Title	Examining the dynamical transition in the Dow Jones Industrial Index from Bull to Bear market using Recurrence Quantification Analysis
Author(s)	Moloney, Kitty; Raghavendra, Srinivas
Publication Date	2012-11-20
Item record	http://hdl.handle.net/10379/3063

Examining the dynamical transition in the Dow Jones Industrial Index from Bull to Bear market using recurrence quantification analysis

Kitty Moloney[*]and Srinivas Raghavendra[†]

Abstract

We present evidence of phase transitions (periodic to chaotic and chaotic to chaotic) in the Dow Jones Industrial Index as it transitions from Bull to Bear market. There is also evidence of a completely unpredictable (i.e., nondeterministic) regime just as the market peaks. The noisy trader theory is suggested as the economic explanation for this unpredictability i.e. rational but uninformed traders chase noise rather than the usual macro economic and financial variables. We suggest that the collapse in determinism allows the dynamics of the market to break from the past and that the market is in fact piecewise deterministic. A principal component series is developed and named the random market indicator, (RMI). This can be used to indicate when the market is transitioning. The RMI indicator could be used by market participants, financial regulators and policy makers as an indicator of market crisis. During times of crises, quantitative risk estimation techniques such as stationary value at risk models, will give misleading results and should not be used.

1 Introduction

The twentieth century provides us with a rich repository of data in terms of financial crises. We have had numerous crashes but the 1929 crash and 2007 crash stand out in terms of their persistence and their effect on the real economy. However, the frequency of crashes and the vulnerability of the financial system

---

[*]Corresponding author. J.E.Cairnes School of Business and Economics, National University of Ireland Galway, Galway. Email: kitty.moloney@nuigalway.ie.

[†]J.E.Cairnes School of Business and Economics, National University of Ireland Galway, Galway.

has increased during the last three decades of the twentieth century. Even though this can be attributed to the nature of financial globalization in the last part of the twentieth century, the nature and subsequent impact of the latest crash seems to resemble the 1929 crash. In this context, it is pertinent to ask are there any quantifiable markers that we can derive from the '29 and '07 crashes that will help us understand the transition from a bubble state of the market to the crash. The objective of this paper is to precisely analyse the properties of such a transition. The following section summarizes the theoretical background to the methodology. Section 3 describes the empirical data to be used in the paper. Section 4 presents and examines the empirical findings and section 5 concludes the paper.

2   A brief review of Recurrence Analysis

Dynamical recurrence indicates a recurring trajectory and suggests the existence of an attractor. To analyse recurrence in a time series, we first refer to the embedding theorems (Sauer et al 1991) noting that the motion of a dynamical deterministic system can be reconstructed using the phase space of a single time series variable. Thus the scalar time series, s_n is embedded into a vector, choosing an embedding dimension m and time delay τ (Kantz and Screiber 2003), as follows:

$$x_i = (s_i, s_{i+\tau}, \ldots, s_{i+(m-1)\tau}) \tag{1}$$

The method of false nearest neighbours is used to find m, (Kantz and Screiber 2003). Assuming the time series is continuous, the nonlinear mutual information function (Fraser and Swinney 1986) is used to assess the value of τ. If the process is discontinuous or discrete, Webber and Zbilut (2005) advise that the delay is best set equal to 1 as no points in the time series will be skipped.

For low dimensional deterministic systems a 3 dimensional phase space reconstruction of the embedded time series may reflect the motion of the system. For higher dimensional systems such as financial data, (Moloney and Raghavendra forthcoming), an alternative approach must be used. One approach was recently suggested by Eckmann et al, (1987). By analyzing the recurrence in the embedded time series we can evaluate if the system is revisiting certain areas of the plane.

2.1   Recurrence Plots (RP)

A recurrence plot is a visual representation of recurrence in the system. The recurrence plot can be represented as follows:

$$R_{i,j} = I(\epsilon_i - \| x_i - x_j \|), \, x_i \in \Re^m, \, i.j = 1 \ldots n \tag{2}$$

where n is the number of considered states x_i, ϵ_i is a threshold distance, $\|.\|$ a norm and $I(.)$ the Heaviside function, (Marwan 2003). In the recurrence

plot, 2 close vectors are represented by a black dot, whereas all other vectors are represented by a white space.

2.1.1 The threshold, ϵ

Many methods have been suggested for the correct selection of the threshold value. One such method is to choose the threshold equal to 5% of the maximal phase space diameter (Basto and Caiado 2011). We found this approach lead to too large a threshold. Instead we followed the advice of Webber and Zbilut (2005) and applied the following guidelines;

(i) Threshold must fall within the linear scaling region of the double logarithmic plot of recurrences versus threshold.

(ii) Threshold must be such that the number of recurrences is kept low, ($\lesssim 2\%$)

Guideline (i) above has strong theoretical underpinnings as we choose a threshold such that the box counting dimension, (Webber and Zbilut 2005) can be estimated. Thus this threshold should illustrate the deterministic structure of the system. The second guideline ensures that we maintain the level of recurrences to a low value so that we can be confidant we are focusing on the deterministic structure rather than the noise. Thus the choice of threshold was made by applying both guidelines to all data sets and averaging the resultant estimated threshold.

2.1.2 Interpreting recurrence plots

RPs can be interpreted by following the pattern in the plot as the eye moves from the bottom left corner up to the top right corner following the central diagonal line, (known as the line of identity, LOI). By familiarizing ourselves with the recurrence plots of known deterministic systems, such as the Lorenz system for example, we can then search for similar patterns in empirical data (Kantz and Screiber 2003). Below we illustrate the phase space reconstruction of the Lorenz system of equations and the corresponding recurrence plot.

The Lorenz system of equations are;

$$\frac{dx}{dt} = \sigma \left(y - x \right) \tag{3}$$

$$\frac{dy}{dt} = -xz + rx - y \tag{4}$$

$$\frac{dz}{dt} = xy - bz \tag{5}$$

As noted by Eckmann et al. (1987), the recurrence plot "checkerboard type" structure indicates that the trajectories of the system are spinning around attractors.

Figure 1: (a) Three dimensional phase space reconstruction of a Lorenz curve with parameters $\sigma = 10, r = 28, b = \frac{8}{3}$ (b) its corresponding recurrence plot.
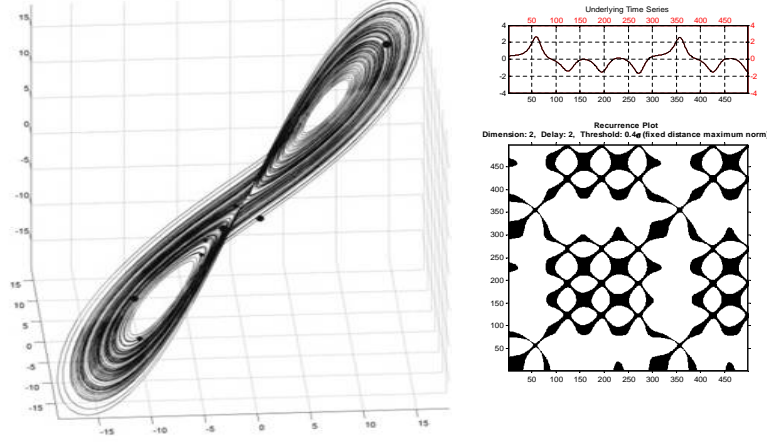
The analysis of recurrence plots can indicate if we are indeed analyzing a periodic, chaotic or random process. In figure 2, we present three such recurrence plots.

Figure 2: Recurrence plots of (a) periodic, (b) chaotic and (c)random processes
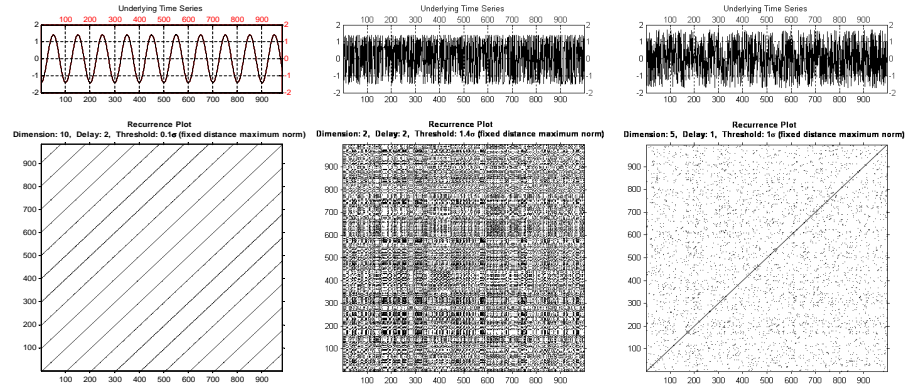
The periodic process is iterated from the following equation;

$$y = \sin\left[\frac{\pi}{50}\left(t\right)\right] \tag{6}$$

The chaotic process is iterated from the logistic map with $r = 3.999$;

$$x_{n+1} = rx_n\left(1 - x_n\right) \tag{7}$$

4

The random process is generated from random numbers. The periodic process is represented by long diagonal lines indicating the recurrence of the trajectories in the same areas of the phase plane. The chaotic process can be represented by a number of different structures; such as the checkerboard structure above for the Lorenz curve, or by a much more subtle structure; such as for the logistic map with r=3.999. At this parameter value the logistic map is extremely chaotic, yet small square-like structures can be seen in the recurrence plot. This indicates that the trajectories are spinning around, moving close to or being repelled from attractors. A completely random process will show no structure and will be represented by a series of random black dots. As RPs make no assumptions about the process, they can be used to analyse nonstationary systems and the "phase transition" (Marwan 2003) of a process into differing states. For these reasons, recurrence plots are particularly useful in the analysis of bull and bear markets in financial time series. With this in mind, we simulated an ARGARCH process, from random Gaussian numbers and the S&P 500, (details of the simulation to be found in the data section). The ARGARCH process represents the following equations;

$$r_t = \mu_t + \theta r_{t-1} + \varepsilon_t \tag{8}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{9}$$

$$\varepsilon_t \sim N(0,1) \tag{10}$$

This model was first suggested by Bollerslev (1986) and is commonly used in financial econometrics as it recognises the autocorrelation in returns and in the variance of the financial time series (Jondeau et al 2007). It has been found to be most useful in quantitative risk estimation, although there is significant evidence to show that financial data residuals are non-Gaussian, (Jondeau et al 2007). Even if a non-Gaussian distribution (such as a generalised extreme value distribution, (GEV)) is applied, the ARGARCH model still underpredicts risk in the financial markets (Jondeau et al 2007). Applying the recurrence plot methodology, we present the recurrence plot in figure 3.

