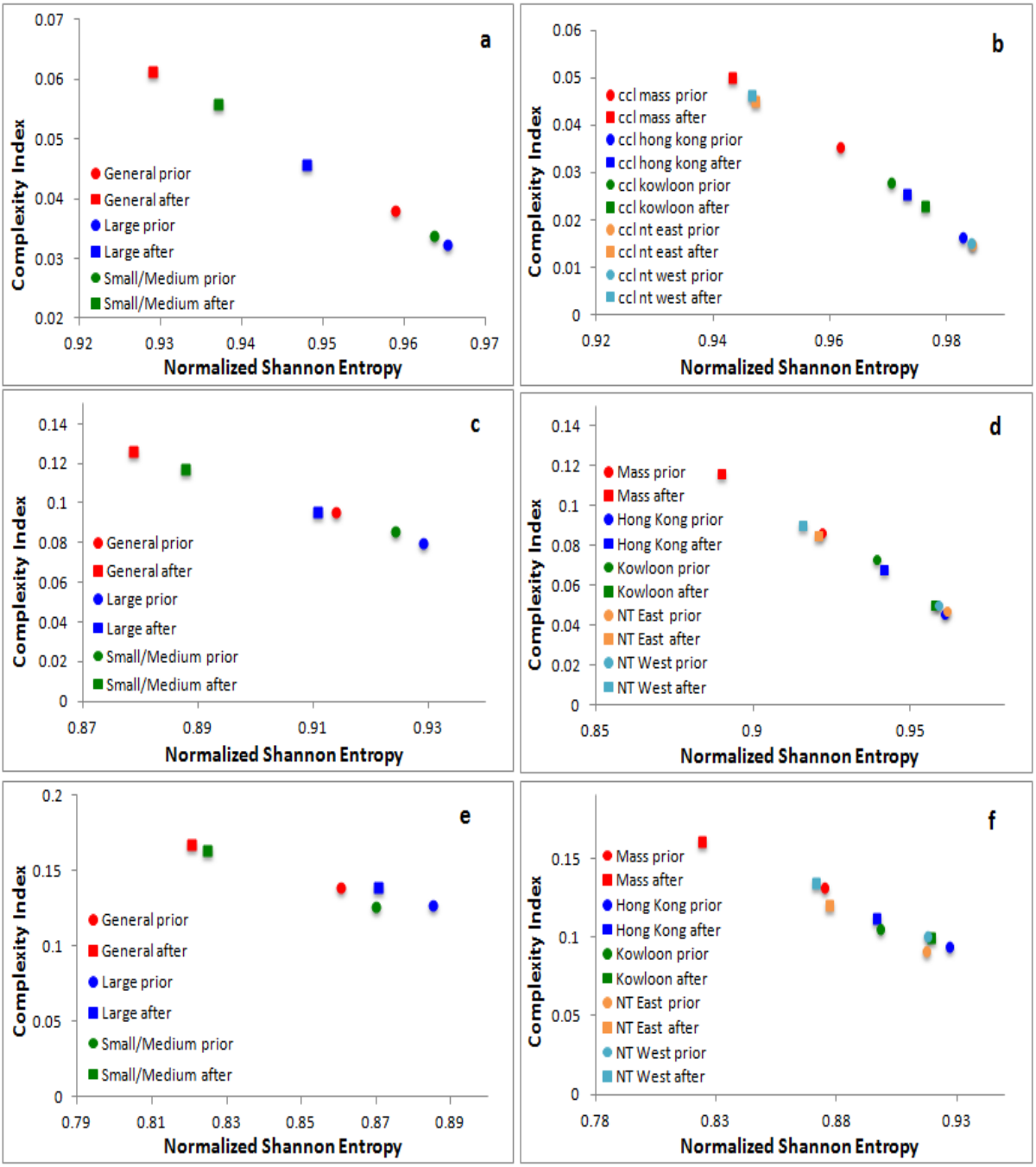


Figure 19: Location of the various indices before and after the outbreak of the crisis, in terms of size and region in the CECP with embedding dimensions D=3 (a and b panels), D = 4 (c and d panels), D = 5 (e and f panels), and time delay $\tau = 1$.

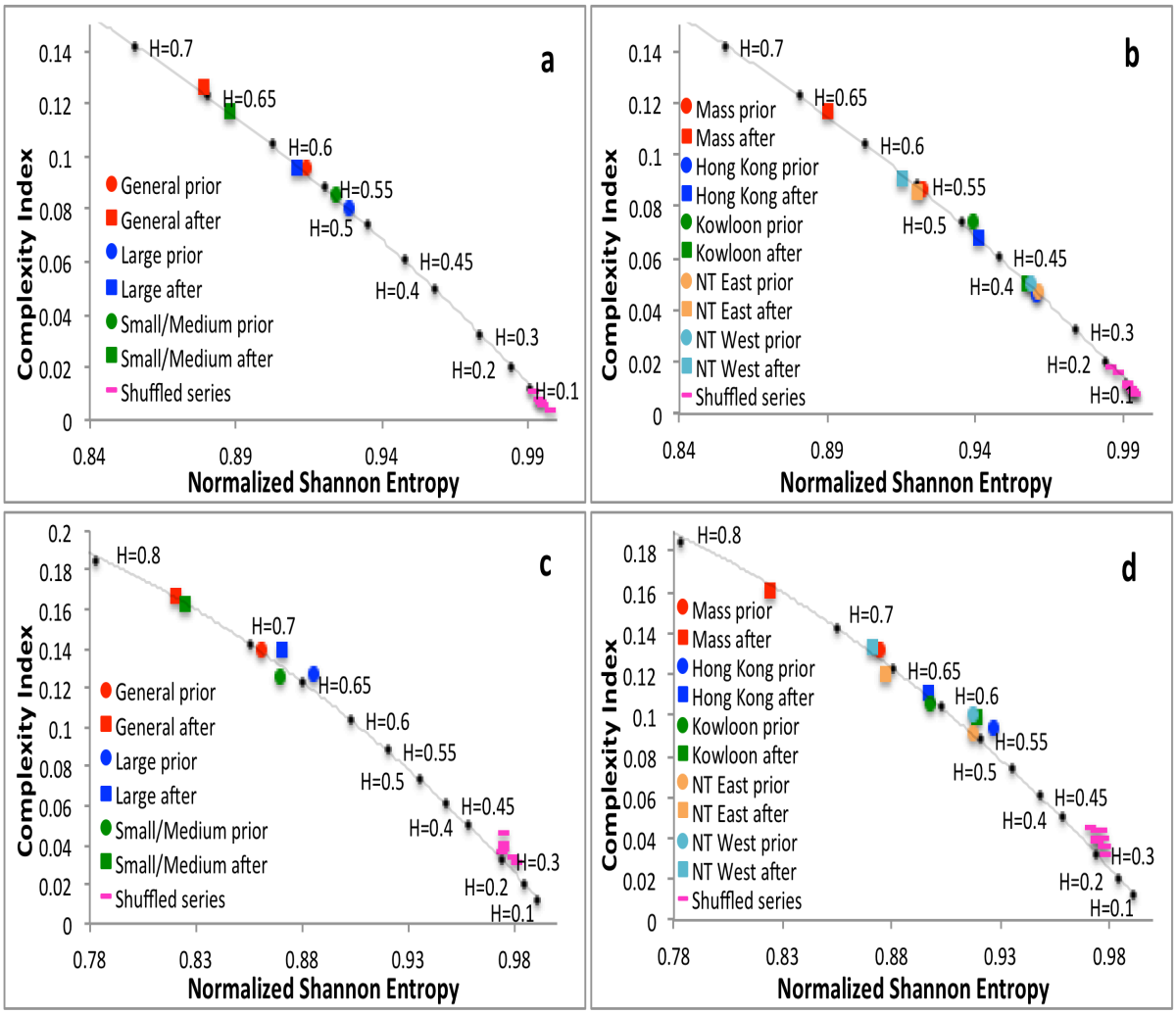


Figure 20: Location of the various indices with the Hurst exponents and shuffled series before and after the outbreak of the crisis, in terms of size and region in the CECP with embedding dimensions D = 4 (a and b panels), D = 5 (c and d panels), and time delay τ = 1.

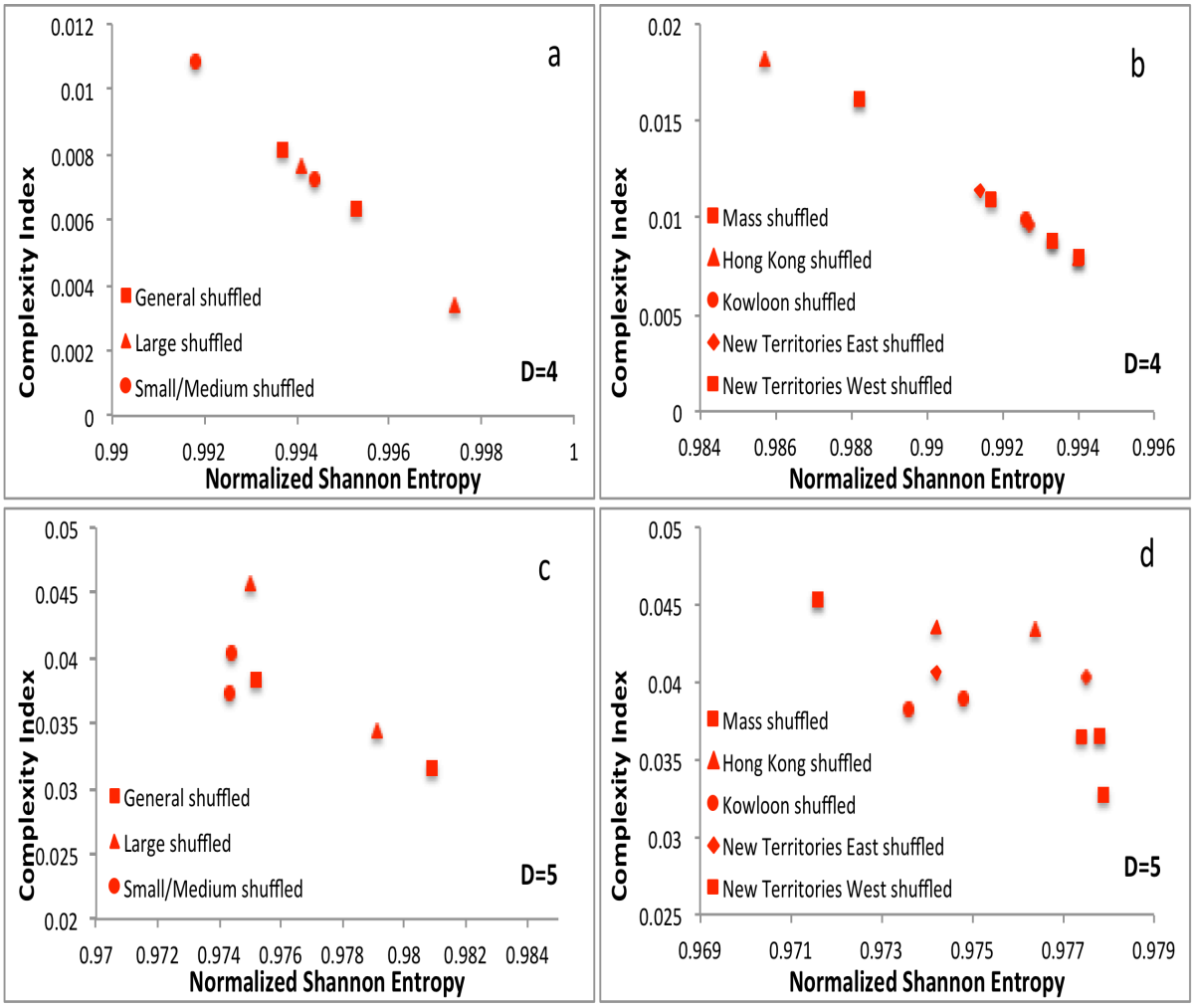


Figure 21: Location of the shuffled indices before and after the outbreak of the crisis, in terms of size and region in the CECP with embedding dimensions D = 4 (a and b panels), D = 5 (c and d panels), and time delay τ = 1. All results are depicted with complexity closed to 0, and entropy closed to 1.

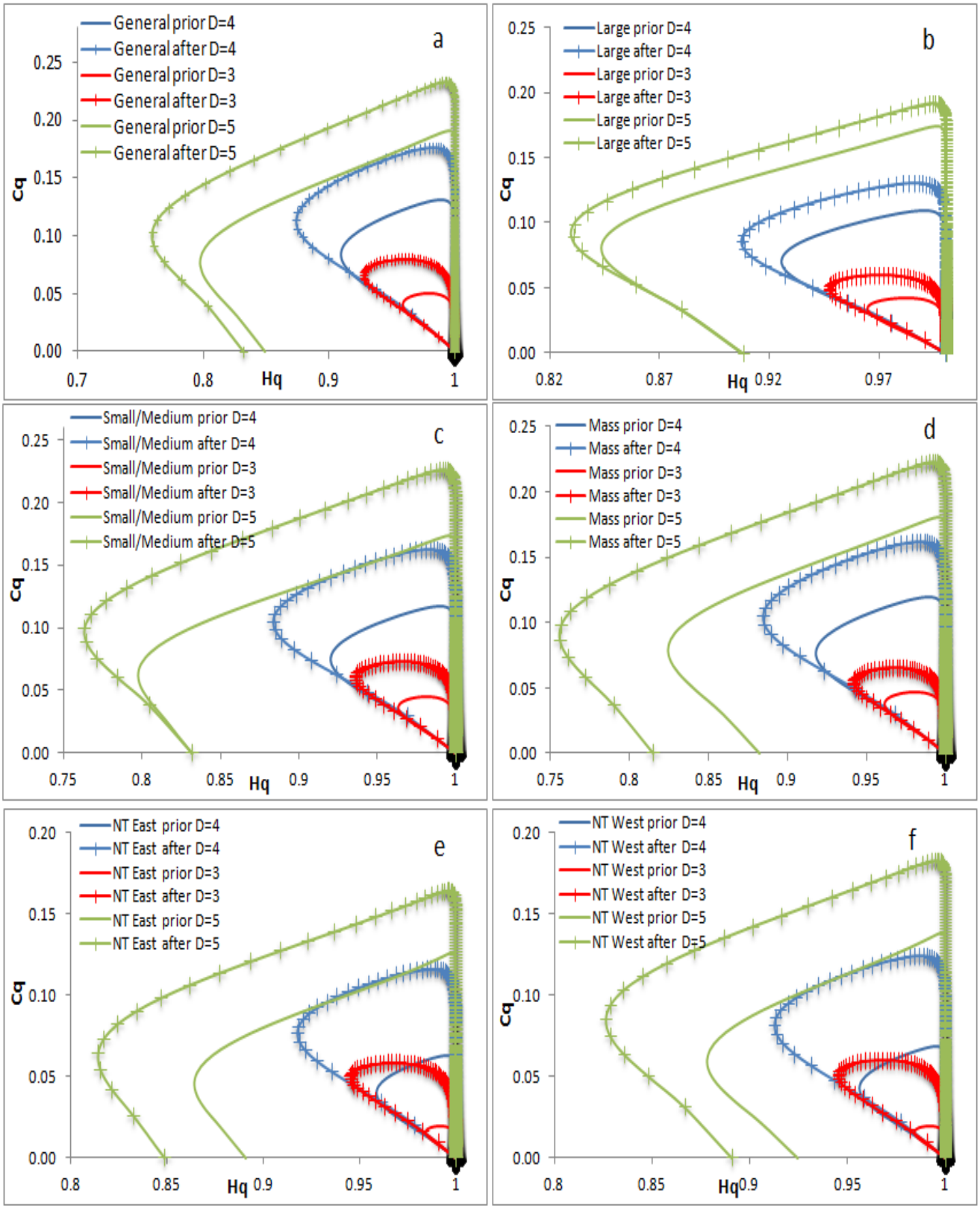


Figure 22: Dependence of the entropy *Hq* and complexity *Cq* on the parameter *q* for all indices, and embedding dimensions (D = 3, 4, 5). The star markers indicate the points (Hq ,Cq) for q = 0+, while the open circles are the same for q → ∞.

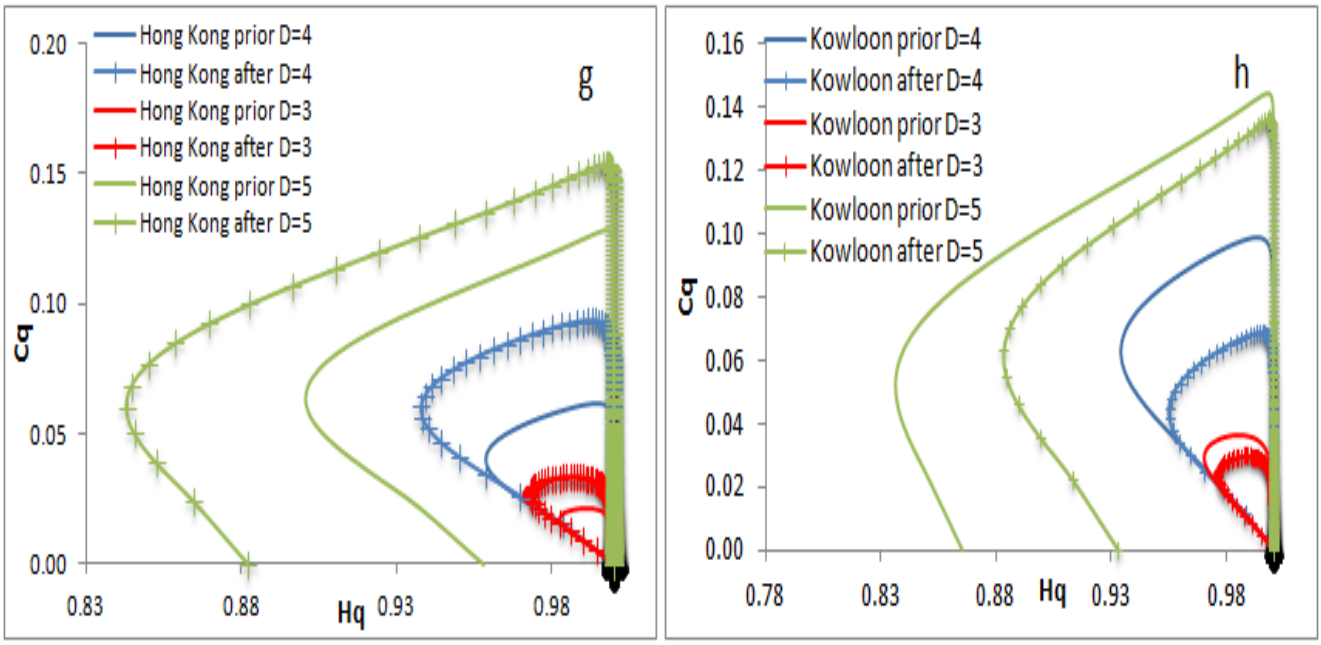


Figure 23: Dependence of the entropy *Hq* and complexity *Cq* on the parameter *q* for all indices, and embedding dimensions (D = 3, 4, 5). The star markers indicate the points (Hq ,Cq) for q = 0+, while the open circles are the same for q → ∞.

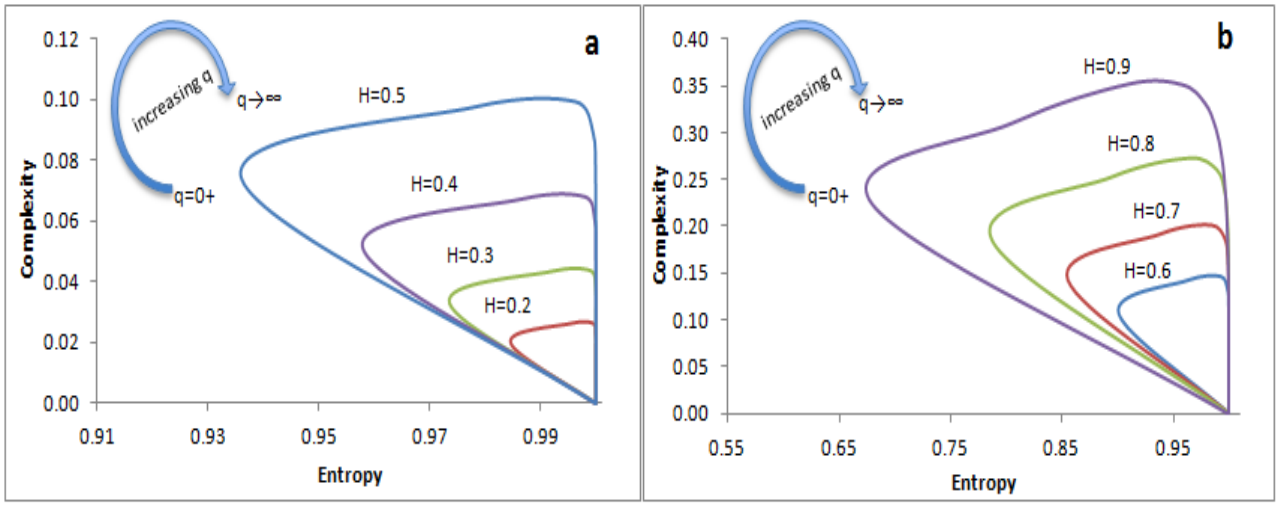


Figure 24: Panels (a) and (b) show the dependence of the entropy *Hq* and complexity *Cq* on the parameter *q* for the fractional Brownian motion with D = 4 and values of the Hurst exponent *h* = (0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9). The values of *q* are increasing (from *q* = 0+ to *q* = 1000 with 10^3 evenly spaced points between them).

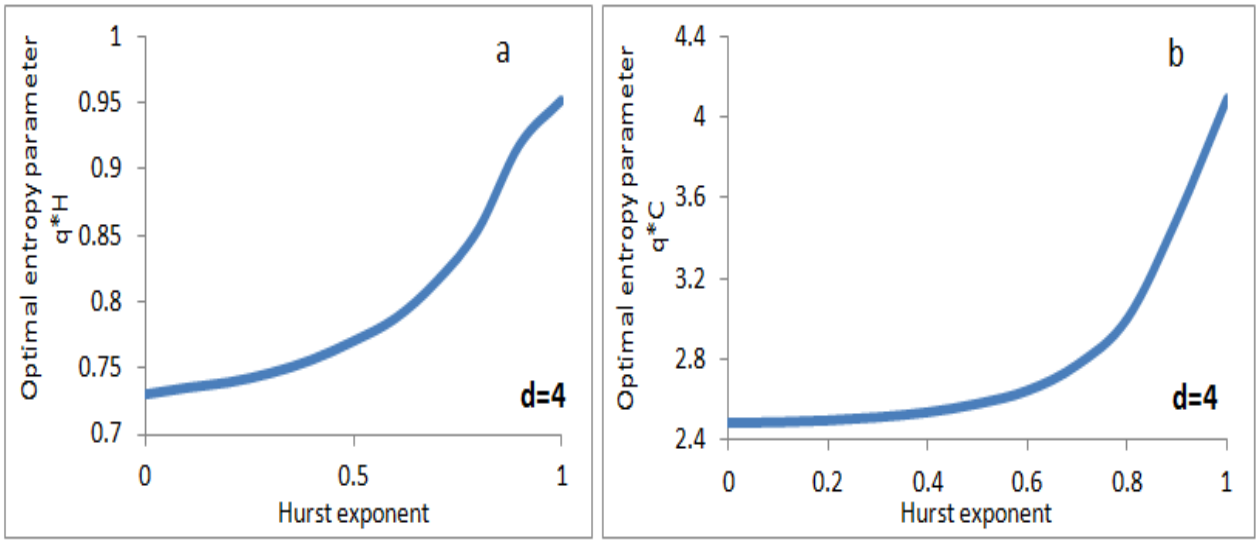


Figure 25: Values *q* H* and *q* C* obtained from the simulations and the results for the fractional Brownian motion. Panel (a) shows the values of *q = q* H* for which *Hq* reaches a minimum as a function of the Hurst exponent *h*, while Panel (b) shows the values of *q = q*C* for which *Cq* reaches a maximum as a function of the Hurst exponent *h* and *D = 4*.

