Statistical Modelling of Financial Crashes

John Michael Fry

Thesis submitted for the degree of Doctor of Philosophy

Department of Probability and Statistics,
School of Mathematics and Statistics,
University of Sheffield

September 2007

Abstract

As the stock market came to the attention of increasing numbers of physicists, an idea that has recently emerged is that it might be possible to develop a mathematical theory of stock market crashes. This thesis is primarily concerned with statistical aspects of such a theory.

Chapters 1-5 discuss simple models for bubbles. Chapter 1 is an introduction. Chapter 2 describes a skeleton exploratory analysis, before discussing some economic interpretations of crashes and a rational expectations model of financial crashes – a slightly simplified version of that in *[52]*. This model assumes that economic variables undergo a phase transition prior to a crash, and we give some empirical support of this idea in Chapters 4 and 5.

Chapter 3 discusses SDE models for bubbles. We describe maximum likelihood estimation of the model of *[94]* and refine previous estimation of this model in *[2]*. Further, we extend this model using a heavy-tailed hyperbolic process, *[34]*, to provide a robust statistical test for bubbles. In Chapter 4 we examine a range of volatility and liquidity precursors. We have some evidence that crashes occur on volatile illiquid markets and economic interpretation of our results appears interesting. Chapter 5 synthesises Chapters 2-4.

In Chapter 6 we develop calculations in *[55]*, to derive a generalised Pareto distribution for drawdowns. In addition, we review a method of using power-laws to distinguish between endogenous and exogenous origins of crises *[100]*. Despite some evidence to support the original approach, it appears that a better model is a stochastic volatility model where the log volatility is fractional Gaussian noise.

[6] makes a distinction between *insurance crisis* and *illiquidity crisis* models. In Chapter 7, focusing upon illiquidity crises, we apply the method of *[71]* to evaluate contagion in economics. Chapter 8 summarises the main findings and gives suggestions for further work.

Acknowledgements

I would like to thank my first supervisor/mathematical dad Professor Nick Bingham. Ace. I would also like to acknowledge the support and working relationship I have had with my second supervisor Professor Dave Applebaum. I would also like to acknowledge the contribution of Professor John Biggins who read through an earlier version of this thesis. I haven't always appreciated criticisms made by Dave and John, but I am grateful to them as I think they've each helped me to improve the quality of this work by several notches. THANK YOU!!!

I would also like to say a big thank you to my old lecturer Professor Robin 'dynamo' Johnson at the University of Newcastle-upon-Tyne who was simply inspirational. I would also like to thank my old tutor Professor Richard Boys, also from the University of Newcastle-upon-Tyne, for looking out for me when I was in trouble. I would also like to acknowledge the support of Dr Eleanor Stillman in her previous role as MSc admissions tutor.

I would also like to take this opportunity to say how much I have enjoyed the work of numerous authors, in subjects as diverse as physics and complex systems, stochastic differential equations and economics. In particular, I would like to express my personal appreciation of the following:

[68], [77], [83], [6], [62], [44], [22], [84], [103], [71], [52], [55], [100], [94] [2] [95] [107], [75], [38], [92], [109], [26], [80].

Finally, I would like to thank various pub-going friends and I would also like to gratefully acknowledge the love and support of my parents, especially their support in the final few months.

J.J.

Contents

Chapter 1

Introduction

The aim of this thesis is to build towards a mathematical theory of financial crashes. The first chapter is essentially a summary and a review of relevant aspects of statistical physics models in finance, particularly log-periodic models *[93]*. Chapter 2 will demonstrate that simpler models will suffice, before additional models are developed in Chapters 3-4. Statistical models and tests for bubbles are synthesised in Chapter 5.

Chapters 6 and 7 are intended to provide further additions to such a theory. Chapter 6 discusses power laws for drawdowns *[55]*, and power laws for volatility decay associated with external/internal shocks *[100]*. In Chapter 7 we discuss, based primarily upon *[71]*, Chapter 6, statistical approaches to evaluating contagion in economics.

The particular aim of this chapter is to provide a general introduction to one of the main themes of this thesis – the application of ideas from physics and complex systems in economics and finance. We discuss this in more detail in the next section. Of particular interest are possible applications of statistical mechanics and self-organized criticality. We motivate these considerations further with an Ising-type model of financial crashes. For monograph treatments of statistical mechanics and self-organized criticality, we refer to *[44]* and *[50]*.

Key to the application of physics techniques in finance have been a number of statistical physicists essentially using financial markets as a testing ground for a range of models describing a large number of independent units with nonlinear interactions *[75]*. It is largely considerations along these lines which have motivated the highly contentious subject of log-periodic precursors to financial crashes *[98]*. This was originally motivated by modelling of acoustic emissions – stress waves produced by the sudden internal stress redistribution of materials caused by changes in their internal structure – in relation to destructive testing of kevlar tanks *[4]*. The intuition here is that complex systems may exhibit universal "fingerprints" prior to failure. This is due to an analogy with statistical mechanics, discrete scale invariance, and a hypothesised universal fibre-bundle mechanism. We discuss log-periodicity and log-periodic models in Section 3. In Section 4, we apply these models to real data. Section 5 is a summary.

1.1 Introduction to physics and complexity in finance

According to *[38]*, physics has traditionally been defined as the study of matter and energy and the interaction between these two complementary concepts. However, both the analytical tools and underlying empirical outlook can be applied to non-traditional areas. As a result the influence of physicists and physical ideas has gone beyond traditional boundaries. Consequently there has been a great deal of interest among physicists in economics and this has lead to the emergence of the so-called field of "econophysics". For a monograph treatment see *[22]*.

There are two major reasons for this development. One factor is the thawing of the cold war, leading to decreased funding opportunities for research into traditional areas of physics. However, another important reason is that the areas of finance and physics can be said to share broadly the same aims, namely, matching theory with empirical observation. This is in contrast to the purely empirical nature of some statistical models. *[38]* describes a process of cross-fertilization between economics and physics. One of the most prominent examples of this is Bachelier's famous doctoral thesis of 1900, which used Brownian motion as a model of the Paris stock exchange. It is also interesting to note that Fisher Black – of Black-Scholes fame – was originally a physicist *[75]*. However, there are some key differences between economics and physics. Much of physics is concerned with building a body of theory that is capable of being tested by observation or experimentation. However, this feature does not translate particularly well to economics as a whole. Firstly, *[38]* describes economics as typically being quite data poor, with finance a notable exception. Secondly, human beings are not inanimate and are immensely complicated. Describing this behaviour mathematically is challenging, and various approaches to tackling this problem are also discussed in *[38]*.

This process of cross-fertilisation continued in the 1960s with the publication of a milestone paper by the French mathematician Benoit Mandelbrot *[73]*. Here, for the first time, the assumption of Gaussian distributed price variations was rejected – in favour of the heavier tailed α stable or Lévy stable laws. Today there has been a proliferation of Lévy type models in the mainstream financial literature, see for instance *[31]*, with these models not restricted to the stable processes originally considered. Meanwhile, continued work by Mandelbrot and others brought attention to possible features of scale invariance and power-laws in price dynamics *[72]*. However, the comment is made in *[12]* that such apparent scaling in finance could simply be an artefact of naturally occurring semi-heavy tails.

The late 1980s and early 1990s witnessed an immense growth in the world's financial markets, and also an increased level of automatisation. The result was vast amounts of readily available financial data. This attracted the increasing attention of statistical physicists, interested both in the empirical statistics of financial time series and also in applying novel nonlinear methods. These nonlinear models attempt to model large numbers of units, subject to (usually very simple) interactions. Alongside the application of these nonlinear models to markets, strong analogies became apparent between speculative markets and statistical physics, especially features such as universality, spin systems, self-organized criticality and complexity. For details and references we refer to the discussion in *[75]*.

Unfortunately, it is not possible to give a rigorous mathematical definition of what exactly constitutes a complex system as no such definition is thought to exist. Typically, what is meant by a complex system is a system governed by relatively simple – usually nonlinear – equations. However, these equations can exhibit rich behaviour on temporal and spatial scales, that is not explicitly contained in the system's constituent parts nor their associated equations. Non-trivial collective phenomena can emerge from a series of seemingly trivial interactions between single components.

According to *[75]* it is not difficult to identify indicators for nonlinearity – and hence the use of complex systems – in finance. These include incidence of speculative bubbles, financial crashes, panic, etc. Also apparent are signs of *universal* features, with the existence of a number of remarkably consistent stylized empirical facts about financial markets, which have been observed over a wide range of asset types and across different time scales *Cont and Tankov*, Chapter 7, page 210. With regard to bubbles and crashes, there is an important class of so-called *autocatalytic* processes, whereby small stimuli can be strongly amplified by means of the internal dynamics of the system. Also important here is the concept of self-organized criticality, whereby the nonlinear interactions of a system can place the system in an almost permanent critical state.

1.2 The analogy with Statistical Physics

Recall that we have defined a complex system, somewhat loosely, as a system whose macroscopic behaviour is governed by a multitude of microscopic (usually nonlinear) interactions. Further, recall that (roughly speaking) statistical mechanics is the study of how macroscopic properties emerge from a series of microscopic interactions. There are some classical statistical mechanics systems in (thermal) equilibrium which can be solved exactly. A key example here is the 2-d Ising model in the presence of zero magnetic field *[76]*, *[15]*. However, this is very much the exception. It is more often the case that even when analyzing models in thermal equilibrium an exact solution cannot be found and practical analysis centres around more informal techniques like Monte-Carlo simulation, mean field theory and epsilon (series) expansions *[107]*. Useful introductory treatments of statistical mechanics are given by *[44]* and *[107]*. We explore the extension from these classical statistical models to complicated real world systems, like stock markets, which are *far from equilibrium*. We assume that these systems will share some general features, such as the emergence of power laws as an indication of co-operative phenomena for example. A mathematical exposition of these types of ideas is given by *[48]*.

Two key features make the comparison between stock markets and some statistical mechanics models pertinent. Firstly, the notion from statistical mechanics of macroscopic behaviour emerging from the small-scale interactions. Secondly, phase transitions in statistical mechanics. In thermodynamics, a phase transition occurs when there is a singularity in the free energy or one of its derivatives in some thermodynamic system *[107]*, page 1. Typically, one sees a visibly sharp change in the properties of the system such as a transformation from liquid to gas. However, the notion of phase transitions is not restricted solely to change of state. Yeomans gives an example where the magnetization of beta-brass is affected by the ordering of zinc and copper atoms in a molecular lattice. Moreover, there are (applied probability rather than physics) papers on phase transitions in social networks *[47]* and computer information networks *[74]*. These articles are perhaps best viewed as describing the *collective phenomena* that emerge as the result of numerous and very complicated interactions between people. It is precisely this feature that we seek to model. An introduction to possible physical applications in social sciences is given by *[9]*.

In this thesis, the key idea is to model a stock market crash as occurring when there is a disorder-order transition from a roughly equal mix of buyers and sellers to a situation whereby the number of sellers outweighs the number of buyers. Here there is a striking metaphor with the physical phenomenon of *ferromagnetism*. Ferromagnetic materials can exhibit spontaneous magnetisation in the absence of an applied external field. As the temperature is lowered past the (critical) Curie temperature T_c, magnetic spins which can take the values ± 1 align in the same direction. The net result is magnetisation which occurs spontaneously, and in the absence of an applied field. Informally, if we imagine agents on a stock market taking the values $+1$ meaning sell and -1 meaning buy, we can imagine a crash occurring when all the agent spins spontaneously co-ordinate themselves to align in a $+1$ position. Realising that the price cannot rise any further, agents place sell orders at the same time, and the price plummets.

Here, we discuss the Cont-Bouchaud percolation model *[30]*. Our aim is motivate physical models incorporating large numbers of microscopic interactions. The account here is also based on that in *[85]*, Chapter 5, and *[22]*, Chapter 20. According to Paul and Baschangel purely statistical fluctuations govern a stable market when supply and demand are well tempered. However, the situation is more complicated prior to a crash when an immense number of traders sell spontaneously. According to this interpretation, crashes are triggered by the spontaneous development of long-range correlations between market traders. Thus understanding crashes, and perhaps more general financial fluctuations, necessitates an understanding of traders' behaviour and leads naturally to models incorporating a large number of microscopic interactions.

Let $P(t)$ denote the price at time t. Consider a market of N microscopic traders of identical size. At time, $t+1$, trader i can buy ($\phi_i = 1$), sell ($\phi_i = -1$), or remain inactive ($\phi_i = 0$). In this model, the difference between supply and demand is given by the sum of all the ϕ_i. It is assumed that the price change is proportional to this difference between supply and demand:

$$\Delta P(t+1) = \frac{1}{\lambda}\sum_{i=1}^{N}\phi_i(t+1), \tag{1.1}$$

where $\Delta P(t+1) = P(t+1) - P(t)$ and λ is a measure of how sensitive the market prices are to variations in supply and demand *[85]*, Chapter 6. Suppose all traders in the market can 'communicate' with each other in some way. This communication may make two traders adopt a common market position – i.e. they both decide to buy, sell, or be inactive. If this happens, we say that a 'bond' is created between the traders. The creation of such a bond should happen with a small probability:

$$p_b = \frac{b}{N},$$

since any trader can bind to any other, and the average number of bonds per trader is $(N-1)p_b$. Thus, p_b is determined by a single parameter b, which characterises the rate at which traders comply with each other. Thus the N traders are divided at random into 'clusters' of inter-connected traders. These clusters in effect act as one large trader, linking as they do traders who trade identically. The decision of each cluster is assumed independent of its size, and also the decisions taken by other clusters so that s_c and ϕ_c in (1.2) are assumed to be independent.

Let $\phi_c(t+1)$ denote the state of cluster c at time $t+1$. If there are N_c clusters in total, and s_c denotes the size of each cluster, then (1.1) can be re-written as

$$\Delta P(t+1) = \frac{1}{\lambda}\sum_{c=1}^{N_c} s_c \phi_c(t+1). \tag{1.2}$$

Thus, we see from (1.2) that the distribution of $\Delta P(t+1)$ is determined by the distribution $p_c(S)$ of the group sizes s_c. However, the distribution of the group size is well known in percolation theory as the cluster-size distribution of an infinite-range bond percolation model, a classical problem *[101]*:

$$p_c(s) \sim \frac{1}{s^{1+3/2}} e^{-(1-b^2)s} \qquad (1 << s << N), \tag{1.3}$$

if $b \leq 1$. In this case the price changes are approximately given by

$$\Delta P(t+1) = \frac{1}{\lambda}\sum_{c=1}^{N_c} p_c(s)\phi_c(t+1), \tag{1.4}$$

where the $p_c(s)$ in (1.4) refer to random draws from the distribution in (1.3) and the $\phi_c(t+1)'s$ are an i.i.d. sequence of random variables taking the values $\{-1, 0, 1\}$ with probabilities $\{a, 1-2a, a\}$ independent of the $p_c(s)$. (Here, we have assumed that $1-2a$ is the probability that a given cluster is inactive at time $t+1$, and that each given cluster buys or sells with probability a). $b = 1$ corresponds to a so-called percolation threshold. In this case (1.3) becomes a power law with exponent $3/2$ ($p_c(s) \sim s^{-1-3/2}$), and *[22]* states that in this case the central limit theorem for non-Normal random variables applies so that for large N_c the distribution of the price increment approaches a symmetric Stable distribution with index $\mu = 3/2$. If $b < 1$, (1.3) shows that the variance of the $p_c(s)$ exists. So in this case, the usual central limit theorem applies and the distribution of the price increments approaches a Gaussian. If $b > 1$, increasing numbers of traders join a spanning cluster which dominates this system. This situation corresponds to a crash. Thus, this simple model is intended to provide a plausible and basic mechanism generating heavy-tailed distributions and crashes in financial markets.

According to *[85]* there is some suggestion of self-organized criticality, with some internal market dynamics permanently maintaining b around 1 meaning that the market is preserved at an almost critical state. This is the essence of self-organized criticality *[50]*. *[22]* make some criticisms of this model. Firstly, this model is static and cannot describe how the clustering of agents evolves over time. Secondly, there is a discrepancy between the derived power law with exponent $\mu = 3/2$, and the empirically observed $\mu \sim 3$. There is a vast literature regarding so-called agent-based models. Some useful introductory references are given at the end of Chapter 20 in *[22]*.

1.3 Log-periodic models in finance

1.3.1 Log-periodicity

Log-periodicity occurs in connection with the functional equation $f(\mu s) = f(s)$, where μ is a real constant. Solutions to this equation are known variously as *multiplicatively periodic* or *logarithmically periodic*, the function f being periodic after a logarithmic change of argument. Letting $g(x) = f(e^x)$, we can see that

$$\begin{aligned} g(\log s) &= g(\log(\mu s)) \\ &= g(\log \mu + \log s), \end{aligned}$$

so g is periodic in the usual sense with period $\log(\mu)$. Log-periodicity occurs naturally within the context of fractals and branching processes, see for example *[16]*. Many fractals or approximate fractals in applied sciences are created by a recursive procedure. The passage from the n^{th} stage to the $n+1^{th}$ stage of the approximating pre-fractal corresponds to the transition of the n^{th} generation to the $n+1^{th}$ generation of a branching process. This also corresponds naturally, in a more concrete setting, to models with a hierarchical structure. In the context of finance, such a hierarchical structure might imply cascades of information from a global level to a local level, i.e. from the top downwards.

The search for log-periodicity in real-world systems was motivated by the study of phase transitions *[38]*. The proto-typical example of a phase transition is the simple ferromagnet which consists of magnetic spins, each of which can take the value ± 1. In most solids, these spins are disordered and cancel out in aggregate, so as to produce zero net magnetization. However, in a ferromagnet, the spins can spontaneously align to produce a nonzero aggregate magnetization. There is thus a phase transition to an ordered (magnetised) state from a disordered (not magnetised) state.

The magnetic susceptibility of a magnetic system is defined as

$$\chi = \left.\frac{\partial M}{\partial B}\right|_{B=0},$$

where M is the magnetisation and B is the magnitude of an external magnetic field. The susceptibility of the ferromagnet follows a power law for temperatures $T > T_c$:

$$\chi(T) = (T - T_c)^{\alpha}, \tag{1.5}$$

where $\alpha < 0$, and T_c is the critical (Curie) temperature which was briefly introduced in the previous section *[38]*. This power law can be attributed to the continuous scaling symmetry of the underlying lattice. The spins can be grouped into blocks of spins, each of which corresponds to a composite spin equal to the sum of the spins in the block. Feigenbaum states that, according to renormalization group theory, we can find a model involving interactions between these composite spins that replicates properties of the true model *[45]*. Scaling symmetry and the power law behaviour in (1.5) emerge as a result of the fact that the macroscopic predictions of the block model do not depend on the size of the blocks. (Ignoring finite size effects of spins (atoms), we can rescale the blocks arbitrarily.)

Feigenbaum states that similar considerations carry through in a magnetic system that exhibits a discrete scaling symmetry. *The system remains unchanged only if we rescale by integer powers of some length scale.* Feigenbaum gives the example of a Sierpinski gasket as an object with discrete scaling symmetry. Discrete scaling symmetry is less restrictive than continuous scaling symmetry, and the exponent of the power law is generalized and can now be complex:

$$\chi(T) = Re[(T - T_c)^{\alpha + i\theta}], \tag{1.6}$$

where $\alpha < 0$ and θ is real. The magnetic susceptibility thus satisfies

$$\begin{aligned} \chi(T) &= (T - T_c)^{\alpha}[Re\,(\exp(i\theta \ln(T - T_c)))] \\ &= (T - T_c)^{\alpha}[\cos(\theta \ln(T - T_c))]. \end{aligned} \tag{1.7}$$

According to Feigenbaum, the recognition that discrete scale invariance could lead to log-periodicity led to a search for log-periodicity in the natural world. Rupture events, like earthquakes, were found to be a very fruitful source of log-periodic oscillations *[90]*. Stock market crashes have also been likened to a rupture event. There is an analogy here between the formation of microcracks and a spread of pessimism in the market place. A pessimistic trader going against the trend is like an isolated microcrack in a solid. When these microcracks reach a critical density, they combine, and the result is a catastrophic failure. As a result of the search for log-periodicity in the real world, possible indications of log-periodicity in finance were found independently by *[41]* and *[98]*.

Continuing the discussion along physical lines, there is a suggestion of a seemingly universal structure behind these so-called rupture events. According to Feigenbaum, the hypothesis is that each of these systems exhibiting ruptures have a *fibre-bundle-like structure with discrete scale invariance*. The system contains N fibres, each of which contains N smaller fibres etc. Rupture may then be seen to correspond to a phase transition, from a connected to a disconnected state, and it is anticipated that certain variables may then display log-periodic oscillations by analogy with the susceptibility in (1.7).

We discuss log-periodic models used in finance in subsections 3 and 4. The advantage of these models is that they enable one to make predictions under what are anticipated to be rather general circumstances. This practical aspect should not be understated. However, one of the key themes of this thesis is that these log-periodic models actually prove to be rather ill-equipped as financial models which have to be estimated statistically, and it seems that somewhat simpler methods will suffice.