Figure 3: ARGARCH process
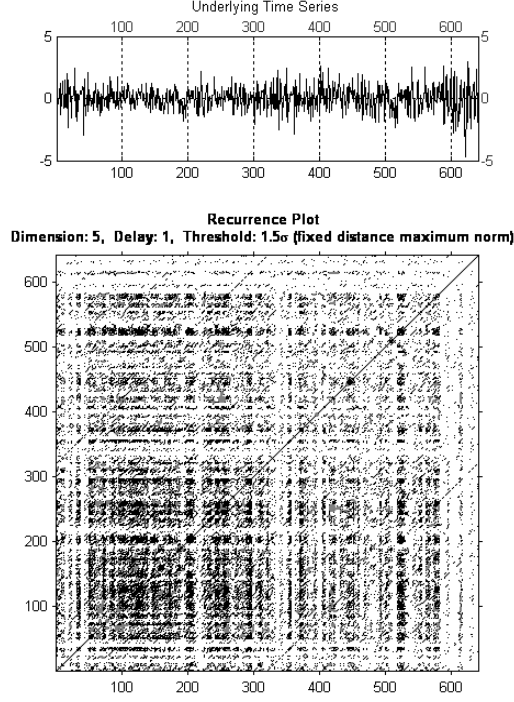
The structure of the recurrence plot of an ARGARCH model shows square structures. Zbilut (2004) noted similar squares in an RP of the S&P 500 and suggested they indicate the autocorrelations in the system. The ARGARCH model allows for periodic heteroscedastic variance, which is a stylised fact of the financial markets (Jondeau et al 2007). In this model the variance is increasing as we move towards observation 600. This is taken to be a stationary model based on the Bollerslev constraints; that is that the two parameters α_1, β must sum to less than or equal to one, (in this simulation that is. $0.083 + 0.916 = 0.999$) (Bollerslev 1986). The RP fades as we move from left to right up the centre line, (LOI), this indicates a trend in the system (Eckmann et al 1987) and may suggest remaining nonstationarities. To further analyse these dynamical systems we need to introduce recurrence quantification analysis.

2.2   Recurrence Quantification Analysis

The quantification of the patterns in recurrence plots came from the field of physiology (Zbilut and Webber 1992, Webber and Zbilut 1994), where statistical

values known as recurrence quantification analysis (RQA) were first proposed. These statistical values can indicate many surprising characteristics of a time series and can facilitate distinction between random, periodic and chaotic processes. More than fifteen interrelated statistical values have been proposed. we will focus on three of these values, that is determinism, (DET), the maximum diagonal line length, (Lmax) and the maximum vertical line length, (Vmax). We will also consider the proportion of recurrences, (RR). These measures can be used to extract meaningful information from dynamical systems which have no supporting mathematical theory or conceptualisation (Webber et al 2009). Physiological signals are nonlinear, nonstationary and noisy (Webber et al 2009) so too is financial data (Moloney and Raghavendra 2011). Hence the RQA approach allows us to analyse this data without imposing parametric or modeling constraints.

The first of the statistical values is the recurrence rate (RR) which is the proportion of recurrence points in the matrix relative to the total number of points. This can be expressed mathematically as;

$$RR_{(\epsilon)} = \frac{1}{n^2} \sum_{i,,j=1}^{n} R_{i,j(\epsilon)} \tag{11}$$

This measures the density of the recurrence points in the RP. It corresponds to the definition of the correlation integral, (Marwan et al 2007), except that the main diagonal line, (LOI) is not included. At the limit, $n \to \infty$, it gives the probability that a state recurs in the ϵ-neighbourhood of the phase space (Marwan et al 2007).

This density measure is then used as a benchmark for examining the dynamics of the recurrences. For example, determinism, (DET) is the proportion of diagonal lines relative to the recurrence rate. To measure this we develop a histogram $P(l)$ of diagonal lines of length l. As we see above in figure 2 (a) , periodic processes have long diagonal lines parallel to the main diagonal, (LOI) and no isolated recurrence points. Whereas random processes, (figure 2, (c)) have no or very few diagonal lines and many isolated recurrence points. Thus analysis of the proportion of diagonal lines, (of at least length l_{min}) relative to the total number of recurrences can indicate if a deterministic process is present. The measure DET can be expressed mathematically as;

$$DET = \frac{\sum_{l=l_{min}}^{n} lP(l)}{\sum_{l=1}^{n} lP(l)} \tag{12}$$

Marwan et al (2007) describe this as a measure of the predictability of the system, as the higher DET, the higher the number of recurrences that are in diagonal lines, indicating that through time the trajectory of the system is recurring. A recurring dynamical trajectory suggests the existence of attractors and that the system is deterministic. We note in the chaotic system (see figure 2(b)) there are many short diagonal lines. Therefore additional measures such as the maximum diagonal line length can also be useful in system diagnosis. The maximum diagonal line length can be expressed as;

$$Lmax = max\left(\{l_i\}_{i=1}^{n_l}\right) \tag{13}$$

Lmax has been associated with the speed of the divergence of the phase space trajectory. The faster the divergence, the smaller Lmax will be, we note, as above, the chaotic system in figure 2(b) has short diagonal lines. Eckmann et al. (1987) suggested that the length of the diagonal lines is related to the dynamical invariant, the Lyapunov exponent. Following from this, Trulla et al. (1996) studied the bifurcation behaviour of the logistic equation (equation 7) and noted a strong linear correlation, ($\rho = 0.912$) between the inverse of Lmax and the Lyapunov exponent, during the chaotic regions of the logistic map, (i.e. $r > 3.5688$). Marwan et al. (2007) developed this analysis further and showed that in fact the inverse of Lmax, (known as DIV) is related to the K_2 entropy, (that is the lower limit of the sum of the positive Lyapunov exponents). By analysing DIV, the inverse of Lmax, we can examine the speed of expansion of nearby trajectories into new areas of the state space. The higher DIV is, the more chaotic the system and the less predictable.

The final measure we will be using is the maximum vertical line length, (Vmax). This value is similar to Lmax except it measure the maximum of all the vertical line lengths,(v).This can be expressed as;

$$Vmax = max\left(\{v_l\}_{l=1}^{n_v}\right) \tag{14}$$

Marwan et al. (2007) describe a vertical line as a time when the state of the system is 'trapped' and suggest that this is typical of laminar states, (i.e., intermittency). Intermittencies occur when the system alternates between chaotic and periodic phases (Sprott 2004). Thus analysis of Vmax can facilitate the discovery of phase transitions and laminar states.

### 2.2.1   Interpreting RQA

Although some information can be gathered by measuring each statistical value for the whole data set, analysis of moving window (epoch) RQA can reveal far more about the dynamical nature of the system. We begin, by presenting moving window RQA results for the periodic and chaotic processes presented above in figure 2. The epoch window size is 100 observations and the step size is 1. The embedding dimensions, (m, τ, ϵ) are $(10, 2, 0.1)$ for the periodic process and $(2, 2, 0.5)$ for the chaotic process. As the random numbers used above in figure 2 (c) are computer generated, (from an algorithm), the RQA for the 'random' process are similar to that of a chaotic process. A truly random process would have very low or zero values for DET, Lmax and Vmax. Firstly we will present the periodic RQA, in figure 4.

Figure 4: RQA for sine wave
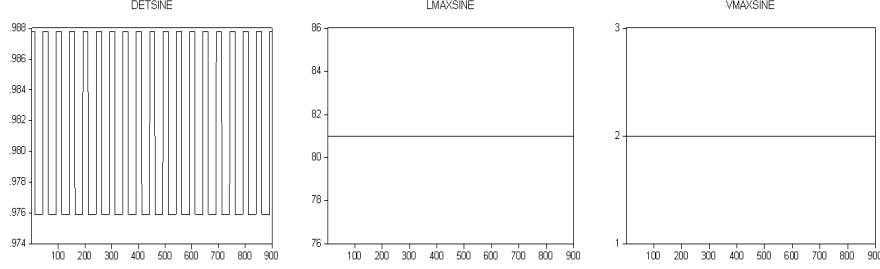
DET is ≈ 1, Lmax is fixed at 81 and Vmax is fixed at 2, indicating the periodic nature of the process. Marwan et al. (2002) notes that periodic states are associated with vanishing Vmax. The RQA for the chaotic process are presented in figure 5.