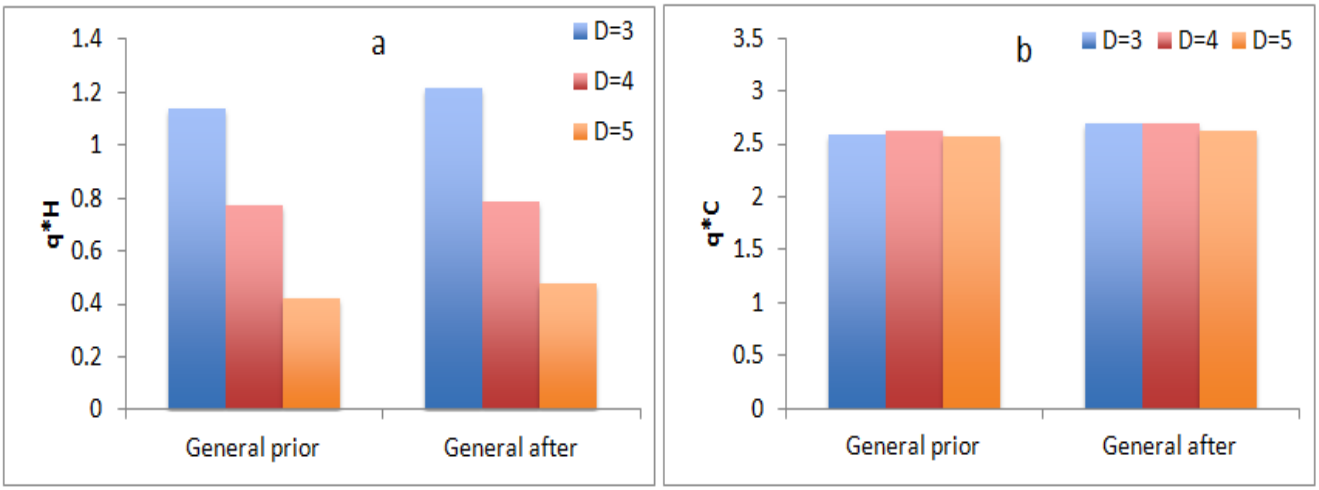


Figure 26: The extreme values *q* H* and *q*C* obtained for time series with *D=3, 4, 5*

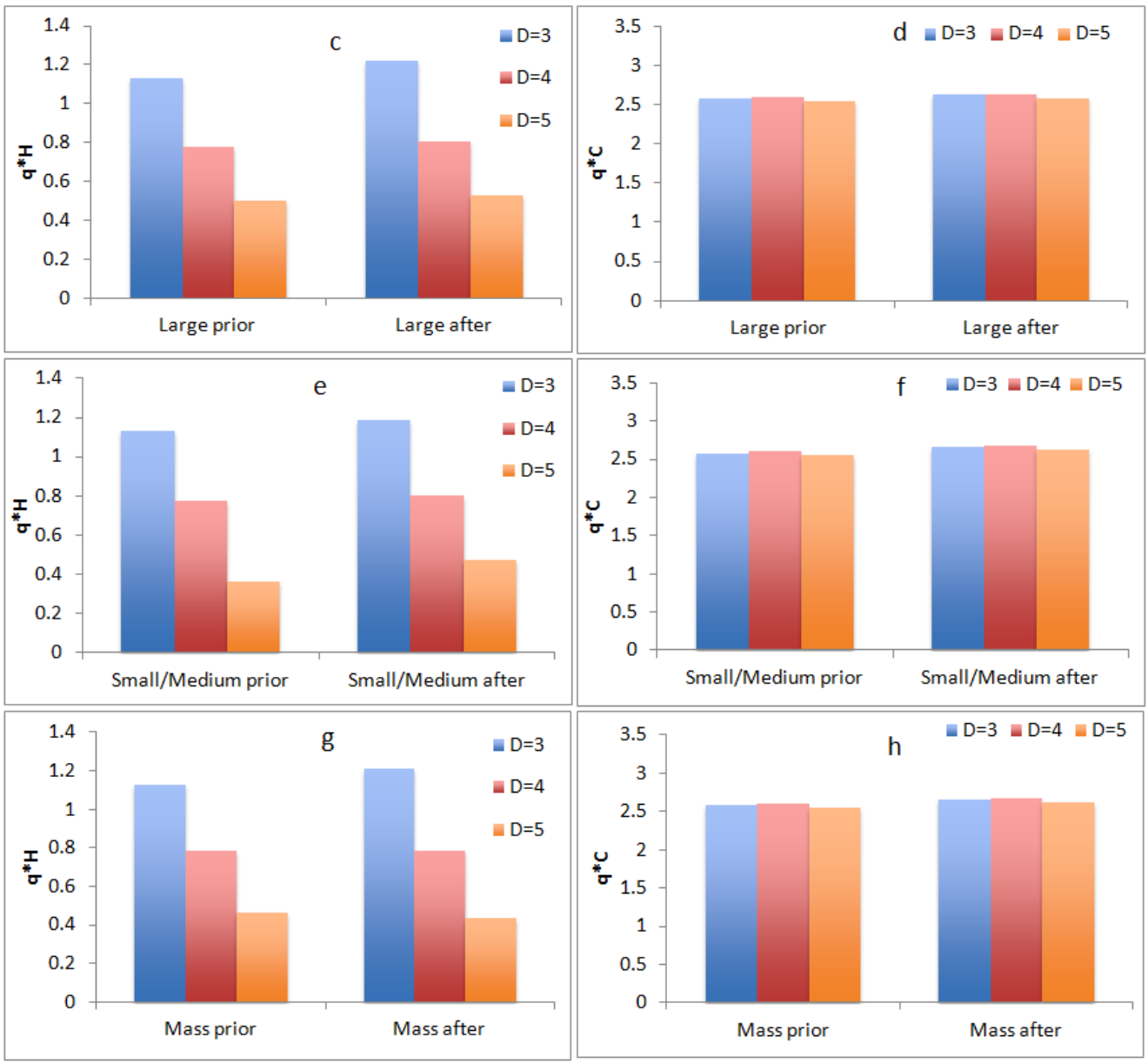


Figure 27: The extreme values *q* H* and *q* C* obtained for time series with *D=3, 4, 5*

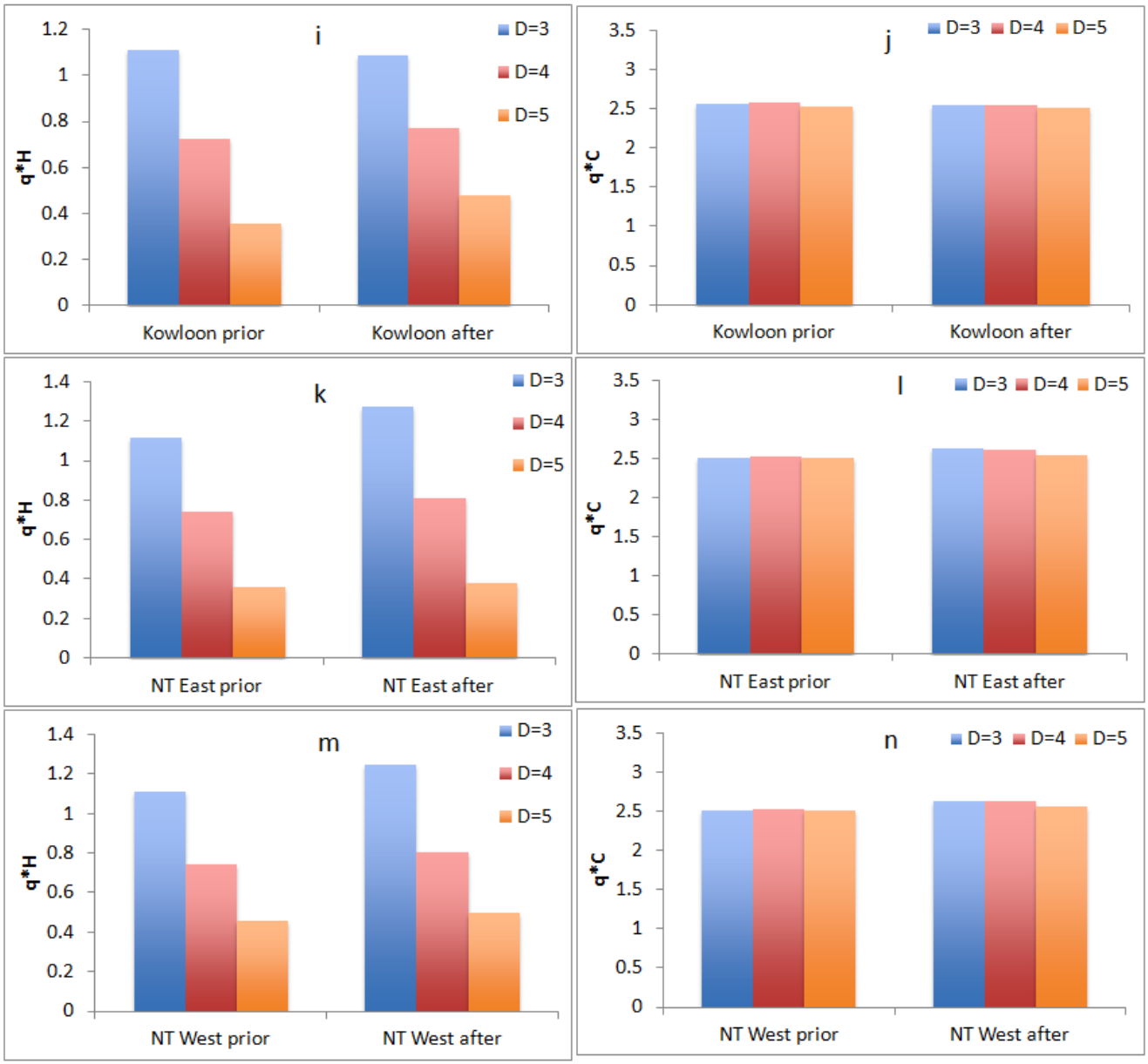


Figure 28: The extreme values *q* H* and *q*C* obtained for time series with *D=3, 4, 5*

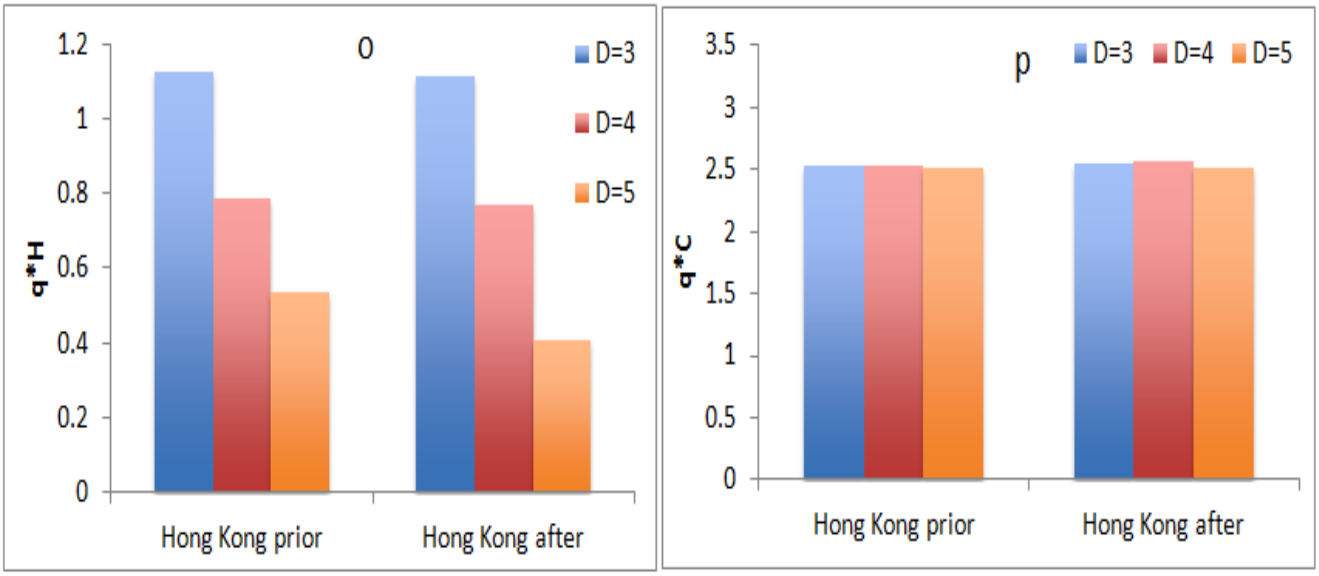


Figure 29: The extreme values *q* H* and *q* C* obtained for time series with *D=3, 4, 5*

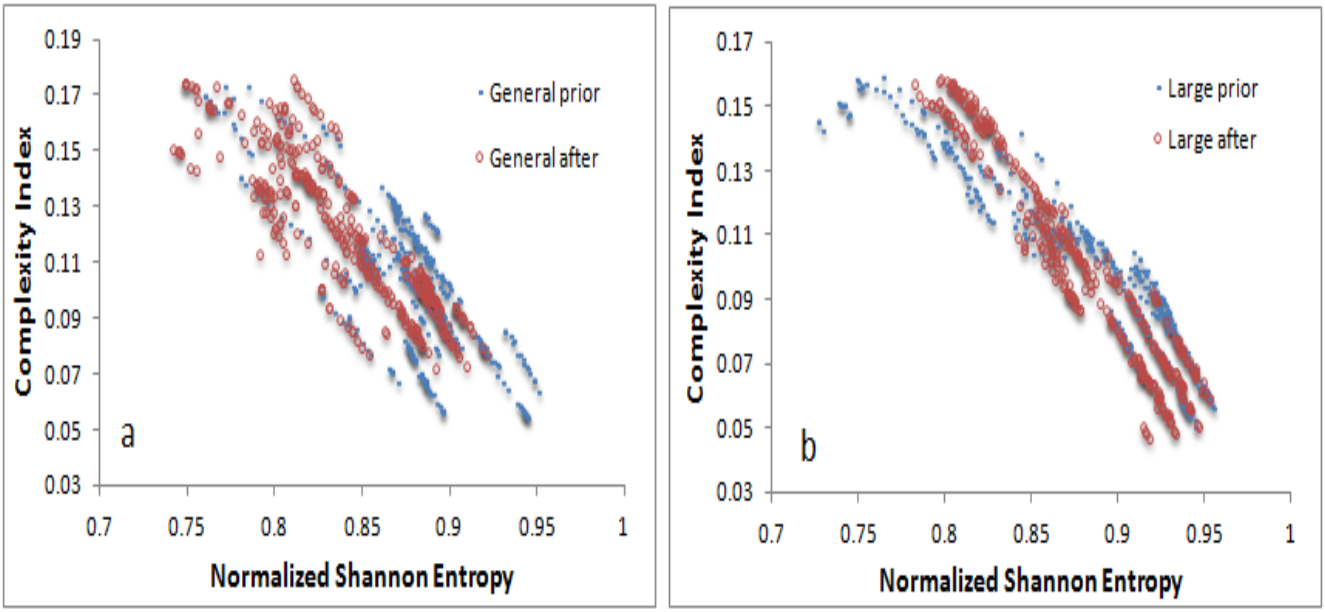


Figure 30: Complexity entropy causality plane based on Shannon entropy methods for different time periods. Movement of permutation quantifiers for prior and after indices with D=4, t=1 window

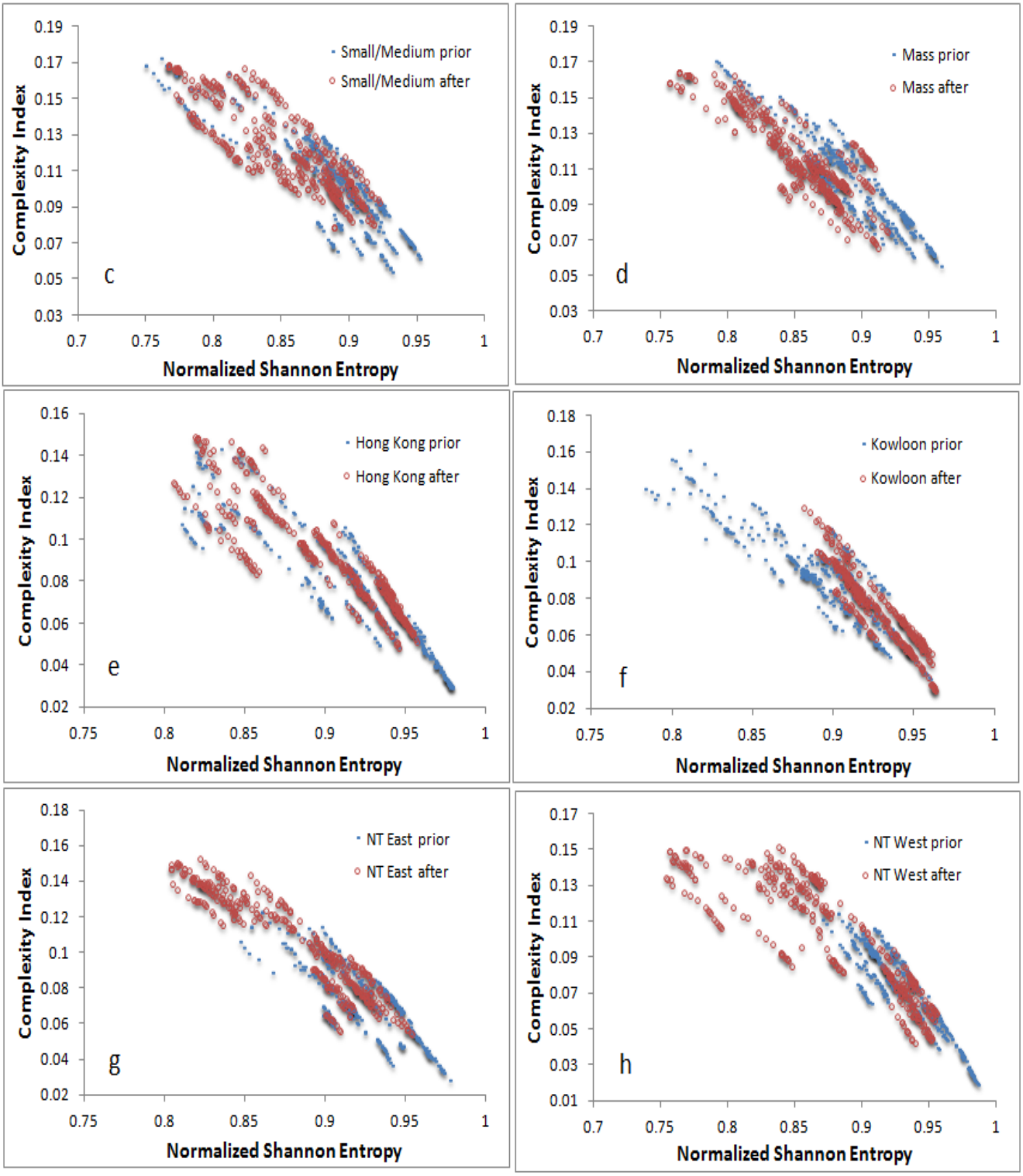


Figure 31: Complexity entropy causality plane based on Shannon entropy methods for different time periods. Movement of permutation quantifiers for prior and after indices with D=4, t=1 window

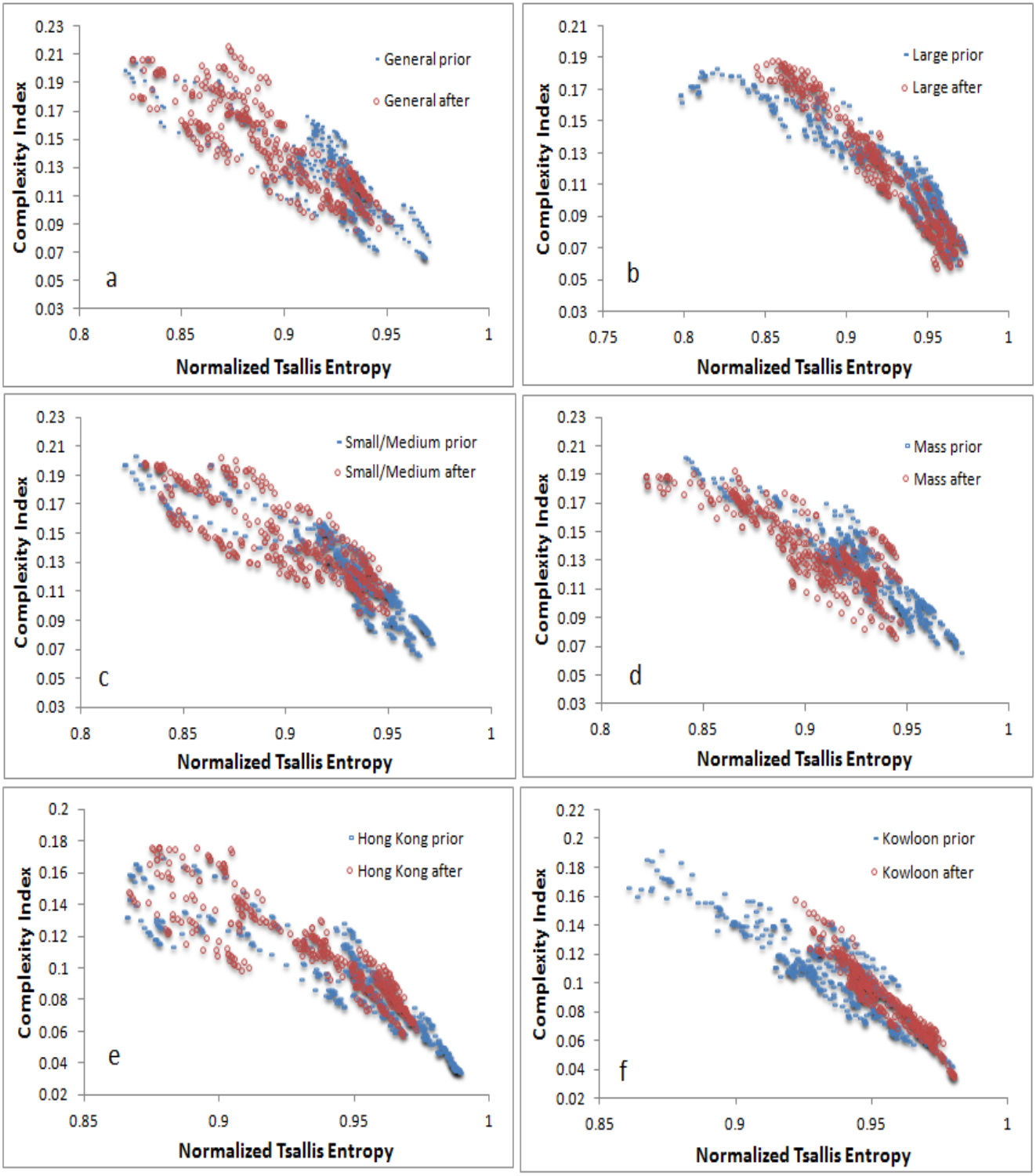


Figure 32: Complexity entropy causality plane based on Tsallis entropy methods for different time periods. Movement of permutation quantifiers for prior and after indices with D=4, t=1 window