1.3.2 Simple Power Law model

In the canonical 2-d Ising model, as the temperature T approaches the critical point T_c, the magnetization M goes to zero as

$$M \sim (T_c - T)^{\beta},$$

where β=1/8. *[70]*, page 233, describes this kind of behaviour as being quite general for critical phenomena. Given this generality of power-law type behaviour, a reasonable first order approximation suggests fitting the following model to the log-price $y(t)$:

$$y(t) = A + B(t_c - t)^{\alpha} + \epsilon_t, \tag{1.8}$$

where A and B are constants and ϵ_t is a zero-mean error term. This model simply predicts an explosive power law growth prior to the crash.

1.3.3 Simple log-periodic fracture model

The argument here follows that in *[4]*. Consider a mechanical system under strain. Let dE/dt be the instantaneous energy rate of mechanical waves in the system, as a function of the applied internal pressure p up to the rupture threshold p_r. Let $x = (p - p_r)$, and set $F(x) = (dE(x)/dt)^{-1}$. By the phase transition analogy, we anticipate that the instantaneous energy rate becomes infinite precisely when the system fails. Thus we are left with $F(x) = 0$ precisely when $x = 0$. By the analogy with a statistical mechanical system, we assume that we have a generic scaling relation

$$\mu F(x) = F(\phi(x)), \tag{1.9}$$

for some real constant μ and $\phi(\cdot)$ a differentiable function known as the flow map. By (1.9) and the considerations above, we have that $\phi(0) = 0$. Expanding ϕ up to first order in a Taylor series, we obtain $\phi(x) = \lambda x + O(x^2)$. This approximation is made for the sake of tractability but, since we are concerned with x small corresponding to the region of the rupture threshold, this may be physically reasonable. Under the linear approximation

$\phi(x) = \lambda x$ we obtain the functional equation for log-periodicity:

$$\mu F(x) = F(\lambda x). \tag{1.10}$$

It can easily be verified that a solution to (1.10) is given by the simple power-law $F_0(x) = x^\alpha$, where $\alpha = \log(\mu)/\log(\lambda)$. The general solution to (1.10) is

$$F(x) = F_0(x)p(\log(F_0(x))), \tag{1.11}$$

where $p(x)$ is a periodic function of x, with period $\log(\mu)$ *[4]*, and is also without any zeros since by assumption $F(x) = 0$ precisely when $x = 0$. Since $p(x)$ is periodic, so is $1/p(x)$, and this can be expanded in a Fourier series. Since we are only interested in $x \geq 0$, corresponding to the time up to and including rupture, it is sufficient to assume that both $p(x)$ and $1/p(x)$ are even functions of x, and we can arrive at the series expansion

$$\frac{1}{p(\log(F_0(x)))} = \sum_{n=0}^{\infty} a_n \cos\left(\frac{2n\pi \log(x)}{\log(\lambda)}\right), \tag{1.12}$$

since $\frac{\alpha}{\log(\mu)} = \frac{1}{\log(\lambda)}$. Retaining only the first order term in (1.12), we obtain

$$\frac{1}{p(\log(F_0(x)))} = \left(a_0 + a_1 \cos\left(\frac{2\pi \log(x)}{\log(\lambda)}\right)\right). \tag{1.13}$$

In order to find an expression for $dE/dt(x)$, we invert our expression for $F(x)$ in (1.11) to obtain

$$dE(x)/dt = Bx^{-\alpha}\left(1 + C\cos\left(\frac{2\pi \log(x)}{\log(\lambda)}\right)\right), \tag{1.14}$$

where $B = a_0$ and $C = a_1/a_0$. The form taken by (1.14) is assumed to be very general and, in anticipation of a "universal" log-periodic acceleration, dE/dt is identified with the derivative of the log-price $y(t)$. In practice, as a pragmatic step, a phase constant ϕ is added to compensate for the change of units between x and t_c *[28]*. This leaves us with

$$dy/dt = B(t_c - t)^{-\alpha}\left(1 + C\cos\left(\omega \log(t_c - t) + \phi\right)\right), \tag{1.15}$$

where $\omega = 2\pi/\log(\lambda)$. Now the function

$$f(t) = F_1(t_c - t)^{1-\gamma}[1 + F_2\cos(\omega \ln(t_c - t) + \phi) + F_3\sin(\omega\ln(t_c - t) + \phi)], \tag{1.16}$$

has derivative

$$\begin{aligned} f'(t) &= -(1-\gamma)F_1(t_c - t)^{-\gamma} \\ &\times [1 + F_2\cos(\omega\ln(t_c - t) + \phi) + F_3\sin(\omega\ln(t_c - t) + \phi)] \\ &+ \omega F_1(t_c - t)^{-\gamma}[F_2\sin(\omega\ln(t_c - t) + \phi) - F_3\cos(\omega\ln(t_c - t) + \phi)]. \end{aligned}$$

Equating coefficients, we see that this can be rewritten in the form

$$f'(t) = -F_1'(t_c - t)^{-\gamma}[1 + F_2' \cos(\omega \ln(t_c - t) + \phi) + F_3' \sin(\omega \ln(t_c - t) + \phi)],$$

where

$$\begin{aligned} F_1' &= (1-\gamma)F_1, \\ F_2' &= F_2 + \frac{\omega}{1-\gamma}F_3, \\ F_3' &= F_3 - \frac{\omega}{1-\gamma}F_2. \end{aligned}$$

These relations can be trivially inverted to give

$$\begin{aligned} F_1 &= \frac{F_1'}{1-\gamma}, \\ F_2 &= \frac{1}{1+\left(\frac{\omega}{1-\gamma}\right)^2}\left[F_2' - \frac{\omega}{1-\gamma}F_3'\right], \\ F_3 &= \frac{1}{1+\left(\frac{\omega}{1-\gamma}\right)^2}\left[F_3' + \frac{\omega}{1-\gamma}F_2'\right]. \end{aligned}$$

Thus, by setting $F_1' = -B, F_2' = C, F_3' = 0$ we can see that the solution to (1.15) takes the form of (1.16) with $F_1 = \frac{-B}{1-\alpha}, F_2 = \frac{1}{1+(\frac{\omega}{1-\alpha})^2}C, F_3 = \frac{1}{1+(\frac{\omega}{1-\alpha})^2}\left[\frac{\omega}{1-\alpha}C\right]$. This allows us to write
$y(t) = A + \frac{-B}{1-\alpha}(t_c - t)^{1-\alpha}\left[1 + F_2\left(\cos(\omega \ln(t_c - t) + \phi) + \frac{\omega}{1-\alpha}\sin(\omega \ln(t_c - t) + \phi)\right)\right].$

We perform a trick used in *[28]* which enables some simplification. If we set $\theta = \tan^{-1}\left(\frac{\omega}{1-\alpha}\right)$, the trigonometric terms in the round brackets can then be written as

$$\frac{\cos(\theta)\cos(\omega \ln(t_c - t) + \phi) + \sin(\theta)\sin(\omega \ln(t_c - t) + \phi)}{\cos(\theta)}.$$

Rearranging, using the trigonometric formula $\cos(A - B) = \cos(A)\cos(B) + \sin(A)\sin(B)$, this can be seen to be equal to

$$\frac{\cos\left(\omega \ln(t_c - t) + \phi - \theta\right)}{\cos(\theta)}$$

These considerations lead to the log-price being proportional to the functional form given in (1.16) with $F_3 = 0$. Thus we are led to the following model for the log-price $y(t)$:

$$y(t) = A + B(t_c - t)^{1-\alpha}\left(1 + C\cos(\omega \ln(t_c - t) + \phi)\right), \tag{1.17}$$

where $A, B, C, t_c, \phi, \omega, \alpha$ are constants independent of time t with $B = \frac{-B}{1-\alpha}$ and $C = \frac{F_2}{\cos(\theta)}$.

1.3.4 Sornette-Johansen nonlinear log-periodic model

This section deals mainly with the Sornette-Johansen nonlinear log-periodic formula which is fitted to log-prices. The method takes the view that the form of (1.17) is a reasonable approximation, but seeks a simple refinement to increase the accuracy of the original model. The intention is to build a model which, under generic but largely phenomenological considerations, better accounts for the underlying oscillations of financial and physical reality. The methodology is contained in *[96]*, and builds on an analogy with Landau expansions in second-order phase transitions in physics. This theory assumes analyticity in the vicinity of a critical point (see for example *[66]*). For instance, the thermodynamic potential Φ of a uniform sample of a pure substance at constant temperature is modelled by an *analytic* function of temperature T and the square η of the order of the sample[1], so that

$$\Phi(T,\eta) = \Phi_0(T) + A_1(T)\eta^2 + A_2(T)\eta^4 + ..., \tag{1.18}$$

where $A_1(T) \sim (T - T_c)$ as $T \to T_c, A_2 > 0, a > 0$ and T_c is the transition temperature of the substance. For discussion related to the appropriate physical context we must refer to *[66]*, Chapter XIV. It is the series expansion in (1.18), and the underlying assumption of analyticity in the vicinity of the critical point, that motivates the method considered here. The first model considered the functional equation

$$\mu F(x) = F(\lambda x), \tag{1.19}$$

which generates log-periodicity. A simple solution to this functional equation is the power law $F(x) = x^{\alpha}$, where $\alpha = \frac{\log\mu}{\log\lambda}$. Such a power law relationship can also be generated from the equation

$$\frac{dF(x)}{d(\log(x))} = \alpha F(x). \tag{1.20}$$

In order to generalise (1.20), Sornette and Johnansen consider

$$\frac{dF(x)}{d(\log(x))} = (\alpha + i\omega)F(x)) + (\eta + i\kappa)|F(x)|^2 F(x) + O(F^5), \tag{1.21}$$

which is assumed to constitute a second-order series expansion with (1.20) the corresponding first-order condition. In the sequel we identify x with $t_c - t$. We write $F(x) = B(x)e^{i\psi x}$, and consider the modulus and argument parts separately. We can

[1] Thermodynamic potentials are parameters of a thermodynamic system which have the dimensions of energy, such as the internal energy or the Helmholtz free energy. The order of a sample takes the values 0 and 1 corresponding to disordered and ordered states.

derive the following pair of differential equations

$$\begin{aligned}\frac{\partial B}{\partial(\log(x))} &= \alpha B + \eta B^3, &(1.22)\\ \frac{\partial \psi}{\partial(\log(x))} &= \omega + \kappa B^2. &(1.23)\end{aligned}$$

Sornette and Johansen give the solution to this set of differential equations as

$$B^2 = B_0^2 \frac{(x/x_0)^{2\alpha}}{1+(x/x_0)^{2\alpha}}, \tag{1.24}$$

$$\psi = \omega \log(x/x_0) + B_0^2 \frac{\kappa}{2\alpha} \log(1+(x/x_0)^{2\alpha}). \tag{1.25}$$

Here $B_0 = \alpha/\eta$, and x_0 is an arbitrary constant. If we retain solely the real parts, the solution for $F(x)$ is given by

$$F(x) = B_0 \frac{(x/x_0)^\alpha}{\sqrt{1+(x/x_0)^{2\alpha}}} cos[\omega \log(x/x_0) + B_0^2 \frac{\kappa}{2\alpha} \log(1+(x/x_0)^{2\alpha})]. \tag{1.26}$$

This suggests the following pragmatic generalisation of (1.17) *[96]*:

$$y(t) = A+B\frac{(t_c-t)^\alpha}{\sqrt{1+((t_c-t)/\Delta t)^{2\alpha}}}[1+C\cos[\omega\log(t_c-t)+\frac{\Delta\omega}{2\alpha}\log(1+((t_c-t)/\Delta t)^{2\alpha})+\phi]], \tag{1.27}$$

with the trigonometric formula (1.25) correcting first-order behaviour of the form $y(t) = A + B\frac{(t_c-t)^\alpha}{\sqrt{1+((t_c-t)/\Delta t)^{2\alpha}}}$. However, it appears that there may be a slight mistake in these calculations, which suggests a slightly different form for equation (1.27).

1.3.5 Amendment to nonlinear log-periodic model

We show in this subsection that the form of (1.27), should be slightly amended. We demonstrate the suggested form of modifications by a direct integration approach, but the result can be equivalently verified by performing the reverse differentiation. We start with the pairs of differential equations (1.22), (1.23). Let $Y = B^2$. Then (1.22) becomes

$$\int \frac{dY}{Y+Y^2/B_\infty^2} = \int (2\alpha d(\log(x)), \tag{1.28}$$

where $B_\infty^2 = \alpha/\eta$. The integral on the left-hand side becomes

$$ln\left(\frac{Y}{Y+B_\infty^2}\right) + C.$$

Exponentiating the resulting integrals in (1.28), we obtain

$$\left(\frac{Y}{Y+B_\infty^2}\right) = (x/x_0)^{2\alpha},$$

where x_0 is an arbitrary constant. Rearranging, we finally obtain that

$$B^2 = B_\infty^2 \frac{(x/x_0)^{2\alpha}}{1-(x/x_0)^{2\alpha}}, \tag{1.29}$$

The form of (1.29) also means that we have a slightly different solution to (1.25). Writing things out in full, we have

$$\frac{\partial\psi}{\partial log(x)} = \omega + \kappa B^2. \tag{1.30}$$

Performing the integration directly, we can see that we have

$$\psi = \omega log(x/x_{0,2}) - B_\infty^2 \frac{\kappa}{2\alpha} log(1-(x/x_0)^{2\alpha}), \tag{1.31}$$

for some arbitrary constant $x_{0,2}$. Working in the same vein as Sornette and Johansen, and retaining solely the real parts, the solution for $F(x)$ is given by

$$F(x) = B_0 \frac{(x/x_0)^\alpha}{\sqrt{1-(x/x_0)^{2\alpha}}} cos[\omega \log(x/x_0) - B_\infty^2 \frac{\kappa}{2\alpha} \log(1-(x/x_0)^{2\alpha})]. \tag{1.32}$$

This suggests as a pragmatic generalization of (1.17)

$$y(t) = A + B\frac{(t_c-t)^\alpha}{\sqrt{1-((t_c-t)/\Delta t)^{2\alpha}}}[1 + C\cos[\omega\log(t_c - t) - \frac{\Delta\omega}{2\alpha}\log(1-((t_c-t)/\Delta t)^{2\alpha}) + \phi]]. \tag{1.33}$$

This form is almost identical to (1.27) except for two changes of sign from plus to minus.

1.4 Data analysis

We apply the models discussed here to three major indices; the S&P500 index prior to the crash of 1987, the Nasdaq index prior to the crash in prices observed in April 2000, and the Hong Kong Hang Seng index prior to a crash observed in early 1994. Thus we look at a variety of different crashes, in different markets, and at different times. We look at data from both four years and two years before the crash, and look for any differences between the two. Specifically, for the S&P 500 we look at daily data from both September 1st 1983 and September 3rd 1985 to September 30th 1987. For the Nasdaq we look at data from March 1st 1996 and March 2nd 1998 to March 31st 2000, and for the Hang Seng from December 1st 1989 and December 2nd 1991 to December 31st 1993.

Parameter estimation is not easy. Here no constraints were imposed on parameters, and estimation was performed by the method of least squares. Equations (1.8), (1.17), and (1.33) are "partially linear", with respect to A, B, and BC. In principle, these equations can be estimated using the Golub-Pereyra algorithm, for partially linear non-linear regression models, without the need to state starting values for the partially linear parameters A, B

and *BC* [104], Chapter 9, Page 275. However, in practice we found that the easiest and most robust approach was to use the Nelder-Mead method whilst using the Golub-Pereya algorithm to provide choices of starting values. The Nelder-Mead method is a gradient search based method which can be used to solve complicated numerical optimisation problems without specifying derivatives. In using this method we took care to consider a range of randomly generated starting values for the algorithm as nonlinear problems with multiple local minima may be extremely sensitive to the choice of starting value. For a reference on optimisation see [87].

From simple exploratory plots, it becomes clear that it does indeed appear reasonable to fit the models discussed in Section 3. A fit of the logarithm of the Hang Seng index by a simple power law (1.8) and the linear fracture model (1.17) both appear reasonable (see Figure 1.1). However, there may also be a rather subtle suggestion that the fit of the power law model is such that the log-periodic model may just be simply over-fitting noise.

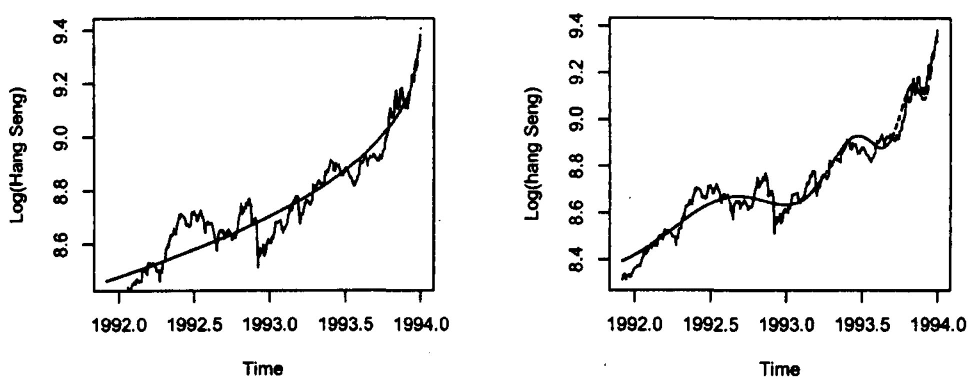


Figure 1.1: Left panel: Plot of Log(Hang Seng) from December 2nd 1991 to December 31st 1993 and a fit by a power law model with estimated parameters $t_c = 1994.00, \alpha = 0.38, A = 9.44, B = -0.74$. Right panel: Fit by the fracture model with estimated parameters $t_c = 1994.13, \alpha = 0.2, \omega = 7.76, \phi = 2.78, A = 10.32, B = -1.59, C = 0.07$.

The crash of 1987, infamously occurred on Black Monday, October 19th 1987 which corresponds to 1987.8. We take the time of the crash on the Hang Seng to be January 1994, and the crash on the Nasdaq to be sometime in April 2000. Plots over the relevant time scales in question are shown in Figure 1.2. Reasonable estimates of the crash-time might be April 10th 2000 or 2000.276 for the Nasdaq crash, which is the seventh trading day shown. Similarly, a reasonable estimate for the Hang Seng crash appears to be the fifth trading day shown which corresponds to January 6th 1994 or 1994.016.

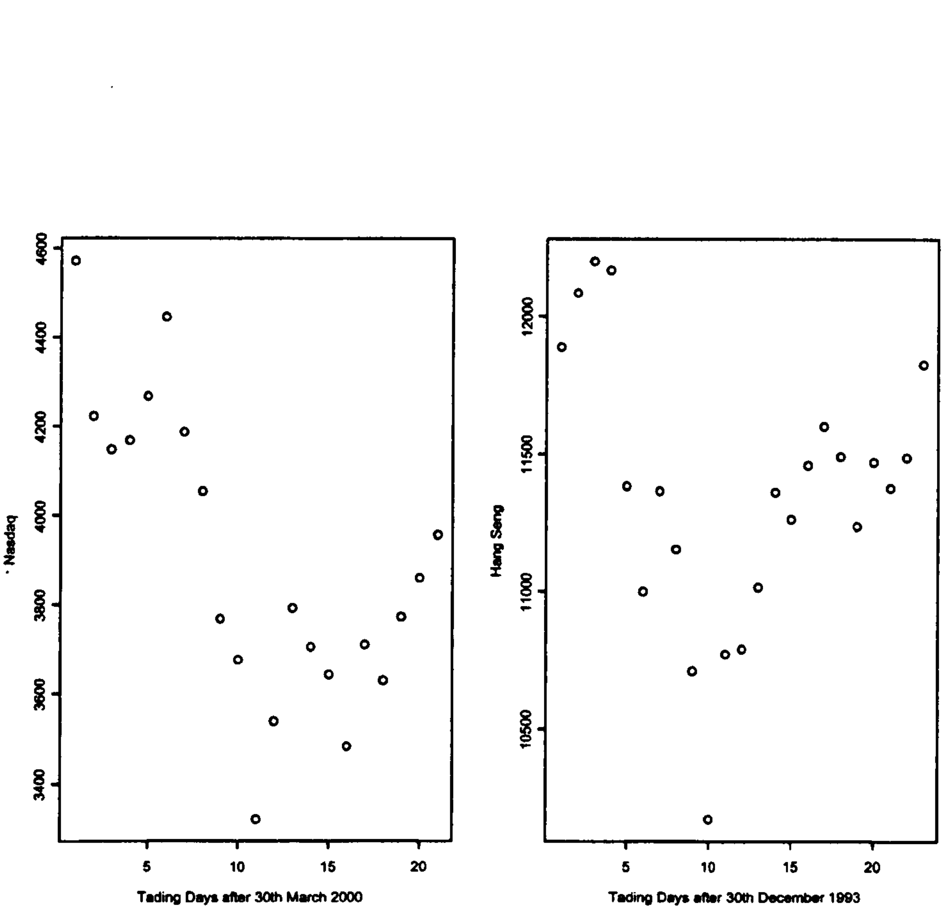


Figure 1.2: Plot of Nasdaq and Hang Seng indices over crash period

Results for the linear fracture model ((1.17), *lfm*) and the amended nonlinear log-periodic model ((1.33), *nlp*) are shown in Table 1.1. At least superficially these models appear to fit the data well with large R^2 values suggesting that these models explain a large degree of the variability in the data. The results also appear to give reasonable and interpretable estimates of the parameter t_c, the time when the market is adjudged most likely to crash. The estimated standard errors for the estimates of t_c are generally smaller for the more complicated nonlinear model.