Figure 5: RQA for logistic map with r=3.999
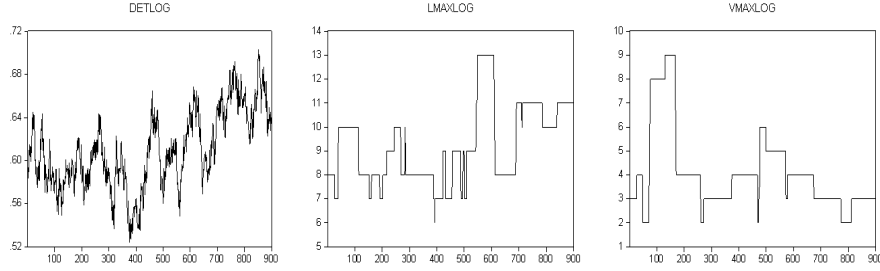
The level of determinism, (DET) is much lower for the chaotic process and the value varies over time. This indicates the chaotic nature of the determinism. Lmax also varies and is much lower than for the periodic sine wave. This indicates the lower predictability of the chaotic process. A low value for Lmax indicates a high value for DIV i.e. 1/Lmax. This correlates with a high positive Lyapunov exponent, suggesting a chaotic regime. Vmax is higher and variable for the chaotic process. In general we would expect Vmax to be low or zero for a periodic process.

Moving window RQA can be used to indicate transitions in the system, such as periodic to chaotic transitions or chaotic to chaotic transitions. Marwan et al. (2002) use moving window RQA to indicate bifurcation points i.e. period doubling. Bifurcation points occur as the system changes, the structurally stable attractors become unstable and new attractors are born, (Sprott 2004). The main RQA characteristics which indicate transitions and bifurcation points are summarized in table 1 below.

We will use table 1 to interpret the empirical findings presented in section

9

Table 1: Using RQA measures to identify phase transitions.

Author	Year	RQA	Characteristic of transition
Trulla et al.	1996	DET	Sharp rise indicates chaotic-periodic transitions
		Lmax	Should collapse prior to bifurcation points
		1/Lmax	Is related to the Lyapunov exponent during chaotic windows
Marwan et al.	2002	Lmax	Maxima at periodic-chaotic/chaotic-periodic transitions
		Vmax	Minima at periodic-chaotic transitions and maxima at chaotic-chaotic transitions, vanishes during periodic windows.
Marwan et al.	2007	Lmax,	Peaks at periodic-chaos transitions. Lmax finds all periodic-chaos transitions. Size of Lmax is related to the predictability of the underlying system.
		DET	Peaks at periodic-chaos transitions, although DET only finds some of the transitions.
		Vmax	Peaks at chaos-chaos transitions, (for example, laminar states/intermittencies, band merging and inner crisis). Transitions can still be identified for small noise levels.
Zbilut & Webber	2008	Lmax	Should decrease precipitously prior to the inception of a laminar state, and be at a minimum just prior to the laminar areas., (unstable singularities, i.e. repellors)
Marwan	2010	DET	DET rising to 1 implies a chaotic-periodic transition.

4. The RQA measures of an ARGARCH simulation will now be presented in figure 6. The epoch window size is 100 observations and the step size is 1. The embedding dimensions, (m, τ, ϵ) are $(5, 1, 0.45)$.

Figure 6: RQA for ARGARCH simulation
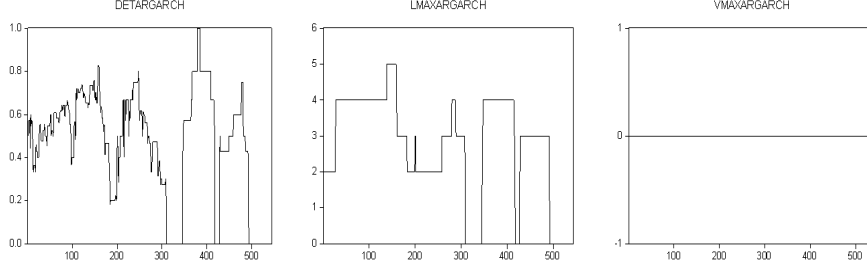
We note that just after observation 300 DET and Lmax collapse to zero, and that Vmax is zero throughout. We present in figure 7 a scaled moving window variance (σ_i) for the ARGARCH process, calculated with a window size of 100 and a step size of 1. The moving window variance is scaled by dividing it by the average variance for the entire series, such that;

$$\sigma_i = \frac{\sigma_{i...i+99}}{\sigma}, \, i = 1 \ldots (n - 100) \tag{15}$$

Looking at the time series graph, (above the RP in figure 3) we can see that this is around the time that the volatility is rising and clustering. It appears as if the rise in the dynamic variance is cancelling out the autoregressive signal and the system appears to be completely random for some time. As σ_i falls again, the signal returns.

Collapses in RQA have been noted around the time of stock and currency market collapse (Basto and Caiado 2001, Marwan et al 2007, Zbilut 2004) but this has not been explained or modeled. It appears from the representation of the ARGARCH simulation in figure 6 and 7, that the collapse in the RQA is connected to the volatility clustering in the data. We will further test this assertion by analysing empirical time series of the Dow Jones Industrial Index.

11

Figure 7: Moving window scaled variance of the ARGARCH model

VARARGARCH
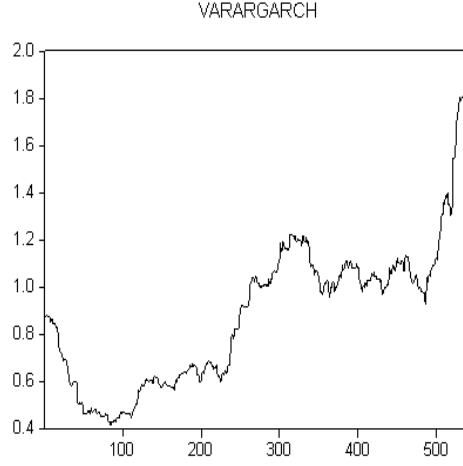
3   Data

The data to be used in this analysis is the Dow Jones Index of 30 top ranking industrial equities trading on the US Stock Exchange. This index is chosen as it is one of the longest data samples available for the equity markets. The objective of the research is to analyse the transition from bubble to collapse, therefore we focus on the time just before and just after the market has reached a localized peak. The four peaks and subsequent crashes chosen are 1929, 1973, 2000 and 2007. These samples are chosen due to the economic significance of each crash in terms of rising unemployment and hardship and also due to the large overall fall in the Dow Jones Index during each crash. The crash of 1973 and 2000 are generally agreed to be attributable to the oil and technology sectors respectively, the crashes of 1929 and 2007/2008, have been attributed to banking and financial market collapse. The details of each crash are outlined in table 2 below;

Table 2: Dow Jones Industrial Average peaks and troughs

Date of Peak	Date of Trough	Peak value	Trough value	Decline %
03/09/1929	8/7/1932	381.17	41.22	-89%
08/01/1973	6/12/1974	1047.86	577.6	-45%
14/01/2000	9/10/2002	11722.98	7286.21	-38%
10/09/2007	9/3/2009	14164.53	6547.05	-54%

According to Johansen and Sornette (2000) there are distinguishable periodic oscillations in the log of the time–to-crash. They present a Log Periodic Power

Law model (LPPL) and suggest this can be used to predict a market crash. The LPPL model has been tested and questioned by Brée et al. (2011) who conclude that the model is not reliable. We also wish to review the LPPL model. To do so we will simply test to see if there is evidence of a periodic regime prior to a market peak. Johansen and Sornette (2000) note the oscillations can be exhibited for up to 100 days prior to the market peak. Thus a sample size of 12 months (250 trading days) prior to and 12 months (250 trading days) after the market peak is taken in order to analyze the lead up to the peak and the turbulence or crash following on from the peak. The focus is on the dynamics of the transition from bull into bear market. If for example, we note a rise in DET up to 1, this will indicate the existence of a periodic phase prior to the peak. The sample size in each case is 502 observations. One observation is lost as the data is converted to log differentials, giving us 501 observations with the market peak in the middle at observation 251. By placing the peak in the middle of the sample, the changing dynamics prior to and after the event can be analyzed (Marwan 2010). We will focus on the bubble prior to the peak with the objective of assessing the possibility of early warning signals of the oncoming crash.

This sample neither take into account the entire period of the market crash which can take as long as 3 ¼ years to occur (as in the case of 1929); nor the entire period of market recovery after the crash. Specifically we are focusing on the lead up to the market peak. The details of the four samples taken are outlined in table 3.