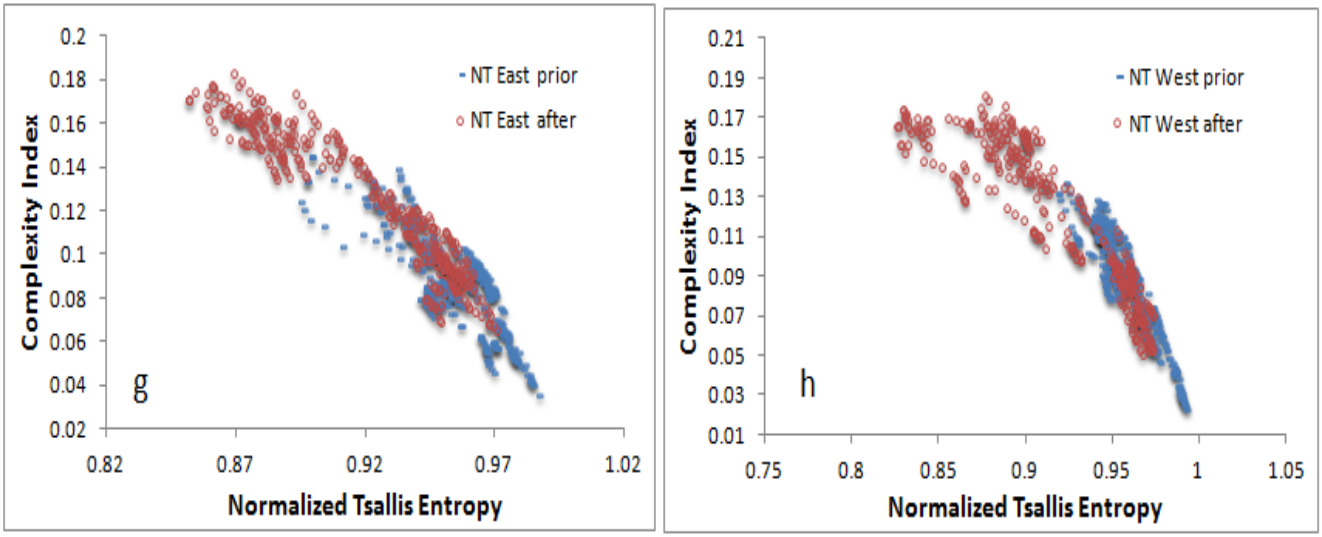


Figure 33: Complexity entropy causality plane based Tsallis entropy methods for different time periods. Movement of permutation quantifiers for prior and after indices with D=4, t=1 window

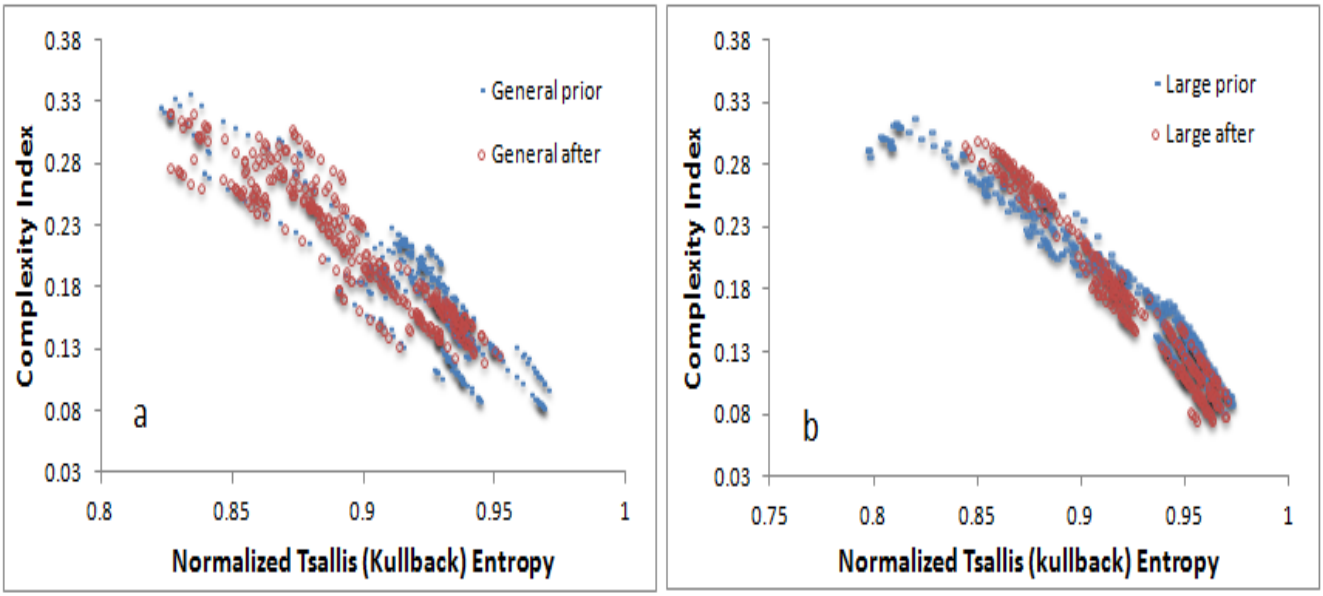


Figure 34: Complexity entropy causality plane based on Tsallis (Kullback) entropy methods for different time periods. Movement of permutation quantifiers for prior and after indices with D=4, t=1 window

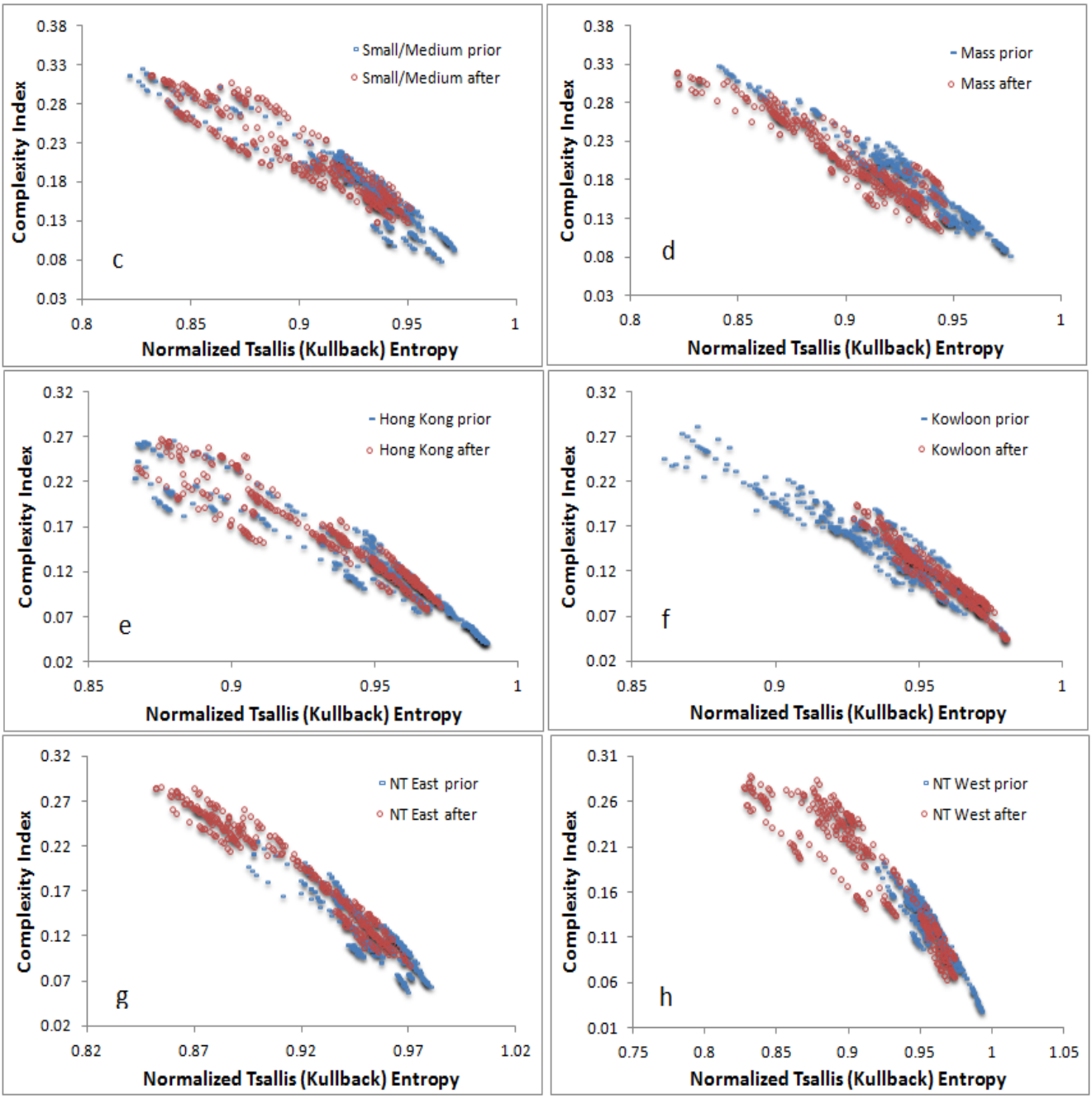


Figure 35: Complexity entropy causality plane based Tsallis (Kullback) entropy methods for different time periods. Movement of permutation quantifiers for prior and after indices with D=4, t=1 window.

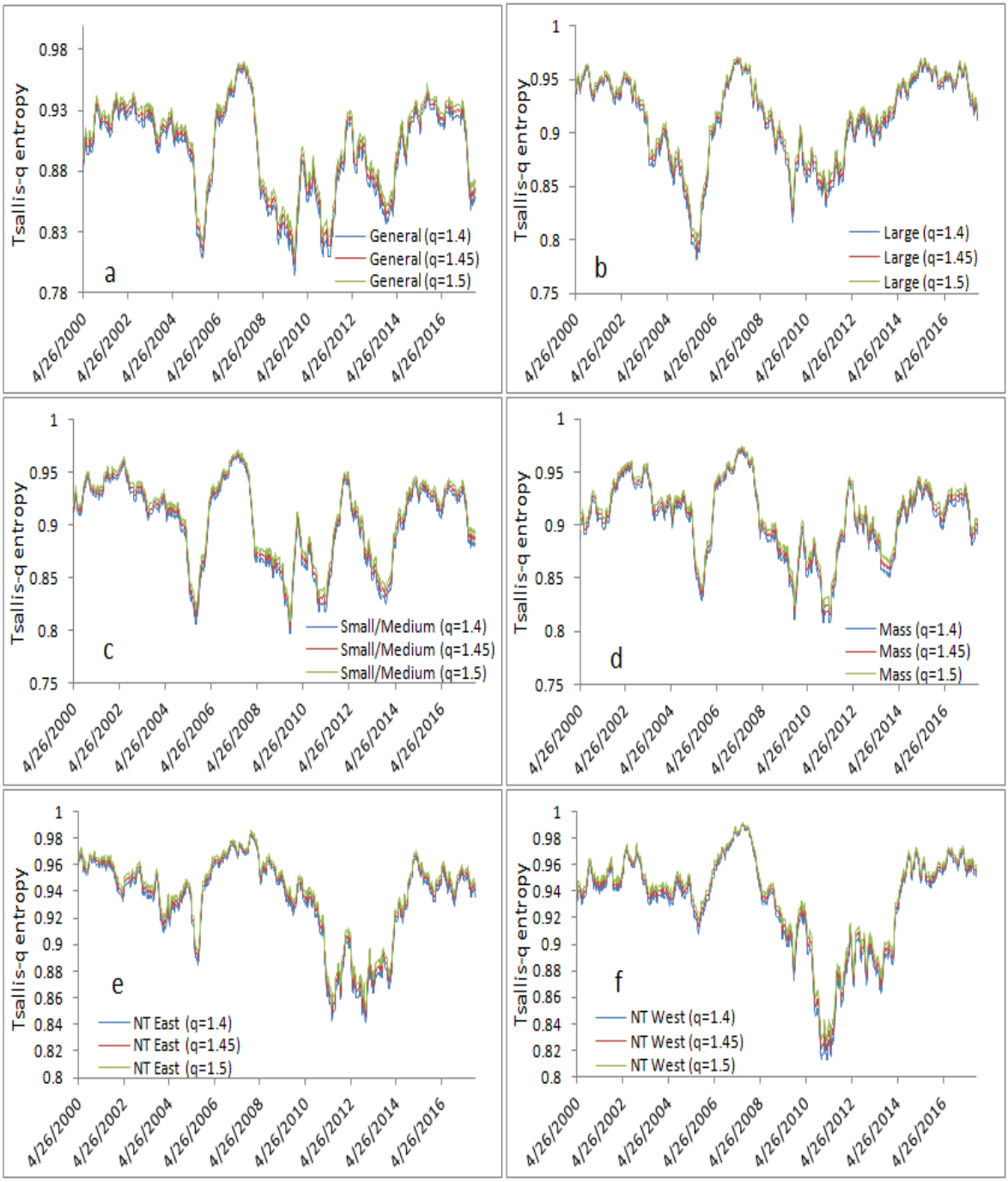


Figure 36: Tsallis-q entropy for all housing indices

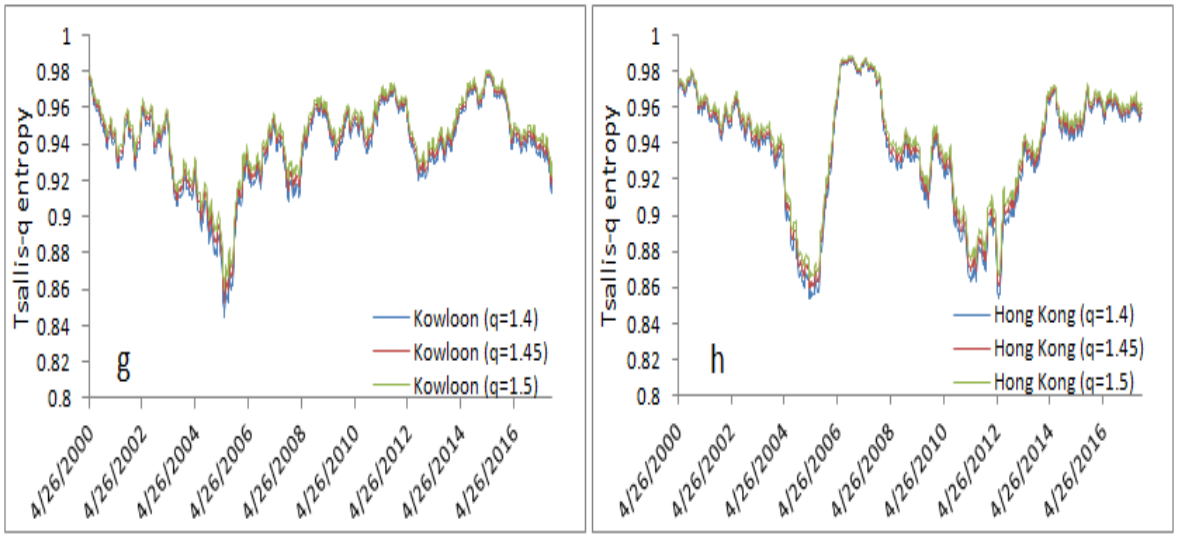


Figure 37: Tsallis-q entropy for all housing indices

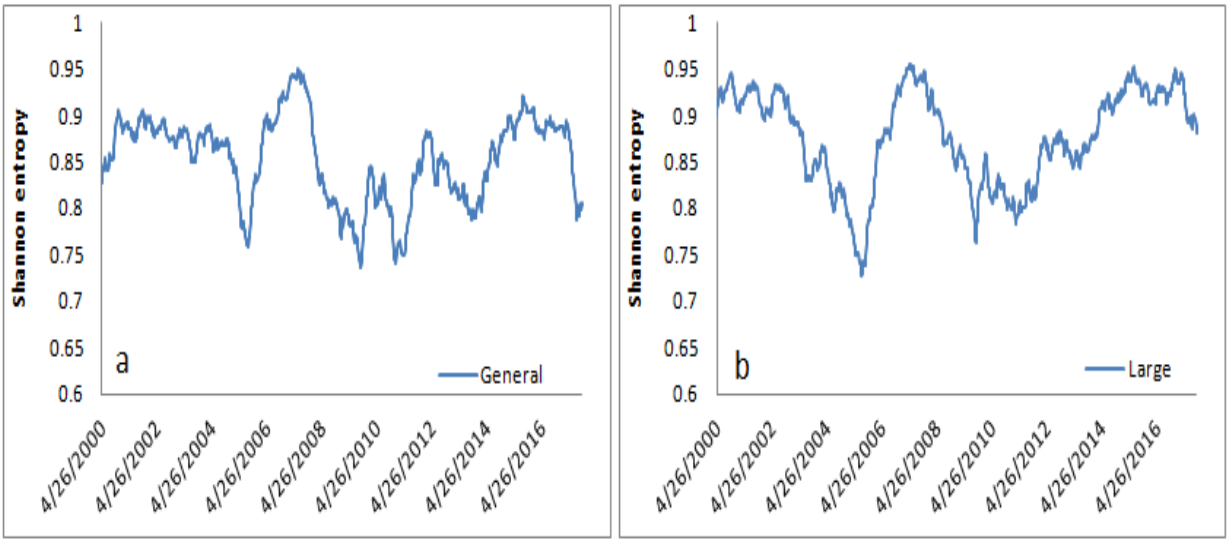


Figure 38: Shannon entropy for all housing indices

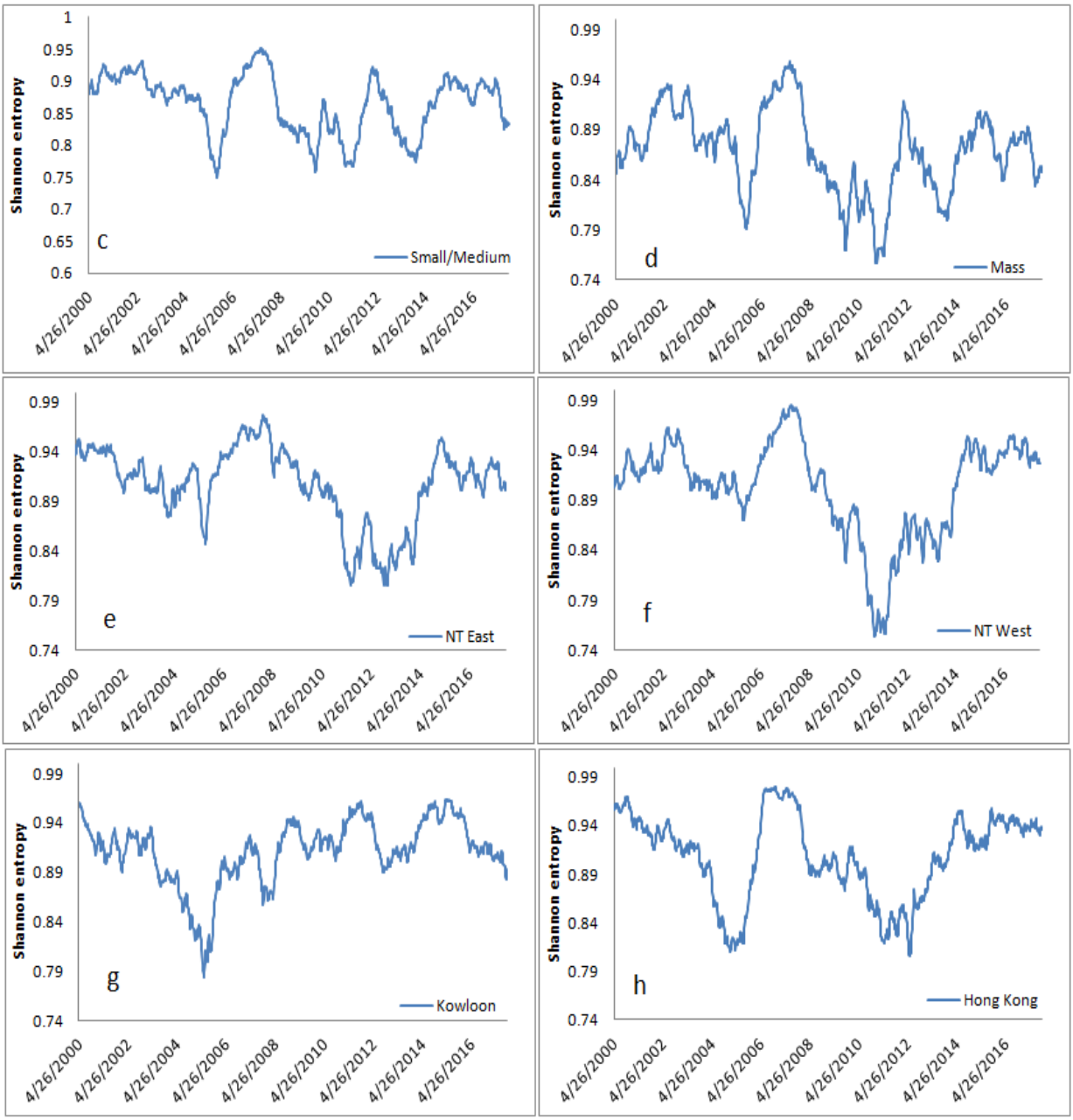


Figure 39: Shannon entropy for all housing indices

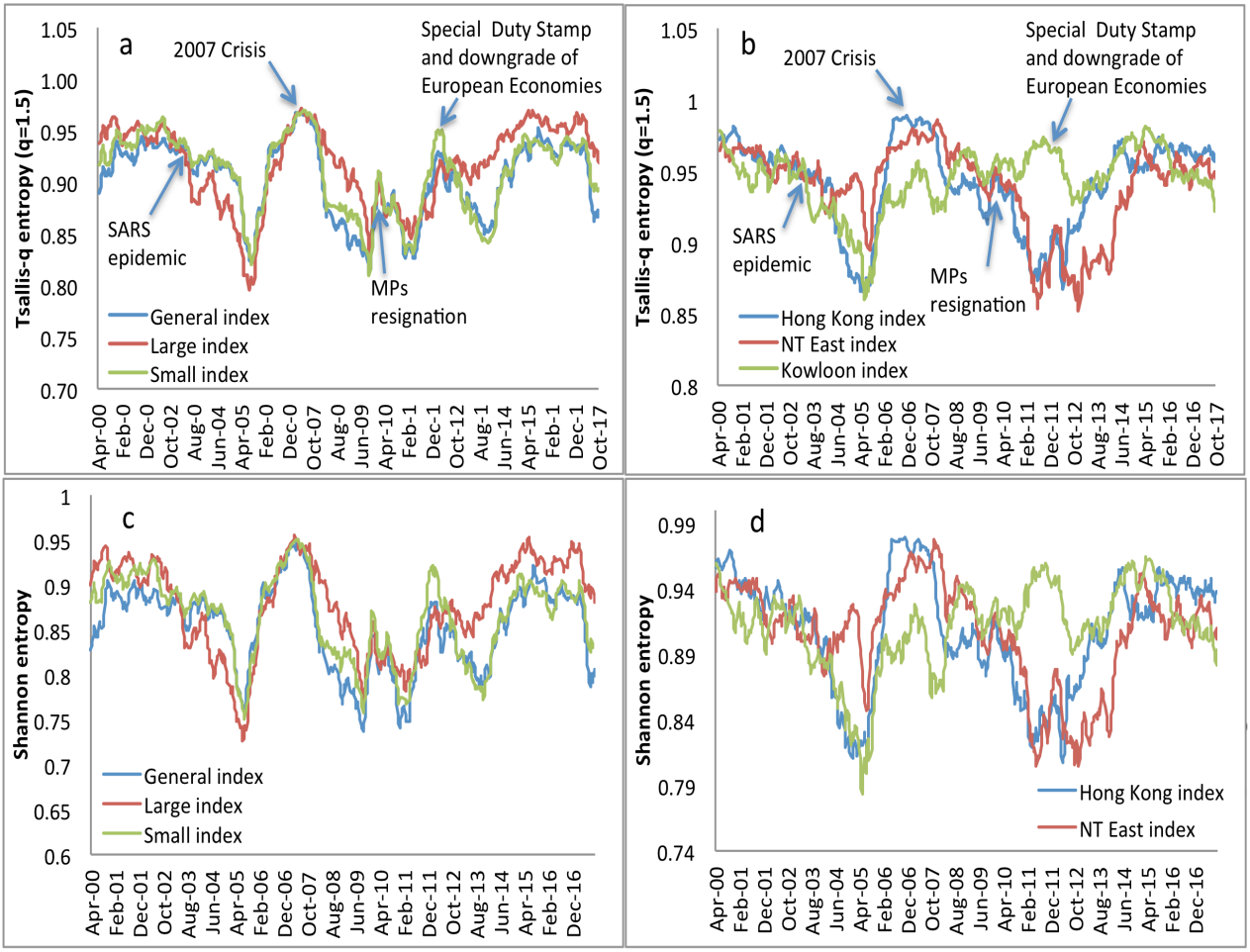
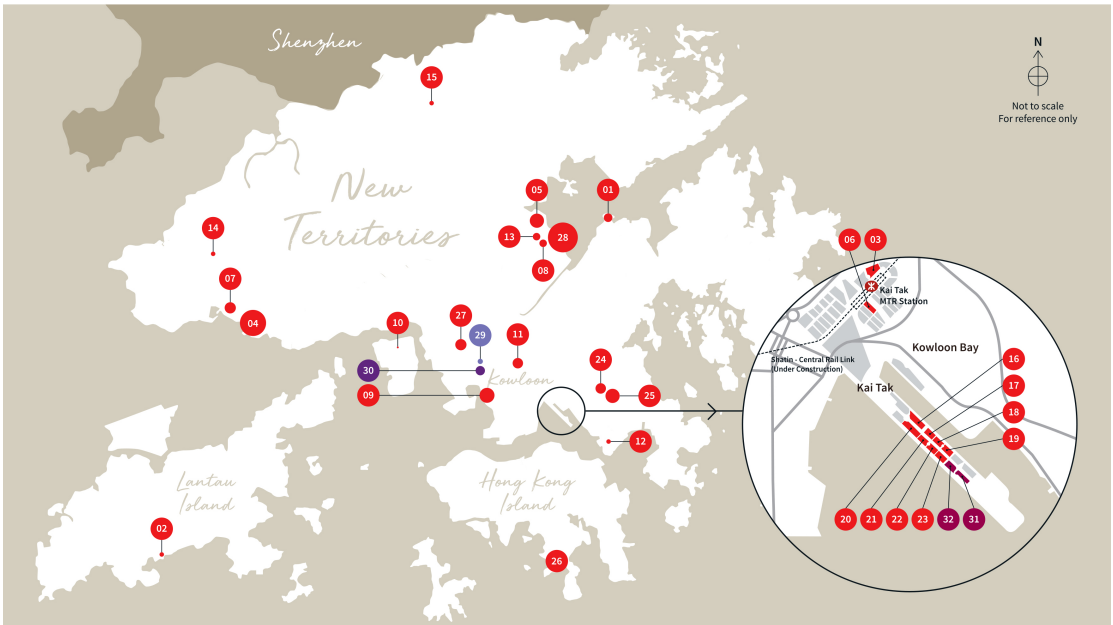