1.5 Conclusions

Interpretable crash time estimates and a generally good fit to data suggest that the models considered here are capable of detecting genuine features in financial data. The generalisation proposed by Sornette and Johansen in (1.33) performs well and seems to provide better estimates of t_c, than the linear fracture model (1.17). Moreover, the estimated standard errors are generally smaller for the more complicated nonlinear model. There is however some suggestion that despite such an apparently good fit the models considered here might be susceptible to over-fitting noise, with some suggestion from Figure 1.1 that a power law growth dominates any apparent log-periodic oscillations. In

Market	Time Span	Model	R^2	t_c (e.s.e)
S&P 500	4 years	nlp	98.24	1988.00 (0.030)
	4 years	lfm	97.59	1990.35 (0.148)
	2 years	nlp	98.23	1987.99 (0.007)
	2 years	lfm	97.87	1988.00 (0.043)
Nasdaq	4 years	nlp	97.95	2000.33 (0.007)
	4 years	lfm	97.75	2000.87 (0.055)
	2 years	nlp	97.48	2000.40 (0.019)
	2 years	lfm	97.34	2005.59 (0.974)
Hang Seng	4 years	nlp	97.88	1994.16 (0.329)
	4 years	lfm	97.56	1995.43 (0.299)
	2 years	nlp	96.04	1994.00 (0.010)
	2 years	lfm	95.06	1994.13 (0.011)

Table 1.1: Results for log-periodic models

Chapters 2 we consider a more simple expressions for the log-price in a non-linear regression formulation of the problem, before examining more realistic SDE models in Chapters 3-5.

Chapter 2

Simple models for bubbles

2.1 Motivation

As discussed in Chapter 1.3.4-1.3.5, the Sornette-Johansen methodology entails fitting the following expressions to the log-price:

$$y(t) = A + B(t_c - t)^{1-\alpha}(1 + C\cos(\phi + \omega log(t_c - t))) + \epsilon_t, \tag{2.1}$$

and

$$\begin{aligned} y(t) &= A + B\frac{(t_c - t)^{\alpha}}{\sqrt{1 + ((t_c - t)/\Delta t)^{2\alpha}}} \times \\ & \left[1 + C\cos[\omega log(t_c - t) + \frac{\Delta\omega}{2\alpha} log(1 + ((t_c - t)/\Delta t)^{2\alpha}) + \phi]\right] + \epsilon_t, \end{aligned} \tag{2.2}$$

where the ϵ_t are i.i.d. zero-mean errors, $y(t)$ is the log-price and the parameter t_c represents the most likely time for the market to crash. Further details and methodology are given in *[52]* and *[97]*. For critical reviews see *[39]*, *[40]*, and *[38]*. *[93]*, Chapter 6, describes log-periodic oscillations as enabling one to "lock in" on oscillations relevant to the critical data t_c (the time when a crash is most likely). These oscillations represent periodic corrections to departures from plain power-law type behaviour. Plotting log-prices against time, what is arguably more noticeable is an approximate power-law growth with any apparent log-periodic oscillations appearing somewhat secondary, see Figure 1.1.

The Sornette-Johansen method has come under some criticism. *[39]* criticizes Sornette (see also *[40]*) essentially for a perceived continued lack of rigour. *[65]* makes three criticisms of Sornette's work. A criticism is made that (2.2) is not based on a convincing underlying theoretical model. A further criticism was made that the method requires estimation of a large number of parameters (in this case nine), and a further concern that these formula may simply be over-fitting noise. We have some sympathy with this point of view. A plot of the logarithm of the S&P 500 index (y_i) together with the fitted values of the simple model $y_i = A + B(t_c - t_i)^\alpha + \eta_i$, where the η_i follow an AR(2) process is shown in Figure 2.1. One might argue that this is not really a serious candidate model for the price process as many models for financial processes such as the random walk model are non-stationary unit root processes. However, this plot seems sufficient to argue that simple statistical fluctuations might be enough to generate an apparently log-periodic signal and we should perhaps look for a more simple model. By a simple exploratory analysis it seems that rather than log-periodicity, we should perhaps instead seek to model super-exponential growth[1]. If the underlying model for the price is geometric Brownian motion, then the log-returns (first difference of the log-price) will be an i.i.d. sequence of normal random variables. Table 2.1 gives the empirical values of the mean and standard deviation of log-returns. Higher log-return means are observed in the two-years immediately prior to the crash. Higher standard deviations in the two years immediately before crashes also seem to point to an increase in volatility prior to crashes.

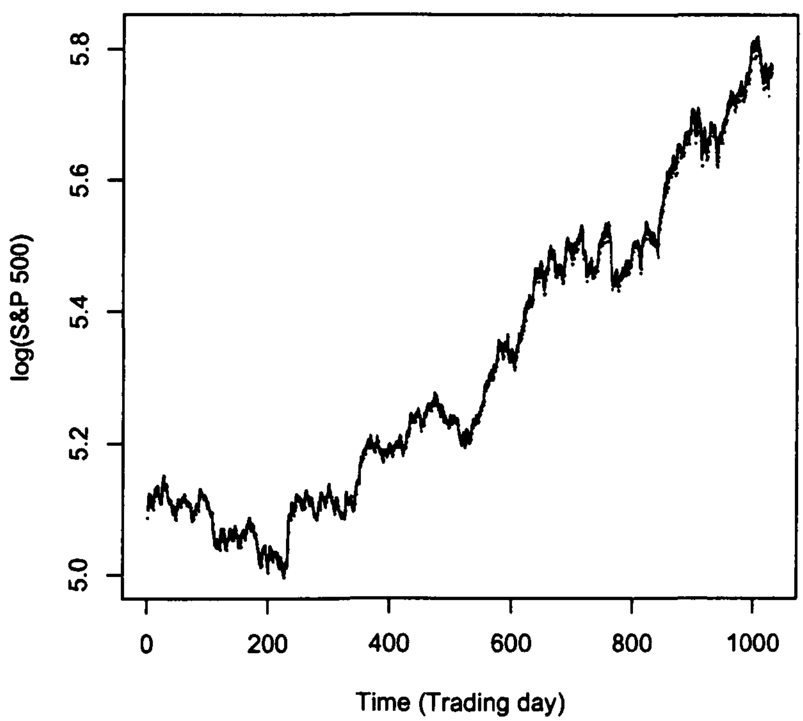


Figure 2.1: Plot of Log(S&P500) and fit of power-law model with autocorrelated errors

Financial crashes are often viewed along the lines of the title of Shiller's *Irrational*

[1]For reference, the comment is made in *[26]*, Chapter 7, that to a reasonable approximation, most time series in economics tend to exhibit roughly exponential growth on average.

Index	Time span	Mean (standard deviation)
S&P 500	Four years	0.000653 (0.00826)
	Two years	0.00102 (0.00916)
Nasdaq	Four years	0.00139 (0.0151)
	Two years	0.00181 (0.0184)
Hang Seng	Four years	0.00144 (0.013009)
	Two years	0.00205 (0.01408)

Table 2.1: Means and standard deviations of log-returns

Exuberance *[89]*. However, some recent work suggests that this might in fact not be the case, and there may be some mileage in regarding the process as obeying principles of rationality. An economic mechanism may exist whereby, as stated by *[93]*, Chapter 5 page 137, sufficiently many traders behave in such a way that prices tend to reflect all available information. As this happens, the inherent risks are approximately fairly priced. The aim of this chapter is to produce a simple model *without* log-periodicity, but which is interpretable econometrically and is still able to account for super-exponential growth.

In the second section we present, for illustration, an admittedly brief review of some of the theories of financial crashes that can be found in the economics literature. In the third section we describe the Johansen-Ledoit-Sornette (JLS) model of financial crashes, as presented in *[52]*. The authors' original intention was to produce a simple model where, at least in the regime leading to a financial crash, log-prices exhibit a power-law acceleration. Log-periodicity in finance is then motivated by analogy with a statistical mechanics model – solved by *[32]* – in which log-periodicity occurs and generalises the plain power-law. We present a simple extension to the original formulation. The rather striking conclusion is that, rather than exhibiting log-periodic oscillations before a crash, log-prices should be the sum of a linear trend and a super-exponential growth term. We are thus lead to an elegant nested model formulation of the problem. Section 4 describes backward predictions corresponding to observed crashes. Section 5 is a brief section on residual model checks. Section 6 is a conclusion.

2.2 Some economic theories of financial crashes

[84] disputes the idea of irrational roots of crises. According to *[84]*, investors may be drawn to the market at times of great uncertainty, by the possibility of large potential profits. Under this interpretation, there is thus no requirement for mass irrationality in order for a stock market to crash. The view of *[84]* is that since, at least to a reasonable approximation, traders can heed all available information and still participate in such dangerous markets, crashes in fact perform a vital house-cleaning function. Crashes are seen as allowing the market – somewhat brutally – to sort truly profitable investments

from merely illusory ones.

[108] and *[21]* stress the importance of economic fundamentals. According to *[108]* the key economic mechanism is *informational overshooting.* Market fundamentals change over an unknown period of time, and the economy undergoes a boom. A crash finally occurs as information is processed when the boom comes to an end. Informational overshooting occurs when the market expands to a new capacity, which is unknown until this capacity is finally reached. This overshooting can occur in two ways, either increased productivity brought about by rapid technological progress, or by an influx of new investors into the stock market. *[108]* goes on to state that the fact that financial liberalisations (literally economic freedoms) are typically followed by booms and crashes, reflects that it is possible to identify an underlying economic process behind the mass panics and hysterias.

[21] also focus on the role played by so-called economic fundamentals. However, in contrast to Zeira, they focus more on *technology shocks* – factors affecting production – than on the nature of the flow of information. According to *[21]*, when an old mode of production cannot be improved further, it is dismissed and progressively replaced by a newer one. The dismissal of the old technology and the introduction of new machines trigger recessions. These recessions may continue for a few periods while the new technology is being developed, but end before the development of the new machines is completed. According to *[21]*, their model predicts large and sudden stock market crashes which follow periods of regular growth. These crashes are the result of fundamental factors and not the result of market inefficiencies, nor irrational behaviour of investors. According to *[21]*, their model predicts market crashes of the order of 30%. However, the comment is made in *[86]*, that a slight computational error occurs in this paper, and in fact the model can generate falls of the order of 25-50%.

[67] takes a similar point of view to *[108]* and emphasises the role played by informational aggregation. However, in contrast to *[108]*, *[67]* does not pay explicit attention to the role of economic fundamentals. According to *[67]*, hidden information in the market may be released by a small trigger and in this model information aggregation can cause a crash, even when investors behave rationally. Irrationality of bubbles is also at odds with the rational bubble models considered by *[60]*.

2.3 JLS Power-Law model

[93], Chapter 5 page 137, assumes that sufficiently many traders behave in such a way that "prices tend to reflect available information and risk is adequately and approximately fairly remunerated". It is this economic mechanism – of inherent risks faithfully reflected by

prices – which we hope will prove insightful here. In this section we discuss the Johansen-Ledoit-Sornette (JLS) power law model as presented in *[52]*. In the original paper this model is simply used as a stepping stone to motivating a log-periodic model in finance.

We consider a purely speculative asset that pays no dividends. This simple model assumes that only one crash can occur and so the effects of any past crashes are ignored. JLS assume for the sake of simplicity that in the case of a crash, the price drops by a fixed percentage $\kappa \in (0, 1)$. We introduce the filtration $\mathcal{F}_t, (t > t_0)$ to model the flow of relevant information. To model a crash we consider a jump-process $j(t)$ which takes the values $\{0, 1\}$. We define a non-negative continuous random variable X which corresponds to the time of the crash so that $X = \inf\{t : j(t) \neq 0\}$. Johansen et al. start by ignoring the effects of the interest rate, risk aversion, information asymmetry and the market-clearing condition. Johansen et al. then state that in this dramatically simplified framework, rational expectations are simply equivalent to a martingale hypothesis for the price process:

$$E[p(t')|\mathcal{F}_t] = p(t), \tag{2.3}$$

where $t' > t$ and $j(t)$ and $p(t)$ are adapted to the filtration $\mathcal{F}_t, (t > t_0)$. Ignoring a zero-mean noise term, the dynamics of the asset price up to and including the crash is given by

$$dp(t) = \mu(t)p(t)dt - \kappa p(t)dj(t), \tag{2.4}$$

where $\mu(t)$ is a time-dependent drift chosen in order to satisfy the martingale condition and the $dj(t)$ term corresponds to the jump process $j(t)$. Equation (2.4) states that the price exhibits a general growth trend, which is offset by the crash when it finally happens. (2.4) is a very simple separable stochastic differential equation, which can be solved exactly to give

$$p(t) = p(t_0) \exp\left\{ \int_{t_0}^{t} \mu(s)ds - \kappa j(t) \right\}. \tag{2.5}$$

If we impose a martingale condition, this has implications for $\mu(\cdot)$ in (2.5) and helps to give the model empirically testable content.

2.3.1 First-order martingale approximation

Here, we introduce the hazard rate $h(t)$ and consider the non-negative continuous random variable X which corresponds to the time of the crash. The hazard rate[2] $h(t)$ of X is

[2]If F is the absolutely continuous cdf of a random variable X with density f, the hazard rate is simply given by $h(t) = f(t)/(1 - F(t))$.

defined as

$$h(t) = \lim_{\delta \downarrow 0} \frac{1}{\delta} Pr(X \le t + \delta | X > t) \quad t \ge 0.$$

This suggests that if the crash has not happened by time t, the conditional probability that the crash will occur within a small period of time $[t, t+\delta]$ is $\delta h(t) + o(\delta)$. Next we replace the continuous time dynamics in (2.4) with the discrete approximation

$$p(t+\delta) - p(t) = \mu(t)p(t)\delta - \kappa p(t)[j(t+\delta) - j(t)]. \tag{2.6}$$

Suppose that the crash has not happened by time t. Invoking the martingale condition when taking conditional expectations in (2.6) we get

$$\begin{aligned} E[p(t+\delta) - p(t)|\mathcal{F}_t] &= \mu(t)p(t)\delta - \kappa p(t)E[j(t+\delta) - j(t)|\mathcal{F}_t], \\ &= \delta p(t)[\mu(t) - \kappa h(t) + o(\delta)/\delta] \\ &= 0, \end{aligned} \tag{2.7}$$

where we have used $Pr(j(t+\delta) - j(t) = 1 | j(t) = 0) = Pr(X \in (t, t+\delta) | X > t) = \delta h(t) + o(\delta)$. Since the price $p(t)$ is assumed to remain positive, we obtain $\mu(t) - \kappa h(t) + o(\delta)/\delta = 0$, and on taking the limit $\delta \downarrow 0$ we get

$$\mu(t) - \kappa h(t) = 0. \tag{2.8}$$

Thus with (2.8) in mind, it is probably equally valid to interpret κ as some kind of coefficient of risk aversion, rather than its slightly artificial introduction as the pre-determined size of a crash. The important economic intuition behind (2.8) is that as the probability of the crash increases, the price must increase on average in order to induce traders into holding the asset. Inserting (2.8) into (2.5) predicts that we have

$$\log(p(t)) = \log(p(t_0)) + \kappa \int_{t_0}^{t} h(u)du \tag{2.9}$$

as our pre-crash dynamics. JLS build on (2.9) by assuming that the stock market evolves – to a good approximation – according to a statistical mechanics model similar to the 2-d Ising model. In models of this kind, there is a critical point K_c that determines the properties of the system. When $K < K_c$ disorder dominates and sensitivity to a small global influence is small. Clusters of like spins are small in size and imitation of signs propagates only between close neighbours. Here the susceptibility χ, defined in Chapter 1, is finite. As K increases towards K_c, order starts to appear, and the system becomes extremely sensitive to a small global perturbation. Large clusters of spins with the same size occur, and imitation propagates over large distances. In this case the susceptibility χ tends to infinity. This is one of the characteristics of critical phenomena. One of the hallmarks of criticality is the power law increase of the susceptibility:

$$\chi \sim A(K_c - K)^{-\alpha}, \tag{2.10}$$

where A and α are positive constants. For the 2-d Ising model $\alpha = 7/4$ *[52]*. We do not know the dynamics that drive the key parameter K of our toy model of the stock market. JLS are merely content to assume that the parameter K evolves as a sufficiently smooth function of time t, so that we have $(K_c - K(t)) \approx \tilde{B}(t_c - t)$, where $\tilde{B}$ is some constant independent of t. Using this approximation, JLS posit that the hazard function behaves in a manner comparable to the susceptibility as defined above. Thus we have

$$h(t) \approx B(t_c - t)^{-\alpha}, \tag{2.11}$$

for some constant $B = A\tilde{B}^{-\alpha}$. If we assume that this condition holds exactly, we note that integrating (2.11) would give

$$-\ln(1 - F(t)) = \frac{-B}{1-\alpha}(t_c - t)^{1-\alpha} + \text{Const.}, \tag{2.12}$$

and (2.12) enables us to recover $F(t)$. However, we note that $\alpha = 1$ would also be a valid exponent. In this case integrating (2.11) leads to

$$F(t) = 1 - A(t_c - t)^B,$$

where A is a constant determined by the initial conditions. Assuming equality in (2.11) and using (2.9) leaves us with the following prediction for the pre-crash dynamics:

$$\log(p(t)) = \log(p(t_0)) + C(t_c - t)^{1-\alpha}, \tag{2.13}$$

where $C = \frac{-\kappa B}{1-\alpha}, (\alpha \neq 1)$. The model of *[52]* assumes that α is constrained to lie in the region $(0, 1)$. Under the specification $\kappa, B \geq 0$ and $0 < \alpha < 1$, it follows that C is negative. The original JLS paper merely used these workings to motivate a log-periodic model in finance. In certain systems, as discussed in Chapter 1, log-periodicity and complex exponents occur and generalise power laws with purely real exponents. Based on this analogy, *[52]* insert a log-periodic hazard function $h(t)$ into (2.9) in place of a plain power law hazard function.

Here, we take the view that the original JLS power law model can be improved upon very simply *without* log-periodicity. We do this by introducing the interest rate. In particular, we assume that it is the discounted price process rather than the raw price process that satisfies the martingale condition. Thus we replace (2.3) by

$$E[e^{-rt'}p(t')|\mathcal{F}_t] = e^{-rt}p(t), \tag{2.14}$$

where $r > 0$ represents a constant interest rate. We claim that (2.14) is more realistic than (2.3) and allows us to incorporate a "baseline" exponential growth into the model. Upon taking conditional expectations, (2.7) now becomes

$$p(t)(1 + r\delta + o(\delta)) - p(t), \tag{2.15}$$

since r is a constant interest rate. To see this we note that if $q(t) = e^{-rt}p(t), E(q(t + \delta)|\mathcal{F}_t) = q(t) = e^{-rt}p(t)$. So $E(p(t + \delta)|\mathcal{F}_t) = e^{r\delta}p(t)$. (2.15) leaves us with

$$\delta p(t)[r + o(\delta)/\delta] = \delta p(t)[\mu(t) - \kappa h(t) + o(\delta)/\delta],$$

which implies

$$\mu(t) - \kappa h(t) - r + o(\delta)/\delta = 0,$$

and taking the limit $\delta \downarrow 0$ gives

$$\mu(t) - r - \kappa h(t) = 0. \tag{2.16}$$

(2.16) has been previously derived in the literature, see for instance *[99]*, although it does not seem to have been estimated statistically. (2.16) inserted into (2.5) now predicts as our pre-crash dynamics

$$\log(p(t)) = A + rt + C(t_c - t)^{1-\alpha}, \tag{2.17}$$

where A is a constant determined by the unknown initial conditions. In practice when estimating (2.17) we find $1 - \alpha \approx 0$. This suggests logarithmic behaviour and that we may reasonably take $\alpha = 1$ in (2.11). In this case (2.17) becomes

$$\log(p(t)) = A + rt + C\ln(t_c - t), \tag{2.18}$$

where $C = -\kappa B$. Equations (2.17) and (2.18) are interesting as they describe a hypothesised super-exponential growth associated with risky markets. We refer to (2.17) and (2.18) as the SEG model. The super-exponential growth acts as bait which enables rational investors to be attracted to risky markets. Note that in a 'fundamental' regime where $h(t) \approx 0$, we have

$$\log(p(t)) = A + rt.$$

Thus, if (2.17) and (2.18) are to be associated with bubbles, then the fits obtained should show clear deviations from a straight line.

2.4 Backward predictions

In this section we fit the SEG model (2.18) against the price series prior to the observed crashes that we examined in the previous chapter, using ordinary least squares. The fit of the SEG model is generally good and explains around 98% of the variation in the log-price series. A plot of the fit obtained is shown in Figure 5.3. It seems that looking at four years of data may be preferable to looking at two years of data, as looking at four years of data it appears easier to visually identify the super-exponential growth that our simple model predicts. We discuss results for each of the markets in detail.