Table 3: Dow Jones Industrial Average peak sample details

Sample Name	Start Date	Finish Date	Start value	Finish value
1929	31/08/1928	4/09/1930	240.41	236.04
1973	7/01/1972	07/01/1974	910.37	876.85
2000	29/01/1999	04/01/2001	9358.8	10912.4
2007	09/10/2006	07/10/2008	11857.8	9955.5

In experimental sciences it is customary to compare the relevant sample to a control sample. In Marwan et al. (2002) samples of data are taken just prior to the onset of a ventricular tachyarrhythmia (VT) and at a control time i.e. when the patient is at rest, thus without a life threatening arrhythmia. In this paper, we are interested in developing indicators of an upcoming transition from bubble to collapse. As economic data is collected not from a controlled laboratory environment but from the actual market data, a reliable control is difficult to obtain. We propose taking a sample just prior to the peak sample under question. This control sample will be similar to the peak sample in terms of the number of observations and will occur in the same general epoch, but will differ as a significant localized peak of historical economic importance does not occur at the middle point of the control sample. This methodology has some success although we note that market volatilities can occur in the control period. As there is a 20% collapse in the market from April to May 1970, we remove this period from the 1971 sample to maintain its control characteristics. Thus

1971 sample is 90 observations less than the others. We use the full sample for all other samples including 1998. This sample may not be fully appropriate as a control as 1998 saw significant turbulence in the markets due to the Russian currency crisis and the collapse of LTCM (Long Term Capital Management). The details of the control samples are outlined in table 4.

Table 4: Dow Jones Industrial Average control sample details

Sample Name	Start Date	Finish Date	Start value	Finish value
1927	03/09/1926	30/08/1928	163.75	238.85
1971	27/05/1970	6/01/1972	663.2	908.49
1998	12/02/1997	28/01/1999	6961.63	9281.32
2005	12/10/2004	06/10/2006	10002.32	11866.69

The prices are transformed into log differentials. The log differentials are a customary transformation in financial econometrics, used to remove the nonstationary trend in the data (Patterson 2000). We note that the transformation of the prices to log differentials may have its drawbacks, as Sprott (2004) argues; the nonstationarity may be the interesting feature of the data. The transformation is applied in order to ensure our results are comparable to existing financial econometric results, and to ensure that any evidence of determinism is not merely as a result of a linear trend in the data. We present in table 5, a summary of the chosen parameters for financial data to be found in recent literature. Note that in the table ws is window size, ss is step size, ld is logged differentials, td is transformed differences, n is normalized i.e. $\frac{x_i - x_{min}}{x_{max} - x_{min}}$ and ns is not specified. We note the significant heterogeneity in the parameters chosen, this can lead to difficulties in comparison of results.

As discussed in section 2, the false nearest neighbours method is used to estimate the embedding dimension, m, and it is found to be equal to 5. The time delay τ, is kept equal to 1, as we are dealing with discrete data. The threshold is found (using the methodology outlined above in section 2.3.1) to be 0.73. As suggested by Zbilut (2004) when he analyzed the S&P 500, in all cases we take a window size of 90 observations and a step size of 1, giving 411 estimates of each RQA. We note that Sornette and Johansen (2000) suggest that on average it takes 30 days for the log periodic model to lead to the critical time that is market collapse. Thus 90 days should give enough time to allow the dynamics to unfold, without being so large as to average out any small scale dynamical changes.

The ARGARCH simulation applied in section 2, is generated from random normally distributed numbers that have been transformed using the saved ARGARCH variance series and the estimated coefficients of a sample of the S&P 500 from 8th May 2006 until 26th November, 2008, (x_o is set equal to the value of the log differential on 8th May 2006). This simulation is also used in Moloney and Raghavendra (2011) to compare phase portraits of Gaussian and non-Gaussian series.

For illustrative purposes we will also analyse a series of data from 9th October 2006 until 5th October 2010, named the 2008 sample. This data set extends the 2007 data set by a further 502 observations.

Table 5: Embedding dimensions for financial data

Author	Year	Data	m	tau	ϵ/σ	ws/ss	ld/td/n	No. of obs
Fabretti and Ausloo	2005	Nasdaq and Dax	5	10	ns	100/10	ld	1482
Marwan et al.	2007	Exchange rates	20	1	0.048-4.350/σ	-	ld	168
Strozzi et al.	2007	High frequency exchange rates	11	260	0.043	336/48	td	268554
Crowley	2008	EU and US GDP	4	1	0.009/σ	ns	ld	147
Guhatha-kurta et al.	2010	Dow, Nifty Hong Kong Index	3	10	0.1/σ	200-100/100	n	766-921
Karagia-nni and Kyrtsou	2011	CPI and Dow Jones	1	1	10	100/10	ld	561
Basto & Caiado	2011	46 MSCI Indices	11	1	0.16/σ	260/ns	n	3914
Aparicio et al.	2011	Simulations and equity data	1-6	1	ns	-	ld	2211

4 Empirical Findings

Initially we will present the RPs of the four peak samples in figure 8 and the four control samples in figure 9, to compare and contrast the dynamic structures.

Figure 8: RPs of the Dow Jones Index peak samples (a) 1929 (b) 1973 (c) 2000 (d) 2007
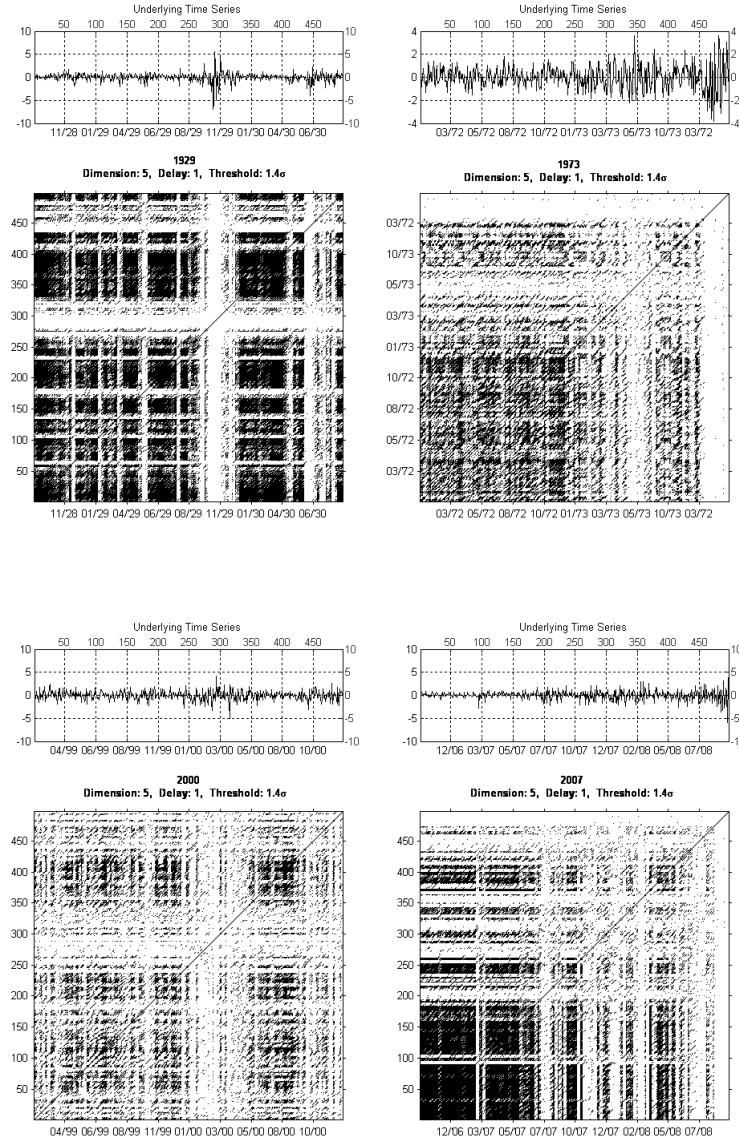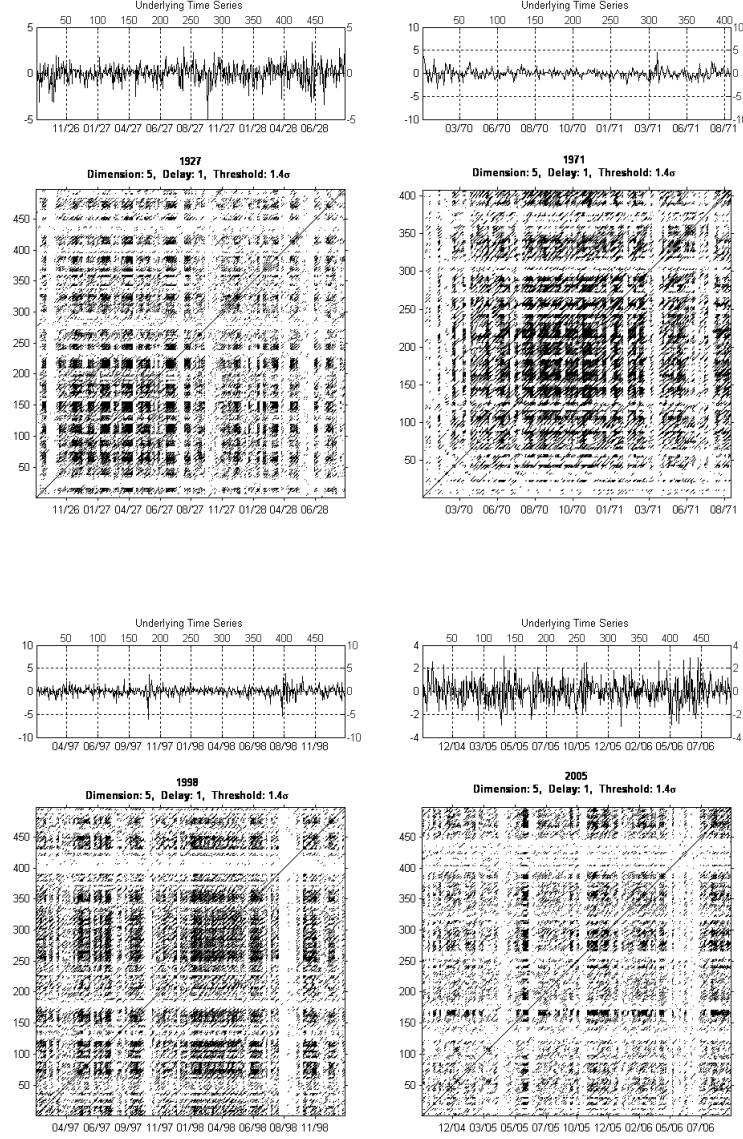
Figure 9: RPs of the Dow Jones Index control samples (a) 1927 (b) 1971 (c) 1998 (d) 2005