Figure 40: Time variation of Shannon and Tsallis entropy, and major events that took place in that period

Land Sale Programme

Sites to be released by the Hong Kong SAR Government
FY 2017/18 (April 2017 - March 2018)

JLL 仲量聯行



Residential Sites

			Site Area (sq ft)[1]
01	STTL 611	Whitehead, Ma On Shan, Sha Tin	49,700
02	Lot 763 in DD 332 Lantau	South Lantau Road, Cheung Sha, Lantau Island	26,700
03	NKIL 6568	Kai Tak Area 1F Site 1	178,200
04	TMTL 546	Castle Peak Road - So Kwun Wat, Area 48, Tuen Mun	292,800
05	TPTL 231	Yat Yu Avenue, Tai Po Kau, Tai Po	82,500
06	NKIL 6567	Kai Tak Area 1K Site 1	104,600
07	TMTL 520	Kwun Chui Road, Area 56, Tuen Mun	131,300
08	TPTL 234	Kon Hang, Cheung Shue Tan, Tai Po	67,300
09	NKIL 6549	Hing Wah Street West, Cheung Sha Wan	207,700
10	TYTL 192	Junction of Liu To Road and Hang Mei Street, Tsing Yi	14,900
11	NKIL 6579	Junction of Lung Cheung Road and Lion Rock Tunnel Road, Kowloon Tong	121,600
12	NKIL 6593	Ko Chiu Road, Yau Tong	39,800
13	TPTL 241	Tai Po Road – Tai Po Kau, Kon Hang, Tai Po	48,500
14	Lot 2883 in DD 130	Fuk Hang Tsuen Lane, Lam Tei, Tuen Mun	20,000
15	FSSTL 270	San Lok Street, Sheung Shui	2,000
16	NKIL 6577	Kai Tak Area 4A Site 1	176,500
17	NKIL 6576	Kai Tak Area 4B Site 1	103,100
18	NKIL 6575	Kai Tak Area 4B Site 2	97,400
19	NKIL 6574	Kai Tak Area 4B Site 3	104,500
20	NKIL 6554	Kai Tak Area 4A Site 2	197,000
21	NKIL 6553	Kai Tak Area 4C Site 1	102,000
22	NKIL 6552	Kai Tak Area 4C Site 2	105,200
23	NKIL 6551	Kai Tak Area 4C Site 3	117,900
24	Lot 1068 in SD 3	Anderson Road, Kwun Tong	57,600
25	Lot 1069 in SD 3	Anderson Road, Kwun Tong	114,100
26	RBL 1204	Cape Road, Stanley	322,900
27	KCTL 514	Wah King Hill Road, Kwai Chung	111,900
28	TPTL 244	Junction of Yau King Lane and Pok Yin Road, Pak Shek Kok, Tai Po	357,400

Business / Commercial Sites

			Site Area (sq ft)[1]
29	NKIL 6572	Junction of Wing Hong Street, Yu Chau West Street and Wing Ming Street, Cheung Sha Wan	30,900

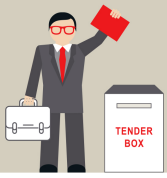
Hotel Sites

			Site Area (sq ft)[1]
30	NKIL 6550	Hing Wah Street West, Cheung Sha Wan	52,400

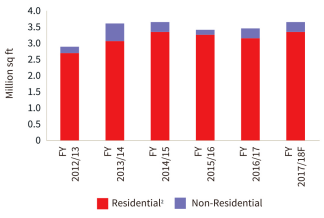
Hotel / Commercial Sites

			Site Area (sq ft)[1]
31	NKIL 6547	Kai Tak Area 4C Site 5	102,000
32	NKIL 6546	Kai Tak Area 4C Site 4	115,100

Note: [1] Site areas quoted (approx.) above are for reference only. Site details will be included in the respective Conditions of Sale.

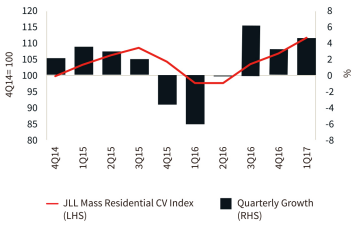


Site Area of Sites Disposed via Government Land Sales

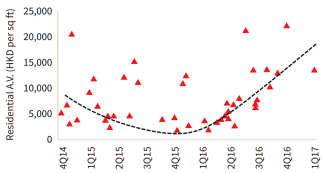


Note: [2] Including Mixed-Use with Residential components

Mass Residential Capital Values Growth



Residential A.V. Achieved in Government Land Sales



Note: Cut-off date 31 March 2017

Sources: Lands Department, Hong Kong SAR Government & JLL

Figure 41: Land Sale Program 2017-2018

	Pre crisis period				After crisis period			
Real Estate Ir	Hs Shannon	Hq Tsallis	Cjs Shannon	Cq Tsallis	Hs Shannon	Hq Tsallis	Cjs Shannon	Cq Tsallis
General	0.914	0.942	0.095	0.115	0.879	0.915	0.126	0.152
Large	0.929	0.952	0.080	0.096	0.911	0.937	0.096	0.115
Small/Mediu	0.924	0.949	0.086	0.103	0.888	0.920	0.117	0.141
mean	0.922	0.948	0.087	0.105	0.892	0.924	0.113	0.136
st. deviation	0.008	0.005	0.008	0.009	0.016	0.012	0.016	0.019
Mass	0.922	0.947	0.087	0.105	0.890	0.923	0.117	0.140
Hong Kong	0.961	0.974	0.046	0.055	0.941	0.961	0.068	0.082
Kowloon	0.939	0.961	0.074	0.088	0.958	0.972	0.050	0.061
New Territor	0.962	0.975	0.047	0.057	0.921	0.944	0.085	0.102
New Territor	0.959	0.974	0.050	0.061	0.916	0.940	0.091	0.109
mean	0.948	0.966	0.061	0.073	0.925	0.948	0.082	0.099
st. deviation	0.017	0.012	0.018	0.022	0.026	0.019	0.025	0.030

Table 6: Normalized Shannon and Tsallis permutation entropy and complexity for the Real Estate Indices in Hong Kong with τ = 1 and D = 4.

	Entropy			Complexity		
q values	D=3	D=4	D=5	D=3	D=4	D=5
General prior	1.138	0.773	0.422	2.593	2.620	2.569
General after	1.219	0.790	0.477	2.694	2.701	2.628
Large prior	1.130	0.779	0.500	2.575	2.591	2.547
Large after	1.222	0.801	0.528	2.631	2.641	2.574
Small/Medium prior	1.132	0.770	0.358	2.579	2.599	2.549
Small/Medium after	1.183	0.802	0.467	2.665	2.683	2.629
Mass prior	1.127	0.784	0.464	2.584	2.603	2.557
Mass after	1.208	0.786	0.437	2.647	2.680	2.624
Kowloon prior	1.113	0.724	0.355	2.559	2.571	2.519
Kowloon after	1.084	0.770	0.478	2.542	2.537	2.508
Hong Kong prior	1.125	0.784	0.533	2.524	2.536	2.512
Hong Kong after	1.112	0.772	0.409	2.551	2.564	2.521
NT East prior	1.119	0.742	0.359	2.520	2.531	2.504
NT East after	1.274	0.811	0.379	2.638	2.622	2.548
NT West prior	1.110	0.745	0.459	2.520	2.531	2.505
NT West after	1.246	0.805	0.495	2.638	2.629	2.570

Table 7: Values of q that optimize H_q and C_q for the Real Estate Indices in Hong Kong, with D = 3,4,5.

CHAPTER 4

4.1 Introduction

In the history of economics, the determination of the arrow of causality between two time series is an important subject of the literature, and the empirical findings of the transmission channels could support both the investors and the institutions. The goal of the causality analysis is to understand the influence of one variable to another, and some of the most characteristic examples in economics are the relationships between inflation and growth rate, inflation and GDP, exchange rate and stock prices, interest rate and stock market returns, returns and volatility, bond yields and stock prices, and finally advertising and consumption.

A very simple yet effective measure of causality in the paper of Sornette and Zhou (2005) is the lagged cross correlation function $C_{x,y}(\tau) = \langle x(t)y(t+\tau)\rangle / \sqrt{Var[x]Var[y]}$, where the brackets $\langle x\rangle$ denotes the statistical expectation of the random variable x. Then according to the theory, the knowledge of x during t could give information about the value of y at time t+τ. Nevertheless, we are still dealing with correlations that do not necessarily imply causation.

Furthermore, Granger causality tests present temporal relationships in terms of predictability. This method is similar to the recursive approaches and the reduced-form VARs. As stated in Becketti (2013), we can confirm whether the lagged values of a variable z can improve the forecastability of a variable x, whose lagged values have already taken into account. The mathematical form of Granger causality is $y_t = \alpha_0 + \sum_{i=1}^{k} \alpha_{t-i}\, y_{t-i} + \sum_{i=1}^{k} b_{t-i}\, x_{t-i}$. Since correlation does not imply causation, we should disclaim that the results are suggestive, but not definite (Sargent, 1976). For more information about Granger causality, please refer to (Chen, Rangarajan, Feng, & Ding, 2004), (Ashley, Granger, & Schmalensee, 1980).

Fact of the matter is that neither of the two previous presented methods gives an answer to the real causality links between two or more variables, and in order to conclude about the real causality relation it is important to distinguish between exogenous and endogenous variables. The former can be causal variables only. The methodology applied here could be part of the same group of causality testing like

cross-correlation and Granger causality, because it does not address the notion of endogenous versus exogenous variables.

There are some other methods developed in the physical community as well. Based on event synchronization, Quiroga et al. (2002) proposed a fast and simple method to measure time delay patterns and synchronization between two time series. Similar to the notion of the recurrence plots, which help for the diagnosis of dynamical systems are (Eckmann, Kamphorst, & Ruelle, 1987), (Marwan & Kurths, 2002), and (Marwan, Thiel, & Nowaczyk, 2002). These papers presented the cross recurrence plots (CRP), which enables the analysis of time differences or synchronization in two time series.

In this part, we utilize a new method called the 'optimal thermal causal path', a non-parametric method capable of detecting a priori arbitrary non-linear dependence structures. It is considered as an adaptation to the time evolution of the dependence structure. This method extends the notion of 'time distance' measure, which compares trend lines upon two time series horizontal differences (Granger & Yeon, 1997). The 'optimal thermal causal path' supports the already existed Granger causality method, because this new method of dependences detection between two time series actually extends the standard lagged-correlation method.

According to Sornette and Zhou (2005), the optimal thermal causal path is deduced from the field of computational statistics, starting from the problem of recognizing related sequences, see (Needleman & Wunsch, 1970). A general computer method was introduced in order to find matches in the amino acid sequences of two proteins. Thus, the two authors could assess whether the proteins presented significant homology, and trace the possible evolutionary development of the proteins. Sankoff and Kruskal (1983) explain in their book how to measure distance among sequences and compute it effectively, and how to identify related sequences.

Application examples are found in speech recognition, molecular biology, error correcting in computer software and analysis of bird songs. Similar to our approach is the concept of 'time-warping', presented in the second part of Sankoff and Kruskal (1983). In this technique we compare functions by altering the time axis that has been developed for speech recognition (Myers, Rabiner, & Rosenberg, 1980). Dynamic Time Warping (DTW) has been applied to data mining (Keogh & Pazzani, 2000), (Yi, Jagadish, & Faloutsos, 1998), and (Berndt & Clifford, 1994), robotics

(Schmill, Oates, & Cohen, 1999), gesture recognition (Gavrila & Davis, 1995), medicine (Caiani, et al., 1998), manufacturing (Gollmer & Posten, 1995), and speech processing (Rabiner & Juang, 1993). More information is also in (Shinde & Pawar, 2014), (Keogh & Pazzani, 2001), (Silva & Batista, 2016), and for the implementation of global constraints in DTW, please see (Itakura, 1974), and (Sakoe & Chiba, 1978).

Furthermore, we should mention a series of papers by Hwa and co-workers (Hwa & Lassig, 1996), (Drasdo, Hwa, & Lassig, 1998), (Hwa & Lassig, 1998), (Olsen, Hwa, & Lassig, 1999), (Drasdo, Hwa, & Lassig, 2001). Similar to our plan about time series and using concepts of statistical physics, they focused of how to detect similarities between long DNA and protein sequences. While they take into account that there are four basic units in DNA sequences to detect similarities, our aim is to recover the lag among two time series in a dynamical way.

As we will explain in detail later, our goal is to normalize first the time series dataset, and create a distance matrix that contains every possible distance pair of increments of the time series given. Hwa and co-workers above study the results of noise and energy landscapes between two similar time series, and showed the 'fidelity' of the alignment, which is a localization-delocalization phase transition (as a function of a control parameter). In real estate time series, the well-matched data would apply to a localized optimal path, meaning that there are bounded fluctuations of the lag, while the optimal path is delocalized if the data are not well matched. In the findings presented below, we support the idea of adding a temperature in the optimal paths, mainly because we believe that it provides robustness to the results, and also it helps to smooth the idiosyncratic noises of the time series.

As far as the Granger causality is concerned, numerous papers are under consideration. Granger causality tests were applied by Myer & Webb (1994), where they find that the EREIT returns 'caused' the NCREIF direct property returns. Also in Barkham and Geltner (1995), price discovery in the USA and UK is under consideration. In their findings there is price discovery from the indirect to the direct market in the UK. On the other hand, price discovery in the US market seems to be weaker, since price information does not fully transmit for a year or more from the securitized real estate market.

As a result, the securitized market in the US is not as linked to the direct investments as the UK securitized market. Granger causality is also used to investigate the existence of housing market bubbles, see (Shen, Hui, & Liu, 2005) or

the habit formation in consumption (Fuhrer, 2000). In the paper of Chen (2001) the observed price fluctuation between 1973 and 1992 showed that equity prices do Granger cause the Real Estate assets. Furthermore, Kapopoulos and Siokis (2005) found that an increase in Real Estate prices could stimulate stock market prices. This transmission is tested through channels of Granger causality.

Extensions of Granger causality can be found in Hristu-Varsakelis and Kyrtsou (2008). Instead of the linear VAR specification, the authors used a noisy Mackey-Glass model and the appropriate returns, in order to test for "assymetric" causality. The results confirmed that the assymetries of positive and negative returns give a significant predictive power to the Mackey-Glass model.

Furthermore, Hristu-Varsakelis and Kyrtsou (2010) proposed a bivariate noisy Mackey-Glass model to find causal relationships in macroeconomic and financial data, when complex dynamics like chaotic non-linearities are present. Their applied test, which is similar to Granger causality aims to capture forms of causality that the latter cannot. The results showed the effectiveness of the test even when the generated dynamics are a combination of stochastic and chaotic series. In the paper of Panana et al. (2017), the authors found the interrelationships between 21 international stock indexes. For comparative reasons the standard linear Granger causality (CGCI) and the partial mutual information on mixed embedding (PMIME) were used. Results suggest that PMIME outperforms CGCI, while both methods verify the strong influence of the US stock market over the rest indexes, and its ability to predict how other global equity markets perform.

Also, in many empirical financial studies the main goal was to identify long term correlations in single or multiple time series data (Granger & Ding, 1996), (Ding, Granger, & Engle, 1993), (Liu, Gopikrishnan, Cizeau, Meyer, Peng, & Stanley, 1999), (Podobnik, Wang, Horvatic, Grosse, & Stanley, 2010), (Wang, Podobnik, Horvatic, & Stanley, 2011). Additional sources of the theory of Granger causality are the papers of (Chau, Macgregor, & Schwann, 2001), (Granger & Newbold, 1974), (Engle & Granger, 1987), (Granger, 1969).

4.2 Methodology

The thermal optimal causal path was initially presented in 2005 by Sornette and Zhou (2005), Zhou and Sornette (2006). The authors used the method to find the lead-lag dependence between the US Treasury bond yields and the US stock market for three years (2000 to 2003). In the same paper, they also revealed non-trivial lag relationship between inflation change, inflation, unemployment rate and GDP growth. The results confirmed their previous study of the stock market preceding the yield rates at short maturities, as well as the Federal Reserve Funds' adjustments (Zhou & Sornette, 2004). Professor Zhou and Sornette applied the thermal optimal path with adjustments in the distance matrix finding a more effective path capable of recognizing possible regime shifts between negative and positive correlations (Zhou & Sornette, 2007). Their results confirmed that this method could perform even if the data contain substantial noise.

By using the economic growth rate and the US inflation rate, they concluded that the new TOP method could outperform standard cross-correlation methods. What is more, Hu et al. (2017) explored the lead-lag relationship between the offshore CNH and onshore CNY exchange rates during the period 2012-2015. In the analysis of daily data we see that the rates have a weak lead-lag relationship that depends on the prevailing market factors. They also found that the minute-scale interaction pattern does not remain constant over time because of market factors like the US dollar appreciation. More specifically, the former is connected with an offshore to onshore lead-lag relationship, while a Renminbi appreciation is associated with the opposite relationship.

Another TOP method is applied in 2011 by Guo et al. (2011), where the lead-lag relationship between the S&P500 stock market index and the US short-term and long-term maturities like Federal Funds Rate, 3 month and 20 year maturities is presented. The research included both monthly and weekly data. Findings confirmed that yield changes and stock market variations are correlated, and the central bank rate changes being a proxy of the monetary policy cannot be implied as a stock market predictor.

In addition, Meng et al. (2017) presented another alternative of thermal optimal path, named symmetric thermal optimal path (TOPS) to define the time-

dependent lead-lag relationship of two time series, focusing in periods under economic stress. This modified TOP method was designed to eliminate various inconsistencies of the previous method by imposing that even after a change of sign the new relationship should be invariant regarding a time reversal of the time series. The authors took for the 1991-2011 period the lead-lag relationship of the housing market and monetary policy of the United Kingdom and the United States. They resulted a leading housing price index over interest rate until the Great Recession. Interestingly after the crisis the roles changed and the central banks lead and causally influenced the housing market, presenting the importance of a change of regime.

The One Belt and One Road Exchange Rate index (OBORR) is under investigation along with the RMB Effective Exchange Rate index (CNYX) (Lai & Guo, 2017). In this paper, it is concluded that in the medium term there is a dynamic lead-lad relationship, while in the long term OBORR has higher increasing velocity than CNYX. Another research with the use of TOP was conducted in daily and one-minute returns by Wang et al. (2017). The lead-lag relationship between CSI 300 index future and spot markets between 2010-2014 revealed for the daily data case that there is a volatile relationship between the future and spot markets and no dominant place is found. In the case of one-minute returns, the cash return lags the future return by 0-5 minutes, irrespective of the price trend of the market.

Considering the dynamical evolution and the relationship between stock index futures and stock index, the research of Gong et al. (2016) analyzes the HSI, S&P500, and CSI 300 stock index with their future indices respectively. The results support that HSI and S&P500 stock index futures lead their stock indices, while the opposite result was found in the case of CSI 300, an index of a developing market.

4.2.1 Thermal Optimal Path

The key ideas behind the optimal thermal causal path method can be summarized as follows.

(a) The formation of a distance matrix is the comparison of the first time series values $x(t_1)$ with all the values of the second time series $y(t_2)$, introducing the distance $d(x(t_1), y(t_2))$.

(b) The dependence relationship concerning the two time series is examined in the

form of a one-to-one mapping $t_2 = \varphi(t_1)$ between the times of the first time series and the times of the second time series in a way that the two time series are the closest in a certain sense, for instance to check if $X(t_1)$ and $Y(\varphi(t_1))$ match best. Additionally, we apply a kind of smoothness requirement, which is equivalent in the majority of cases to continuity and monotonicity of the mapping φ. Nevertheless, the 'optimal thermal causal path' method allows us to find situations where the lag can jump and perform in an arbitrary way as a function of time.

(c) A weighted average over many potential mapping is introduced in the previous step, so as to eliminate potential influence from non-informative noises in the time series. There is an exact mapping of this problem to a well-known problem in statistical physics known as the directed polymer in a quenched random potential landscape at non-zero temperature, hence the name 'optimal thermal causal path'. For more information, see (Wang, Havlin, & Schwarz, 1999)

(d) The final mapping shows the lag between the time series under analysis as a function of time that matches them in the best way possible. Thus, we are allowed to figure the time evolution of the special dependence relationship among the two time series. In the next part, we will explain in more detail the implementation of the ideas above.

4.2.2 Distance Matrix

We start by considering the two input time series data. They should be presented in discrete time, in units of some elementary discretization step, taken as unity without loss of generality. We apply the first time series as $\{X(t_1): t_1 = 0, \dots, N_1 - 1\}$ and the second time series as $\{Y(t_2): t_2 = 0, \dots, N_2 - 1\}$. We disclaim that the two time series can have different length, as the results will not be affected, nevertheless we present the case where $N_1 = N_2 = N$. Although the time series used as an input can differ in nature with large meanings and differing units, we still can use them by adjusting them with the use of normalization in order to become comparable. In this Thesis, $\{X(t_1)\}$ and $\{Y(t_2)\}$ will always denote normalized time series.

We introduce a distance matrix $E_{X,Y}$ between X to Y with elements defined as

$$\epsilon(t_1, t_2) = |X(t_1) - Y(t_2)| \quad (27)$$

The value $|X(t_1) - Y(t_2)|$ describes the distance between the realization of the first time series at time t_1 and the realization of the second time series at time t_2. Even though we can consider other distances as well, the method applied in this research applies without modifications in the distance value above. But this is not always the case, as the nature of the time series given could demand to put more weight on large differences $|X(t_1) - Y(t_2)|$ by taking $|X(\square_1) - Y(t_2)|^q$, where $q > 1$. As we aforementioned, we use the discrepancies where $q = 1$.

For clarity, if $X(t_1)$ is the same as $Y(t_2)$, then in a matrix created using (27), we introduce a threshold so that entries of the matrix (27) larger (respectively smaller) than the threshold are set to 1 (respectively 0), has been introduced under the name of 'recurrence plot' to examine complex chaotic time series (Eckmann, Kamphorst, & Ruelle, 1987). The binary matrix deduced from (27) using a threshold for two different time series is called in the physical literature a cross-recurrence plot. This matrix and some of its statistical properties have been used for the characterization of the cross-correlation structure between time series (Marwan & Kurths, 2002), (Marwan, Thiel, & Nowaczyk, 2002), (Quiroga, Kreuz, & Grassberger, 2002), (Strozzia, Zaldvarb, & Zbilutc, 2002).