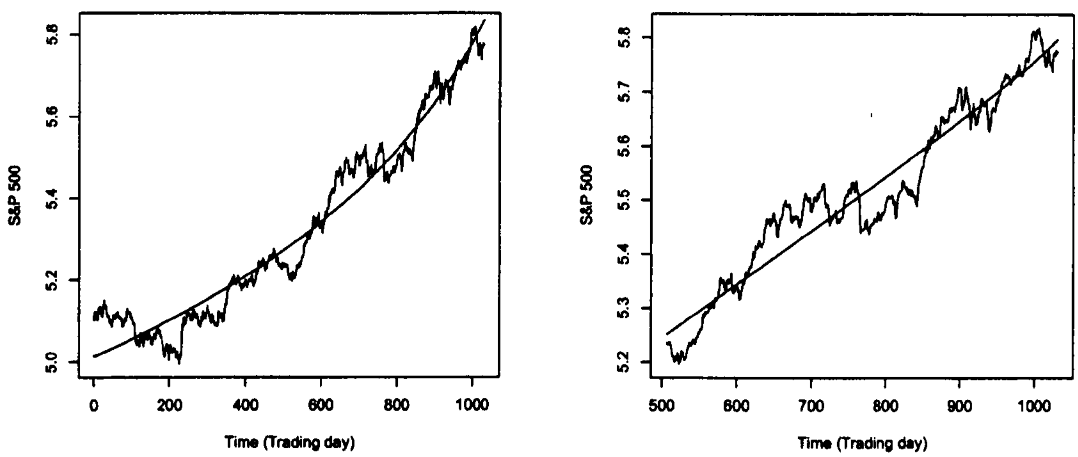


Figure 2.2: The SEG model as fitted to S&P500. Left panel: 1983 to 1987. Right panel: 1985 to 1987

2.4.1 S&P 500

The results for the S&P 500 index are shown in Table 2.2. The values of t_c correspond to January 1989 and early November 1987 respectively. The numbers in brackets denote estimated standard errors. However, note that the estimated standard error for the C parameter suggests that for the final two years the SEG terms are non-significant. These results reinforce the impression from Figure 2.2 that when testing for bubbles, four years of data may be preferable to two years of data. Note also from changing parameter estimates that there is some inherent non-stationarity that our simple model is unable to account for. These results also suggest that it would be reasonable to consider models for changing expectations in order to describe crashes. However, this is not something we pursue further here. Models for change-point detection in financial data are discussed in *[60]*.

Parameter	Four years	Two years
A	8.941 (0.129)	4.860 (0.103)
r	0.000 (0.000)	$9.589 * 10^{-4}(4.469 * 10^{-5})$
C	-0.546 (0.017)	-0.015 (0.012)
t_c	1325.097 (19.604)	1061.816 (41.319)

Table 2.2: Results for SEG model on S&P 500

2.4.2 Nasdaq

The results for the Nasdaq index are shown below in Table 2.3. Here estimated standard errors suggest that the SEG term is significant for both time periods. The estimates of t_c correspond to early May 2000 and early June 2000. Combined with associated estimated standard errors of around 5.583 days and 6.904 days, these results seem reasonable. Changing parameter estimates again suggest some inherent non-stationarity that our simple model is unable to account for. A plot of the fit obtained is shown in Figure 2.3 and appears reasonable in both cases. Note also that in both cases the fit obtained shows a clear departure from a straight line. Thus whilst this model is unable to handle some of the inherent non-stationarity, it is nonetheless interesting as a phenomenological description of the super-exponential growth prior to the crash.

Parameter	Four years	Two years
A	9.522 (0.150)	10.680 (0.120)
r	$3.546 * 10^{-4}(3.454 * 10^{-5})$	0.000 (0.000)
C	-0.360 (0.021)	-0.510 (0.037)
t_c	1068.853 (5.583)	1093.634 (6.904)

Table 2.3: Results for SEG model on Nasdaq

2.4.3 Hang Seng

The results for the Hang Seng index are shown below in Table 2.4. This time we seem to have rather less evidence for non-stationarity as the results obtained appear to be similar for both periods. The predicted crash times correspond to early January 1994 which seems reasonable. From the estimated standard errors all terms in the model appear significant and we seem to have strong evidence for super-exponential growth prior to the crash. A plot of the fit obtained is shown in Figure 2.4 and appears reasonable.

2.5 Brief model checks

The SEG model is primarily of interest as it is a simple theoretical model which describes the phenomenology of super-exponential growth prior to financial crashes. The regression

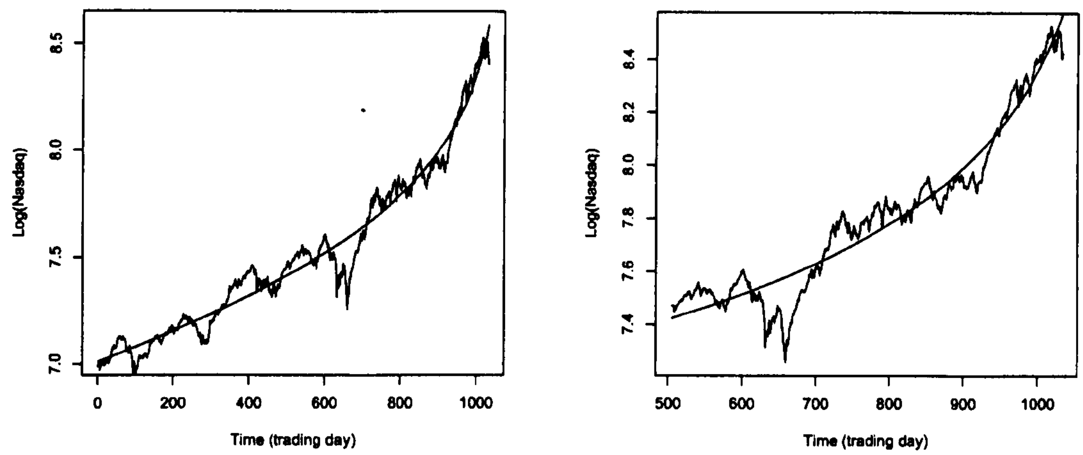


Figure 2.3: The SEG model as fitted to the Nasdaq. Left panel: 1996 to 2000. Right panel: 1998 to 2000

Parameter	Four years	Two years
A	8.580 (0.050)	8.923 (0.108)
r	$9.009 * 10^{-4}(1.902 * 10^{-5})$	$6.375 * 10^{-4}(5.818 * 10^{-5})$
C	-0.102 (0.007)	-0.126 (0.012)
t_c	1016.050 (1.270)	1017.226 (1.731)

Table 2.4: Results for SEG model on Hang Seng

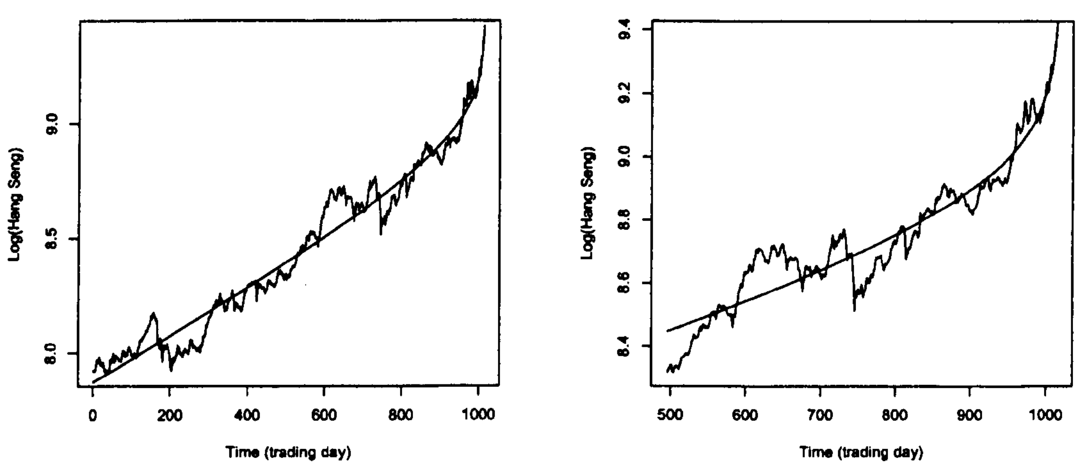


Figure 2.4: The SEG model as fitted to the Hang Seng. Left panel: 1989 to 1993. Right panel: 1991 to 1993

analysis in Section 4 assume that the residuals $e_i = y_i - A - rt_i + C\ln(t_c - t_i)$ should constitute an i.i.d. sequence from $N(0, \sigma^2)$. There are two key ways in which these model assumptions are violated. A normal plot of the residuals, see Figure 2.5, appears reasonable but does suggest some deviation from Normality in the tails. Most importantly, an ACF plot in Figure 2.6 shows that the residuals appear to be very highly correlated. (We can also detect signs that the residuals may be correlated by looking at the plots of best fit in Section 4.) One might consider trying to model the residuals as a low-order autoregressive process, see for instance *[88]*, Chapter 6, *[29]*. However, in practice, this method suffers from the fact that the estimated autoregressive models are close to non-stationary unit root processes.

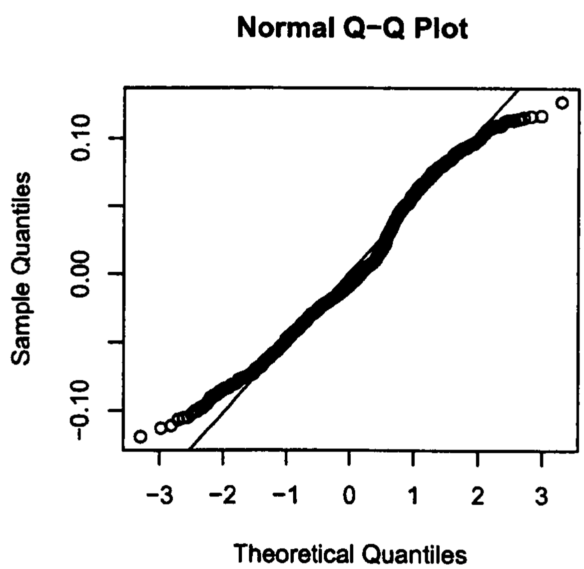


Figure 2.5: Sample normal probability plot for residuals of the SEG model

Both these problems can be rectified by considering stochastic differential equation models for financial bubbles. See Chapter 3. As a consequence of this, we can develop a formal statistical test for bubbles which can then be used in out-of-sample applications. For the prediction problem, which we discuss in Chapter 5, we propose to combine the SDE models in Chapter 3 with the simple SEG model in this chapter.

2.6 Conclusions

A very brief exploratory analysis suggested that very simple stochastic behaviour might produce an apparent log-periodic signal. Thus despite encouraging results in the previous chapter there is some suggestion that the log-periodic formulae may be susceptible to over-

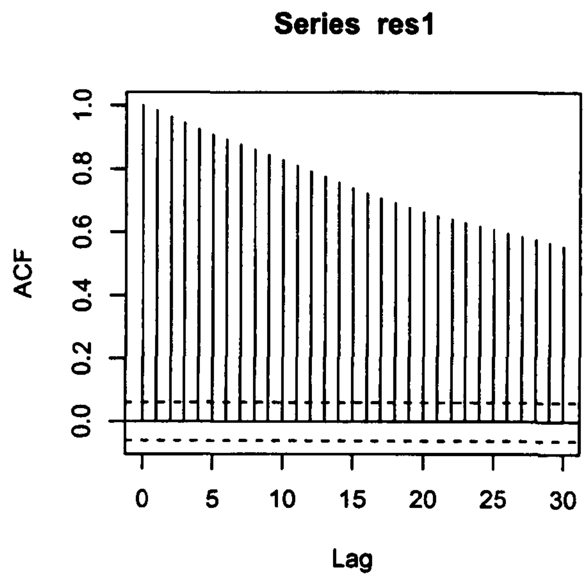


Figure 2.6: Sample ACF plot for residuals of the SEG model

fitting noise when applied to financial time series. Looking for something more simple, a very basic exploratory analysis suggested that rather than log-periodicity we could seek to model super-exponential growth. Some economic interpretations of financial crashes were discussed. In particular, the most interesting concept was that a crash might have a rational root, with investors induced into holding risky assets if the potential rewards were sufficiently high *[84]*. The JLS power-law model introduced in Section 3, is an interesting piece of applied mathematics and was intended to illustrate this feature.

A simple extension of the first-order martingale approximation in the JLS model was discussed. A slight modification of the original approach suggests that the log-price experiences a logarithmic singularity of the form $\propto \ln(t_c - t)$ prior to a crash. Further, the model leads to a possible framework for both nested models comparisons and simple graphical tests for bubbles. We find evidence of super-exponential growth prior to observed crashes on the S&P 500, Nasdaq and the Hang Seng. However, the super-exponential growth terms were non-significant for the S&P 500 when restricting the analysis to just two years of data and hence it appears that in order to detect bubbles and super-exponential growth, four years of data rather than just two years might be appropriate. The model was used to provide interesting back-testing results prior to observed crashes on the S&P500, the Nasdaq and the Hang Seng. However, changing parameter estimates suggest some inherent non-stationarity that our simple model is unable to account for. There are also two ways in which model assumptions appear to be violated. A normal probability plot gives some suggestion of non-normality of residuals. Further, it appears that residuals for this model are very highly correlated. We note that both of these problems can be rectified by the SDE models in the next chapter.

For completeness, the exact definition of the hazard rate is given in Appendix A. It might be of interest to incorporate this directly into the model. One way to do this might involve a Bayesian formulation of the problem (see *[28]*).

Chapter 3

SDE models for bubbles

3.1 Motivation

In this chapter we begin to consider tests for out-of-sample application of the SEG model. Hence, we are naturally drawn to consider stochastic differential equation (SDE) models for bubbles as we try to avoid spuriously regressing a random walk. We use the model developed in practical applications in later chapters. In Chapter 4 we use a novel application of the Sornette-Andersen model, [94], to gauge which factors remain significant once we take into account super-exponential growth. In Chapter 5 we suggest use of an SDE model, formulated in this chapter, alongside material developed in Chapter 2 for out-of-sample work.

In addition to practical considerations, there are also interesting theoretical considerations which motivate material in this chapter. The discussion in [2] suggests that it might be possible to have a hierarchy of models for bubbles, see also [94]. There is a proposed division between *fearless* and *fearful* bubbles. In the JLS power law model, a crash is preceded by super-exponential growth only. This corresponds to a fearless bubble regime. In the model of [94], super-exponential growth can be accompanied by an increase in volatility – a fearful bubble regime.

Initially, in Section 1, we restrict attention to the case where the background noise is Brownian motion. We have three competing models; geometric Brownian motion as a proxy for the absence of bubbles, an SDE formulation of the JLS power law model, and the model of [94]. We reject the proposed fearless bubble formulation, and instead the main point of interest appears to be that the Sornette-Andersen model represents an empirically reasonable SDE extension of earlier ODE models for bubbles. We discuss the solution of the Sornette-Andersen model [94] which constitutes a nonlinear generalisation of the Black-Scholes model of a financial market. By discussing the solution to this model, we show how two identifiability constraints can be made which reduce the apparently six-dimensional estimation problem into a four-dimensional problem. In [2], where an attempt was made to estimate the Sornette-Andersen model statistically, the authors erroneously claim that the likelihood-ratio test cannot be used to reliably choose the correct model for a given time series. However, as we demonstrate, the likelihood function for the Sornette-Andersen model can be written down exactly. Moreover, the first-order maximum likelihood condition further reduces the four dimensional problem to a two-dimensional parameter search, which can be easily solved numerically.

One of the so-called *stylised empirical facts* of financial markets, see for example [31], Chapter 7 page 210, is that financial data, particularly on short time-scales, is seen to exhibit heavy-tailed non-Gaussian behaviour. In Section 2, we consider a reformulation of the Sornette-Andersen model using a heavy-tailed hyperbolic process as the background driving noise. We are thus able to further refine SDE models and empirical tests for bubbles. Section 3 is a brief conclusion.

3.2 No bubble vs. Fearless bubble vs. Fearful bubble

3.2.1 No bubble model: Geometric Brownian Motion

Neglecting heavy-tailed non-Gaussian effects, one might use geometric Brownian as a reasonable non-bubble proxy model of a financial market. This model assumes the following SDE for the price X_t:

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \tag{3.1}$$

where $(W_t, \quad t > 0)$ is standard Brownian motion, $\mu \in \mathbb{R}$ and $\sigma > 0$. Using Itô's formula, it follows that $Y_t = \log(X_t)$ has the stochastic differential

$$\begin{aligned} dY_t &= (\mu - \sigma^2/2)dt + \sigma dW_t, \\ &= \alpha dt + \sigma dW_t, \end{aligned} \tag{3.2}$$

where $\alpha = \mu - \sigma^2/2$. We can recognise (3.2) as the equation for Brownian motion with drift. If we observe y_i from Y_t at time i, the product of these transition densities for Y_t can then be written as

$$\prod_{i=2}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - \alpha - y_{i-1})^2}{2\sigma^2}}.$$

Using the rule for transformations of univariate random variables, we can write the likelihood for the observed prices x_i as

$$\prod_{i=2}^{n} \frac{1}{\sqrt{2\pi}\sigma x_i} e^{-\frac{(\log(x_i)-\alpha-\log(x_{i-1}))^2}{2\sigma^2}},$$

so that the log-likelihood function becomes

$$-\frac{n-1}{2}\log(2\pi) - (n-1)\log(\sigma) - \sum_{i=2}^{n}\log(x_i) - \sum_{i=2}^{n}\frac{(\log(x_i) - \alpha - \log(x_{i-1}))^2}{2\sigma^2}, \quad (3.3)$$

This leads to the maximum likelihood estimates

$$\begin{aligned}\hat{\alpha} &= \frac{\sum_{i=1}^{n-1}\Delta y_i}{n-1}, \\ \hat{\sigma}^2 &= \frac{\sum_{i=1}^{n-1}(\Delta y_i - \hat{\alpha})^2}{n-1},\end{aligned}$$

where $\Delta y_i = y_{i+1} - y_i$.

3.2.2 Fearless bubble model

If one assumes that the noise added on the end of the log-price in the JLS power-law model is Brownian motion, this model can be formulated as

$$dX_t = \mu(t)X_t dt + \sigma X_t dW_t - \kappa dj, \quad (3.4)$$

where the dj term corresponds to a jump term which occurs with fixed amplitude κ. Before a crash, so that the dj term remains zero, a first-order martingale argument given in Chapter 2 necessitates that $\mu(t) = r + \kappa h(t)$, where $h(t)$ is the hazard function of the non-negative continuous random variable X which determines the time of the crash. These considerations suggest that the pre-crash dynamics of the fearless bubble model in (3.4) should be

$$dX_t = X_t(r + B(t_c - t)^{-\alpha})dt + \sigma X_t dW_t, \quad (3.5)$$

where we have assumed that $h(t) = B(t_c - t)^{-\alpha}$. We may solve this model as follows. Let $Y_t = \log(X_t) - rt + \frac{B(t_c-t)^{1-\alpha}}{1-\alpha}$. It follows from Itô's formula that

$$dY_t = \frac{-\sigma^2}{2}dt + \sigma dW_t. \quad (3.6)$$

It follows from (3.6) that if we observe y_i from Y_t at time i, the product of these transition densities for Y_t can then be written as

$$\prod_{i=2}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-\left(y_i+\sigma^2/2-y_{i-1}\right)^2}{2\sigma^2}}.$$

Using the rule for transformations of univariate random variables, we can write the likelihood for the observed prices x_i as

$$\prod_{i=2}^{n} \frac{1}{\sqrt{2\pi}\sigma x_i} \exp\left\{ \frac{-\left(\log(x_i) - \gamma - \log(x_{i-1}) + \frac{B(t_c-i)^{1-\alpha}}{1-\alpha} - \frac{B(t_c-i+1)^{1-\alpha}}{1-\alpha}\right)^2}{2\sigma^2} \right\}, \quad (3.7)$$

where $\gamma = r - \sigma^2/2$. From (3.7), the log-likelihood function is

$$\begin{aligned} l(\Theta|\underline{X}) &= \tfrac{n-1}{2}\log(2\pi) - (n-1)\log(\sigma) - \textstyle\sum_{i=2}^{n}\log(x_i) \\ &- \tfrac{1}{2\sigma^2}\textstyle\sum_{i=2}^{n}\left(\log(x_i) - \gamma - \log(x_{i-1}) + \frac{B(t_c-i)^{1-\alpha}}{1-\alpha} - \frac{B(t_c-i+1)^{1-\alpha}}{1-\alpha}\right)^2. \quad (3.8) \end{aligned}$$

Note that (3.8) contains five parameters, and has to be maximised numerically.