We can see evidence of nonstationary or extreme events indicated by the white bands or lines (Marwan et al. 2002). During the peak samples there appears to be more white bands both in number and width, particularly after

the peak (observation 251). In the recurrence plot figure 8 (a) 1929 and figure 8(d) 2007 we see large black squares, followed by thick white lines indicating that the system swings from having many close trajectories i.e. low turbulence, to extreme turbulence. The RPs of the control samples are more uniform and have small squares indicating small scale correlations throughout the period. We note that figure 9 (c) 1998 is somewhat similar to the peak samples as there are large black squares followed by a large white band around August 1998. This may be explained as the summer of 1998 saw the Russian currency crisis and the LTCM crisis, which was noted earlier. There is a fading of the recurrence plot figure 8 (d) 2007 as the eye moves up the LOI. Similarly to the ARGARCH simulation, this indicates an underlying trend in the system. Analysis of the time series (shown above the recurrence plot) shows a rising variance. With this in mind, the sample is extended as outlined in the data section and we illustrate the recurrence plot for sample 2008 (figure 10). The overall pattern in this recurrence plot is very similar to that of figure 8 (a) 1929. A large white band can be seen around September 2008, which was when Lehman Brothers collapsed. This initial analysis suggests similarities in the dynamics of the market during the peak samples, with particular similarities in the RPs of 1929 and 2008.

Figure 10: RP of the Dow Jones Index 2008 sample

18
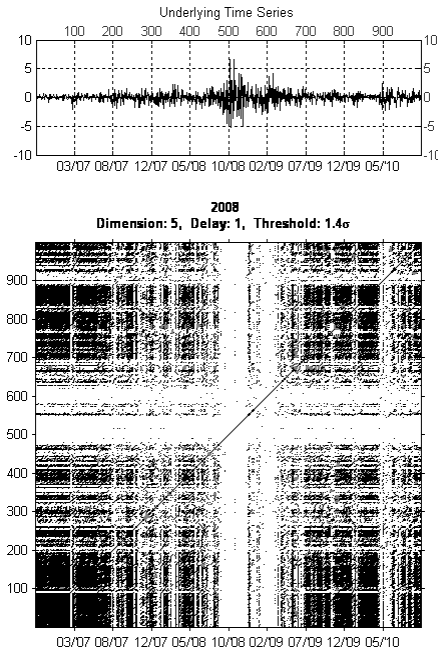
In both cases the time series and the RPs indicate that a rising variance in the logged differentials is coupled with extreme events or nonstationarities i.e. white bands in the RPs.

4.1   Phase transitions

The moving window RQA measures for all eight samples are calculated. The objective of the paper is to examine the transition in the samples from one phase of the market to another, i.e. from bull to bear market. The terminology, phase transition, is borrowed from physics were it is used to define the transition of a substance from one form to another, e.g. from liquid to a gas. In nonlinear dynamics, phase transitions occur as structure of the trajectory changes due to changes in the parameter values of the underlying system. The trajectory can change from periodic to chaotic, or from chaotic to chaotic., for example. Bifurcations points occur at a period doubling. RQA measures have been shown to indicate phase transitions and bifurcation points. Table 1 above indicates that Lmax has been shown to peak during a periodic to chaotic transition and to collapse prior to bifurcation points. Vmax has been shown to peak during a chaotic to chaotic transition. Evidence of phase transitions around the time of a market peak and collapse could be a useful indicator of rising risk and a changing market. In figure 11 and 12 below, we present graphs of the windowed Lmax and Vmax for all eight samples of the Dow Jones Index. Looking at the scales of the graphs it appears as if Lmax peaks at values above 20 in 1929 and 2007. We also note a smaller peak in Lmax (above 15) in January 1998.

Following the methodology of Marwan et al. (2002) (when examining heart-rate variability), we use the Mann-Whitney U test to compare the medians of the peak samples with the control samples. The null hypothesis of the test is that the two samples have the same distribution. As part of the test the two samples are ranked and the medians are calculated and compared.

Marwan et al. (2002) interpret a rejection of the null as evidence of a phase transition in the data and concluded that RQA measures can be used to indicate an oncoming life-threatening arrhythmia. The results are illustrated in table 6 and 7 below. In all cases the null hypothesis is rejected. If we compare the estimates of the median and the mean for Lmax (table 6) we note that only for peak 1929 and 2007 is Lmax(peak) greater than Lmax(control) i.e. Lp>Lc. In the other two samples, although the null is rejected, the median Lmax is the same, and the means are similar. Thus we can only be confident of suggesting evidence of periodic to chaotic transitions in samples 1929 and 2007. Looking at the graphs themselves, we see Lmax collapses around the middle of the peak samples i.e. around observation 251, the market peak. A collapse on Lmax suggests a rise in the Lyapunov exponent and this can indicate bifurcation points around the market peak (see table 1).

19

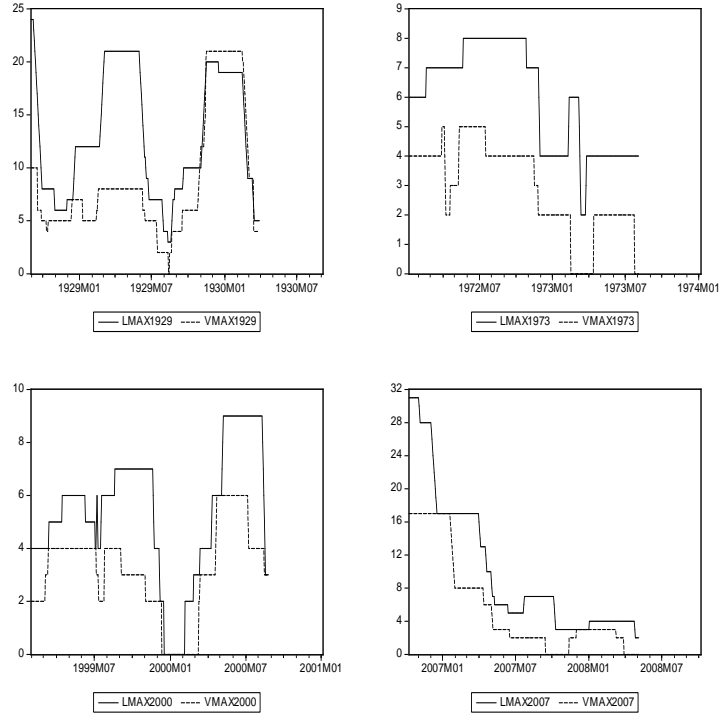Figure 11: Lmax & Vmax (a) 1929 (b) 1973 (c) 2000 (d) 2007

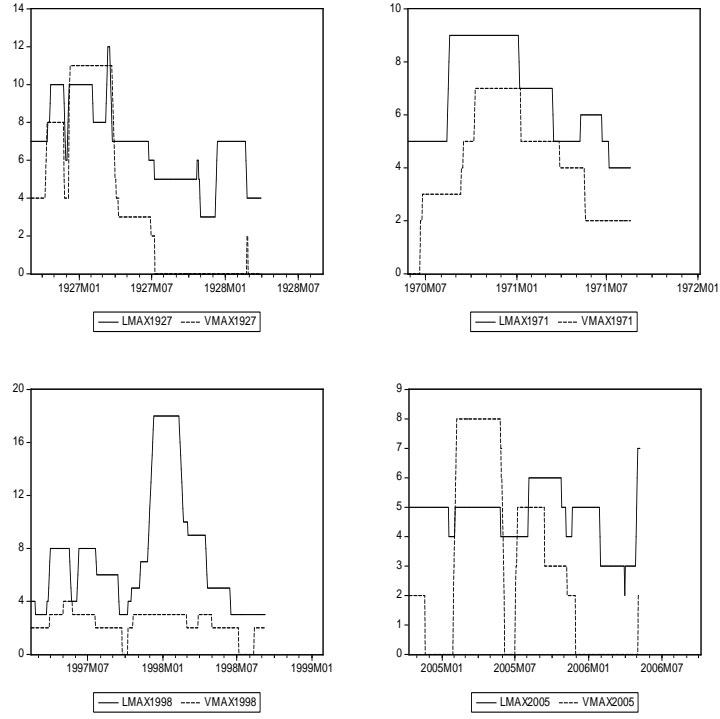Figure 12: Lmax & Vmax (a) 1927 (b) 1971 (c) 1998 (d) 2005
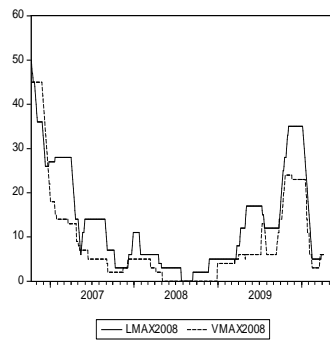