Also, consider an example where $Y(t) = X(t - k)$ with $k > 0$ is a positive constant. Then, $\epsilon(t_1, t_2) = 0$ for $t_2 = t_1 + k$ and is not zero otherwise. By detecting this dependence relationship we find the line with zero values parallel to the diagonal of the distance matrix. This line defines the affine mapping $t_2 = \phi(t_1) = t_1 + k$, corresponding to a constant translation. Our next step is to determine the smallest in value sequence of elements of this distance matrix, as we describe below.

4.2.3 Optimal path at 'zero temperature'

If the relationship is not as simple as the one presented above, and we have something more complex than a naive lead-lag equation $Y(t) = X(t - k)$, then determining the correspondence between $X(t_1)$ and $Y(t_2)$ is not obvious. One first

strategy would be to associate each entry of $X(t_1)$ to every value of $Y(t_2)$, assessing the case where the distance (27) is minimum over all possible t_2 for a fixed t_1. This defines the mapping $t_1 \rightarrow t_2 = \phi(t_1)$ from the t_1-variable to the t_2-variable as

$$\boldsymbol{\phi(t_1) = Min_{t_2}|X(t_1) - Y(t_2)|} \tag{28}$$

In this method each t_1 is independently analyzed. Nevertheless, the main issue with this method is the fact that there are two undesirable properties in the mapping production: (i) numerous large jumps, and (ii) absence of one-to-one matching, which can be regarded as a backward time propagation. More specifically, (i) explains that in the case that there is noise present, the $\phi(t_1 + 1) - \phi(t_1)$ in some values of t_1 is large and of the order of the total duration N of the time series. More specifically, large jumps in the function above are not reasonable. The other property defines that, given large probability, a t_1 might be associated with various t_2, and thus we will have pairs of times where $t_1 < t_1'$ such that $\phi(t_1) > \phi(t_1')$: an event in the future in the first time series is linked with an event in the past in the second time series. It is obvious that the two aforementioned properties do not permit the use of (28) as an ideal construction to correspond the two time series.

The solution to these problems could be solved with the use of a smooth mapping $t_1 \rightarrow t_2 = \phi(t_1)$:

$$\boldsymbol{0 \leq \phi(t_1 + 1) - \phi(t_1) \leq 1} \tag{29}$$

In the continuous time limit, (29) imposes that ϕ should be continuous. Then, the correspondence $t_1 \rightarrow t_2 = \phi(t_1)$ can be interpreted as a time-lead structure of $X(t_1)$ and $Y(t_2)$. For another application it would be desirable to impose more restrictions by ensuring the differentiability of the mapping. This can be done by adding a path 'curvature' energy term; thus generalizing the global optimization problem presented below. Nevertheless, it is out of the scope of this research and our interest. Next, we search the dependence relationship between $X(t_1)$ and $Y(t_2)$ in the form of a mapping $t_2 = \phi(t_1)$, such that $X(t_1)$ and $Y(t_2)$ are the closest in a certain sense, for example $X(t_1)$ and $Y\big(\phi(t_1)\big)$ match taking the two constraints into account.

These ideas are implemented with the proposal of Sornette and Zhou (2005), replacing the mapping (28) by a mapping derived from the global minimization:

$$\boldsymbol{Min}_{\{\boldsymbol{\phi}(\boldsymbol{t_1}),\ \boldsymbol{t_1}=\boldsymbol{0},\boldsymbol{1},\boldsymbol{2},\ldots,\boldsymbol{N}-\boldsymbol{1}\}} \sum_{\boldsymbol{t_1}=\boldsymbol{0}}^{\boldsymbol{N}-\boldsymbol{1}} \left|\boldsymbol{X}(\boldsymbol{t_1}) - \boldsymbol{Y}\left(\boldsymbol{\phi}(\boldsymbol{t_1})\right)\right| \tag{30}$$

under constraint (29). Without that constraint, (30) would simply recover the mapping derived from the local minimization (28). But with the presence of continuity constraint we turn the problem into a global optimization issue.

The problem discussed above has been studied for a long time, especially in the statistical physics sector under the name of '*random directed polymer at zero temperature*', for more information please refer to (Halpin-Healy & Zhang, 1995). In order to understand the notion of distance matrix (27), we can see that it could be interpreted like an energy landscape in the plane (t_1, t_2), where the local distance $\epsilon(t_1, t_2)$ is the energy linked to the node (t_1, t_2). A line, or path, or 'polymer' of equation $\left(t_1, t_2 = \phi(t_1)\right)$ is defined by a mapping taking into account a 'surface tension' to avoid discontinuities.

The fact that $\phi(t_1)$ can only increase explains the fact that the polymer should be directed, which is why it does not turn backwards and there are no overhangs. Using global minimization problem (30), we search for the actual path with he minimum energy. We are interested in non-random time series, because in the case of randomness there are no coherent structures in the distance matrix to detect. We search for a lead-lag structure between the time series that is actually revealed in the structure of the optimal path in the energy landscape.

The global nature of (30) means that in order to create the best path from a hypothetical origin to an end point, we need to consider every point to the left and to the right of every point (t_1, t_2) in the distance matrix $E_{X,Y}$. To solve this optimization problem in polynomial time, the transfer matrix method is used (Derrida, Vannimenus, & Pomeau, 1978), (Derrida & Vannimenus, 1983). Figure 42 shows a typical example of a distance matrix (27) with the optimal path wandering like a river from one edge of the energy landscape to the other.

The transfer matrix method can be formulated as follows. In Figure 43 we

present the (t_1, t_2) plane and some useful notations. The optimal path for two identical time series is the main diagonal. This means that any potential deviation from the diagonal of the plane shows a lead or lag between the time series. The next step is not necessary, yet a convenient tool to visualize and understand the lead-lag path. We use a rotated frame (t, x), where x represents the deviation from the main diagonal, which is the lead or lag time between the time series. A general expectation is to find a path that fluctuates above or below the main diagonal of equation $x(t) = 0$.

According to the theory, the optimal path is constructed is such a way that it can either go horizontally, vertically or diagonally from (t_1, t_2) to $(t_1 + 1, t_2)$, $(t_1, t_2 + 1)$, and $(t_1 + 1, t_2 + 1)$ respectively. Restricting the movement of the path to just those 3 possibilities supports the continuity condition and the one to one mapping. A different movement strategy can be found in (Zhou & Sornette, 2006). Every node (t_1, t_2) carries the distance $\epsilon(t_1, t_2)$ or the 'potential energy' in the two-dimensional lattice under study. Following (30), the cumulative distance denoted as $E(t_1, t_2)$ is the energy that starts from some origin $(t_{1,0}, t_{2,0})$ and ends at (t_1, t_2). The following equation solves the transfer matrix method efficient:

$$\boldsymbol{E(t_1, t_2) = \epsilon(t_1, t_2) + Min[E(t_1 - 1, t_2), E(t_1, t_2 - 1), E(t_1 - 1, t_2 - 1)]} \quad (31)$$

The explanation behind the equation above is that the minimum energy path that reaches every time the point (t_1, t_2) can come only from one of the three preceding points $(t_1 - 1, t_2)$, $(t_1, t_2 - 1)$, $(t_1 - 1, t_2 - 1)$. The minimum energy path reaching (t_1, t_2) is the extension of the minimum energy path reaching one of these three preceding points that is determined in (31). Finally, to determine the global optimal path, one needs to consider the sub-lattice $(t_{1,0}, t_{2,0}) \times (t_1, t_2)$ only, and determine the forenode of each node. To make things clear, and without loss of generality, we assume that the $(t_{1,0}, t_{2,0})$ is the origin $(0,0)$. Initially, we perform a bottom-to-up and left-to-right scanning. The $(\tau_1 - 1, 0)$ is the forenode of the bottom nodes $(\tau_1, 0)$, where $\tau_1 = 1, \dots, t_1$. Next, we determine the forenodes of the nodes in the second-

layer at $t_2 = 1$, based on the results of the first (or bottom) layer. We reiterate the procedure for $t_2 = 2$, $t_2 = 3$, and so on.

4.2.4 Optimal path at finite temperature

The optimal path procedure in the previous section, has a potential disadvantage: it does not take into consideration that there might be some noise contained in the time series under analysis, thus the distance matrix $E_{X,Y}$ (27) is not made only of useful information but contains noise too. It is very possible that $X(t_1)$ and $Y(t_2)$ contain noise and irrelevant structures deriving from random realizations.

As a result the global optimal path made by the procedure in section 4.2.3 Optimal path at 'zero temperature' could be found to be different than the potential without noise structures. Different noise in the dataset could lead to another distance matrix and different global optimal path. As a result, the presence of noise in the data imposes to test the sensitivity of the optimal path, when changes in the distance matrix occur. This issue has been regarded in previous studies; see for example (Halpin-Healy & Zhang, 1995). It is clear that even a small perturbation because of noise could lead to large jumps in the optimal path. Therefore, we need a method capable of absorbing the error of noise patterns and show a truly informative structure.

In a realistic scenario, we expect to have noise in the dataset, yet with clear and coherent patterns, so that the optimal path will contain the genuine lead-lag relationship in the energy landscape. Still, it is not certain whether the existing noise could affect the optimal path and cause false spurious wanderings. A modification of the aforementioned method was initially proposed by Sornette and Zhou (2005), which makes the determination of the mapping less vulnerable of existing noise and more robust in total. In general, it is a demanding project to 'filter' the useful information out of noise. The main advantage of this method is the fact that the underlying dynamics must not be known a priori.

The idea of the Thermal Optimal Path (TOP) is based on the following process. By taking into account the path, which is built on the energy landscape minimizing its energy E we allow 'thermal' excitations around the optimal path so that paths around it with higher global energies are allowed to be taken into

consideration as well, with probabilities decreasing with their energy. The probability to use each candidate path is specified with energy $\Delta \in$ above the absolute minimum energy path by a Boltzmann weight that is proportional to $exp\left[{}^{-\Delta \in}/_{T}\right]$. T is the "temperature" that presents how much deviation from the minimum energy is allowed. If we use zero temperature *(T=0)*, then the probability to select a path with energy $\Delta \in$ above the global minimum becomes zero, as we have seen the optimization problem at zero temperature.

On the other hand, by increasing the temperature T we can use more paths around the one with minimum energy. In this way we are able to eliminate the existence of the noise, and wash out any potential idiosynchratic dependencies of the path conformation. Nevertheless, the temperature T has to be used with prudence, because for temperatures higher than those in the publications presented here, we lose all information and the distance matrix is meaningless. For these reasons, there is a compromise between the information that we need to wash out from the signal and the level of noise to be extracted. The TOP can be found by adjusting higher temperature that shows an actual 'average' path over more and more path conformations. More information over the use of the temperature T can be found in (Andersen, 2006).

4.3 Empirical Results

Given the aforementioned theory, we created an algorithm to assess the causal relationship between the time series of our interest. According to the theory of (Sornette & Zhou, 2005) we designed a Dynamic Time Warping path. We used this technique in order to find the optimal alignment between time series, the one that minimizes the total distance between them. This is an algorithm that will also produce sufficient results even when one time series is wrapped non-linearly along its time axis. DTW was initially used for speech recognition (Myers, Rabiner, & Rosenberg, 1980). This first step was completed, taking into consideration the maximum capability of our electronic equipment. We confirmed the validity of our algorithm because we found the same results with another MATLAB algorithm, found available online (Ellis, 2003). At the end of this section both MATLAB codes can be found.

Nevertheless, the next step was not complete. We planned to find the paths around the one with the lowest energy in order to eliminate potential noise that time series could carry, according to 4.2.4 Optimal path at finite temperature. But in order to do this, it is apparent that we should have first found all possible paths from the bottom left element of the distance matrix, to the top right element, in other words all the routes through the grid made by the two time series we use, which is a time complexity project. After the assessment of the total energy of each path separately - a plan that was not successfully operated, we wanted to use the ones with the lowest energy. Unfortunately, this is a memory intensive procedure that either our experiment gear cannot provide, or the efficiency of our algorithm is questionable.

In the second attempt, we focused to find all possible paths by applying a programming method called recursion. With the use of recursion, the algorithm can call itself until it finds the final result, in other words it solves a problem by solving first smaller cases of the same problem, meaning that we can find the results efficiently and faster. The algorithm we wrote can also be found in detail at the end of the section. In this case it was possible for the algorithm to calculate matrices no larger than 7 by 7 (time series of seven records), meaning that we could not use this method as well.

There are already some known ways to speed up the dynamic time warping. Firstly, various constraints are used to limit the evaluated cells in the cost matrix. Secondly, there is the method that could minimize the times DTW must run called indexing. Thirdly, there is also another method called abstraction that performs the same algorithm in a reduced time series representation (Salvador & Chan, 2007), (Bylund, Danielsson, Malmquist, & Markides, 2002), (Keogh & Pazzani, 2000), (Keogh & Smyth, 1997).

We abstract our time series with two methods; the piecewise aggregate representation (Chu, Keogh, Hart, & Pazzani, 2002), and the piecewise linear representation (Keogh & Pazzani, 1999). In the former method we divide the time series into equal length segments. We supposed that if we were able to find faster a shorter than the original path that is similar to the best path of the matrix using the abstraction, and taking into account the fact that we minimized the length of our time series, then the rest of the paths could be found the same way for the minimized matrix by using recursion. We would then imply that the other paths from the abstract matrix we created follow the same pattern of the original difference matrix. After a

series of experiments with the time series and the calculation of the original DTW, we concluded that the path found after the abstraction is far from similar with the original one, meaning that this would make impossible for us to find the rest of the paths needed to calculate the optimal path. Unfortunately, the length of our time series (circa 1500) is large enough to stop the procedure.

Due to the limitations of our MATLAB algorithm which is not capable of calculating any matrix larger than 7×7, and due to our limited computer programming experience to apply more advanced programming methods like memoization (Cormen, Leiserson, Rivest, & Stein, 2009), we will show in this third part only the thermal optimal path for zero temperature. The algorithm did not go according to the plan, and we assume that the time series should contain some potential noise. Nevertheless we think that there is still some space for useful results.

As already reported in CHAPTER 3, we use weekly indices containing contract prices from transactions taken in Centaline Property Agency Limited. In this part we shall refer again to CCL as the "General index". The Large and Small/Medium housing indices are both sub-indices of the General index, showing the trend of flats over and under 1076 sq. ft. (100 sq.m.). We also analyze the Mass index, a general trend of all regions of Hong Kong, and its four sub-indices named from the regions chosen: Hong Kong Island, Kowloon, New Territories East and New Territories West.

In Figure 44, panel we presented the causality relationship between the General index and its two sub-indices, named Large and Small/Medium index. We can see the evolution of the lead-lag structure between them, taking as the first time series the General index. The most important observation is that the General index leads both Large and Small/Medium index during the period under consideration, since *x(t)* is always positive. Furthermore, their fluctuations follow the same path from beginning to the end. There is an increase of the index during the 2007 crisis starting from the mid 2006 and falls to the end of 2008, meaning that both sub-indices decrease the level of lag to the General index.

In the case of Mass index (Figure 44, panel b), we see that it lags Hong Kong and NT West, while NT East fluctuates from negative to positive *x(t)*, and at the end lags the main index again. But the most striking feature is the relationship of Mass index with Kowloon. In contrast with the other results, it lags Mass index during the whole period under analysis. This result agrees with the findings of the third part of

Thesis where the Kowloon index managed to hold its level of informational efficiency for the reasons already discussed in Empirical results of CHAPTER 3, depicting even with this analysis a different trend from the rest. Thus we see that different methods support the special characteristics of Kowloon as a region concerning the Hong Kong real estate market.

Furthermore we examine whether the Thermal Optimal Path for zero temperature and Granger causality show the same causality relationship between the Hong Kong housing indices. In order to find Granger causality tests, we should first apply unit root tests on the logarithms of the time series, and on the first order differences to check for stationary. We used two methods, the Augmented Dickey Fuller test (ADF), and the Phillips-Perron (PP) to confirm the results. Our null hypothesis is that the time series has a unit root. According to the results presented in Table 8 and Table 9, the logarithmic time series are not stationary, and we cannot reject the null hypothesis. In contrast, all logarithmic differences are found to be stationary at the 1% significance level. This confirms the use of normalized logarithmic differences in order to avoid spurious signals from potential non-stationary data.

Lastly, we use the Granger test to find whether the explanatory variable Granger cause the dependent variable. The results are presented in Table 10. The first thing that we notice is that the cases of General-Large, Mass-Kowloon, Mass-NT East and Mass-NT West show bidirectional Granger cause. That means all indices Granger cause the other one, and vice versa. The two cases left, surprisingly, are in contrast with our initial DTW findings. In the General-Small/Medium index, it is found that Small/Medium index Granger cause the General index, while the opposite applies in the Mass-Hong Kong index, where Mass Granger cause the Hong Kong.

4.4 Conclusion

We attempted to reveal the causality effect between real estate market indices of Hong Kong. In the case of our initial hypothesis that TOP does not follow Granger causality findings, we cannot formally express our decision, but we will suggestively report that we cannot reject it. First of all because even though in four out of six cases we found similar causation analysis, and more specifically in General-Large, Mass-Kowloon, Mass-NT East and Mass-NT West, it is a two way causation, which could be interpreted as a mechanism of price adjustment of one index from the other.

The two remaining cases, named General-Small and Mass-Hong Kong, do not follow the DTW findings. A first explanation for the one-way causation in the Granger analysis about the General-Small case is that the Small and Medium market form a large percentage of the total market of Hong Kong. The city is well known for the lack of available space, and its small housing assets, meaning that the Small/Medium index could possibly causes any change in the General index. In any case, since the Thermal Optimal Path is not completed, we cannot be sure for the validity of our results.

Once again we should clarify that we used two different methods to test the causality of the Hong Kong Real Estate indices, and the implementation of the TOP at finite temperature is not complete. Those findings could be taken into consideration as a first step for further research, in order to delve deeper in the TOP approach. Therefore, we provide to the reader all the algorithms that we used in this chapter as a recommendation tool for further analysis in order to guide and suggest information that could be used as an extra step for the understanding of the TOP theory.

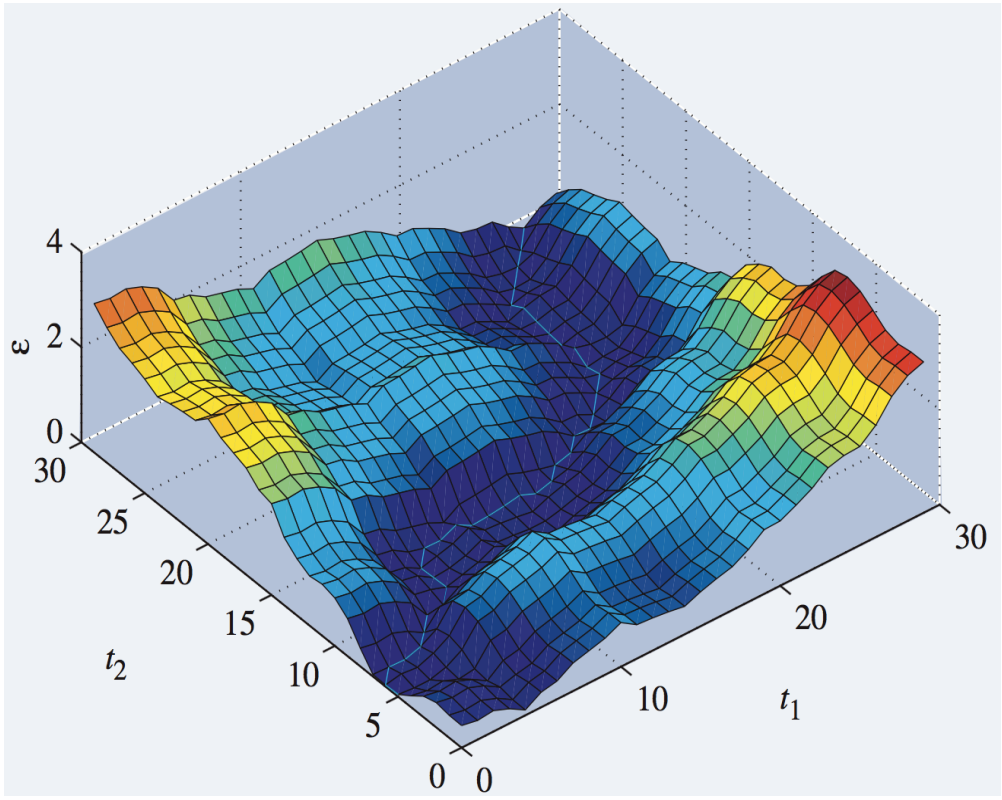


Figure 42: Energy landscape $E_{X,Y}$ given by (27) for two noisy time series and their corresponding optimal path wandering at the bottom of the valley similarly to a river.

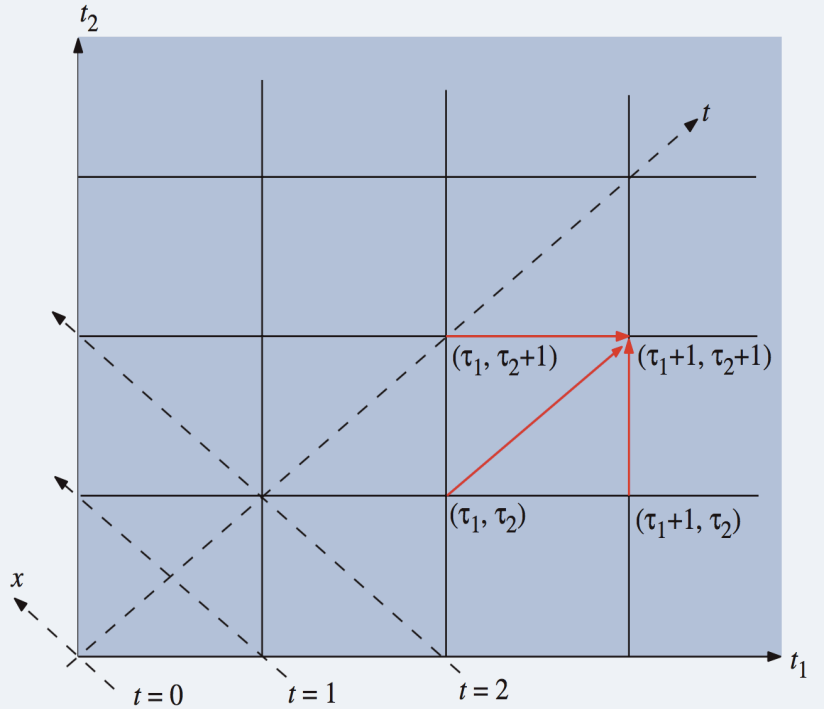


Figure 43: Representation of the lattice (t_1, t_2) and of the rotated frame (t, x). We refer to the (t_1, t_2) coordinate system as the square system, and the (x, t) system as the triangle system. The three arrows depict the three moves allowed from any node in one step, in accordance with the continuity and monotonicity conditions (29).