3.2.3 Fearful bubble model

This model is discussed in *[94]* and *[2]*. Not only is the model an SDE alternative to the ostensibly ODE-based JLS power-law model, it is intended to account for the possibility that prior to a crash we may have both super-exponential growth *and* an increase in volatility. The starting point is to choose a nonlinear SDE for the bubble price B_t to generalise (3.4):

$$dB_t = \mu(B_t)B_t dt + \sigma(B_t)B_t dW_t - \kappa B_t dj(t), \quad (3.9)$$

where the $dj(t)$ term corresponds to a jump process as before and κ is assumed to be an observation from some distribution with finite mean which determines the size of the crash. Suppose that the bubble process obeys the martingale property and let $B(t)$ and $j(t)$ be adapted to the filtration $\mathcal{F}_t$, $t > 0$. A simple linearisation argument suggests that if a crash has not happened by time t

$$\begin{aligned} E[B_{t+\delta} - B_t|\mathcal{F}_t] &\approx \delta\mu(B_t)B_t - B_t(E(\kappa(j(t+\delta) - j(t))|\mathcal{F}_t), \\ &= \delta\mu(B_t)B_t - B_t < \kappa > (\delta h(t) + o(\delta)), \\ &= \delta B_t\left[\mu(B_t) - < \kappa > (h(t) + o(\delta)/\delta)\right], \\ &= 0, \end{aligned}$$

where we have assumed that κ is independent of the filtration $\mathcal{F}_t$. If B_t is non-zero, we must have $\mu(B_t) - <\kappa> h(t) + o(\delta)/\delta = 0$, and taking the limit $\delta \downarrow 0$ leaves us with

$$h(t) = \frac{\mu(B_t)}{<\kappa>}. \tag{3.10}$$

(3.10) is intended to illustrate the risk-return interplay inherent within the bubble process. However, in contrast to the JLS power law model, *[52]*, it is the price which drives the hazard rate rather than the reverse. In *[94]* (3.10) is referred to as the variable hazard-rate and gives a possible way of applying the model to summarise risk levels in financial applications. (3.9) suggests that prior to a crash the bubble dynamics should be

$$dB_t = \mu(B_t)B_t dt + \sigma(B_t)B_t dW_t. \tag{3.11}$$

[94] use

$$\mu(B_t)B_t = \frac{m}{2B_t}[B_t\sigma(B_t)]^2 + \mu_0[B_t/B_0]^m, \tag{3.12}$$
$$\sigma(B_t)B_t = \sigma_0[B_t/B_0]^m, \tag{3.13}$$

where m, B_0, σ_0, μ_0 are all positive constants. The exact form of (3.12) is chosen as a convenient device to simplify the solution of the stochastic differential equations when applying Itô's formula. (3.12) and (3.13) are intended to represent a nonlinear generalisation of geometric Brownian motion which is recovered for $m = 1$. The power laws in (3.12) and (3.13) are assumed to be a simple and meaningful way of incorporating nonlinearity. As we discuss later, this nonlinearity creates a singularity in finite time which becomes stochastic in this SDE framework. This nonlinearity is also interesting as a phenomenological description of an explosive bubble phase that continues to feed on itself. We note that the power B_t^m can be decomposed as $B_t^m = B_t^{m-1} \times B_t$ so that B_t^{m-1}, which plays the role of a growth rate, is a function of the bubble itself. This is intended to capture a positive feedback effect whereby a larger bubble B_t feeds a larger growth rate, which in turn leads to a larger bubble, and so on.

As a statistical formulation of the model, *[2]* use

$$P(t) = Fe^{rt}B(t), \tag{3.14}$$

where F denotes a constant fundamental value and r denotes a constant interest rate. Direct application of the model in (3.14) then requires estimation of six parameters $(F, r, m, B_0, \sigma_0, \mu_0)$ and a seventh if one estimates the unknown constant determined by the initial conditions in (3.11). Here we choose the identifiability constraint $F = 1$, since it follows from (3.11-3.13) that $Y_1(t) = e^{rt}B(t)$ and $Y_2(t) = \lambda e^{rt}B(t)$ satisfy the same stochastic differential equation. In the sequel we show how to solve this model and demonstrate how a further identifiability constraint can be made. Let $P(t) = e^{rt}B(t)$. It follows from Itô's formula that

$$dP(t) = (\mu(B(t)) + r)P(t)dt + \sigma(B(t)))P(t)dW_t.$$

Next, define $Y(t) = P(t)^{1-m}e^{r(m-1)t}$. Using Itô's formula, it follows that

$$\begin{aligned} dY(t) &= (1-m)P(t)^{-m}e^{r(m-1)t}[\mu(B(t))P(t)dt + \sigma(B(t))P(t)dW_t] \\ &- \frac{(1-m)m}{2}P(t)^{-m-1}e^{r(m-1)t}\sigma^2(B(t))P(t)^2dt. \end{aligned} \tag{3.15}$$

In order to solve (3.15), we note that from (3.12) and (3.13)

$$\begin{aligned} \mu(B(t)) &= \frac{m}{2}(\sigma(B(t)))^2 + \frac{\mu_0 B(t)^{m-1}}{B_0^m}, \\ \sigma(B(t)) &= \frac{\sigma_0 B(t)^{m-1}}{B_0^m}. \end{aligned} \tag{3.16}$$

Plugging these expressions into (3.15) leaves us with

$$\begin{aligned} dY(t) &= \frac{\mu_0(1-m)P(t)^{1-m}e^{r(m-1)t}B(t)^{m-1}}{B_0^m}dt \\ &+ \frac{(1-m)P(t)^{1-m}e^{r(m-1)t}\sigma_0 B(t)^{m-1}}{B_0^m}dW_t. \end{aligned} \tag{3.17}$$

Next, using $B(t) = e^{-rt}P(t)$, (3.17) simplifies to give

$$dY(t) = \frac{\mu_0(1-m)}{B_0^m}dt + \frac{(1-m)\sigma_0}{B_0^m}dW_t. \tag{3.18}$$

From (3.18) it appears that we may impose the identifiability constraint $B_0 = 1$, as including this parameter in the model does not provide any additional information. Further, we can recognise (3.18) as the equation for Brownian motion with drift, which implies the solution

$$Y(t) = Y_0 + \mu_0(1-m)t + (1-m)\sigma_0 W_t.$$

Finally, using $P(t) = [Y(t)e^{r(1-m)t}]^{\frac{1}{1-m}} = e^{rt}[Y(t)]^{\frac{1}{1-m}}$, we can solve for the price $P(t)$ as

$$P(t) = e^{rt}[A(t_c - t) + \sigma_0(1-m)W_t]^{\frac{1}{1-m}}, \tag{3.19}$$

where $A = (m-1)\mu_0$ and $t_c = \frac{-Y_0}{\mu_0(1-m)}$. As in *[94]*, equation (8), (3.19) is intended to illustrate that in the deterministic case $\sigma_0 = 0$ we have a finite-time singularity at a critical time t_c determined by the initial conditions. Two further observations are also made by *[94]*. Firstly, they state that as the term in square brackets in (3.19) goes to zero the bubble price explodes and as a consequence of (3.10), so does the hazard rate.

Thus, as this happens a crash becomes more and more likely and the price is unable to grow without check. Secondly, under the dynamics given by (3.19), a crash is not a certain event and the price can also deflate spontaneously due to the random realisation of $W(t)$ which can bring the term in the square brackets in (3.19) towards zero.

Here, it is the form of (3.18) which allows us to construct the exact likelihood function. If we observe y_i from $Y(t)$ at time i, the product of these transition densities for $Y(t)$ can then be written as

$$\prod_{i=2}^{n} \frac{1}{\sqrt{2\pi}\tilde{\sigma}} e^{\frac{-(y_i-\tilde{\mu}-y_{i-1})^2}{2\tilde{\sigma}^2}},$$

where $\tilde{\mu} = (1-m)\mu_0$ and $\tilde{\sigma} = |(1-m)|\sigma_0$. Using the transformation rule, we can write down the equation for the transition density of the observed prices p_i:

$$\frac{|\,(m-1)|p_i^{-m}e^{r(m-1)i}}{\sqrt{2\pi}\tilde{\sigma}} \exp\left\{\frac{-\left(p_i e^{r(m-1)i} - \tilde{\mu} - p_{i-1}e^{r(m-1)(i-1)}\right)^2}{2\tilde{\sigma}^2}\right\}. \tag{3.20}$$

From (3.20), the log-likelihood function for $m > 1$ can be written as

$$\begin{aligned} l(\Theta|\underline{X}) &= \frac{n-1}{2}\log(2\pi) - (n-1)\log(\tilde{\sigma}) - m\sum_{i=2}^{n}\log(p_i) + (n-1)\log(m-1) \\ &\quad + r(m-1)\sum_{i=2}^{n} i - \frac{1}{2\tilde{\sigma}^2}\sum_{i=2}^{n}\left(p_i e^{r(m-1)i} - \tilde{\mu} - p_{i-1}e^{r(m-1)(i-1)}\right)^2. \end{aligned} \tag{3.21}$$

As it stands maximising (3.21) represents a four-dimensional problem. However, the form of the log-likelihood function is such that maximum likelihood estimates $\hat{m}$ and $\hat{r}$ are sufficient to determine the maximum likelihood estimates of $\tilde{\mu}$ and $\tilde{\sigma}$ and (3.21) reduces to a simple two-dimensional parameter search for $\hat{m}$ and $\hat{r}$. Differentiating the log-likelihood in (3.21) with respect to $\tilde{\mu}$ and equating to zero, we see that

$$\sum_{i=2}^{n}\left(p_i e^{\hat{r}(\hat{m}-1)i} - \tilde{\mu} - p_{i-1}e^{\hat{r}(\hat{m}-1)(i-1)}\right) = 0.$$

Similarly, differentiating with respect to $\tilde{\sigma}$ leaves us with

$$\frac{-(n-1)}{\tilde{\sigma}} + \frac{1}{\tilde{\sigma}^3}\sum_{i=2}^{n}\left(p_i e^{r(m-1)i} - \tilde{\mu} - p_{i-1}e^{r(m-1)(i-1)}\right)^2 = 0,$$

leading to the nontrivial solution

$$\tilde{\sigma}^2 = \frac{1}{(n-1)}\sum_{i=2}^{n}\left(p_i e^{r(m-1)i} - \tilde{\mu} - p_{i-1}e^{r(m-1)(i-1)}\right)^2.$$

In conclusion, the major contribution of this subsection is the imposition of two identifiability constraints and a further reduction of the maximum likelihood problem to a two-dimensional parameter search. The combined effect is to allow practical and efficient calibration of the Sornette-Andersen model to empirical data and a refinement of the analysis in [2]. Further, (3.19) and the decomposition $P(t) = B(t)e^{rt}$ suggest intersting ways of extending the original model (See Chapter 3.3).

3.2.4 Data Analysis

Here, we choose our data analysis to coincide with [2]. We look at six markets which have previously been thought by analysts to incorporate bubbles. Data is from US$:DEM 3/1/1983 to 8/3/1985, Nasdaq 18/6/1999 to 27/3/2000, S&P 500 1/7/1985 to 31/8/1987 and 2/1/1991 to 4/9/1997 and the Hang Seng index (HSI) 2/1/1992 to 6/1/1994 and 3/1/1995 to 3/10/1997.

Suppose that we wish to conduct likelihood inference for nested hypotheses regarding a parameter θ. Suppose $\theta' = (\theta_1', \theta_2')$ where θ_1 has dimension p and θ_2 has dimension $p-q$. We wish to test the hypothesis $H_0 : \theta_1 = \theta_{1,0}$ against the alternative $H_1 : \theta_1 \neq \theta_{1,0}$. Under appropriate regularity conditions, we have the following asymptotic result for $n \longrightarrow \infty$:

$$\chi = 2(l(\hat{\theta}_1, \hat{\theta}_2) - l(\theta_{1,0}, \hat{\theta}_{2,0})) \sim \chi_q^2, \tag{3.22}$$

where n denotes the number of the sample, l the log-likelihood function, and the hats denote maximum likelihood estimates. The fearless bubble model reduces to the Black-Scholes model when $B = 0$. Thus we see that under the null hypothesis of the Black-Scholes model the likelihood ratio statistic in (3.22) should be approximately χ_3^2. Results for these tests are shown in Table 3.1. In each case we conclude that the fearless bubble does not offer a significant improvement over the Black-Scholes model and the SDE formulation (3.5) is rejected.

	Geometric Brownian Motion	SEG diffusion	χ	p-value
US$:DEM	2213.137	2213.976	1.678	0.642
Nasdaq	-1077.357	-1077.045	0.624	0.891
S&P 500 '87	-1189.802	-1189.535	0.534	0.919
HSI '97	-4326.728	-4326.702	0.052	0.997
HSI '94	-2985.627	-2984.016	3.222	0.359
S&P 500 '97	-4539.082	-4537.959	2.246	0.523

Table 3.1: Likelihood ratio tests: Geometric Brownian Motion vs. SEG diffusion

In contrast, the fearful bubble model can be seen to provide a significant improvement over the Black-Scholes model (see Table 3.2). The Black-Scholes model is seen to reduce to the case that $r = 0$ and $m = 1$. We see that under the null hypothesis of the Black-Scholes model, the likelihood ratio statistic in (3.22) should be approximately χ^2_2. These results suggest that we have bubbles on all these markets with the exception of the Hang Seng between 1995 and 1997. After the rejecting the proposed fearless bubble model, we see that the main point of interest appears to be that the Sornette-Andersen is useful in its own right both as an empirically reasonable SDE model for financial bubbles and as an alternative to the ostensibly ODE approach of the JLS power-law model.

	Geometric Brownian Motion	Sornette-Andersen	χ	p-value
US$:DEM	2213.137	2228.119	29.964	0.000***
Nasdaq	-1077.357	-1073.183	8.348	0.015*
S&P 500 '87	-1189.802	-1182.958	13.688	0.001**
HSI '97	-4326.728	-4326.644	0.168	0.919
HSI '94	-2985.627	-2978.439	14.376	0.0001***
S&P 500 '97	-4539.082	-4502.044	74.076	0.000***

Table 3.2: Likelihood ratio tests: Geometric Brownian motion vs. Sornette-Andersen model

3.3 Heavy-tailed extension of Sornette-Andersen model

One of the key stylised empirical facts regarding financial markets, is the prevalence of heavy-tailed non-Gaussian phenomena, particularly, over short time scales. To illustrate this, Figure 3.1 is a plot of the log-densities of the log-returns on the S&P 500. We see that the tails of the Gaussian density are far too thin with respect to the non-parametric kernel density estimate, and we cannot justify empirically the choice of the normal distribution to model the log-returns. In contrast, the fit of the symmetric hyperbolic distribution is seen to better replicate the empirically observed tail behaviour.

The symmetric hyperbolic distribution $h(\alpha, \delta, \mu)$ has density given by

$$f(x) = \frac{1}{2\delta K_1(\delta\alpha)} \exp(-\alpha\sqrt{\delta^2 + (x-\mu)^2}), \tag{3.23}$$

where K_v is the Bessel function of the third kind defined by the formula

$$K_\nu(x) = \frac{1}{2}\int_0^\infty u^{\nu-1} \exp\left\{-\frac{x}{2}\left(u + \frac{1}{u}\right)\right\} du,$$

([5], Chapter 5, [106]). The distribution given by (3.23) has mean μ, and variance $\frac{\delta}{\alpha}\left[\frac{K_1(\delta\alpha)}{K_2(\delta\alpha)}\right]$. Instead of Brownian motion in the Sornette-Andersen model, we replace σW_t by the Lévy process Z_t whose one-step transition densities are given by a symmetric hyperbolic distribution with zero mean [34]. Thus we are able to improve upon the Gaussian model by taking into account kurtosis (heavy-tails), although we retain the simplifying assumption of symmetry. In the original Sornette-Andersen model, the

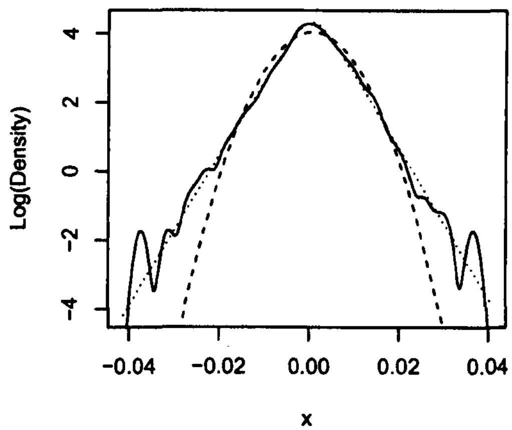


Figure 3.1: Plot of log-densities of log-returns on S&P 500. Solid line: non-parametric kde; dashed line: normal distribution; thin dashed line: symmetric hyperbolic distribution.

equation for a "fundamental" regime corresponding to random variation about a simple geometric trend ($m = 1, r = 0$) is

$$\log(P(t)) = \text{Const.} + \tilde{\mu}t + \sigma W_t, \tag{3.24}$$

where $P(t)$ denotes the price. In a bubble regime, the equation for the price becomes

$$P(t) = e^{rt}\left(\text{Const.} + \tilde{\mu}t + \sigma W_t\right)^{\frac{1}{m-1}}. \tag{3.25}$$

Inserting a hyperbolic noise process, the model for fundamental "geometric" behaviour becomes

$$\log(P(t)) = \text{Const.} + \tilde{\mu}t + Z_t, \tag{3.26}$$

with σW_t replaced by Z_t. Similarly, we change the equation for a bubble to

$$P(t) = e^{rt}\left(\text{Const.} + \tilde{\mu}t + Z_t\right)^{\frac{1}{m-1}}. \tag{3.27}$$

(3.27) thus becomes our model for a bubble, with power-law behaviour emerging as a result of collective phenomena characterising a 'bubble regime' and generalising the more regular, approximately geometric, behaviour (3.26). From (3.23) the log-likelihood function for the geometric model (3.26) becomes

$$-(n-1)\log(2\delta K_1(\delta\alpha)) - \sum_{i=2}^{n}\log(p_i) - \alpha\sum_{i=2}^{n}\left(\delta^2 + (\log(p_i) - \log(p_{i-1}) - \mu)^2\right)^{1/2}. \tag{3.28}$$

Similarly, the log-likelihood for the model (3.27) becomes

$$\begin{aligned} l(\Theta|\underline{X}) &= -(n-1)\log(2\delta K_1(\delta\alpha)) + (n-1)\log(m-1) - m\sum_{i=2}^{n}\log(p_i) \\ &+r(m-1)\sum_{i=2}^{n} i - \alpha\sum_{i=2}^{n}\left(\delta^2 + \left(p_i e^{r(m-1)i} - \mu - p_{i-1}e^{r(m-1)(i-1)}\right)^2\right)^{1/2} \end{aligned} \quad (3.29)$$

where p_i denotes the observed price at time i. Here, (3.28) and (3.29) have to be maximised numerically. We test between the geometric behaviour (3.26) and the bubble model (3.27), by using the χ^2 test for nested models.

The results using this test are shown in Table 3.3. This time we reject evidence of a bubble in the US$:DEM series. Moreover, the results in Table 3.3 are generally less significant than those shown in Table 3.2. This suggests that extending the original formulation by accounting for heavy-tailed non-Gaussian effects is successful in obtaining a more robust test for detecting bubbles in financial markets.

	Geometric Hyperbolic Model	Sornette-Andersen	χ	p-value
US$:DEM	2369.954	2371.791	3.674	0.159
Nasdaq	-1077.152	-1073.184	7.936	0.019*
S&P 500 '87	-1167.754	-1163.987	7.534	0.023*
HSI '97	-4280.854	-4280.854	0.000	1.000
HSI '94	-2954.711	-2948.239	12.944	0.002**
S&P 500 '97	-4468.12	-4441.845	52.55	0.000***

Table 3.3: Likelihood ratio tests: Geometric Hyperbolic model vs. Hyperbolic Sornette-Andersen model

3.4 Conclusions

In this chapter we tested a proposed hierarchy; no bubble vs. fearful bubble vs. fearless bubble. The formulation of the fearful bubble model was rejected, but we were successful in refining the analysis of *[2]*, to provide maximum likelihood tests for a nonlinear SDE model of bubbles. In sum, rather than the proposed hierarchical structure, the main feature of interest appears to be the maximum likelihood estimation and tests for this nonlinear SDE model of bubbles.

However, the original formulation of the model makes the assumption that the background driving noise is Gaussian, which is difficult to justify empirically (Figure 3.1). In Section 2 we re-formulated the model using a hyperbolic process, *[34]*, in order to take into account heavy-tailed non-Gaussian effects in financial markets. Results for the Sornette-Andersen model using the original Gaussian formulation and the new hyperbolic formulation are both interesting, but the hyperbolic formulation appears to be an improvement and seems more robust. In sum, using the results obtained in Section 3, we conclude in favour of bubbles on the Nasdaq 18/6/1999 to 27/3/2000, S&P 500 1/7/1985 to 31/8/1987 and 2/1/1991 to 4/9/1997, and the HSI 2/1/1992 to 6/1/1994. We suggest the absence of bubbles in the US$:DEM series 3/1/1983 to 8/3/1985 and in the HSI 3/1/1995 to 3/10/1997.