Figure 13: Lmax & Vmax (a) 2008

Table 6: Mann-Whitney U test for Lmax

Sample	median	mean rank	Mean	(s.e.)	Mann U	p value	
1927	7	279.3333	6.7664	0.1033			
1929	12	543.6667	13.0827	0.2932	15.9591	0.0000	Lp > Lc
1971	6	409.4783	6.57683	0.1026			
1973	6	333.7202	5.9440	0.0874	4.8073	0.0000	Lp$\lesssim$Lc
1998	6	443.3224	7.2871	0.2247			
2000	6	379.6776	5.3577	0.1253	3.8424	0.0001	Lp$\lesssim$Lc
2005	5	345.8796	4.6813	0.0467			
2007	7	477.1204	10.1800	0.4127	7.9236	0.0000	Lp > Lc

Similarly to the Lmax test, the null hypothesis is rejected in all cases and the median and the mean value for Vmax(peak) is great than the median and mean value for Vmax(control) i.e. Vp>Vc for the samples 1929 and 2007. Therefore it appears as if there is evidence of chaotic to chaotic transitions in these peak samples, as well as evidence of periodic to chaotic transitions.

Table 7: Mann-Whitney U test for Vmax

Sample	median	mean rank	Mean	(s.e.)	Mann U	p value	
1927	3	299.3345	3.6521	0.2092			
1929	7	523.6655	8.8127	0.2936	13.5440	0	Vp>Vc
1971	4	429.7019	4.0155	0.1118			
1973	3	317.8759	2.9273	0.0744	7.0961	0	Vp$\lesssim$Vc
1998	2	337.4647	2.3406	0.0453			
2000	3	485.5353	3.1192	0.0868	8.9397	0	Vp$\gtrsim$Vc
2005	2	352.6156	2.8029	0.1515			
2007	3	470.3844	5.7202	0.2919	7.1102	0	Vp>Vc

This suggests that there is significant and measurable change in the deterministic structure of the trajectory of the market around the time of these economically and historically important market peaks. Without evoking a model, we suggest that these indicators could be used as warning signals of transition in the market. For illustrative purposes we include the 2008 sample in figure 14. We note that for all of the peak samples (including 2008) there is a collapse in Vmax as well as Lmax at or close to the observation 251 i.e. the market peak. This pattern is repeated for DET below. We can see a clear repeated pattern in all the peak samples of a collapse in the RQA measures around the market peak. This collapse may be another indicator of the transition from bull to bear market. The following section further analyzes this pattern and suggests an interpretation.

4.2 Collapse in RQA

To summarize the collapse in the RQA measures, we examine the windowed DET and compare it to the scaled variance, VAR (see equation 15). We note that the scaled variance VAR rises around the same time as DET collapses. We also note that VAR falls as DET recovers. This can be illustrated in all four peak samples (figure 14).

Figure 14: DET & VAR (a) 1929 (b) 1973 (c) 2000 (d) 2007
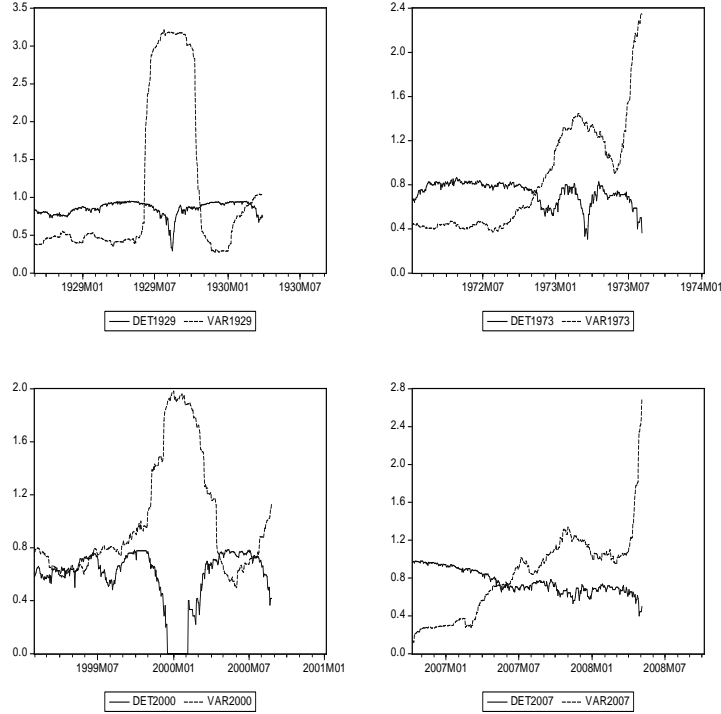

In the peak samples (including 2008) DET varies more widely than in the control samples; at times DET rises to close to 1 and collapses to zero. the change in VAR is particularly notable in 1929 and 2008. In 2008 VAR rises as high as 9.

Figure 15: DET & VAR (a) 1927 (b) 1971 (c) 1998 (d) 2005
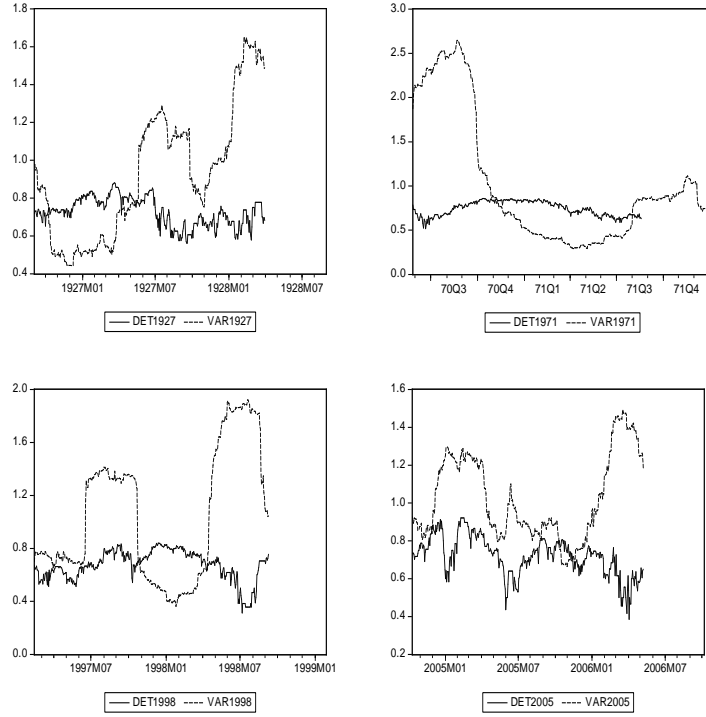
   In the control samples DET varies between 0.5 and 0.9 and VAR stays below 2 (except for a brief period in Q31970). The negative relationship between DET and VAR is not as striking in the control samples.

Figure 16: DET & VAR 2008
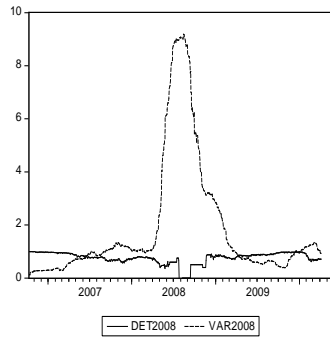
   We note that around the time of the collapse in DET, the scaled variance

24

VAR peaks. Why is this happening? It appears as if an additional force is entering the system, causing the relationship between signal and noise to alter. We examine this proposition in more detail in the next section.

4.3 Random Market Indicator (RMI)

Let us visualize the system as follows. Let us assume that the signal is deterministic with a map f which is not known to us. All that we do have knowledge of are noisy measurements (s_n) of this signal (x_n). Following from the work of Kantz and Schreiber (2003) we present the system using time delay embedding, such that;

$$s_n = f(x_{n-m}, \dots x_{n-1}) + \epsilon_t \tag{16}$$

$$\epsilon_t \sim g\left(\overline{x}, \sigma\right) \tag{17}$$

We assume that the noise variable (ϵ_t) is random and has no correlation with the signal and also that the system is stationary. This is a stochastic model (it could be linear or nonlinear). It is clear from the analysis of VAR that during times of transition from bull to bear market, variance rises considerably. If we take VAR as a proxy of the variance of the noise ϵ_t, we can see that during peak times variance is correlated with f (measured by DET) and cannot be said to be stationary (figure 14 and 16). We suggest that during these times the system is no longer conservative but is experiencing an exogenous force which is causing ϵ_t to rise. So much so, that the deterministic signal f can no longer be measured. All the RQA measures collapse to zero and the system appears completely random.