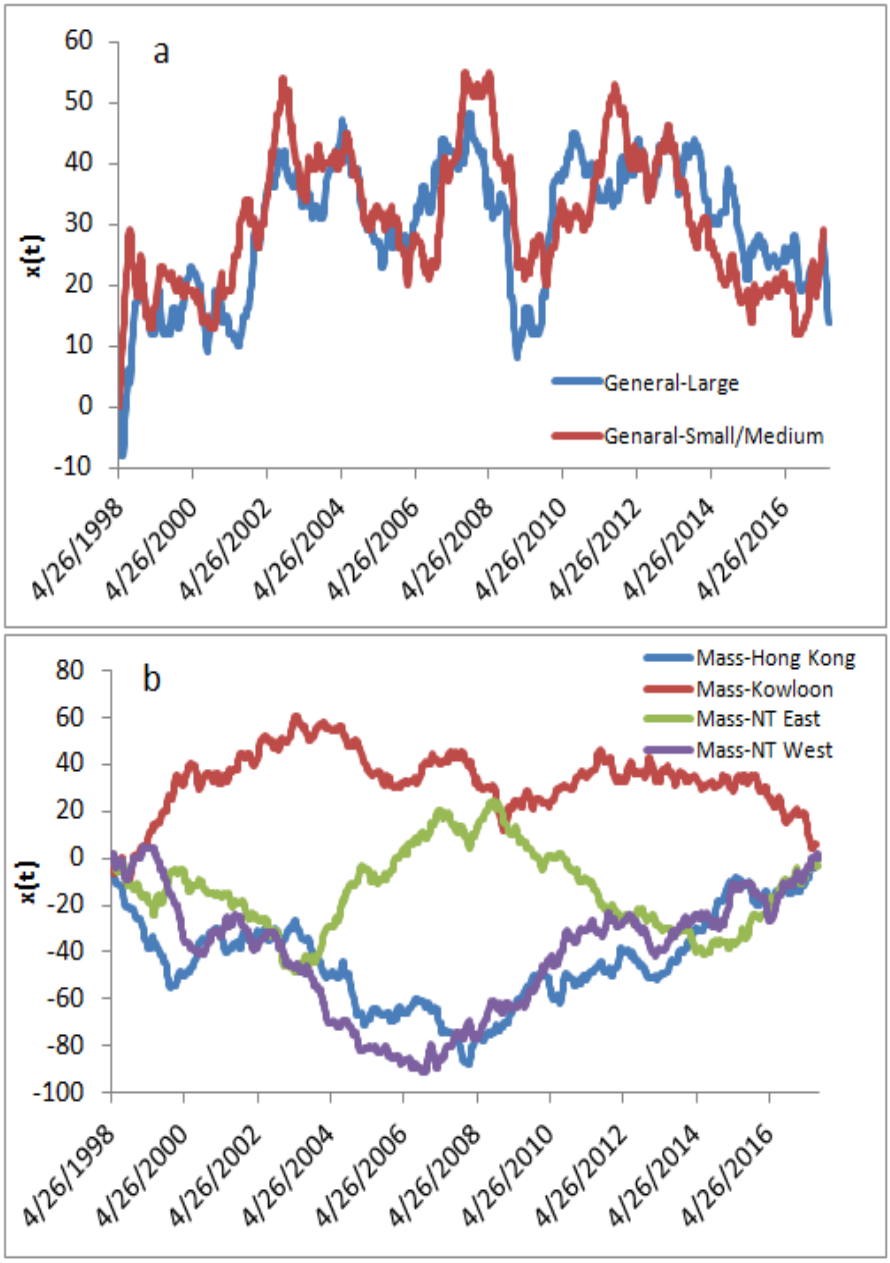


Figure 44: Thermal Optimal Path for Zero Temperature, panel (a) dependence of the lead-lag x(t) between the returns of the General index taken as the first time series and the Large and Small/Medium index series. Panel (b), dependence of the lead-lag x(t) between the returns of the Mass index taken as the first time series and the Hong Kong, Kowloon, New Territories East and New Territories West index series.

				General		Large		Small/Medium	
method	V(1%)	V(5%)	V(10%)	test statistic	p-value	test statistic	p-value	test statistic	p-value
ln differences of data									
ADF	-3.43	-2.86	-2.57	-28.36	0.00	-31.51	0.00	-28.62	0.00
PP	-3.43	-2.86	-2.57	-29.91	0.00	-32.05	0.00	-30.18	0.00
ln data									
ADF	-3.43	-2.86	-2.57	1.71	1.00	0.87	0.99	1.78	1.00
PP	-3.43	-2.86	-2.57	0.96	0.99	0.57	0.99	1.01	0.99

Table 8: The Augmented Dickey Fuller (ADF) and Phillips Perron (PP) tests are adopted for General, Large, and Small/Medium index. V x% is the critical value at the x% significance level, end test statistic and p-value are also available.

				Mass		Hong Kong		Kowloon		NT East		NT West	
method	V(1%)	V(5%)	V(10%)	test statistic	p-value	test statistic	p-value	test statistic	p-value	test statistic	p-value	test statistic	p-value
ln differences of data													
ADF	-3.43	-2.86	-2.57	-29.48	0.00	-36.01	0.00	-38.46	0.00	-34.13	0.00	-35.64	0.00
PP	-3.43	-2.86	-2.57	-30.71	0.00	-35.75	0.00	-38.19	0.00	-34.04	0.00	-35.45	0.00
ln data													
ADF	-3.43	-2.86	-2.57	1.67	1.00	0.57	0.99	0.47	0.98	0.95	0.99	1.44	1.00
PP	-3.43	-2.86	-2.57	1.02	0.99	0.59	0.99	0.71	0.99	0.93	0.99	1.31	1.00

Table 9: The Augmented Dickey Fuller (ADF) and Phillips Perron (PP) tests are adopted for Mass, Hong Kong, Kowloon, New Territories East, and New Territories West index. V x% is the critical value at the x% significance level, end test statistic and p-value are also available.

Granger causality Wald tests

Equation	Excluded	chi2	df	Prob > chi2
generalnorm	smallnorm	18.55	4	0.001
generalnorm	ALL	18.55	4	0.001
smallnorm	generalnorm	4.8142	4	0.307
smallnorm	ALL	4.8142	4	0.307
generalnorm	largenorm	8.2894	4	0.082
generalnorm	ALL	8.2894	4	0.082
largenorm	generalnorm	73.417	4	0.000
largenorm	ALL	73.417	4	0.000
massnorm	hknorm	1.8077	4	0.771
massnorm	ALL	1.8077	4	0.771
hknorm	massnorm	101.73	4	0.000
hknorm	ALL	101.73	4	0.000
massnorm	kowloonnorm	11.557	4	0.021
massnorm	ALL	11.557	4	0.021
kowloonnorm	massnorm	122.15	4	0.00
kowloonnorm	ALL	122.15	4	0.00
massnorm	ntwestnorm	36.623	4	0.00
massnorm	ALL	36.623	4	0.00
ntwestnorm	massnorm	98.333	4	0.00
ntwestnorm	ALL	98.333	4	0.00
massnorm	nteastnorm	23.958	4	0.00
massnorm	ALL	23.958	4	0.00
nteastnorm	massnorm	127.8	4	0.00
nteastnorm	ALL	127.8	4	0.00

Table 10: Granger Tests

4.5 Algorithms used for the Thermal Optimal Path search

The first algorithm finds the path at zero temperature

```
%TOP (Thermal Optimal Path) for zero temperature ONLY

%please change the variable "source" below, and the "x" and "y" a bit lower.
%the most important finding is the COST1, which is the path with the lowest
%energy set on the xx'

source = 'HK.mat';
fprintf(['Loading_data \n']);
%Zhou-Sornette paper---> x = MSFT and y = IBM
load(source)
for rowToDelete = 1 : 1 %start using 1 to enter the first row,second, etc
 load(source) % use it here so every time it enters it can get the right rowToDelete

%the 'basic' variable always on y !
nteastnorm = nteastnorm(rowToDelete : end);
massnorm = massnorm(1 : end);
x = nteastnorm;
y = massnorm;

%note: the Zhou_mail_trimmed not for fliplr, yes to dp2 source code
%% difference matrix creation and inverse_differences
for i = 1 : length(x)
   for j = 1 : length(y)
      E(i,j) = abs((x(i) - y(j)));   % difference matrix
      inverse_differences(i,j) = i - j;  %matrix for x = t2-t1
      matr(i,j) = {[i,j]};          %matrix with the positions

   end
end
inverse_differences = flipud(inverse_differences);

matr = flipud(matr);
%E = flipud(E);  %flipud E matrix to test the results with another code,
%named dp.m, where the min cost starts from top left to bottom right.
clearvars -except kowloon mass adjustedccl1010experiment adjustedcclarge1010experiment
ccllarge1010experiment ccllarge1010experiment ccl1010experiment propertynorm hsinorm
financialnorm GRAMM STIL gramm stil matr Node_move1 source E E_DIFFERENCES1
E_DIFFERENCES2 PATH_DIFFERENCES1 PATH_DIFFERENCES2 E_differences1
E_differences2 inverse_differences rowToDelete snpret ffret Cost1 COST1 COST2 Cost2 R RR
Cost IBM MSFT FF SNP ffrnorm snp500norm snp ffr snpweeklyln ffweeklyln snpweeklynorm
ffweeklynorm ffrnorm snpnorm mo3norm generalnorm kowloonnorm largenorm smallnorm
ntwestnorm nteastnorm massnorm hknorm y20norm

%% tc matrix (the matrix similar to time warp)
r = size(E,1); %matrix dimensions (rows)
c = size(E,2); %matrix dimensions (columns)

tc(1,c) = E(1,c); %the first top-right node in the tc matrix is the same as E matrix


for  i = 2 : r    %fill up the last column from the top to the bottom
   tc(i,c) = tc(i-1,c) + E(i,c);
end
```

```
for j = (c-1) : -1 : 1   %fill up the first row of tc from right to left
   tc(1,j) = tc(1,j+1) + E(1,j);
end

%internal parts of the tc matrix (apart from the first row and the last
%column)
for i = 2 : r
   for j = (c-1) : -1 : 1
      fakelos = [tc(i-1,j); tc(i-1,j+1); tc(i,j+1)];
      tc(i,j) = min(fakelos) + E(i,j);
   end
end
clear c i j r

%% creation of the paths from the first column of the tc matrix (R node), and their respective
differences

[r,c] = size(tc);
N = @(R,C) sprintf('(%d,%d)',R,C);
node_move1 = {};
p=1; %counts the line for Cost1
% for R = (length(snpret)-5):length(snpret)
noumero = 1;
Cost1 = {};
E_differences1 = {};   %path following the nodes onto E matrix
path_differences1 = {};%path following the nodes onto tc matrix
gramm={};         %line of a flipud matrix, where (1,1) is at the left bottom
stil={};          %column of...
for R = r %had it as (r-20) :r  %from 20th line before bottom, until the last line, in total 21
%R=r; % start 20 lines before the end, and the one left is the line that is the next whole part
%Cost={}; %x, which is t2-t1, and its the distance from the diagonal t
%inverse_Cost={};
%the differences across the matrix path
  %R = r;
  %M = {};        %empty cell
   C = 1 ;        %start from the first column
   q=1;
   w=1; %counts the column for Cost1
   %first column of the cell Cost1
   Cost1(p,w) = { inverse_differences(R,C) };
   %Cost(p,q)   = { ((r-R) - C) };
   %inverse_Cost(p,q) = {C - R}; %just for display
   path_differences1(p,q) = {tc(R,C)};
   E_differences1(p,q)    = {E(R,C)};
   Visual1(p,q) = {[R,C]};          %visual assistant of the best path
   node_move1(p,q) = {N(R, C)};      %insert the initial link
   gramm(p,q) = {matr{R,C}(1,1)};
   stil(p,q) = {matr{R,C}(1,2)};


while R > 1 && C < size(tc,2) %as long R has not reached first line and last column
      t = {N(R, C), tc(R-1,C); ...   %keep cell with the 3 choices for the next step, top, right, and
the diagonal one
          N(R, C), tc(R-1,C+1);...
          N(R, C), tc(R,C+1)};
```

```
  if (tc(R-1,C)< tc(R-1,C+1)) && (tc(R-1,C) < tc(R,C+1))
     minimu = tc(R-1,C);  %the actual difference
     minimuE = E(R-1,C);
     linee = R-1;       % t2
     columnn = C;       % t1

  node_move1(p,q+1) = {N(linee, columnn)}; %insert the connection, apart from the initial link
  Visual1(p,q+1) = {[linee,columnn]}; %visual assistant of the best path
  %chain is a cell with the connected nodes
  %since it was a cell already from next_segment, cant {} so []
  %chain = [ next_segment(:,1) next_segment(:,2) ];  %cell for the
  %connection with the second step

  %Cost(p,q+1) = { ((r-linee) - columnn) }; %this is X, the t2-t1, and
  %the results are inverse
  Cost1(p,w+1) = { inverse_differences(linee,columnn) };
  %inverse_Cost(p,q+1) = {columnn - linee}; %just for display
  path_differences1(p,q+1) = {minimu};
  E_differences1(p,q+1) = {minimuE};
  R = linee;
  C = columnn;
%    M = [M; minimu];
  gramm(p,q+1) = { matr{R,C}(1,1)};
  stil(p,q+1)  = { matr{R,C}(1,2)};
  q=q+1;
  w=w+1;
  elseif ( tc(R-1,C+1) < tc(R-1,C)) && (tc(R-1,C+1) < tc(R,C+1)) %diagonios
     minimu = tc(R-1,C+1);
     minimuE = E(R-1,C+1);
     linee = R-1;
     columnn = C+1;

  node_move1(p,q+1) = {N(linee, columnn)}; %insert the connection, apart from the initial link
  Visual1(p,q+1) = {[linee,columnn]}; %visual assistant of the best path
  %chain is a cell with the connected nodes
  %since it was a cell already from next_segment, cant {} so []
  %chain = [ next_segment(:,1) next_segment(:,2) ];  %cell for the
  %connection with the second step

  %Cost(p,q+1) = { ((r-linee) - columnn) }; %this is X, the t2-t1, and
  %the results are inverse

  %If we go diagonally, then we get two of the same step
  Cost1(p,w+1) = { inverse_differences(linee,columnn) };
  Cost1(p,w+2) = { inverse_differences(linee,columnn) };
  %inverse_Cost(p,q+1) = {columnn - linee}; %just for display
  path_differences1(p,q+1) = {minimu};
  E_differences1(p,q+1) = {minimuE};
  R = linee;
  C = columnn;
%    M = [M; minimu];
  gramm(p,q+1) = { matr{R,C}(1,1)};
  stil(p,q+1)  = { matr{R,C}(1,2)};
  q=q+1;
  w=w+2;
  else
     minimu = tc(R,C+1);
```

```
            minimuE = E(R,C+1);
            linee = R;
            columnn = C+1;

        node_move1(p,q+1) = {N(linee, columnn)}; %insert the connection, apart from the initial link
        Visual1(p,q+1) = {[linee,columnn]}; %visual assistant of the best path
        %chain is a cell with the connected nodes
        %since it was a cell already from next_segment, cant {} so []
        %chain = [ next_segment(:,1) next_segment(:,2) ];  %cell for the
        %connection with the second step

        %Cost(p,q+1) = { ((r-linee) - columnn) }; %this is X, the t2-t1, and
        %the results are inverse
        Cost1(p,w+1) = { inverse_differences(linee,columnn) };
        %inverse_Cost(p,q+1) = {columnn - linee}; %just for display
        path_differences1(p,q+1) = {minimu};
        E_differences1(p,q+1) = {minimuE};
        R = linee;
        C = columnn;
%       M = [M; minimu];
        gramm(p,q+1) = { matr{R,C}(1,1)};
        stil(p,q+1)  = { matr{R,C}(1,2)};
        q=q+1;
        w=w+1;

        end

end   %end of while
%Cost{noumero} = Cost1;
noumero =noumero + 1;

        %RR = RR + 1;
        p=p+1;
        clear q M w
        %clear Cost1
end
%Cost = Cost1;
        COST1{rowToDelete} = Cost1;
        E_DIFFERENCES1{rowToDelete} = E_differences1;
        PATH_DIFFERENCES1{rowToDelete} = path_differences1;
        Node_move1{rowToDelete} = node_move1;
        GRAMM{rowToDelete} = gramm;
        STIL{rowToDelete} = stil;
        clear noumero Cost1 E_differences
        clear c C column Fs linee minimu minimuE N  r t y
        clear p E inverse_differences E
        clear tc
        clear gramm stil
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%
load(source)
metritis = 1;
```

This code was used to support the findings of the previous for the TOP path at zero temperature.

```
% function [p,q,D] = dp(M)
% [p,q] = dp(M)
%    Use dynamic programming to find a min-cost path through matrix M.
%    Return state sequence in p,q
% 2003-03-15 dpwe@ee.columbia.edu
% Copyright (c) 2003 Dan Ellis <dpwe@ee.columbia.edu>
% released under GPL - see file COPYRIGHT

% NOTE FROM THE AUTHOR OF THIS THESIS- ARGYROUDIS GEORGIOS
% This code does not belong to me-see above for the Copyright, I have just made some changes in
order to
% compare it with another that I designed

% to start we input the file of our interest in load command below, and
% also we change the variables VAR1 and VAR2, in VAR1 we use the "basic"
% variable
% plot at the end the horizontall

%data processing
load('HK.mat');
VAR1 = massnorm;  %the "basic variable"
VAR2 = kowloonnorm;
%-----------------------------------------------------------------
%now adjust data to become stationary, and remove average and std
% for i = 2 : length(VAR1)
%     logvar1(i) = log(VAR1(i)) - log(VAR1(i-1));
%     logvar2(i) = log(VAR2(i)) - log(VAR2(i-1));
% end
%
% for i = 1 : length(logvar1)
%     variable1(i) = ((logvar1(i) - mean(logvar1))/std(logvar1));
%     variable2(i) = ((logvar2(i) - mean(logvar2))/std(logvar2));
% end
%-----------------------------------------------------------------
clearvars -except VAR1 VAR2
save('variables.mat');
clear

%creates the matrix to operate
load('variables.mat')
y = VAR1;
x = VAR2;

for i = 1 : length(x)
   for j = 1 : length(y)
      M(i,j) = abs((x(i) - y(j)));

   end
end
M = flipud(M);

clearvars -except M
%-------------------------------------------------------
```

```
[r,c] = size(M);
% costs
D = zeros(r+1, c+1);
D(1,:) = NaN;
D(:,1) = NaN;
D(1,1) = 0;
D(2:(r+1), 2:(c+1)) = M;

% traceback
phi = zeros(r,c);

for i = 1:r;
  for j = 1:c;
    [dmax, tb] = min([D(i, j), D(i, j+1), D(i+1, j)]);
    D(i+1,j+1) = D(i+1,j+1)+dmax;
    phi(i,j) = tb;
  end
end

% Traceback from top left
i = r;
j = c;
p = i;
q = j;
while i > 1 & j > 1
  tb = phi(i,j);
  if (tb == 1)
    i = i-1;
    j = j-1;
  elseif (tb == 2)
    i = i-1;
  elseif (tb == 3)
    j = j-1;
  else
    error;
  end
  p = [i,p];
  q = [j,q];
end

% Strip off the edges of the D matrix before returning
D = D(2:(r+1),2:(c+1));

%horizontal graph
clearvars -except p q
for i = 1 : length(p)
   horizontall(i) = p(i) - q(i);
end
```

The third code contains the method of recursion

```
tic
clear all; clc;

% This is the initial array
%load('E.mat')
array_initial = rand(6,6);

% Number of rows and columns
[rows, cols] = size(array_initial);

% Creates a new array changing the order of rows
% To find the paths from the bottom left element to the top right
for i = 1 : rows
   array(i,:) = array_initial(rows-i+1,:);
end


for  i = 1 : length(array)
for j = 1 : length(array)
if (abs(i-j)) > 2
array(i,j) =0;
end
end
end

%array = sparse(array);
% create path cell

% dimensions
% count number of possible paths
% initialize values
num_of_paths = 0;
a = 1;
b = 1;
path = [];
counter = 1;
% num_of_paths = mycount(a, b, path, num_of_paths, array);
max_length_of_path = 2*rows-1;

% initialize path cell
%paths = cell(num_of_paths, max_length_of_path);
paths = cell(20, max_length_of_path);
SUM_PATH = cell(20,1);
% The algorithm starts from 1,1 and continues on its own
min_sum_path = [];

% Initial value for the sum is the sum of the whole matrix
min_sum = sum(sum(array));
[min_sum_path, min_sum, paths, SUM_PATH] = myfun_opt(a, b, array, path,...
paths, min_sum_path, min_sum, SUM_PATH,counter);

toc
%---------------------------------------------------------------------------
function [min_sum_path, min_sum, paths, SUM_PATH] = myfun_opt(a, b,...
   array, path, paths, min_sum_path, min_sum,SUM_PATH,counter)
```

```
[rows, cols] = size(array);


% check if the last point also appears in another position
idx1 = array == array(rows,cols);

% add a very small value to all the elements that are equal to the last
% one so that the algorithm won't stop there
% added value is of very low magnitude (0.000000001)
idx2 = ones(rows,cols);
idx2(rows,cols) = 0;
array = array + idx1.*idx2.*0.000000001;

% If (a, b) is in the matrix
if check_opt(a, b, rows, cols, array)
   % check if we reached the last point.
   if array(a,b) == array(rows, cols)
      % Add the point (a, b) to the path and return
      final_path = [path array(a,b)];
      % Find the sum of the path
      sum_path = sum(final_path);
      % Check if it is the minimum sum
      if sum_path < min_sum
         min_sum_path = final_path;
         min_sum = sum_path;
      end
      % Save path in "paths"cell
      % Find first empty row of "paths" cell


      counter = counter + sum(~isnan([SUM_PATH{:,1}]));
      if counter > size(paths,1)
         if sum_path < max([SUM_PATH{:,1}])
            [maxi,position] = max([SUM_PATH{:,1}]);
             SUM_PATH{position,1}=sum_path;
             paths(position,:) ={0};

         else
            return
         end
      for j = 1:size(final_path,2)
         paths{position,j} = final_path(j);
      end
       for z = 1 : (2*rows-1)
                  if paths{position,z} == 0
                     paths{position,z} = [];
                  end
       end
      return
      end

%          for i=1:size(paths,1)
         if isempty(paths{counter,1}) == 1
            SUM_PATH{counter,1} = sum_path;
            for j = 1:size(final_path,2)
```

```
                paths{counter,j} = final_path(j);
            end
%               break
        end
%       end

      return
    end
  end
  % The next step can be either the next row or the next column.
  % Firstly we check the next column.
  if check_opt(a, b + 1, rows, cols, array) == 1
    % Now, the function is called recursively. The code will return
    % at this line, either if we reach the desired cell or if we get
    % out of the matrix.
    % From now the function starts being executed again, but with
    % different initial values.
    % There might be more recursions fun -> fun -> fun ... when one
    % of these recursions
    % returns then all the others will return fun <- fun <- ...
    [min_sum_path, min_sum, paths,SUM_PATH] = myfun_opt(a, b + 1,...
     array, [path array(a,b)], paths, min_sum_path, min_sum,...
     SUM_PATH,counter);
  end

  %  Now that we have returned at this point we check for the next column
  %  The code remembers the a, b values and remembers at which
  %  fun call we are.
  %  e.g. fun(1, 2) <- fun(2, 3) <-...fun(..)
  if check_opt(a + 1, b, rows, cols, array) == 1
    [min_sum_path, min_sum, paths,SUM_PATH] = myfun_opt(a + 1, b,...
     array, [path array(a,b)], paths, min_sum_path, min_sum,...
     SUM_PATH,counter);
  end

  % diagonal paths
  if check_opt(a + 1, b + 1, rows, cols, array) == 1
    [min_sum_path, min_sum, paths,SUM_PATH] = myfun_opt(a + 1,...
     b + 1, array, [path array(a,b)], paths, min_sum_path,...
     min_sum,SUM_PATH,counter);
  end

  % All function calls end up either at return True, or in the case
  % we are out of the matrix we end up here,
  % because all checks will be false.

  return
end

%---------------------------------------------------------------------------

function flag = check_opt(a, b, rows, cols, array)
  % Checks if we are out of the matrix. The last index of a row is
  % number of rows - 1 because we start from 0.
  if a > rows || b > cols || (array(a,b) == 0)
    flag = 0;
  else
```

```
        flag = 1;
    end
end

%-------------------------------------------------------------------------
```

Conclusion- Outlook

In the first part we used the multifractal theory in order to assess the multifractality and complexity of both 1997 and 2007 crises in Hong Kong market. We divided the Hang Seng Index in fore-crisis and after-crisis period to understand the mechanism of complexity and multifractality after a turbulent period. Also, the Financial and Properties Index were used for comparison purposes and to conclude about the statistical characteristics of the Real Estate market of the country. Two multifractality methods that measure the multifractal spectral of the daily time series are under investigation: Multifractal Detrended Fluctuation Analysis (MFDFA) and Multifractal Detrended Moving Average (MFDMA). We proceeded this way to compare those two methods, and whether the results conclude similar results. The findings show that we are dealing with comparable methods.

The findings suggest that both 1997 and 2007 crises revealed multifractality, while the fore-after methodology supports that the Financial Index was affected by the financial meltdown more than the Hang Seng Index or the Properties Index. On the other side, their shuffled series showed no temporal correlations. Also, the temporal evolution of multifractality that has been used, confirms the above findings.