Chapter 4

Volatility and liquidity precursors

The aim of this chapter is to examine to what extent volatility and liquidity measures may help to predict financial crashes, and attempt to use these measures to draw meaningful economic conclusions. Volatility is a statistical measure of the fluctuations prevalent in a stock price, and is taken here to be equal to the empirical standard deviation of the log-returns. Liquidity is a rather general concept from economics. Market liquidity corresponds to the ability to quickly buy and sell assets, without causing a significant movement in price. In liquid markets cash flows freely, and options and other financial instruments are relatively easy to price. In contrast, in illiquid markets the flow of money is more constricted and single trades may move the price significantly. Illiquid markets might correspond to markets under strain and about to crash. In this chapter we consider various measures of liquidity.

The structure of this chapter is as follows. Section 1 is an introduction, giving a brief review of the literature and describing the various liquidity measures used. Section 2 is an exploratory data analysis. Of direct interest here is the analogy between stock market crashes and complex systems. Section 3 tests for the significance of precursors, using a novel regression application of the Sornette-Andersen model introduced in the previous chapter. Results are seen to offer an improvement over a nonlinear regression application of the SEG model in Chapter 2. Section 4 explores possible evidence of phase transitions in trading volume and liquidity. Section 5 is a conclusion.

4.1 Introduction

4.1.1 Literature review

The comment is made in *[64]* that some kind of volatility-based method may produce a more econometrically interpretable way of predicting crashes than the log-periodic method of Sornette and Johansen. With this aim at least partially in mind, [39] includes a graph of the standard deviation of the log-returns of the S&P 500 over a moving window of fifty trading days from January 1980 to October 1987. Somewhat disappointingly, the results seem to show little more than periodic variation. The comment is made in [102] that volatility is a purely statistical measure of market fluctuations and hence does not possess a meaningful economic interpretation. Moreover, given the apparent long-range dependence in volatility, [102] makes the point that volatility tends to lag behind price. Thus, using volatility measures, there might be a delay while information signifying that a crash may be imminent is incorporated. Hence volatility may not even be the right thing to look at. The relevant concept would then seem to be liquidity rather than volatility. This point seems to be echoed somewhat by [22]. Here, rather than discussing volatility, the comment is made (Chapter 7, page 126) that crashes are known to correspond to illiquid markets where the bid-ask spread increases and the trading volume becomes depleted.

4.1.2 Liquidity measures

Given the above comments by [22], there are two relatively straightforward measures of liquidity. In particular, we use the logarithm of the trading volume and the relative daily spread (RDS). The RDS is defined by

$$\ln(P_{max}) - \ln(P_{min}),$$

where P_{max} and P_{min} denote the day's maximum price, and the day's minimum price respectively. There are two further liquidity proxies that we introduce. [102] makes the comment that [1] and [63] measure market liquidity using the coefficient λ in the regression

$$P_t = P_{t-1} + \lambda Q_t + \epsilon_t, \tag{4.1}$$

where P_t denotes the price, Q_t the trading volume at time t and ϵ_t is a zero mean error term. Since the notion of a negative liquidity does not seem economically reasonable, (4.1) motivates the following *linear* liquidity proxy:

$$\lambda_t = \frac{|P_t - P_{t-1}|}{Q_t}. \tag{4.2}$$

Now (4.2) measures daily liquidity as the number of trades it takes to shift the price by one unit. High values should thus correspond to illiquid markets, as single trades thus cause a greater movement in the underlying price. However, [102] makes the comment that (4.1) and hence (4.2) are essentially measured in an arbitrary scale as price differences tend to be nominally larger for higher stock prices. What arguably makes more sense is a measure of relative change. As such, Tsuji replaces (4.1) by

$$|r_t| = \lambda \ln(Q_t) + \epsilon_t, \tag{4.3}$$

where r_t denotes the log-return at time t, $r_t = \log(P_{t+1}) - \log(P_t)$, and ϵ_t is again a zero-mean error term. (4.3) then suggests the following *logarithmic* measure of liquidity:

$$\lambda_t = \frac{|r_t|}{\ln(Q_t)}. \tag{4.4}$$

4.2 Exploratory Data Analysis

4.2.1 Volatility

Here we define volatility as the standard deviation of the log-returns calculated over a moving time window of 50 days. The results are shown in Figure 4.1. There is some suggestion that volatility might be increasing as the crash time approaches. This is quite clear cut for the Nasdaq, but is less obvious for the S&P 500 and the Hang Seng, where results do seem subject to quite considerable cyclic variations. Further, the scales on the Y-axes in Figure 4.1 suggest that the Nasdaq is a more volatile market than the S&P 500. This is perhaps to be expected since the Nasdaq consists of more 'new economy' stocks, which are usually thought to be more volatile.

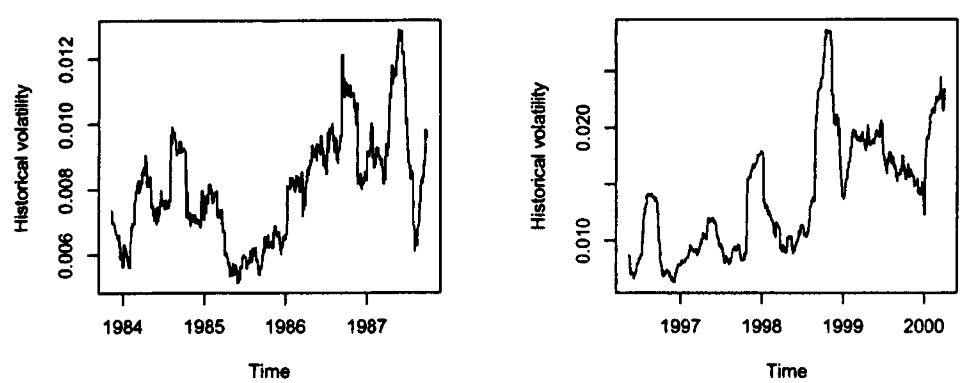


Figure 4.1: Historical volatilities for the S&P 500 (left panel) and for the Nasdaq (right panel).

4.2.2 Trading volume

Unfortunately, trading volume figures were available only for the S&P 500 and the Nasdaq. As one might expect, the trading volume decreases as the crash time approaches. The picture obtained is qualitatively similar for both the S&P 500 and the Nasdaq. It seems reasonable to try to fit a simple power-law model to the logarithm of the trading volume. However, some level of cyclic behaviour in the trading volumes is also apparent (see Figure 4.2).

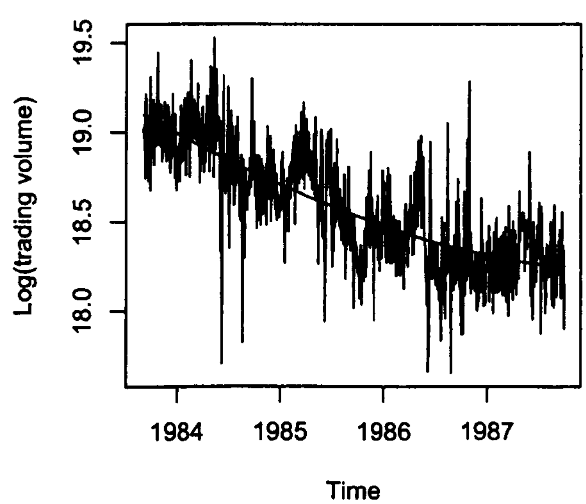


Figure 4.2: Plot of Log(Trading Volume) for S&P 500 and OLS fit by simple power-law model

4.2.3 Relative Daily Spread

The RDS values seem to be generally decreasing as the crash time approaches. This feature seems quite clear for the Nasdaq, though somewhat less clear for both the S&P500 and Hang Seng indices (see Figure 4.3).

4.2.4 Linear liquidity measure

In this section we examine the linear liquidity proxy given by (4.2). In the absence of available trading volume data for the Hang Seng index, we concentrate attention on the Nasdaq and S&P 500 indices. It seems rather less obvious to identify a smooth underlying trend for the S&P 500, although generally increasing values do suggest some enhanced illiquidity as the crash time approaches. However, when we look at the Nasdaq we have

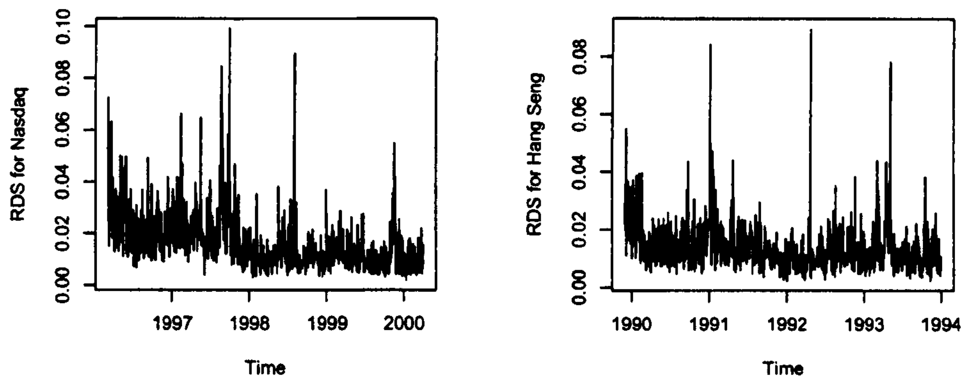


Figure 4.3: RDS from 4th March to 31st March 2000 for the Nasdaq and from 1st December 1989 to 31st December 1993 for the Hang Seng

a clear suggestion of approximate power-law behaviour. This illustrates a substantial increase in illiquidity prior to the crash, and the approximate power-law behaviour highlights the analogy between complex systems and financial markets (see Figure 4.4).

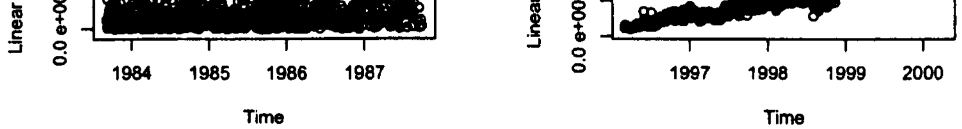


Figure 4.4: Plot of linear liquidity measures for S&P 500 and Nasdaq

4.2.5 Logarithmic liquidity measure

The results for the logarithmic liquidity measures seem a little harder to interpret than was the case for the linear liquidity measures (Figure 4.5). There seems to be some evidence of increased illiquidity on the Nasdaq as the crash time approaches, but it appears less easy to identify a smooth underlying trend. In contrast, it is not clear that these values are generally increasing for the S&P 500 index.

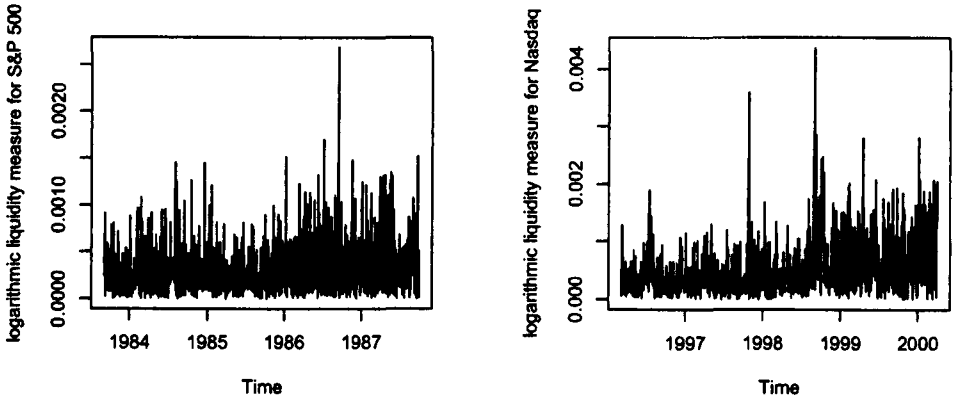


Figure 4.5: Plot of logarithmic liquidity measures for S&P 500 and Nasdaq

4.3 Significance of precursors

In the Gaussian formulation of the Sornette-Andersen model we have that $y(t) = p(t)^{1-m}e^{r(m-1)t}$ is Brownian motion with drift. We can use this result to produce a regression test to determine which factors are significant once we take into account super-exponential growth. This approach is seen to avoid problems with a nonlinear regression approach using the SEG model in Chapter 2, which is too sensitive to the effects of correlated random error terms.

Consider a series x_t. Let $y(t) = p(t)^{1-m}e^{r(m-1)t}$ and $\Delta y(t) = y_t - y_{t-1}$. Under the Sornette-Andersen model we have

$$y_t = A + \mu t + \sum_{i=1}^{t} \epsilon_i, \tag{4.5}$$

where A is a constant independent of time and the ϵ_i are an i.i.d. sequence from $N(0, \sigma^2)$. Suppose we build on (4.5) and consider the model

$$y_t = A + \mu t + \sum_{i=1}^{t} \epsilon_i + \gamma x_t. \tag{4.6}$$

The corresponding condition at time $t-1$ is

$$y_{t-1} = A + \mu(t-1) + \sum_{i=1}^{t-1} \epsilon_i + \gamma x_{t-1}. \tag{4.7}$$

Subtracting (4.7) from (4.6), we obtain

$$\Delta y(t) = \mu + \gamma \Delta x_t + \epsilon_t. \tag{4.8}$$

(4.8) suggests that we may test the hypothesis $\gamma = 0$ by fitting a simple OLS regression of Δy_t against Δx_t. We report results for each of the three markets before we present a brief comparison with a nonlinear regression formulation using the SEG model of Chapter 2. In addition to considering the instantaneous-effects model of (4.6), we also consider a lag-1 version of this model whereby variables influence the price one day in advance. The condition (4.6) becomes

$$y_t = A + \mu(t) + \sum_{i=2}^{t} \epsilon_i + \gamma x_{t-1},$$

and the regression equation (4.8) is replaced by

$$\Delta y(t) = \mu + \gamma \Delta x_{t-1} + \epsilon_t. \tag{4.9}$$

Extensions to other lagged regression are possible but seem less relevant.

4.3.1 S&P 500

Here we fit the models using data from September 4th 1984 to September 30th 1987. The results are shown in Table 4.1. In addition to the super-exponential growth (SEG) model shown in (4.10), regressors used are *volatility (vol.), log trading volume (LTV), relative daily spread (RDS), linear liquidity measure (linliq.),* and *logarithmic liquidity measure (logliq.).* Once we take into account super-exponential growth the only factors that remain significant are historical volatility and the logarithmic liquidity measure. Thus, we have some evidence to support the hypothesis that crashes occur on volatile, illiquid markets. Further, the logarithmic liquidity measure seems useful in generalising the purely linear measure of liquidity, which is found to be non-significant for both the contemporaneous regression (4.8) and the lag-1 regression (4.9). The results for the Nasdaq are shown in the next subsection.

Variable	Contemporaneous p-value	lag-1 p-value
Historical volatility	0.000***	0.009**
Log trading volume	0.134	0.936
Relative daily spread	0.128	0.867
Linear liquidity measure	0.361	0.189
Logarithmic liquidity measure	0.002**	0.051 (·)

Table 4.1: Results for the S&P 500

4.3.2 Nasdaq

Here we fit the models using data from 3rd March 1997 to 31st March 2000. The results for the Nasdaq index are shown in Table 4.2. We have some suggestion that crashes occur on volatile, illiquid markets, as volatility and the logarithmic and linear liquidity measures are significant under the instant-effects model (4.8) and volatility and logarithmic liquidity remain significant or borderline significant under the lag-1 model (4.9) . We also have some evidence that the nature of the illiquidity on the Nasdaq prior to the 2000 crash is different in nature to the illiquidity present on the S&P 500 prior to the 1987 crash, as different variables are significant once we account for super-exponential growth.

Variable	Contemporaneous p-value	lag-1 p-value
Historical volatility	0.000***	0.001**
Log trading volume	0.668	0.471
Relative daily spread	0.822	0.524
Linear liquidity measure	0.011*	0.573
Logarithmic liquidity measure	0.002***	0.079 (·)

Table 4.2: Results for the Nasdaq

4.3.3 Hang Seng

Here we fit the models using data from 4th December 1990 to 31st December 1993. The results for the Hang Seng index are shown in Table 4.3. Here, once we account for super-exponential growth and using the instant effects model (4.8), the historical volatility is found to be significant whilst the relative daily spread is found to be non-significant. In contrast, using the lag-1 model (4.9), neither term is found to be significant. This result leads us to conclude that the nature of the illiquidity present on the Hang Seng index prior to the 1994 crash is different in nature to that present on the Nasdaq prior to the crash in April 2000 and on the S&P 500 prior to the crash in 1987.

Variable	Contemporaneous p-value	lag-1 p-value
Historical volatility	0.000***	0.472
Relative daily spread	0.170	0.804

Table 4.3: Results for the Hang Seng

4.3.4 Comparison with a simple nonlinear regression approach

In Chapter 2 we introduced the SEG model, which stated that in the regime prior to a financial crash, the log-price y should satisfy

$$y_i = A + Bt_i + C\ln(t_c - t_i) + \epsilon_i, \tag{4.10}$$

where the ϵ_i is a zero mean error term. We now test for the significance of each of the precursors x by fitting the model

$$y_i = A + Bt_i + C\ln(t_c - t_i) + \gamma x_i + \epsilon_i, \tag{4.11}$$

estimating (4.10) and (4.11) using ordinary least squares. We might test for the significance of possible crash precursors by using approximate F-tests based on the extra sum-of-squares principle for nested models (see for example Chapter 3 in *[13]*). There is some similarity between this approach and the simple regression analysis contained in *[109]*. However, we note from Figure 2.6 in Chapter 2 that the residuals for this model are highly correlated. We obtain different results for this simple approach assuming ordinary least squares and our novel co-integration approach given by (4.8). We conclude that model violation, namely serially correlated residuals, mean that the results suggested by this simple approach are not valid. Tables 4.4-4.6 summarise the results. The results for the Hang Seng return the same conclusions as the previous section, but we note some differences for the S&P 500 and the Nasdaq. In particular, the simple ordinary least squares approach seems to overlook the impact of volatility and logarithmic liquidity for the S&P 500, and seems to exaggerate the significance of RDS for the Nasdaq.

Model	SEG + vol.	SEG + LTV	SEG + RDS	SEG + linliq.	SEG + logliq.
RSS	1.129	1.110	1.129	1.127	1.128
SEG RSS	1.129	1.129	1.129	1.129	1.129
Extra SS	0.001	0.020	0.001	0.002	0.001
RMS	0.001	0.001	0.001	0.001	0.001
F-value	0.363	13.742	0.391	1.656	0.794
P-value	0.547	0.000***	0.531	0.198	0.373

Table 4.4: SEG nested models comparisons for S&P 500

Model	SEG + vol.	SEG + LTV	SEG + RDS	SEG + linliq.	SEG + logliq.
RSS	4.005	4.458	4.432	4.217	4.250
SEG RSS	4.463	4.463	4.463	4.463	4.463
Extra SS	0.458	0.006	0.031	0.247	0.213
RMS	0.005	0.006	0.006	0.006	0.005
F-value	88.499	1.001	5.429	45.253	38.857
P-value	0.000***	0.317	0.020*	0.000***	0.000***

Table 4.5: SEG nested models comparisons for Nasdaq

Model	SEG + volatility	SEG + RDS
RSS	2.996	3.214
SEG RSS	3.219	3.219
Extra SS	0.222	0.004
RMS	0.004	0.004
F-value	55.046	1.014
P-value	0.000***	0.314

Table 4.6: SEG nested models comparisons for Hang Seng

4.4 Crash prediction using method of critical points

In this section we fit the simple power-law model,

$$y(t) = A + B(t_c - t)^{\alpha}, \tag{4.12}$$

by ordinary least squares to time series of log trading volume and the linear liquidity measure. Our aim is to see whether this admittedly simple-minded approximation seems able to predict the time of the crashes studied. Here we are trying to link the t_c parameter – the time when the market is deemed most susceptible to a crash – to meaningful economic variables. In so doing we are exploring, quantitatively, the analogy between market crashes and phase transitions in statistical mechanics (as discussed in earlier chapters).

4.4.1 Log trading volume

For ease of computation we fit (4.12) to the logarithm of the trading volume for the S&P 500. Here, we fit the model using simple ordinary least squares. The fit of the power-law seems reasonable, see Figure 4.2, and gives an R^2 value of 60.8%. However there are clear suggestions that there appear to be roughly cyclic fluctuations in the logarithm of the trading volume series that this simple model is unable to account for. Results are sufficient to suggest there may be at least some similarities between stock market crashes and phase transitions in complex systems although estimated standard errors, obtained using a bootstrap test based on 10,000 simulated values, are rather large. Here, since the the estimate of t_c is so close to the sampled dates, a near singularity occurs in the observed Fisher's Information matrix. Nonetheless, the bootstrap test leads to a p-value of 0.00 that B is non-zero. An F-test based on the extra sum of squares principle gives a p-value of 0.000, suggesting a significant improvement over the simple model $y(t) = A + \epsilon_t$. The predicted crash time corresponds to the beginning of October 1987, though the estimated standard error associated with this estimate is rather large. The regression results in Table 4.1, suggest that log trading volume is non-significant once we take into account super-exponential growth. However, the analogy between phase transitions and market crashes is brought out rather better by results for the Nasdaq index in the next subsection.