Classical Newtonian dynamics assume the existence of derivatives at all times i.e. Lipschitz conditions This allows for the use of differential equations as a mathematical framework. Zbilut (2004) and Zbilut and Webber (2008) suggest that under 'non-Lipschitz dynamics' if a system experiences an exogenous force, this will cause unstable singularities and a 'stochastic' repellor to be born. The trajectory of the system becomes probabilistic and the dynamics of the system are irreversible. Zbilut asserts that non-Lipschitz dynamics allow for a 'dynamic pause' as the system adapts to its environment. In physiological or social systems, non-Lipschitz dynamics would allow the system to 'learn'. By definition, chaotic systems require that the dynamical structure of the system is dependent on initial conditions (Sprott 2004). Zbilut (2004) argues that for physiological and social systems this is unlikely. For example, it is unlikely that the behaviour of Coca Cola's share price today is determined by the IPO (Initial Public Offering) price back in 1919. Although it is also true that social systems do not appear to be completely random. They appear to have deterministic structure. It is argued by Trulla et al. (1996), Zbilut (2004), and Zbilut and Webber (2008) that these systems should be modeled so that they can adapt or learn. Zbilut's approach allows a dynamical pause in the system which allows the system to forget and adapt. Zbilut (2004) describes this as "piecewise determinism". We

assert that the collapse in the RQA measures indicated in the above graphs illustrates a transition in the stock markets from a deterministic to a random system. After the market peak, the RQA measures recover as the variance falls, and the deterministic relationships within the system reappear. This collapse in the RQA measures when VAR increases has also been illustrated by the analysis of the ARGARCH model (figure 6 and 7).

With this is in mind, we use principal component analysis (PCA) to linearize the relationship between DET and VAR and to create a principal component series. We suggest that this series could be useful in indicating a market peak and collapse. PCA allows us to compose the principal components of a set of variables by computing the eigenvalue decomposition of the observed variance matrix. We will use the first principal component, which is the unit length linear combination of the original variables with maximum variance.

In table 8, we present the eigenvalues for the eight samples, the 2008 sample, and the ARGARCH sample. The proportion is the eigenvalue divided by the number of data series (in this case two). It can be interpreted as the percentage of variance between the variables i.e. DET and VAR, which is accounted for by the first principal component.

Table 8: Principal Component Analysis of DET & VAR

Sample	Eigenvalue	Proportion	Correlation
1927	1.4326	0.7163	-0.4326
1929	1.4409	0.7205	-0.4409
1971	1.1859	0.5929	-0.1859
1973	1.6729	0.8364	-0.6729
1998	1.5213	0.7607	-0.5213
2000	1.8356	0.9178	-0.8356
2005	1.1990	0.5995	-0.1990
2007	1.9257	0.9629	-0.9257
2008	1.7699	0.8849	-0.7699
ARGARCH	1.5535	0.7767	-0.5535

In general the proportion is higher in the peak samples, indicating the appropriateness of the principal component series as a linear summation of the relationship between DET and VAR. This point is also highlighted by comparing the correlations, which in all cases are negative, but which are higher in the peak samples than the control. The correlations are particularly strong in the later samples of 2000 2007, and 2008. This suggests that the linear relationship between DET and VAR is strengthening through the twentieth century and into the twenty first century. We note that the correlation for the ARGARCH model is not as strong as for the actual market samples. This could be due to the fact that the ARGARCH model is by design stationary and thus will not allow for the extreme changes we see in the actual market. This point is highlighted in figure 17 which presents the squared return series for the S&P 500 (as a proxy

for realized volatility) with the forecast variance from the ARGARCH Gaussian model. The ARGARCH Gaussian model reflects the volatility of the actual data but consistently underestimates it. Therefore a principal component series based on the ARGARCH model would underestimate the variance of the market. Figures 18, 19, and 20 illustrate graphs of the principal component series, which we have named the random market indicator (RMI) for the eight samples and for 2008. Comparing the samples, we can see that RMI is higher during the peak samples and peaks around the time of the market peak. This is particularly noticeable for figure 18 (a) 1929, figure 18 (c) 2000 and figure 20 2008. Comparing the absolute value of the RMI, it appears as if a value well above 3 indicates a transition in the market; variance is rising high and DET is collapsing to close to or to zero. During these times the signal f can no longer be read and the variance VAR is high. To use a stationary stochastic model, such as that in equation 16 and 17, would lead to poor estimates of risk. The system appears nonstationary and the covariance between f and ϵ_t is nonzero. We suggest that stationary risk models should be switched off as the market is behaving in a nonstationary and unpredictable manner. Once the RMI falls back to acceptable levels the models can be used again.

One central question remains: what is this (exogenous) force? Much work has been done in the economic literature on the relationship between sentiment and noise (Brown and Cliff 2004, Malcolm and Wurgler 2006, Mendel and Shleifer, forthcoming). The 'noisy trader theory' suggests that 'noise traders' are rational but uninformed and as a result chase noise. By chasing noise, these traders cause the market to fall into and remain in disequilibrium for a period of time. If we consider this exogenous force as sentiment or ignorance or fear, we can explain the collapse in the RQA measures as a result of noise traders. When sentiment recovers, trading becomes more predictable and the market calms down; variance falls, and the RQA measures recover.
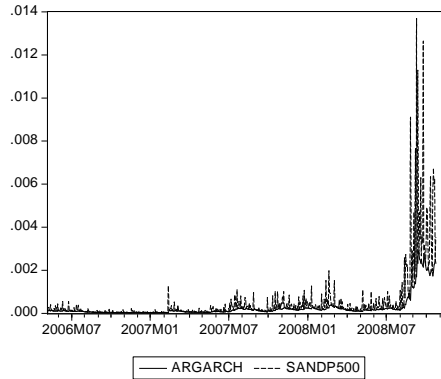
Figure 17: Comparison of ARGARCH model with S&P 500

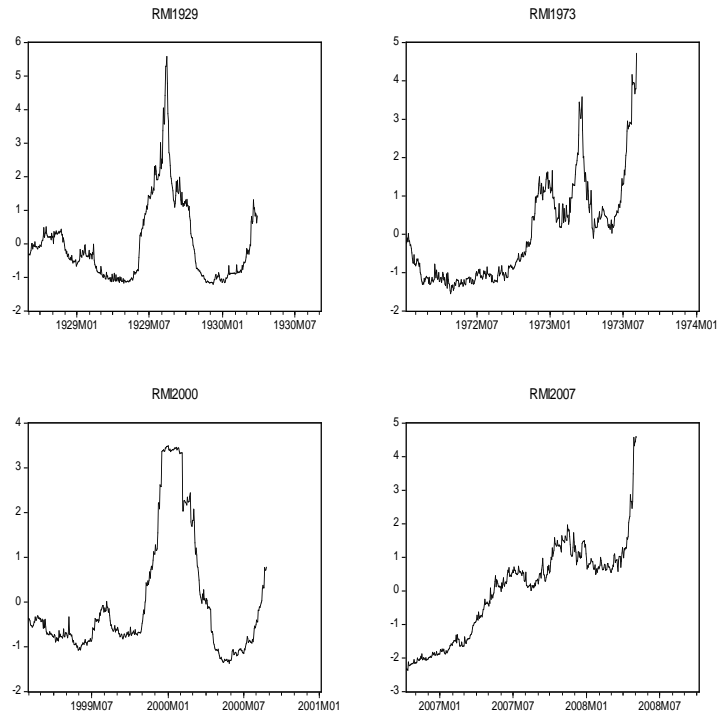Figure 18: RMI (a) 1929 (b) 1973 (c) 2000 (d) 2007
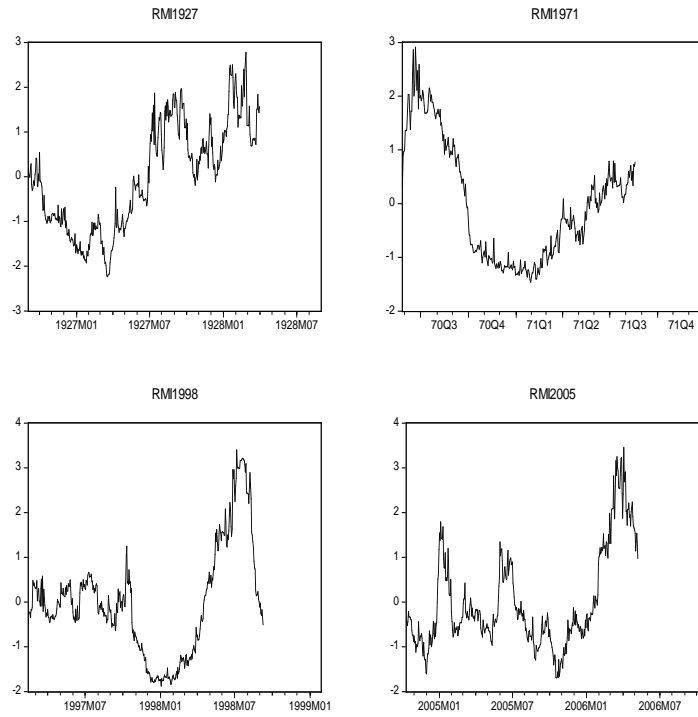
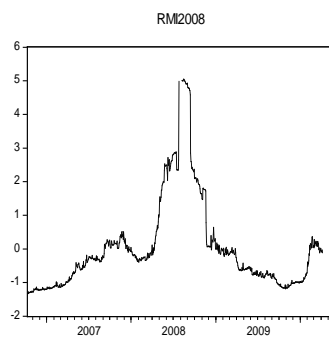Figure 19: RMI (a) 1927 (b) 1971 (c) 1998 (d) 2005

Figure 20: RMI 2008

29

5 Conclusions

The comparison of recurrence plots and recurrence quantification analysis measures for peak and control samples of the Dow Jones Industrial Index have illustrated some similarities across the time periods, particularly the 1929 and 2007 samples. We find some evidence of periodic to chaotic transitions and chaotic to chaotic transitions prior to the market peak in the 1929 and 2007 time series. Secondly we find that the RQA measures collapse just prior to or around the time of the market peak. This would indicate that the trajectory of the market loses its deterministic structure at this time. As the markets do not appear to be periodic, our results do not support Johansen and Sornette's LPPL model (2000).