In the second part of the research we use the concept of statistical entropy. We focus again in the Real Estate housing market of Hong Kong, using a monthly time series data set. The data set is divided in region-based and size-based categories. The aim of the analysis is to examine their informational efficiency before and after the 2007 downturn, and to understand whether we are able to discredit the celebrated Efficient Market Hypothesis (EMH), and to assess any change in statistical characteristics during anxious times.

We used two methods, the Shannon and Tsalllis-q entropy in order to test and compare the validity of each other. For both types of entropic measurement, we studied the complexity entropy causality plane (CECP) for all selected indices, assessing their position before and after the crisis. Furthermore, we shuffled the series to confirm their position in the CECP, and we verify the relation of the findings with the fractional Brownian motion series. Likewise, we showed the evolution in time of both Shannon and Tsallis entropy quantifiers, proving that there are changes in the

entropic measurement due to various important events presented in our analysis. Additionally, considering a 100 weeks window through time, we assess the evolution and the progress of the indices in order to draw dynamically the inefficiency level.

The results showed that the majority of the markets clearly present a higher inefficiency level after the 2007 mortgage market crisis. Nevertheless, signs of stability before and after the crisis period are found in one specific index (Large market index), while the most striking finding is the reaction of the index for the Kowloon region, which shows higher efficiency level after the 2007 crisis.

In the third part of the research we used the Thermal Optimal Path method (TOP). The focus is again on the Real Estate market of Hong Kong, and the data set is divided in region-based and size-based categories. The aim of this part was to understand the causality relationship between the housing market indices, and if possible compare it with the Granger causality. Due to the fact that this is a memory intensive algorithm or it is not efficiently designed, we presented the results about causality between the Hong Kong real estate indices using only the first part of the theory concerning the TOP method; it is the Thermal Optimal Path for zero temperature.

The results showed that the General index leads both Large and Small/Medium index during the period under consideration. In the case of Mass index, we see that the latter lags Hong Kong and NT West, while NT East fluctuates from negative to positive $x(t)$, and at the end lags the main index again. Still the most interesting finding is the Kowloon index that lags Mass index, and supports our previous findings. Nevertheless, the results do not accord with the Granger causality ones. The two out of six Granger causality tests presented the opposite relationship with the TOP, while the rest were found to be bidirectional.

Bibliography

Abry, P., & Sellan, F. (1996). The Wavelet-Based Synthesis for Fractional Brownian Motion Proposed by F. Sellan and Y. Meyer: Remarks and Fast Implementation,. *Applied and Computational Harmonic Analysis , 3*, 377-383.

Adams, C., Mathieson, D. J., Schinassi, G., & Chadha, B. (1998, September). *www.imf.org*. Retrieved June 3, 2017, from International Monetary Fund: http://www.imf.org/external/pubs/ft/icm/icm98/

Alessio E., C. A. (2002). *Eur. Phys. J. B , 27* (197).

Alvarez-Ramirez J., R. E. (2005). *Physica A , 354* (199).

Alvarez-Ramirez, J., Rodriguez, E., & Espinosa-Paredes, G. (2012). Is the US stock market becoming weakly efficient over time? Evidence from 80-year-long data. *Physica A , 391* (22), 5643–5647.

Amigo, J. M., & Keller, K. (2013). Permutation entropy: One concept, two approaches . *THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS , 222*, 263-273.

Amigo, J. M., Zambrano, S., & Sanjuan, M. A. (2007). True and false forbidden patterns in deterministic and random dynamics September 2007 www.epljournal.org . *epljournal , 79*.

Andersen, J. V. (2006). Comments on ‘‘Non-parametric determination of real-time lag structure between two time series: The ‘Optimal thermal causal path’ method with applications to economic data’’. *Journal of Macroeconomics , 28*, 225-227.

Anteneodo, C., & Plastino, A. R. (1996). Some features of the L6pez-Ruiz-Mancini-Calbet statistical measure of complexity . *Physics Letters A , 223*, 348-354.

Arianos S., C. A. (2007). Detrending moving average algorithm: A closed-form approximation of the scaling law. *Physica A , 382*.

Arianos S., C. A. (2008). Detrending Moving Average Variance: a derivation of the scaling law.

Ashley, R., Granger, C. W., & Schmalensee, R. (1980). Advertising and aggregate consumption: an analysis of causality. *Econometrica , 48*, 1149–1167.

Asuncion, A., & Newman, D. (2007). UCI machine learning repository.

Ausloos, M. (2000). Statistical physics in foreign exchange currency and stock markets . *Physica A , 285*, 48-65.

Bachelier, L. (1964). Theory of Speculation, reprint in P. Cootner, "Random Character of Stock Prices".

Bandt, C. (2017). A New Kind of Permutation Entropy Used to Classify Sleep Stages from Invisible EEG Microstructure. *Entropy , 19* (197).

Bandt, C. (2003, 10 2). Ordinal time series analysis. *Preprint submitted to Ecological Modelling* .

Bandt, C., & Pompe, B. (2002, 2 18). Permutation entropy — a natural complexity measure for time series .

Bardet, J. M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S., & Taqqu, M. S. (2003). Generators of long-range dependence processes: a survey, Theory and applications of long-range dependence. 579-623.

Bariviera, A. F., Guercio, M. B., Martinez, L. B., & Rosso, O. A. (2015). The (in)visible hand in the Libor market: an information theory approach. *Eur. Phys. J. B , 88* (8), 208.

Bariviera, A. F., Zunino, L., & Rosso, O. A. (2017). Crude oil market and geopolitical events: an analysis based on information-theory-based quantifiers. *Fuzzy Economic Review , 21* (1), 41-51.

Bariviera, A. F., Zunino, L., Guercio, M. B., Martinez, L. B., & Rosso, O. (2013). Revisiting the European sovereign bonds with a permutation- information-theory approach. *Eur. Phys. J. B , 86* (509).

Bariviera, A., Guercio, M. B., & Martinez, L. B. (2012). A comparative analysis of the informational efficiency of the fixed income market in seven European countries. *Economics Letters , 116*, 426–428.

Barkham, R. J., & Geltner, D. M. (1995). Price discovery in American and British property markets. *Real Estate Economics , 23* (1), 21-44.

Barkoulas, J. T., & Baum, C. F. (1996). Long-term dependence in stock returns. *Economics Letters , 53*, 253–259.

Barunik J., A. T. (2012). Understanding the source of multifractality in financial markets . *Physica A: Statistical Mechanics and its Applications , 391* (17), 4234-4251.

Becketti, S. (2013). *Introduction to Time Series Using Stata.* Texas: STATA PRESS.

Bentes, S. R., & Menezes, R. (2012). Entropy: A new measure of stock market volatility? *Journal of Physics: Conference Series , 394*.

Bentes, S. R., Menezes, R., & Mendes, D. A. (2008, December 4). Long Memory and Volatility Clustering: is the empirical evidence consistent across stock markets?

Bentes, S. R., Menezes, R., & Mendes, D. A. (2008). Stock market volatility: An approach based on Tsallis entropy .

Berndt, D., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series.

Bian, C., Qin, C., Ma, Q. D., & Shen, Q. (2012). Modified permutation-entropy analysis of heartbeat dynamics. *PHYSICAL REVIEW E , 85* (021906).

Bond, S. A., Dungey, M., & Fry, R. (2006). A Web of Shocks: Crises across Asian Real Estate Markets. *Journal of Real Estate Financial Economics , 32*, 253-274.

Bouchaud J.P., P. M. (2000). Apparent multifractality in financial time series . *Eur. Phys. J. B , 13* (595).

Bratko, I. (2000). Prolog Programming for Artificial Intelligence.

Bruzzo, A., Gesierich, B., Santi, M., Tassinari, C., Birbaumer, N., & Rubboli, G. (2008). Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients. A preliminary study. *Neurol. Sci. , 29*, 3-9.

Bylund, D., Danielsson, R., Malmquist, G., & Markides, K. E. (2002). Chromatographic Alignment by Warping and Dynamic Programming as a Pre-Processing Tool for PARAFAC Modeling of Liquid Chromatography-Mass Spectrometry Data . *Journal of Chromatography A* , 237-244.

Caiani, E. G., Porta, A., Baselli, G., Turiel, M., Muzzupappa, S., Pieruzzi, F., et al. (1998). Warped-average template technique to track on a cycle- by-cycle basis the cardiac filling phases on left ventricular volume.

Calbet, X., & Lopez-Ruiz, R. (2001). Tendency towards maximum complexity in a nonequilibrium isolated system. *Physical Review E , 63*.

Calvet, L., Fisher, A., & Mandelbrot, B. (1997). Large Deviations and the Distribution of Price Changes .

Cao, Y., Tung, W., Protopopescu, V., & Hively, L. (2004). Detecting dynamical changes in time series using the permutation entropy. *Physical Review E , 70*.

Carbone, A., Castelli, G., & Stanley, H. E. (2004). Analysis of clusters formed by the moving average of a long-range correlated time series. *Phys. Rev. E , 69*.

Carbone, A., Castelli, G., & Stanley, H. E. (2004). Time-dependent hurst exponent in financial time series. *Physica A , 344*, 267-271.

Carpi, L. C., Saco, P. M., & Rosso, O. A. (2010). Missing ordinal patterns in correlated noises. *Phys. A , 389*, 2020–2029.

Case, K. E., & Shiller, R. J. (1989). The efficency of the Market for Single-Family Homes. *American Economic Review , 79* (1), 125-137.

Chaitin, G. J. (1966, October). On the length of programs for computing finite binary sequences. *J ACM 1S* , 547-569.

Chau, K. W., Macgregor, B., & Schwann, G. M. (2001). Price discovery in the Hong Kong real estate market. *Journal of Property Research , 18* (3), 187–216.

Chen , N.-K. (2001). Asset price fluctuations in Taiwan: Evidence from stock and real estate prices 1973 to 1992. *Journal of Asian Economics , 12*, 215–232 .

Chen, Y., Rangarajan, G., Feng, J., & Ding, M. (2004). Analyzing multiple nonlinear time series with extended Granger causality. *Phys. Lett. A , 324*, 26-35.

Chu, S., Keogh, E., Hart, D., & Pazzani, M. (2002). Iterative Deepening Dynamic Time Warping for Time Series .

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to Algorithms.* The MIT Press.

Crutchfield, J. P., & Young, K. (1989). Inferring Statistical Complexity. *Physical Review Letters , 63*.

Czarneski, L., & Grech, D. (2009, December 17). Multifractal dynamics of stock markets.

Daw, C. S., Finney, C. E., & Tracy, E. R. (2003). A review of symbolic analysis of experimental data. *Rev. Sci. Instrum.* , 915–930.

Derrida, B., & Vannimenus, J. (1983). Interface energy in random systems. *Phys. Lett. B , 27*, 4401–4411.

Derrida, B., Vannimenus, J., & Pomeau, Y. (1978). Simple frustrated systems: chains, strips and squares. *J. Phys. C , 11*, 4749–4765.

Di Matteo, T., Aste, T., & Dacorogna, M. M. (2005). Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *Journal of Banking & Finance , 29* (4), 827-851.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance , 1* (83).

Dionisio, A., Menezes, R., & Mendes, D. A. (2006). An econophysics approach to analyse uncertainty in financial markets: an application to the Portuguese stock market. *Eur. Phys. J. B , 50*, 161–164.

Drasdo, D., Hwa, T., & Lassig, M. (1998). A statistical theory of sequence alignment with gaps. (I. L. J. Glasnow, Ed.) 52-58.

Drasdo, D., Hwa, T., & Lassig, M. (2001). Scaling laws and similarity detection in sequence alignment with gaps. *J. Comput. Biol. , 7*, 115–141.

Drozdz, S., Grummer, F., Ruf, F., & Speth, J. (2008). Dynamics of competition between collectivity and noise in the stock market.

Eckmann, J.-P., Kamphorst, S. O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhys. Lett. , 4*, 973–977.

Ellinger, A. G. (1971). The Art of Investment .

Ellis, D. (2003). *Dynamic Time Warp (DTW) in Matlab*. Retrieved from http://www.ee.columbia.edu/~dpwe/resources/matlab/dtw/

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica 55 , 2*, 251–276.

Eom, C., Choi, S., Oh, G., & Jung, W.-S. (2008). Hurst exponent and prediction based on weak-form efficient market hypothesis of stock markets. *Physica A , 387*, 4630–4636.

Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance , 46* (5), 1575–1617.

Fama, E. F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance , 25* (2), 383–417.

Feldman, D. P., & Crutchfield, J. P. (1998). Measures of statistical complexity: Why? *Physics Letters A , 238*, 244-252.

Filimonov, V., & Sornette, D. (2013). A stable and robust calibration scheme of the log-periodic power law model. *Physica A , 392*, 3698-3707.

Fisher, A., Calvet, L., & Mandelbrot, B. (1997). Multifractality of Deutschemark/ US Dollar Exchange Rates.

Fry, R., Martin, V. L., & Tang, C. (2010). A New Class of Tests of Contagion With Applications . *Journal of Business & Economic Statistics , 28* (3), 423-437.

Fuhrer, J. C. (2000). Habit Formation in Consumption and Its Implications for Monetary-Policy Models. *The American Economic Review , 90* (3).

Galbraith, J. K. (1997). *The Great Crash, 1929*. Boston: Houghton Mifflin.

Gavrila, D. M., & Davis, L. S. (1995). Towards 3-d model-based tracking and recognition of human movement: a multi-view approach.

Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long-memory time series models. *Journal of Time Series Analysis , 4*, 221-238.

Gollmer, K., & Posten, C. (1995). Detection of distorted pattern using dynamic time warping algorithm and application for supervision of bioprocesses. On-Line Fault Detection and Supervision in the Chemical Process Industries.

Gong, C.-C., Ji, S.-D., Su, L.-L., Li, S.-P., & Ren, F. (2016). The lead–lag relationship between stock index and stock index futures: A thermal optimal path method. *Physica A , 444*, 63-72.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica , 37*, 424-38.

Granger, C. W., & Ding, Z. (1996). Modelling volatility persistence of speculative returns: a new approach. *Journal of Econometrics , 73*.

Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics . *Journal of Econometrics 2* , 111-120.

Granger, C. W., & Yeon, Y. (1997). Measuring lag structure in forecasting models—the introduction of time distance.

Grech , D., & Mazur, Z. (2004). Can one make any crash prediction in finance using the local hurst exponent idea? *Physica A: Statistical Mechanics and its Applications , 334*, 133-145.

Grech, D., & Pamula, G. (2008). The local Hurst exponent of the financial time series in the vicinity of crashes on the Polish stock exchange market. *ScienceDirect , 387*, 4299-4308.

Greenspan, A. (2007). *The Age Of Turbulence.* The Penguin Press.

Grosse, I., Bernaola-Galvan, P., Carpena, P., Roman-Roldan, R., Oliver, J., & Stanley, E. H. (2002). Analysis of symbolic sequences using the Jensen-Shannon divergence. *Physical Review E , 65*.

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review , 70* (3), 393–408.

Gu, G.-F., & Zhou, W.-X. (2010). Detrending moving average algorithm for multifractals. *Physical Review E , 82* (011136).

Guo, K., Zhou, W.-X., Cheng, S.-W., & Sornette, D. (2011, August 10). The US Stock Market Leads the Federal Funds Rate and Treasury Bond Yields. *PLos ONE* .

Halpin-Healy, T., & Zhang, Y.-C. (1995). Kinetic roughening phenomena, stochastic growth directed polymers and all that. *Phys. Rep. , 254*, 215–415.

Harras, G., & Sornette, D. (2011). How to grow a bubble: a model of myopic adapting agents. *J. Econ. Behav. Organ. , 80* (1), 137-152.

Hasan, R., & Mohammad Salim, M. (2015). Multifractal analysis of Asian markets during 2007–2008 financial crisis. *Physica A , 419*, 746-761.

Hatemi-J, A., & Roca, E. (2011). How globally contagious was the recent US real estate market crisis? Evidence based on a new contagion test . *Economic Modeling , 28*, 2560-2565.

Holschneider, M. (1988). *J. Stat. Phys , 50* (953).

Hosking, J. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resour. Res. , 20* (12), 1898–1908.

Hou, Y., Liu, F., Gao, J., Cheng, C., & Song, C. (2017). Characterizing Complexity Changes in Chinese Stock Markets by Permutation Entropy. *entropy , 19* (514).

Hristu-Varsakelis, D., & Kyrtsou, C. (2008). Evidence for Nonlinear Assymetric Causality in US Inflation, Metal, and Stock Returns. *Discrete Dynamics in Nature and Society* .

Hristu-Varsakelis, D., & Kyrtsou, C. (2010). Testing for Granger Causality in the Presence of Chaotic Dynamics. *BRUSSELS ECONOMIC REVIEW , 53* (2).

Hu, H.-C., Zhou, W.-X., & Sornette, D. (2017). Time-dependent lead-lag relationship between the onshore and offshore Renminbi exchange rates. *Journal of International Financial Markets, Institutions & Money , 49*, 173-183.

Hu, K., Ivanov, P. C., Chen, Z., Carpena, P., & Stanley, E. H. (2001). *64* (011114).

Huang, J., Yong, W.-A., & Hong, L. (2016). Generalization of the Kullback–Leibler divergence in the Tsallis statistics. *Journal of Mathematical Analysis and Applications , 436*, 501-512.

Hui, E. C.-M., Ng, I., & Lau, O. M.-F. (2011). Speculative bubbles in mass and luxury properties: an investigation of the Hong Kong residential market. *Construction Management and Economics , 29*, 781-793.

Hurst, H. E., Black, R., & Sinaika, Y. M. (1965). Long- Term Storage in Reservoirs: An experimental Study.

Hurst, H. (1951). Long-term storage capacity of reservoirs. *116*, 770-808.

Hwa, T., & Lassig, M. (1998). Optimal detection of sequence similarity by local alignment. (P. P. S. Istrail, Ed.) 109–116.

Hwa, T., & Lassig, M. (1996). Similarity-detection and localization. *Phys. Rev. Lett. , 76*, 2591–2594.

Ihlen, E. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. *frontiers in Physiology , 3*.

Itakura, F. (1974). Minimum Prediction Residual Principle Applied to Speech Recognition.

Ito, M., & Sugiyama, S. (2009). Measuring the degree of time varying market inefficiency. *Economics Letters , 103*, 62-64.

Ivanova, K., & Ausloos, M. (1999). Low q-moment multifractal analysis of Gold price, Dow Jones Industrial Average and BGL-USD . *European Physical Journal B , 8*, 665-669.

Jauregui, M., Zunino, L., Lenzi, E., Mendes, R., & Ribeiro, H. (2018). Characterization of time series via Renyi complexity-entropy curves. *Physica A , 498*, 74-85.

Jiang, Z. Q., & Zhou, W. X. (2008). Multifractal analysis of chinese stock volatilities based on the partition function approach. *Physica A: Statistical Mechanics and its Applications , 387* (19).

Jiang, Z. Q., & Zhou, W. X. (2008). Multifractality in stock indexes, Fact or Fiction? *Physica A , 387* (3605-3614).

Johansen, A., Ledoit, O., & Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance , 3* (2), 219-255.

Johansen, A., Ledoit, O., & Sornette, D. (1998). Crashes at Critical Points. *Internationan Journal of Theoretical and Applied Finance* .

Jordan, D., Stockmanns, G., Kochs, E. F., Pilge, S., & Schneider, G. (2008). Electroencephalographic Order Pattern Analysis for the Separation of Consciousness and Unconsciousness . *Anesthesiology , 109*, 1014 –22 .

Kalimeri, M., Papadimitriou, C., Balasis, G., & Eftaxias, K. (2008). Dynamical complexity detection in pre-seismic emissions using nonadditive Tsallis entropy . *Physica A , 387*, 1161-1172.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A , 316* (1-4), 87-114.

Kapopoulos, P., & Siokis, F. (2005). Stock and real estate prices in Greece: wealth versus 'credit-price' effect. *Applied Economics Letters , 12*, 125-128.

Karlin, S., & Brendel, V. (1993). Patchiness and Correlations in DNA sequences. *Science , 259*, 677-680.

Keller, K., & Lauffer, H. (2003). Symbolic Analysis of High-dimensional Time Series. *International Journal of Bifurcation and Chaos , 13* (9), 2657–2668.

Keller, K., Mangold, T., Stolz, I., & Werner, J. (2017). Permutation Entropy: New Ideas and Challenges. *Entropy , 134*.

Keller, K., Sinn, M., & Emonds, J. (2007). TIME SERIES FROM THE ORDINAL VIEWPOINT . *Stochastics and Dynamics , 7* (2), 247–272 .

Keogh, E. J., & Pazzani, M. J. (2001). Derivative Dynamic Time Warping.

Keogh, E. J., & Pazzani, M. J. (1999). Scaling up dynamic time warping to massive dataset In PKDD '99: Proceedings of the Third European Conference on Principles of Data Mining and Knowledge Discovery.

Keogh, E. J., & Pazzani, M. J. (1999). Scaling up Dynamic Time Warping to Massive Datasets . 1-11.

Keogh, E., & Pazzani, M. (2000). Scaling up dynamic time warping for datamining applications.

Keogh, E., & Smyth, P. (1997). A Probabilistic Approach to Fast Pattern Matching in Time Series Databases.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of Information. *Problemy Peredachi Informatsii , 1* (1), 3-11.

Kowalski, A. M., & Plastino, A. (2012). The Tsallis-complexity of a semiclassical time-evolution . *ELSEVIER , 391*, 5375-5383.

Kurz-Kim, J.-R. (2012). Early warning indicator for financial crashes using the log periodic power law. *Applied Economics Letters , 19*, 1465-1469.

Kwapien, J. P. (2005). Components of multifractal- ity in high-frequency stock returns. *Physica A: Statistical Mechanics and its Applications , 350* (2).

Lai, L., & Guo, K. (2017). The performance of one belt and one road exchange rate: Based on improved singular spectrum analysis. *Physica A , 483*, 299-308.

Lam, K. C., Yu, C. Y., & Lam, K. Y. (2008). An Artificial Neural Network and Entropy Model for Residential Property Price Forecasting in Hong Kong. *Journal of Property Research* , 321-342.

Lamberti, P. W., & Majtey, A. P. (2003). Non-logarithmic Jensen–Shannon divergence. *Physica A , 329*, 81-90.

Lamberti, P. W., Majtey, A. P., Borras, A., Casas, M., & Plastino, A. (2013). Metric character of the quantum Jensen-Shannon divergence.

Lamberti, P. W., Martin, M. T., Plastino, A., & Rosso, O. A. (2004). Intensive entropic non-triviality measure. *Physica A , 334* (119), 119-131.

Latora, V., & Baranger, M. (1999). Kolmogorov-Sinai Entropy Rate versus Physical Entropy. *Phys. Rev. Lett. , 82*.

Leite, F. S., da Rocha, A. F., & Carvalho, J. (2010). MATLAB SOFTWARE FOR DETRENDED FLUCTUATION ANALYSIS OF HEART RATE VARIABILITY.

Levin , E. J., & Wright, R. E. (1997b). Speculation in the housing market? *Urban Studies , 34*, 1419-1437.

Levin, E. J., & Wright, R. E. (1997a). The impact of speculation on house prices in the United Kingdom. *Economic Modelling , 14*, 567-585.

Li, D., Li, X., Liang, Z., Voss, L. J., & Sleigh, J. W. (2010). Multiscale permutation entropy analysis of EEG recordings during sevoflurane anesthesia. *J. Neural Eng. , 7*.

Li, D., Liang, Z., Wang, Y., Hagihira, S., Sleigh, J. W., & Li, X. (2013). Parameter selection in permutation entropy for an electroencephalographic measure of isoflurane anesthetic drug effect. *J. Clin Monit Comput , 27*, 113-123.

Li, J., Yan, J., Liu, X., & Ouyang, G. (2014). Using Permutation Entropy to Measure the Changes in EEG Signals During Absence Seizures. *Entropy* , 3049-3061.

Li, X., & Richards, D. A. (2007). Predictability analysis of absence seizures with permutation entropy. *Epilepsy Research* .