Parameter	Estimate (e.s.e)
A	18.248 (0.085)
B	$1.844 * 10^{-5}(1.287 * 10^{-5})$
t_c	1031.0 (437.302)
α	1.548 (0.446)

Table 4.7: Results for power-law model fitted to the logarithm of the trading volume for S&P 500

4.4.2 Linear liquidity measure

On the basis of Figure 4.4, we fit the simple power law model to the time series of linear liquidity measures for the Nasdaq index only. In line with the plot obtained, the fit of this model is good. This time the fit of the model is improved, estimated standard errors are much reduced and the analogy with phase transitions in statistical mechanics is brought out more clearly. Further, from Table 4.2 we see that the linear liquidity measure is also seen to remain significant once we take into account super exponential growth. The predicted crash time corresponds to early May 2000, with an estimated standard error of around 4 trading days. Results obtained are shown in Table 4.8. The negative α exponent indicates explosive growth, with the market becoming increasingly illiquid as the crash approaches.

Parameter	Estimate (e.s.e.)
A	$-4.610 * 10^{-5}(4.546 * 10^{-6})$
B	$6.469 * 10^{-5}(4.374 * 10^{-6})$
t_c	1058.237 (3.733)
α	-0.047 (0.004)

Table 4.8: Results for power-law model for linear liquidity measure for Nasdaq

4.5 Conclusions

From the statistical analysis in Section 4 we have at least some evidence to support the hypothesis that crashes occur on volatile, illiquid markets. This feature holds for all three markets examined even after we take into account super-exponential growth (SEG). The analysis in Section 4 also suggests differences in the precise form of illiquidity on these markets. This might be anticipated from the general context, although it is interesting that our novel co-integrated regression approach can nonetheless return this interpretation. This approach, based on the Sornette-Andersen model, is also seen to offer an improvement over a simple F-tests based on the extra sum of squares principle under the assumption of ordinary least squares.

In addition to a purely statistical analysis, we have also explored the analogy between stock markets and complex systems. In particular, the simple power-law model of Section 4 produces interesting results when applied to (a) the S&P 500 crash of 1987 and (b) the Nasdaq crash of 2000 when

(a) applied to the log-trading volume,

(b) applied to the linear liquidity measure introduced in Section 2.

The suggestion is that there is at least some evidence for phase transition behaviour in real economic variables prior to crashes and scope for an economically meaningful prediction mechanism, though the evidence in support of this is much stronger for the Nasdaq than for the S&P 500.

One of the key themes of *[102]* is that using liquidity measures, risk-management may switch focus from the purely statistical notion of volatility to measures with a more meaningful economic interpretation. Here, our statistical analysis reinforces the view that liquidity is indeed an important notion, with results suggesting interesting ways to incorporate the notion of liquidity into our modelling.

Based upon the results of this chapter, we may extend the conclusions of the previous chapter slightly. In order to predict a crash we anticipate super-exponential growth, accompanied by some additional signals of illiquidity. The various liquidity measures considered here suggest close analogies between market crashes and complex systems, with liquidity measures seen to remain significant even after taking into account super-exponential growth. We discuss prediction and detection methodologies using these results in Chapter 5.

Chapter 5

Simple models for bubbles: a synthesis

In this chapter, we attempt to synthesise Chapters 2-4 and provide a robust mechanism for detecting bubbles in financial markets. In particular, we use the previous chapter's hyperbolic re-formulation of the Sornette-Andersen model to test for bubbles. If the Sornette-Andersen model is found to be significant we can then use the SEG model of Chapter 2, and the empirical power-laws of Chapter 3 to predict the most likely time of a crash. In sum, our methodology follows the heuristic shown in Figure 5.1. See Section 4.

The outline of this chapter is as follows. In Section 1, based on original analysis in *[93]* Chapter 7, we discuss tests for super-exponential bubbles in foreign exchange (FX) markets. In Section 2 we discuss two historically observed log-periodic "false-alarms" as identified in *[93]*, Chapter 9. For comparison with the analysis in Section 2, we discuss a regime-switching regression model presented in *[103]*, and find some degree of agreement between interpretation of results in Section 2 and this regime-switching regression model. Finally, in Section 4 we discuss a case study to illustrate the approach suggested by Figure 5.1 in order to predict crashes. Section 5 is a brief conclusion.

5.1 SEG bubbles in FX markets

The use of log-periodic predictions is mentioned in *[93]*, Chapter 7, in relation to detecting speculative bubbles in the world's FX markets. *[93]* concludes that there is evidence of two speculative bubbles in the US$. The first one ending in March 1985 and the second one in August 1998, both after periods of strong growth in the preceding years. Closely following the methodology in *[93]* we look at two sets of data, one immediately preceding March 1985 and one immediately preceding August 1998. We take the price of the US$ as $y(t)$=US$ expressed in Deutchmarks (DEM) for the first period and $y(t)$=US$ expressed

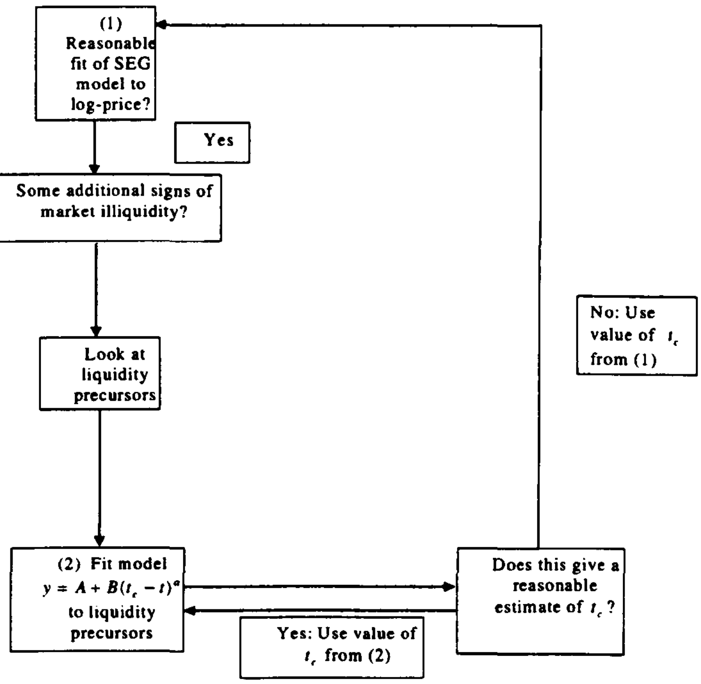


Figure 5.1: Suggested crash prediction methodology

in Canadian dollars (CAD) in the second period. Based on the results for the SEG model in Chapter 2 Section 4 we restrict attention to a single four-year period before the crash, stopping roughly a month prior to the observed crash. We analyse data from February 1st 1981 to February 28th 1985 and from July 1st 1994 to July 31st 1998. Results using the χ^2 test for SEG bubbles are shown in Table 5.1. We have evidence for a bubble prior to 1985, but we reject the presence of a bubble prior to 1998. The results for the fitted SEG model are shown in Table 5.2. The estimated crash time corresponds to mid-August 1985, with an estimated standard error of 17 days. A plot of the fit obtained is shown in Figure 5.2, and appears reasonable.

	Geometric Hyperbolic Model	Sornette-Andersen	χ	p-value
US$:DEM	4375.153	4385.677	21.048	0.000***
US$:CAD	6855.384	6857.438	4.108	0.128

Table 5.1: Likelihood ratio tests: Geometric Hyperbolic model vs. Hyperbolic Sornette-Andersen model

Parameter	Estimate (e.s.e)
A	2.051 (0.040)
r	0.000 (0.000)
C	-0.167 (0.005)
t_c	1656.534 (17.076)

Table 5.2: Results for SEG model on US$:DEM series

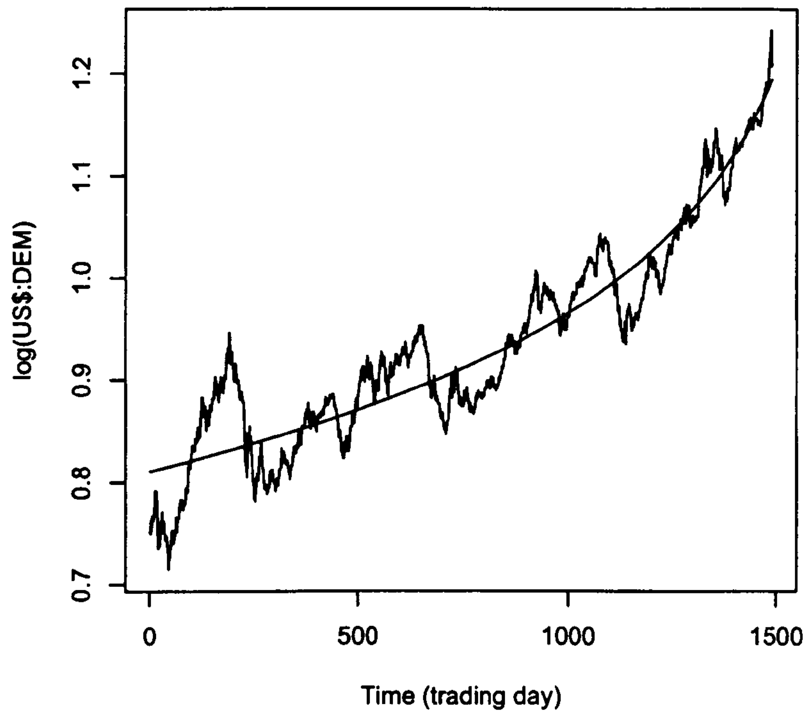


Figure 5.2: Plot of Log(US$:DEM) and best fit of SEG model

5.2 Log-periodic false predictions

Two log-periodic false predictions are listed in *[93]*. Firstly, after using data from the S&P 500 up to Friday, November 21st, 1997, *[93]*, Chapter 9 Page 342, describes a prediction of "a decrease in the price in approximately mid-December 1997". Similarly, later on in the same chapter and using data to the end of September 1999, a prediction is made of a bubble on the Nasdaq due to end in October 1999. *[93]* goes on to describe this as "an aborted event, which turned into a precursor of the large crash in April 2000". χ^2 tests are highly significant, indicating the presence of a bubble in both cases. See Table 5.3. Results for the SEG model of Chapter 2 are shown in Table 5.4. The predicted crash-time for the S&P 500 corresponds to mid-July 1999, roughly 9 months in advance of the April 2000 crash on the Nasdaq and some way after the initial prediction of December 1997. For the Nasdaq index, the predicted crash-time corresponds to early April 2000, very close to the actual timing of the crash. We might suggest that although some level of unstable price acceleration is present, the bubble has not yet reached its most explosive phase when a crash is imminent. We are able to obtain a similar picture by considering a simple regime-switching model in the next section. A plot of the fit obtained from the SEG model is shown in Figure 5.3 and appears reasonable.

	Geometric Hyperbolic Model	Sornette-Andersen	χ	p-value
S&P 500	-2954.055	-2911.215	85.68	0.000***
Nasdaq	-4580.372	-4552.742	55.26	0.000***

Table 5.3: Likelihood ratio tests: Geometric Hyperbolic model vs. Hyperbolic Sornette-Andersen model

Parameter	S&P 500	Nasdaq
A	11.152 (0.172)	8.934 (0.466)
r	0.000 (0.000)	$4.751 * 10^{-4}(6.498 * 10^{-5})$
C	-0.699 (0.022)	-0.286 (0.064)
t_c	1456.178 (25.082)	1182.665 (46.910)

Table 5.4: Results for SEG model on S&P 500 and Nasdaq

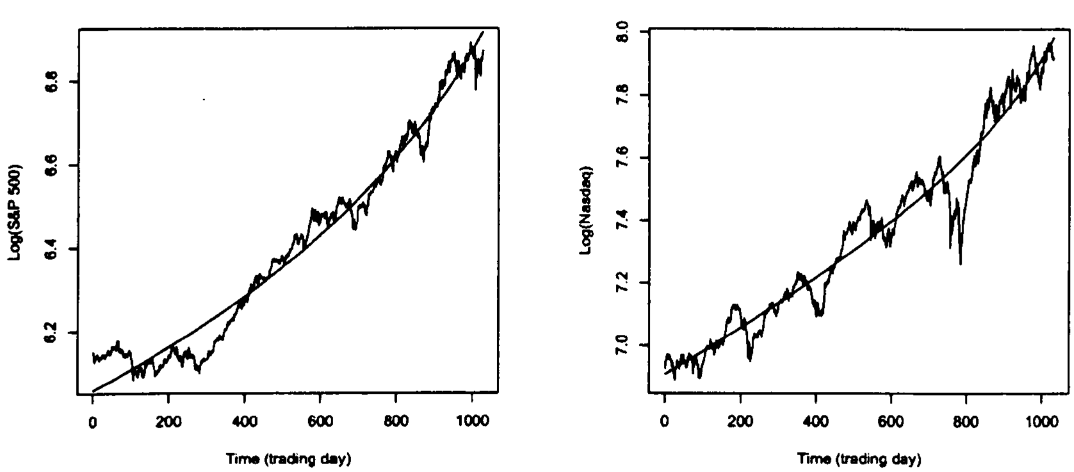


Figure 5.3: Fit of SEG model. Left panel: S&P 500. Right panel: Nasdaq.

5.3 Comparison with a regime-switching regression model

Regime-switching regression models are a valuable tool in providing a simple statistical approach to answer questions like "To what extent does the stock market appear to deviate from a fundamental price?", and "What is the probability of a stock market crash next month?" These models, as formulated in *[103]* and *[23]*, require a decomposition of the price of a dividend paying stock into a bubble component and a fundamental price component. Here, we take "fundamental value" to mean price levels that would appear reasonable given observed dividend payments D_t at time t. Thus, we have

$$B_t = P_t - P_t^*, \tag{5.1}$$

where B_t denotes the bubble price at time t, P_t denotes the observed price at time t, and P_t^* is the fundamental price at time t. We define the relative bubble size $b_t = B_t/P_t$ and the returns $R_{t+1} = (P_{t+1} + D_{t+1})/P_t$. The simplest way to construct fundamental values is to use $P_t^* = \rho D_t$, for some constant ρ. In empirical work ρ is estimated by the mean price-dividend ratio. This approach is considered by both *[103]* and *[23]*. Both papers also consider a more complicated vector autoregressive method, described in *[25]*, to achieve this fundamental value-bubble decomposition. However, both *[103]* and *[23]* report that results seem to be similar for the two methods.

Both *[103]* and *[23]* apply regime-switching regression models to financial markets, with modelling of stock market crashes in mind. Essentially the message of these papers is some suggestion of predictability in stock market returns. Here, we focus on the simpler van Norden-Schaller model *[103]*. We have two regimes, a speculative regime S, and a collapsing bubble regime C. The transition from the speculative bubble regime to the collapsing bubble regime corresponds to a crash, with expected returns higher in regime S. The model of Brooks and Katsaris *[23]* is more complicated as it includes a third 'dormant' regime D, within which the bubble is expected to grow at a constant rate rather than simply at an explosive rate. *[23]* also allow the probability of switching between these three regimes to depend on abnormal trading volume, rather than in *[103]* where the switching probabilities depend only on the absolute relative size of the bubble. Here, we refer to *[103]*, particularly the calculations given in the appendix, for details of the derivation both of the simple method used to construct fundamental prices and the original formulation of their switching regression model. For the sake of completeness we provide details of both derivations in Appendix B.

As a purely statistical formulation of the model, *[103]* use

$$R_{t+1}|S \quad = \quad \beta_{S0} + \beta_{Sb} b_t + \epsilon_{S,t+1},$$

$$
\begin{aligned}
R_{t+1}|C &= \beta_{C0} + \beta_{Cb}b_t + \epsilon_{C,t+1}, \\
q_{t+1} &= \Phi(\beta_{q0} + \beta_{qb}|b_t|),
\end{aligned} \tag{5.2}
$$

where $\Phi(\cdot)$ denotes the standard normal CDF. Given that we are in the speculative state S at time t, $q_{t+1}(\cdot)$ denotes the probability of remaining in S at time $t+1$. This formulation of the model ensures that the estimated q-probabilities lie in the range $[0,1]$. β_{qb} is constrained to be negative so that as b_t increases, the probability of a crash becomes more likely. Further, we see that in this simple model the probability of a market crash depends solely on the relative bubble size b_t. We have the additional constraint $\beta_{Sb} > \beta_{Cb}$, which is intuitive as this suggests higher average returns when in the speculative phase. The $\epsilon_{S,t}$ and $\epsilon_{C,t}$ terms are assumed to be independent sequences of normal white noise errors. Here, since the data analysed was low-frequency monthly data, there was very little suggestion of heavy-tailed non-Gaussian behaviour. Given the normality of the ϵ, the likelihood function can be calculated as

$$
\begin{aligned}
l(\Theta|\underline{X}) &= \prod_{i=1}^{n}\left[\Phi(\beta_{q0} + \beta_{qb}|b_t|)\phi\left(\frac{R_{t+1} - \beta_{S0} - \beta_{Sb}b_t}{\sigma_S}\right)\sigma_S^{-1}\right. \\
&+ \left.\Phi(-\beta_{q0} - \beta_{qb}|b_t|)\phi\left(\frac{R_{t+1} - \beta_{C0} - \beta_{Cb}b_t}{\sigma_C}\right)\sigma_C^{-1}\right],
\end{aligned} \tag{5.3}
$$

where ϕ denotes the density of a $N(0,1)$ random variable. One of the attractive features of the van Norden-Schaller model is that it enables one to construct one-step-ahead probabilities of stock market crashes or stock market rallies. For a given K we can calculate

$$
\begin{aligned}
Pr(R_{t+1} \leq K) &= \Phi\left(\frac{K - \beta_{S0} - \beta_{Sb}bt}{\sigma_S}\right)\Phi(\beta_{q0} + \beta_{qb}|b_t|) \\
&+ \Phi\left(\frac{K - \beta_{C0} - \beta_{Cb}bt}{\sigma_C}\right)\Phi(-\beta_{q0} - \beta_{qb}|b_t|).
\end{aligned} \tag{5.4}
$$

For $K = \mu_{R_t} \pm 2\sigma_{R_t}$ (5.4) is given the interpretation by van Norden and Schaller as the probability of a market rally or the probability of a market crash. The data analysed are from R. J. Shiller's webpage *http://www.econ.yale/~shiller/data.html* and constitute monthly values for the S&P 500 index and the associated dividend component from January 1980 to June 2007, standardised using the consumer price index. Maximum likelihood estimates, obtained using the Nelder-Mead method, are shown in Table 5.5, and seem broadly consistent with those given in Tables 1 and 2 of *[103]*. Estimated standard errors are also given and were calculated from the square root of the diagonal entries of the observed Fisher's information matrix. Here the observed Fisher's information matrix was calculated numerically using a finite difference method – the function *fdHess* in R.

Using (5.4) and $K = \mu_{R_t} - 2\sigma_{R_t}$ we can calculate the one-step ahead probabilities of a market crash as shown in Figure 5.4. According to this interpretation, the probability of a

Parameter	Estimate	Estimated Standard Error
β_{S0}	1.032	0.002
β_{Sb}	-0.023	0.004
β_{C0}	1.011	0.014
β_{Cb}	-0.027	0.014
β_{q0}	1.517	0.350
β_{qb}	-1.110	0.444
σ_S	0.026	0.002
σ_C	0.061	0.009

Table 5.5: Results for Regime-Switching Regression Model

market crash has increased from 1995 onwards as stock prices seem to have diverged from values which would appear reasonable given the dividend. These results fail to predict the crash of 1987, but peak sharply from 1995 until the crash in internet stocks in early 2000. The probability of a market crash seems to have generally decreased from initial highs in the early 2000's, despite an increase in 2003. However, the probability of a market crash remains higher than pre-1995 levels.

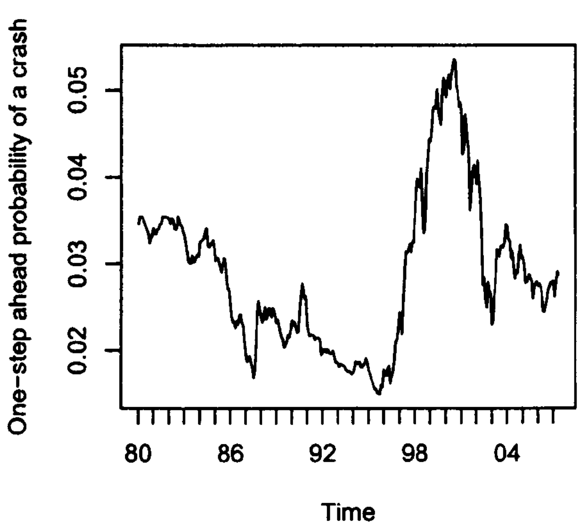


Figure 5.4: One-step-ahead probability of a market crash using monthly data for the S&P500.

5.3.1 Comparison with SEG models/log-periodicity

From Figure 5.4 the probability of a market crash is seen to increase from 1995 until 2000, around the time of the crash in internet stocks. Thus, under this interpretation, it is perhaps to be expected that we may find empirical evidence for SEG or log-periodic precursors from 1995 onwards as the probability of a market crash increases. However, it does appear that the inherent risks peak around April 2000, and not at the earlier times indicated in *[93]*. This coincides with the interpretation from the previous section that there is some suggestion of a bubble from around 1998, but that the probability of a market crash peaks in 2000.