There is also evidence of a collapsing Lmax and thus a rising Lyapunov exponent prior to each of the peaks, which indicates a chaos to chaos phase transition. This suggests that the dynamics of the system are changing prior to the market collapse and that the system becomes structurally unstable. In the middle of the peak samples, there is evidence of collapsing RQA measures, which implies that the market is losing all its deterministic structure and is behaving in a completely random manner.

Following from Zbilut (2004) and Zbilut and Webber (2008) the system can be thought of as piecewise deterministic. We argue that as noise increases (due to an exogenous force) the determinism in the model is lost and the system transitions from chaotic to nonstationary random. We suggest that the cause of the exogenous force is due to rational but uninformed noise traders. The relationship between the RQA measures and the variance can be summarized through a principal component series named the random market indicator (RMI). If RMI increases above 3 this is an indication of a transition in the data to a nonstationary random process. Quantitative risk estimation techniques should be turned off during these times until the market has transitioned back to a more predictable deterministic process.

One criticism of recurrence quantification analysis is that it is presented without a model. In fact applying recurrence quantification analysis allows us to understand the dynamics of the system without evoking a model. This paper suggests that the system is in fact piecewise deterministic: applying any stationary, fixed parameter model onto the entire data series will lead to misleading forecasts. Recurrence quantification analysis allows us to reconsider the methodological steps taken during the modeling procedure as well as allowing us to question the assumptions made as part of the mathematical framework imposed.

This methodology has been used by many disciplines to better understand their dynamics: including biology, physiology, psychology, and meteorology. For example, Trulla et al. (1996) suggest that:

biological processes are high-dimensional entities, living on transients amidst a field of relatively weak attractors

(Trulla et al 1996)

This paper suggests that social systems such as the stock market are similar such entities. Allowing for the market to be piecewise deterministic allows it to adapt and learn from its environment. Over-modeling the data can be misleading and dangerous, as it may ignore phase transitions, intermittencies, transients and dynamical pauses. Quantitative risk estimation techniques need to be accurate, and when they are not accurate, they should not be used. As a general principal, 'known' unknowns should be distinguished from 'unknown' unknowns, that is we should highlight when we are confident in the application of our models and when we are not. This

suggests that at times of high uncertainty and variance, the predictability of the markets collapses. Quantitative risk modeling should not be used at these times. To generalize these results, further work needs to be done with other time series, other embedding values and other frequencies. Also the reliability of the RMI should be back-tested by matching failures of stochastic quantitative risk estimation techniques with different values for the RMI.

References

[1] Aparicio, T., Pozo, E. F. and Saura, D. (2011) 'Detecting Determinism Using Recurrence Quantification Analysis: A Solution to the Problem of Embedding', Studies in Nonlinear Dynamics and Econometrics, 15(1), Algorithm 1.

[1] Bastos, J. A. and Caiado, J. (2011) 'Recurrence quantification analysis of global stock markets', Physica A: Statistical Mechanics and its Applications, 390(7), 1315-1325.

[2] Bollerslev, T. (1986) 'Generalized autoregressive conditional heteroskedasticity', Journal of Econometrics, 31(3), 307-327.

[3] Brown, G. W. and Cliff, M. T. (2004) 'Investor sentiment and the near-term stock market', Journal of Empirical Finance, 11(1), 1-27.

[4] Brée, D. S., Challet, D. and Peirano, P. P. (2011) 'Prediction accuracy and sloppiness of log-periodic functions', Quantitative Finance, 1-6.

[5] Crowley, P., M. (2008) 'Analyzing convergence and synchronicity of business and growth cycles in the euro area using cross recurrence plots', The European Physics Journal Special Topics, 164, 67-84.

[6] Daníelsson, J. (2008) 'Blame the models', Journal of Financial Stability, 4(4), 321-328.

[7] Eckmann, J. P., Kamphorst, S. O. and Ruelle, D. (1987) 'Recurrence Plots of Dynamical Systems', Europhysics Letters, 4(9), 973-977.

[8] Fabretti, A. and Ausloo, M. (2005) 'Recurrence Plot and Recurrence Quantification Analysis techniques for detecting a critical regime. Examples from

Financial Market Indices.', Computational Physics and Physical Computation, 16(5), 671-706.

[9] Fraser, A. M. and Swinney, H. L. (1986) 'Independent coordinates for strange attractors from mutual information', Physical Review A, 33(2), 1134.

[10] Guhathakurta, K., Bhattacharya, B. and Chowdhury, A. R. (2010) 'Using recurrence plot analysis to distinguish between endogenous and exogenous stock market crashes', Physica A: Statistical Mechanics and its Applications, 389(9), 1874-1882.

[11] Johansen, A. and Sornette, D. (2000) 'The Nasdaq Crash of April 2000: Yet Another Example of Log- Periodicity in a Speculative Bubble Ending in a Crash', European Physical Journal B., 17, 319-328.

[12] Jondeau, E., Rockinger, M. and Poon, S.-H. (2007) Financial Modeling Under Non-Gaussian Distribution, London: Springer-Verlag.

[13] Kantz, H. and Schreiber, T. (2003) Non Linear Time Series Analysis, 2nd ed., Cambridge: Cambridge University Press.

[14] Karagianni, S. and Kyrtsou, C. (2011) 'Analysing the Dynamics between U.S. Inflation and Dow Jones Index Using Non-Linear Methods', Studies in Nonlinear Dynamics & Econometrics, 15(2).

[15] Malcolm, B. and Wurgler, J. (2006) 'Investor sentiment and the cross-section of stock returns', Journal of Finance, 61(4), 1645-1680.

[16] Marwan, N. (2003) Encounters With Neighbours-Current Developments of Concepts Based On Recurrence Plots And Their Applications., unpublished thesis University of Potsdam.

[17] Marwan, N. (2010) 'How to avoid potential pitfalls in recurrence plot based data analysis', eprint arXiv:1007.2215, available: [accessed 5th April, 2011].

[18] Marwan, N., Carmen Romano, M., Thiel, M. and Kurths, J. (2007) 'Recurrence plots for the analysis of complex systems', Physics Reports, 438(5-6), 237-329.

[19] Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A. and Kurths, J. (2002) 'Recurrence-plot-based measures of complexity and their application to heart-rate-variability data', Physical Review E, 66(2), 026702.

[20] Mendel, B. and Shleifer, A. 'Chasing noise', Journal of Financial Economics, (0).

[21] Moloney, K. and Raghavendra, S. (2011) 'Testing for nonlinear dependence in the Credit Default Swap Market', Economics Research International, 2011(Article ID 708704), 12.

[22] Moloney, K. and Raghavendra, S. (forthcoming) 'A linear and nonlinear review of the Arbitrage-Free Parity Theory' in Miller, J., ed. Topics in Numerical Methods for Finance:Withe Special Emphasis on Energy-Commodities-Risk Management and Credit Derivatives-Credit Risk Modeling-Liquidity., London: Springer Verlag.

[23] Parker, T. S. and Chua, L. O. (1989) Practical Numerical Algorithms for Chaotic Systems, London: Springer Verlag.

[24] Sauer, T., Yorke, J., A. and Casdagli, M. (1991) 'Embedology', 65(579-616),

[25] Sprott, J., Clinton. (2004) Chaos and Time Series Analysis, 2nd edition ed., Oxford: Oxford University Press.

[26] Trulla, L. L., Giuliani, A., Zbilut, J. P. and Webber, C. L. (1996) 'Recurrence quantification analysis of the logistic equation with transients', Physics Letters A, 223(4), 255-260.

[27] Webber Jr, C. L., Marwan, N., Facchini, A. and Giuliani, A. (2009) 'Simpler methods do it better: Success of Recurrence Quantification Analysis as a general purpose data analysis tool', Physics Letters A, 373(41), 3753-3756.

[28] Webber Jr, C. L. and Zbilut, J. P. (1994) 'Dynamical assessment of physiological systems and states using recurrence plots strategies', Journal of Applied Physiology, 76(2), 965-973.

[29] Zbilut, J. P. (2004) Unstable singularities and randomness: Their importance in the complexity of physical, biological and social sciences, Amsterdam: Elsevier.

[30] Zbilut, J. P. and Webber Jr, C. L. (1992) 'Embeddings and delays as derived from quantification of recurrence plots', Physics Letters A, 171(3-4), 199-203.

[31] Zbilut, J. P. and Webber Jr., C. L. (2008) 'Laminar recurrences, max-line, unstable singularities and biological dynamics', The European Physical Journal Special Topics, 164, 55-65.