Li, X., Cui, S., & Voss, L. J. (2008). Using permutation entropy to measure the electroencephalographic effects of sevoflurane. *Anesthesiology , 9*, 448-456.

Lim, K.-P., Brooks, R.-D., & Kim, J.-H. (2008). Financial crisis and stock market efficiency: Empirical evidence from Asian countries. *ScienceDirect International Review of Financial Analysis , 17*, 571-591.

Liu, Y., Gopikrishnan, P., Cizeau, P., Meyer, M., Peng, C.-K., & Stanley, H. E. (1999). Self-organized complexity in economics and finance. *Physical Review E , 60*, 1390–1400.

Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica , 59*, 1279-1313.

Lopez-Ruiz, R., Mancini, H. L., & Calbet, X. (1995). A statistical measure of complexity. *Physics Letters A , 209*, 321-326.

Mali P., M. A. (2016). Multifractal detrended moving average analysis of particle density functions in relativistic nuclear collisions . *Physica A , 450*, 323-332.

Malkiel, B. (1973). A Random Walk Down Wall Street.

Mandelbrot B.B., V. N. (1968). SIAM Rev.

Mandelbrot, B. (1999). A multifractal walk down Wall Street. *Scientific American , 298* (7).

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business , 36*, 392–417.

Mandelbrot, B. B., & Wallis, J. (1969). *5* (242).

Mandelbrot, B. B., & Wallis, J. R. (1969). *Water Resour. Res. , 5* (228).

Mandelbrot, B. B., & Wallis, J. R. (1969). *Water Resour. Res. , 5* (260).

Mandelbrot, B. B., & Wallis, J. R. (1969). *5* (967).

Mandelbrot, B. (1997). *Fractals and Scaling in Finance.* New York: Springer.

Mandelbrot, B. (1983). *The Fractal Geometry of Nature.* New York: Freeman.

Mandelbrot, B., Fisher, A., & Calvet, L. (1997). A Multifractal Model of Asset Returns . *Yale University* .

Martin, M. T., Plastino, A., & Rosso, O. A. (2006). Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A , 369*, 439-462.

Martina, E., Rodriguez, E., Escarela-Perez, R., & Alvarez-Ramirez, J. (2011). Multiscale entropy analysis of crude oil price dynamics. *Energy Econ. , 33* (5), 936–947.

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Phys. Lett. A , 302*, 299-307.

Marwan, N., Thiel, M., & Nowaczyk, N. R. (2002). Cross recurrence plot based synchronization of time series. *Nonlinear Process. Geophys. , 9*, 325–331.

Matilla-Garcia, M. (2007). A non-parametric test for independence based on symbolic dynamics. *Journal of Economic Dynamics & Control , 31*, 3889-3903.

Matilla-Garcia, M., & Marin, M. R. (2008). A non-parametric independence test using permutation entropy . *Journal of Econometrics , 144*, 139-155.

McCarthy, J., & Peach, R. (2004). Are home prices the next bubble? *Economic Policy Review , 10* (3).

McCarthy, J., & Peach, R. (2005). Is There a 8 Bubble in the Housing Market Now? *Networks Financial Institute Policy Brief* .

McCluskey, W. (1996). Predictive accuracy of machine learning models for the mass appraisal of residential property. *New Zealand Values' Journal , 16* (1), 41-47.

Mendes, R. S., Evangelista, L. R., Thomaz, S. M., Agostinho, A. A., & Gomes, L. C. (2008). A unified index to measure ecological diversity and species rarity . *Ecography , 31*, 450-456.

Meng, H., Xu, H.-C., Zhou, W.-X., & Sornette, D. (2017). Symmetric thermal optimal path and time-dependent lead-lag relationship: novel statistical tests and application to UK and US real-estate and monetary policies. *Quantitative Finance , 17* (6), 959-977.

Micco, L. D., González, C. M., Larrondo, H. A., Martin, M. T., Plastino, A., & Rosso, O. A. (2008). Randomizing nonlinear maps via symbolic dynamics. *Physica A , 387* (14), 3373-3383.

Moody, J., & Wu, L. (1996, October). Improved Estimates for the Rescaled Range and Hurst exponents. 537-553.

Muellbauer, J., & Murphy, A. (1997). Booms and Busts in the UK Housing Market. *The Economic Journal , 107*, 1701-1727.

Muzy J.-F., B. E. (1991). *Phys. Rev. Lett. , 67* (3515).

Muzy J.-F., B. E. (1993). *J. Stat. Phys. , 70* (635).

Muzy J.-F., B. E. (1993). *Phys. Rev. E , 47* (875).

Muzy J.-F., B. E. (1994). *Int. J. Bifurcation Chaos Appl. Sci. Eng. , 4* (245).

Myer, F. N., & Webb, J. R. (1994). Retail stocks, retail REITs and retail real estate. *Journal of Real Estate Research , 9*, 65-84.

Myers, C., Rabiner, L., & Rosenberg, A. (1980). Performance tradeoffs in dynamic time warping algorithms for isolated word recognition,. *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions , 28* (6), 623-635.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applic- able to the search for similarities in the amino acid sequence of two proteins. *J. Molec. Biol. , 48* (3), 443-453.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica , 55* (3), 703-708.

Oh, G., Kim, S., & Eom, C. (2007). Market efficiency in foreign exchange markets. *Physica A , 382* (1), 209–212.

Olofsen, E., Sleigh, J. W., & Dahan, A. (2008). Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. *British Journal of Anaesthesia* , 810-21.

Olsen, R., Hwa, T., & Lassig, M. (1999). Optimizing Smith–Waterman alignments. 302–313.

Ortiz-Cruz, A., Rodriguez, E., Ibarra-Valdez, C., & Alvarez-Ramirez, J. (2012). Efficiency of crude oil markets: Evidences from informational entropy analysis. *Energy Policy , 41*, 365–373.

Oswiecimka, P., Kwapien, J., Drozdz, S., Gorski, A. Z., & Rak, R. (2008). Different fractal properties of positive and negative returns.

Ott, E., Sauer, T., & Yorke, J. A. (1994). *Coping with Chaos.* New York: Wiley.

Ouyang, G., Li, J., Liu, X., & Li, X. (2013). Dynamic characteristics of absence EEG recordings with multiscale permutation entropy analysis. *Epilepsy Research , 104*, 246-252.

Panana, A., Kyrtsou, C., Kugiumtzis, D., & Diks, C. (2017). Financial Networks based on Granger causality: A case study. *Physica A , 482*, 65-73.

Peng, C. K., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M., et al. (1992). London: Nature.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E , 49* (2).

Peng, C.-K., Havlin, S., Stanley, E. H., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *CHAOS , 5* (1).

Percival, D. B., & Walden, A. T. (2000). Wavelet Methods for Time Series Analysis .

Podobnik, B., Wang, D., Horvatic, D., Grosse, I., & Stanley, H. E. (2010). Time-lag cross-correlations in collective phenomena. *EPL , 90*.

Poeschel, T., Ebeling, W., & Rose, H. (1995). Guessing probability distributions from small samples.

Powell, G. E., & Percival, I. C. (1979). A spectral entropy method for distinguishing regular and irregular motion of Hamiltonian systems. *J. Phys. A: Math. Gen. , 12*, 2053–2071.

Quigley, J. M. (1999). Real Estate Prices and Economic Cycles . *INTERNATIONAL REAL ESTATE REVIEW , 2* (1), 1-20.

Quiroga, R. Q., Kreuz, T., & Grassberger, P. (2002). Event synchronization: a simple and fast method to measure synchronicity and time delay patterns. *Phys. Rev. E , 66*.

Rabiner, L., & Juang, B. (1993). Fundamentals of speech recognition.

Raj, A., & Wiggins, C. H. (2008). A non-negative expansion for small Jensen-Shannon Divergences.

Ribeiro, H. V., Jauregui, M., Zunino, L., & Lenzi, E. K. (2017). Characterizing time series via complexity-entropy curves . *PHYSICAL REVIEW E , 95*.

Riedl, M., Mueller, A., & Wessel, N. (2013). Practical considerations of permutation entropy: A tutorial review. *The European Physical Journal Special Topics , 222*, 249-262.

Risso, W. A. (2008). The informational efficiency and the financial crashes. *Res. Int. Bus. Finance , 22* (3), 396–408.

Rosso, O. A., & Mairal, M. L. (2002). Characterization of time dynamical evolution of electroencephalographic epileptic records. *Physica A , 312*, 469–504.

Rosso, O. A., Carpi, L. C., Saco, P. M., Ravetti, M. G., Larrondo, H. A., & Plastino, A. (2012). The Amigo´ paradigm of forbidden/missing patterns: a detailed analysis . *THE EUROPEAN PHYSICAL JOURNAL B , 85* (419).

Rosso, O. A., Carpi, L. C., Saco, P. M., Ravetti, M. G., Plastino, A., & Larrondo, H. A. (2012). Causality and the Entropy-Complexity Plane: Robustness and Missing Ordinal Patterns. *391*, 42-55.

Rosso, O. A., Larrondo, H. A., Martin, M. T., Plastino, A., & Fuentes, M. A. (2007). Distinguishing Noise from Chaos . *Physical Review Letters , 99*.

Ruan, Y.-P., & Zhou, W.-X. (2011). Long-term correlations and multifractal nature in the intertrade durations of a liquid Chinese stock and its warrant. *Physica A , 390*, 1646-1654.

Ruiz, M., Guillamon, A., & Gabaldon, A. (2012). A New Approach to Measure Volatility in Energy Markets . *entropy , 14*, 74-91.

Sakoe, H., & Chiba, S. (1978). Dynamic Programming Algorithm Optimization for Spoken Word Recognition .

Salvador, S., & Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intell. Data Anal. , 11* (5), 561–580.

Sankoff, D., & Kruskal, J. B. (1983). Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison.

Sargent, T. J. (1976). The observational equivalence of natural and unnatural rate theories of macroeconomics. *Journal of Political Economy , 84*, 631-640.

Schmill, M., Oates, T., & Cohen, P. (1999). Learned models for continuous planning.

Schmitt, F. G., Angounou, L., & Angounou, T. (2011). Multifractal analysis of the dollar–yuan and euro–yuan exchange rates before and after the reform of the peg. *Quantitative Finance , 11* (4), 505-513.

Serinaldi, F., Zunino, L., & Rosso, O. A. (2014). Complexity–entropy analysis of daily stream flow time series in the continental United States. *Stoch. Environ. Res. Risk Assess. , 28*, 1685-1708.

Shao, Y. H., Gu, G. F., Jiang, Z. Q., Zhou, W. X., & Sornette, D. (2012). Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series. *Scientific Reports , 2* (835).

Shen, Y., Hui, E., & Liu, H. (2005). Housing price bubbles in Beijing and Shanghai. *43* (4).

Shiller, R. (2015). Irrational Exuberance.

Shimizu Y., T. S. (2002). Multifractal spectra as a measure of complexity in human posture. *Fractals , 10* (01), 103-116.

Shinde, R. B., & Pawar, V. P. (2014). Dynamic time Warping using MATLAB & PRAAT. *International Journal of Scientific & Engineering Research , 5* (5).

Silva, D. F., & Batista, G. E. (2016, June). Speeding Up All-Pairwise Dynamic Time Warping Matrix Calculation.

Sing, T. B., Tsai, I.-C., & Chen, M.-C. (2006). Price dynamics in public and private housing markets in Singapore. *Journal of Housing Economics , 15*, 305-320.

Siokis, F. M. (2018). Credit market Jitters in the course of the financial crisis: A permutation entropy approach in measuring informational efficiency in financial assets . *Physica A , 499*, 266-275.

Siokis, F. M. (2014). European economies in crisis: A multifractal analysis of disruptive economic events and the effects of financial assistance. *Physica A , 395*, 283-292.

Siokis, F. M. (2017). Financial markets during highy anxious time: Multifractal fluctuations in asset returns.

Siokis, F. M. (2013). Multifractal analysis of stock exchange crashes. *Physica A , 392*, 1164-1171.

Siokis, F. M. (2012). The dynamics of a complex system: The exchange rate crisis in Southeast Asia. *Economic letters , 114*, 98-101.

Solomonoff, R. J. (1964). A Formal Theory of Inductive Inference. Part II. *INFORMATION AND CONTROL , 7* (2), 224-254.

Soriano, M. C., Zunino, L., Larger, L., Fischer, I., & Mirasso, C. R. (2011). Distinguishing fingerprints of hyperchaotic and stochastic dynamics in optical chaos from a delayed opto-electronic oscillator. *Optics Letters , 36* (12).

Soriano, M. C., Zunino, L., Rosso, O. A., Fischer, I., & Mirasso, C. R. (2011). Time Scales of a Chaotic Semiconductor Laser with Optical Feedback Under the Lens of a Permutation Information Analysis . *Journal of Quantum Electronics , 47* (2).

Sornette, D. (1999). Complexity, catastrophe and physics. *Physics World* .

Sornette, D. (2003). Critical market crashes. *Physics Reports , 378* (1), 1-98.

Sornette, D. (2004). *Critical Phenomena in Natural Sciences - Chaos, Fractals, Self-organization and Disorder: Concepts and Tools.* Berlin: Springer.

Sornette, D. (2004). *Why stock markets crash?* Princeton, New Jersey: Princeton Univeristy Press.

Sornette, D., & Zhou, W.-X. (2005). Non-parametric determination of real-time lag structure between two time series: the 'optimal thermal causal path' method. *Quantitative Finance , 5* (6), 577–591 .

Sornette, D., Johansen, A., & Bouchaud, J.-P. (1996). stock market crashes precursors and replicas. *Journal de Physique I , 6* (1), 167-175.

Sornette, D., Woodard, R., Yan, W., & Zhou, W.-X. (2013). Clarifications to questions and criticisms on the Johansen–Ledoit–Sornette financial bubble model. *Physica A , 392*, 4417-4428.

Sowell, F. B. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics , 53*, 165-188.

Sprott, J. C. (2004). *Chaos and Time Series Analysis.* Oxford University Press.

Staniek, M., & Lehnertz, K. (2007). Parameter Selection for permutation entropy measurements. *Int. J. Bifurcation Chaos , 17* (10), 3729-3733.

Strozzia, F., Zaldvarb, J.-M., & Zbilutc, J. P. (2002). Application of nonlinear time series analysis techniques to high-frequency currency exchange data. *Physica A , 312*, 520–538.

Struzik, Z. R. (2002). Wavelet transform based multifractal formalism in outlier detection and localisation for financial time series. *Physica A: Statistical Mechanics and its Applications , 309* (3).

Sun, X., Chen, H., Wu, Z., & Yuan, Y. (2001). Multifractal analysis of Hang Seng Index in Hong Kong stock market. *Physica A , 291* (553-562).

Sun, X., Chen, H., Yuan, Y., & Wu, Z. (2001). Predictability of multifractal analysis of Hang Seng stock index in Hong Kong. *Physica A , 301* (1-4), 473-482.

Taherkhani, F., Rahmani, M., Taherkhani, F., Akbarzadehd, H., & Abroshan, H. (2013). Permutation entropy and detrend fluctuation analysis for the natural complexity of cardiac heart interbeat signals . *Physica A , 392*, 3106–3112 .

Telesca, L., & Lapenna, V. (2006). Measuring multifractality in seismic sequences. *Tectonophysics ScienceDirect , 423*, 115-123.

Tenreiro Machado, J. A., Costa, A. C., & Quelhas, M. D. (2011). Shannon, Rényie and Tsallis entropy analysis of DNA using phase plane. *Nonlinear Analysis: Real World Applications , 12*, 3135-3144.

Teverovsky, V., Taquu, M. U., & Willinger, W. (1999). A critical look at Lo's modified R/S statistic. *Journal of Statistical Planning an Inference , 80*, 211-227.

Traversaro, F., Risk, M., Rosso, O., & Redelico, O. (2017, May). An empirical evaluation of alternative methods of estimation for Permutation Entropy in time series with tied values .

Tsallis, C. (2016). Approach of Complexity in Nature: Entropic Nonuniqueness . *axioms* .

Tsallis, C. (2017). Economics and Finance: q-Statistical Stylized Features Galore. *entropy , 19* (457).

Tsallis, C. (2014). News and Views: About Complexity and Why to Care. *Braz J Phys , 44*, 283-285.

Tsallis, C. (1988). Possible Generalization of Boltzmann-Gibbs Statistics. *Journal of Statistical Physics , 52*.

Tsallis, C., Anteneodo, C., Borland, L., & Osorio, R. (2003). Nonextensive statistical mechanics and economics. *Physica A , 324*, 89-100.

Tsekouras, G. A., & Tsallis, C. (2004). Generalized entropy arising from a distribution of q-indices.

Turiel, A. P.-V. (2003). Multifractal geometry in stock market time series. *Physica A: Statistical Mechanics and its Applications , 322*.

Unakafova, V. A., & Keller, K. (2013). Efficiently Measuring Complexity on the Basis of Real-World Data. *entropy , 15*, 4392-4415.

Vanderwalle, N., & Ausloos, M. (1997). Coherent and rnadom sequences in financial fluctuations. *Physica A , 246*, 454-459.

Wang, D., Podobnik, B., Horvatic, D., & Stanley, H. E. (2011). Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. *Physical Review E , 83*.

Wang, D., Tu, J., Chang, X., & Li, S. (2017). The lead–lag relationship between the spot and futures markets in China. *Quantitative Finance , 17* (9), 1447-1456.

Wang, D.-H., Yu, X.-W., & Suo, Y.-Y. (2012). Statistical properties of the yuan exchange rate index . *Physica A , 391*, 3503-3512.

Wang, F., Wang, L., & Zou, R. B. (2014). Multifractal detrended moving average analysis for texture representation. *Chaos: An interdisciplinary Journal of Nonlinear Science , 24*.

Wang, X.-H., Havlin, S., & Schwarz, M. (1999). Directed polymers at finite temperatures in 1+1 and 2+1 dimensions.

Wang, Y., Wei, Y., & Wu, C. (2011). Analysis of the efficiency and multifractality of gold markets based on multifractal detrended fluctuation analysis. *Physica A , 390*, 817-827.

Wang, Y., Wu, C., & Pan, Z. (2011). Multifractal detrending moving average analysis on the US Dollar exchange rates. *Physica A , 390*, 3512-3523.

Weck, P. J., Schaffner, D. A., Brown, M. R., & Wicks, R. T. (2015). Permutation entropy and statistical complexity analysis of turbulence in laboratory plasmas and the solar wind. *Physical Review E , 91*.

Wei, Y., & Huang, D. (2005). Multifracral analysis of SSEC in Chinese stock market: A different empirical result from Heng Seng index. *Physica A , 355* (2-4), 497-508.

Wilson, P., & Zurbruegg, R. (2004). Contagion or Interdependence? Evidence from Comovements in Asia Pacific Securitised Real Estate Markets During the 1997 Crisis. *Journal of Property Investment & Finance , 22* (5), 401-413.

Xi, C., Zhang, S., Xiong, G., & Zhao, H. (2016). A comparative study of two-dimensional multifractal detrended fluctuation analysis and two-dimensional multifractal detrended moving average algorithm to estimate the multifractal spectrum. *Physica A , 454*, 34-50.

Xu, H.-C., Zhou, W.-X., & Sornette, D. (2017). Time-dependent lead-lag relationship between the onshore and offshore Renminbi exchange rates . *Journal of International Financial Markets, Institutions & Money , 49*, 173-183.

Xu, L., Ivanov, P. C., Hu, K., Chen, Z., Carbone, A., & Stanley, E. H. (2005). *Phys. Rev. E , 71* (051101).

Yalamova, R. (2006). WAVELET TEST OF MULTIFRACTALITY OF ASIA-PACIFIC INDEX PRICE SERIES. *AAMJAF , 2* (1), 63-83.

Yao, D., Yang, J., Bai, Y., & Cheng, X. (2016). Railway rolling bearing fault diagnosis based on multi-scale intrinsic mode function permutation entropy and extreme learning machine classifier . *Advances in Mechanical Engineering , 8* (10), 1-9.

Yi, B., Jagadish, H., & Faloutsos, C. (1998). Efficient retrieval of similar time sequences under time warping.

Yudong W., C. W. (2012). Efficiency of Crude Oil Futures Markets: New Evidence from Multifractal Detrending Moving Average Analysis . *42*, 393-414.

Yudong, W., Chongfeng, W., & Zhiyuan, P. (2011). Multifractal detrending moving average analysis on the exchange rates. *Physica A , 390*, 3512-3523.

Zanin, M. (2008). Forbidden patterns in financial time series .

Zanin, M., Zunino, L., Rosso, O. A., & Papo, D. (2012). Permutation Entropy and Its Main Biomedical and Econophysics Applications: A Review . *entropy , 14*, 1553-1577.

Zhang, Y. C. (1999). Toward a theory of marginally efficient markets. *Physica A , 269* (1), 30–44.

Zhou , W.-X., & Sornette, D. (2006). Non-parametric determination of real-time lag structure between two time series: The ‘‘optimal thermal causal path’’ method with applications to economic data. *Journal of Macroeconomics , 28*, 195–224 .

Zhou, R., Cai, R., & Tong, G. (2013). Applications of Entropy in Finance: A Review. *entropy , 15*, 4909-4931.

Zhou, W., Dang, Y., & Gu, R. (2013). Efficiency and multifractality analysis of CSI 300 based on multifractal detrending moving average algorithm. *Physica A , 392*, 1429-1438.

Zhou, W.-X. (2009). The components of empirical multifractality in financial returns. *EPL , 88*.

Zhou, W.-X., & Sornette, D. (2003). 2000–2003 real estate bubble in the UK but not in the USA. *Physica A , 329* (1-2), 249-263.

Zhou, W.-X., & Sornette, D. (2004). Causal slaving of the US treasury bond yield antibubble by the stock market antibubble of August 2000 . *Physica A , 337*, 586-608.

Zhou, W.-X., & Sornette, D. (2006). Is there a real-estate bubble in the US? *Physica A , 361* (1), 297-308.

Zhou, W.-X., & Sornette, D. (2007). Lead-lag cross-sectional structure and detection of correlated–anticorrelated regime shifts: Application to the volatilities of inflation and economic growth rates. *Physica A , 380*, 287-296.

Zipf, G. K. (1949). Human Behavior and the Principle of Least Effort.

Zunino, L., Bariviera, A. F., Guercio, M. B., Martinez, L. B., & Rosso, O. A. (2016). Monitoring the informational efficiency of European corporate bond markets with dynamical permutation min-entropy. *Physica A , 456*, 1-9.

Zunino, L., Bariviera, A. F., Guercio, M. B., Martinez, L. B., & Rosso, O. A. (2012). On the efficiency of sovereign bond markets. *Physica A , 391* (18), 4342–4349.

Zunino, L., Perez, D. G., Kowalski, A., Martin, M. T., Garavaglia, M., Plastino, A., et al. (2008). Fractional Brownian motion, fractional Gaussian noise, and Tsallis permutation entropy. *Physica A , 387*, 6057-6068.

Zunino, L., Perez, D. G., Martin, M. T., Garavaglia, M., Plastino, A., & Rosso, O. A. (2008). Permutation entropy of fractional Brownian motion and fractional Gaussian noise. *Physics Letters A , 372*, 4768-4774.

Zunino, L., Soriano, M. C., Fischer, I., Rosso, O. A., & Mirasso, C. R. (2010). Permutation-information-theory approach to unveil delay dynamics from time-series analysis . *Physical Review E , 82*.

Zunino, L., Tabak, B. M., Serinaldi, F., Zanin, M., Perez, D. G., & Rosso, O. A. (2011). Commodity predictability analysis with a permutation information theory approach. *Physica A , 390*, 876-890.

Zunino, L., Zanin, M., Tabak, B. M., & Pérez, D. G. (2009). Forbidden patterns, permutation entropy and stock market inefficiency. *Physica A , 388*, 2854-2864.

Zunino, L., Zanin, M., Tabak, B. M., Pérez, D. G., & Rosso, O. A. (2010). Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency. *Physica A , 389* (9), 1891–1901.

Zunino, L., Zanin, M., Tabak, B. M., Pérez, D. G., & Rosso, O. A. (2009). Forbidden patterns, permutation entropy and stock market inefficiency. *Physica A , 388*, 2854-2864.