5.4 Case study: How to predict crashes if you really must

In this section we give a rough guide on how one might try to predict crashes. For the suggested methodology we refer to Figure 5.1. The key advantage over the raw log-periodic methodology is that here we are trying to base predictions both on a formal statistical hypothesis test, and also on more interpretable econometric features, namely measures of liquidity closely associated with the trading volume.

In order to provide greater clarity and as an accompaniment to Figure 5.1, pseudocode for the proposed prediction methodology is shown below:

1. Test for bubble using the hyperbolic formulation of the Sornette-Andersen model. If non-significant stop.
2. If the test in 1. is significant, test for super exponential growth using the JLS power law model and (2.18). If non-significant stop.
3. If 1-2 significant search for empirical power laws in liquidity measures, particularly those linked to trading volume.
4. If the fit in 3. is significant, use the estimate and estimated standard error for t_c found in 3. Else, the optimal predictions are those found in step 2.

5.4.1 Case study

In this section we look at the Dow Jones Industrial Average index prior to the 1987 crash and the Nasdaq 100 index prior to the April 2000 crash on the Nasdaq. For both indices we find strong evidence for super-exponential growth (see Table 5.6). However, for the DJIA we do not find any power law type behaviour in any of the liquidity measures introduced in Chapter 3. Thus, we are constrained to a prediction of $t_c = 1215.295$ (towards the end of June 1988) based on the fit obtained by the SEG model (Table 5.7). A plot of the fit obtained is shown in Figure 5.5, and appears reasonable. For the Nasdaq 100 index, we also have some evidence of power law behaviour in the logarithm of trading volume. A power law fit, as shown by Figure 5.6, seems reasonable. This model has an R^2 value of 77.8% and an F-test based on the extra sum of squares principle gives a

p-value of 0.000, suggesting a significant improvement on the simple model $y(y) = A + \epsilon_t$. Further, the cointegrated regression tests of Chapter 4 give p-values of 0.089 and 0.957. This borderline significant p-value of 0.089 for the contemporaneous regression gives at least some suggestion that log trading volume remains significant even after we take into account super exponential growth. Further, the fit of the empirical power law to the log trading volume leads us to a prediction of $t_c = 1053.575$ corresponding to early May 2000. Tabulated results for the Nasdaq 100 index are shown in Table 5.8. The best fit of the SEG model seems reasonable (Figure 5.6). However, the SEG model leads to an estimate of $t_c = 1095.195$ corresponding to the end of June 2000, roughly two months behind the prediction of the empirical power law.

	Geometric Hyperbolic Model	Sornette-Andersen	χ	p-value
DJIA	-4097.661	-4089.944	15.434	0.000***
Nasdaq	-4787.48	-4757.429	60.102	0.000***

Table 5.6: Likelihood ratio tests: Geometric Hyperbolic model vs. Hyperbolic Sornette-Andersen model

Parameter	Estimate (e.s.e)
A	10.600 (0.102)
r	0.000 (0.000)
C	-0.507 (0.014)
t_c	1215.295 (12.940)

Table 5.7: Results for SEG model on DJIA

Model	SEG	Empirical (log trading volume)
A	9.738 (0.271)	23.067 (1.048)
r	$6.855 * 10^{-4}(5.184 * 10^{-5})$	·
C	-0.471 (0.038)	-0.981 (0.763)
t_c	1095.195 (10.403)	1053.575 (11.316)
α	·	0.160 (0.062)

Table 5.8: Results for power-law models for the Nasdaq 100 index

5.5 Conclusions

In this chapter we have used a formal statistical hypothesis test, based on the hyperbolic Sornette-Andersen model of Chapter 3, to refine the statistical analysis of cited instances of log-periodic precursors previously discussed in *[93]*. In Section 1, we found strong evidence of a bubble in the US$:DEM series, but found no evidence for a bubble in the US$:CAD series. In Section 2 we analysed two log-periodic false predictions in *[93]*. We found strong evidence for a bubble in each of these two cases, but found that the original

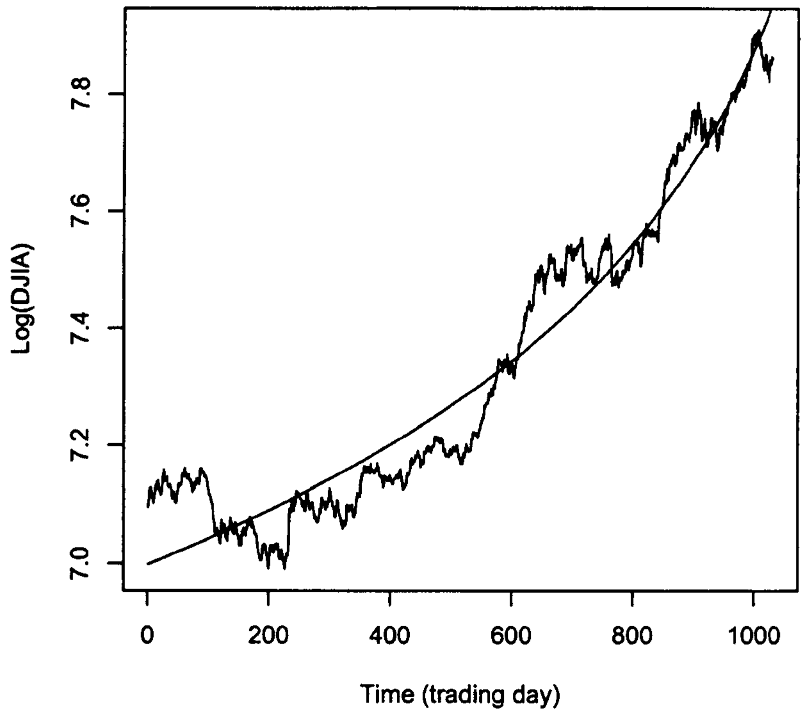


Figure 5.5: Fit of SEG model to DJIA

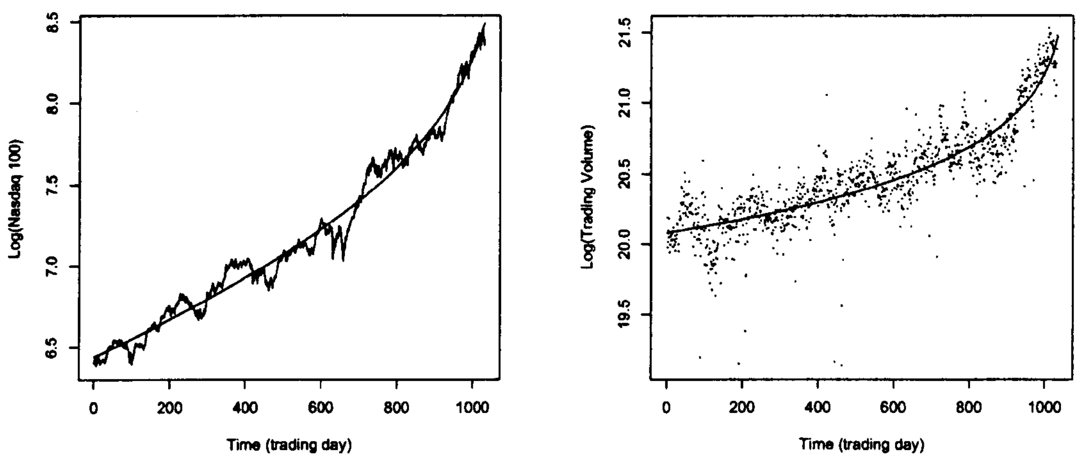


Figure 5.6: Left panel: Fit of SEG model to Nasdaq 100. Right panel: Fit of empirical power-law to log trading volume

crash-time estimates given in *[93]* are too early. In contrast, both our SEG model and the regime-switching regression model of *[103]* suggest that the level of risk appears to peak around the time of the actual crash on the Nasdaq – April 2000. In Section 4 we discussed a prediction methodology, and as a case study discussed the DJIA prior to Black Monday, October 19th 1987, and the Nasdaq 100 index prior to the crash in internet stocks in April 2000. χ^2 tests were highly significant, suggesting strong evidence of a bubble in both cases. The method led to predictions of June 1988 $\pm$ 26 days (Oct. 1987 crash) and early May 2000 $\pm$ 30 days (April 2000 crash), which seem reasonable. In addition, we found further evidence to support the analogy with phase transitions in complex systems and stock market crashes, with an empirical power-law providing a reasonable fit to a log-trading volume series.

Chapter 6

A universal power-law for drawdowns and models for external/internal origins of crises

In Section 1 we discuss drawdowns *[55]*. Expanding on the original calculations in the appendix of this paper, we are able to derive a generalised Pareto distribution for drawdowns. Thus we are able to show that contrary to the title of *[55]*, rather than being outliers large drawdowns in fact obey a universal power law. In Section 2 we introduce a model of exogenous and endogenous shocks in complex systems *[95]*. In Section 3 we discuss the multifractal random walk model *[79]*, *[8]*. Our particular interest is the approach taken by *[100]*, which attempts to relate volatility decay following a particular crises to an endogenous or exogenous cause. The hypothesis of *[100]* is that it should be possible to indicate an endogenous/exogenous root to a crisis by the exponent of the observed power-law decay of volatility. In Section 4 we provide an empirical investigation of this model. It is found that there is at least some evidence to support the hypothesis of *[100]*. However, it seems the observed power law exponents violate the predictions of this model. In Section 5 we show that the empirical results are more consistent with a fractional Gaussian noise model presented in *[95]*. Under this model not only can we derive exponents corresponding to an exogenous shock, but we are able to extend the original approach and derive the power law exponent corresponding to an endogenous shocks. Results for the fractional Gaussian noise model are seen to show a reasonable agreement with the empirical data. Finally, Section 6 is a brief conclusion.

6.1 Universal power law for drawdowns in an exponential Lévy market

Drawdowns are a measure of the percentage fall in the underlying price from one local maximum t_1 to the local minimum t_2, where t_2 is the lowest value the price takes before rising again. Drawdowns are of interest because they are more informative than simply recording the values of a price at fixed time scales, e.g. daily log-returns, and give an added sense of scale and cumulative market loss. Here we look at drawdowns calculated from daily prices. The concept of drawdowns has been advocated in *[55]* and *[53]* as an appropriate quantitative measure of large price drops. In *[56]* the authors have also advocated the concept of ϵ-drawdowns, where the cumulative loss is recorded until a price rise in excess of ϵ occurs, so that small rises in the underlying price cannot mask the true scale of the accumulated losses. This inevitably relies upon a subjective choice of ϵ. However, the authors in *[56]* seem to obtain reasonable results with the choice $\epsilon = \sigma/4$, for some estimate of the standard deviation σ. The results here suggest that the ϵ-drawdowns should obey a similar scaling property for reasonable choices of ϵ.

First we recall and expand upon the calculations in the Appendix of *[55]*. Using Laplace transforms we derive an asymptotic power law for drawdowns which is derived from a first-order approximation. The observed fit to drawdowns corresponding to the DJIA is impressive, and potential applications are briefly discussed.

6.1.1 Drawdowns in an exponential-Lévy market

Let $P^*(t)$ denote the price of a stock at time t. We assume that $P^*(t)$ is given by the exponential of a Lévy process. Here, we consider drawdowns calculated from time series of daily prices. A drawdown is a measure of the relative price drop from a local maximum at $P^*(t_m)$ to a local minimum at $P^*(t_{min})$, so that the price rises again immediately after time t_{min}. The drawdown $D_{t_m,t_{min}}$ is given by

$$D_{t_m,t_{min}} = \frac{(P^*(t_m) - P^*(t_{min}))}{P^*(t_m)} = 1 - \frac{P^*(t_{min})}{P^*(t_m)}.$$

Let $X(t) = \log(P^*(t))$ and let $d_{t_m,t_{min}} = X(t_m) - X(t_{min})$ be the corresponding difference in the log-price. By construction, we have that

$$d_{t_m,t_{min}} = \log\left(\frac{P^*(t_m)}{P^*(t_{min})}\right),$$

or equivalently, that

$$D_{t_m,t_{min}} = 1 - e^{-d_{t_m,t_{min}}}.$$

Next, we consider increments $x_1, \ldots, x_n$ of $X(t)$ corresponding to regularly spaced times $t_1, \ldots, t_{n+1}$ so that $x_1 = X(t_2) - X(t_1), \ldots, x_n = X(t_{n+1}) - X(t_n)$. In our empirical work we use drawdowns calculated from time series of daily returns, so that we imagine that the x_j correspond to daily increments or first-differences of the log-price recorded on consecutive days. By the assumption that $P^*(t)$ is given by the exponential of a Lévy process, the x_j constitute an i.i.d. sample from some common distribution F. We assume further that F has density $p(x)$. Starting with a local maximum $X(t_m)$, the probability density of $d_{t_m, t_{min}}$ is given by

$$f_d(w) = \frac{p_+}{p_-} \sum_{n=1}^{\infty} \int_0^{\infty} p(x_1) \ldots \int_{-\infty}^{0} p(x_n) \delta \left(-w - \sum_{j=1}^{n} x_j \right) dx_1 \ldots dx_n, \tag{6.1}$$

where

$$p_+ = 1 - p_- = \int_0^{\infty} p(x) dx$$

is the probability of observing a positive increment *[55]*, equation (9). The Dirac function $\delta \left(-w - \sum_{j=1}^{n} x_j \right)$ ensures that the summation in (6.1) is over all possible run lengths of length n which sum to w. The equation for the density in (6.1) can be written as

$$f_d(w) = \frac{p_+}{p_-} \sum_{n=1}^{\infty} p^{(*n)}(-w), \tag{6.2}$$

where $p^{(*n)}$ denotes the n-fold convolution of left tail of the probability density of the x_j. Related formulae, with a focus upon random walks and ladder heights, are discussed in *[42]*, Chapter 12. The $\frac{p_+}{p_-}$ is a normalization constant which results from the sum of a G.P. and ensures that the density in (6.2) is proper.

To be precise, we note that d is constructed by sampling initially from the left tail of $p(x)$. We continue to add samples from the left tail of $p(x)$ stopping when we sample from the right tail. Note that since d is constructed in this way, the probability density of d is always proper and this is contrast to the similar-in-spirit ladder height variables, *[42]* Chapter 12, where the mean μ of the incremental distribution of the random walk determines whether or not the distribution of ladder heights is proper[1]. To further illustrate the construction of d, we show how to construct the Laplace transform of (6.2) by summing over run lengths r. Since we have already assumed starting at an initial local maximum $X(t_m)$, a run length of one occurs with probability p_+, since p_+ is the probability of observing a positive increment. This run length distribution is clearly geometric, so that $p_-^{r-1} p_+$ gives

[1]The ascending and descending ladder heights are defined as the cumulative maxima and minima of a random walk starting at the origin. The descending ladder height distribution is only proper for $\mu \leq 0$, and likewise the ascending ladder height distribution is proper only for $\mu \geq 0$. Additional discussion is contained in *[42]* Chapter 12. Figure 2.1 in Chapter 7 in *[7]* shows pictorially how these ladder height variables are defined.

the probability of a run length of length r. Proceeding, we see that the Laplace transform of (6.2) is

$$\begin{aligned}
\hat{P}(s) &= \left(\frac{P(s)}{p_-}\right)p_+ + \left(\frac{P(s)}{p_-}\right)^2 p_-p_+ + \ldots + \left(\frac{P(s)}{p_-}\right)^r p_-^{r-1}p_+ + \ldots, \\
&= \left(\frac{P(s)}{p_-}\right)p_+ \sum_{r=0}^{\infty} P(s)^r, \\
&= \frac{p_+}{p_-}\frac{P(s)}{1-P(s)},
\end{aligned} \tag{6.3}$$

by the geometric sum formula and using

$$P(s) = \int_0^{\infty} p(-x)e^{-sx}dx.$$

In order to proceed, we note that (6.3) can be re-written as

$$\hat{P}(s) = \frac{1}{1 - \frac{1}{p_+}\frac{P(s)-P(0)}{P(s)}}, \tag{6.4}$$

with $P(0) = p_-$. In the sequel we assume that the moment generating function of $X(t)$ exists. This assumption is satisfied by both the generalized hyperbolic distribution *[33]* and the NIG distribution *[11]*, which provide two of the most tractable Lévy models in financial applications. This assumption also guarantees that the Laplace transforms considered are holomorphic functions of s and can thus be expanded in Taylor series of powers of s. Equivalently this means that the moment generating function exists in a neighbourhood of the origin.

6.1.2 A relevant model for market dynamics

In this subsection, we show how a simple first-order approximation suggests an exponential approximation for the density of $d_{t_m,t_{min}}$. We show that an exact exponential left tail of $p(x)$ leads to an exact exponential distribution for $d_{t_m,t_{min}}$. An analogous result for the descending ladder height distribution is given in *[42]*. A heuristic argument suggests that if the left tail of $p(x)$ is approximately exponential then the distribution of $d_{t_m,t_{min}}$ should also be exponential. Further, it is shown that an exponential distribution for $d_{t_m,t_{min}}$ leads to a generalised Pareto distribution for drawdowns on real markets. In the sequel, and for the sake of simplicity, we suppress the $_{t_m,t_{min}}$ subscript.

For small s, corresponding to d large, we can expand $\frac{P(0)}{P(s)}$ up to first order in a Taylor series:

$$\frac{P(0)}{P(s)} = 1 - \left(\frac{P'(0)}{P(0)}\right)s + o(s). \tag{6.5}$$

Plugging the expansion in (6.5) into the equation for the Laplace transform in (6.4) leaves us with

$$\hat{P}(s) \approx \frac{1}{1-\left(\frac{P'(0)}{P(0)}\right)\frac{s}{p_+}},$$

which reduces to

$$\hat{P}(s) \approx \frac{1}{1+\left(\frac{-<x>_-}{p_-p_+}\right)s}, \tag{6.6}$$

with

$$< x >_- = -\int_0^{-\infty} xp(-x)dx.$$

Finally, we note that (6.6) is the Laplace transform of an exponential density with mean $\mu = -\frac{<x>_-}{p_-p_+}$. Suppressing subscripts, the implication is that

$$d \sim \exp(\lambda), \tag{6.7}$$

where $\exp(\lambda)$ denotes the usual exponential density on $[0, \infty)$, with $\lambda = \frac{-p_-p_+}{<x>_-}$. (6.7) was originally stated in *[55]*, albeit subject to a minor mistake in the algebra. However, in *[55]* (6.7) was assumed to hold for the relative price drop D rather than the drop d in log-price. For the moment, the main point of interest is that it appears reasonable to consider the case where the distribution of d is approximately exponential.

Suppose the exponential approximation (6.6) holds exactly. Equating Laplace transforms gives

$$\frac{p_+}{p_-}\frac{P(s)}{1-P(s)} = \frac{1}{1+s\mu},$$

leading to the solution

$$P(s) = \frac{p_-}{1+p_+s\mu} = \frac{p_-}{1-\frac{s<x>_-}{p_-}}, \tag{6.8}$$

where we have defined $\mu = \frac{-<x>_-}{p_-p_+}$. We can see that (6.8) corresponds to an exponential left tail:

$$p(x) = \frac{-p_-^2}{< x >_-}e^{\frac{-p_-x}{<x>_-}}, \quad (x < 0). \tag{6.9}$$

We note that (6.9) is satisfied by the Laplace distribution with non-negative mean μ and variance $2b^2$:

$$f(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}. \tag{6.10}$$

Since we are considering a candidate model for financial log-returns, the condition $\mu \geq 0$ is reasonable and corresponds to generic exponential growth behaviour. In the notation of (6.6), we have that

$$< x >_- = \frac{-e^{-\frac{\mu}{b}}}{2b} \int_0^\infty y e^{\frac{-y}{b}} dy = \frac{-be^{-\frac{\mu}{b}}}{2}. \tag{6.11}$$

Similarly, it follows that

$$p_- = \frac{e^{-\frac{\mu}{b}}}{2b} \int_{-\infty}^0 e^{\frac{x}{b}} dx = \frac{e^{-\frac{\mu}{b}}}{2}, \tag{6.12}$$

with $p_+ = 1 - p_- = 1 - \frac{e^{-\frac{\mu}{b}}}{2}$. Inserting (6.11-6.12) into (6.9), it follows that (6.9) holds for $x \leq 0$ in (6.10) since

$$\frac{-p_-^2}{< x >_-} = \left(\frac{e^{-\frac{2\mu}{b}}}{4}\right)\left(\frac{2e^{\frac{\mu}{b}}}{b}\right) = \frac{e^{-\frac{\mu}{b}}}{2b},$$

and

$$\frac{-p_-}{< x >_-} = \left(\frac{e^{-\frac{\mu}{b}}}{2}\right)\left(\frac{2}{be^{-\frac{\mu}{b}}}\right) = \frac{1}{b}.$$

A related result given in *[42]*, page 387, is that the distribution of the first weak descending ladder height is proportional to $e^{\beta x}$, $x < 0$, if the incremental distribution F has an exponential left tail $F(x) = qe^{\beta x}$ for $x < 0$. If we suppose that empirical distributions for log-returns approximately satisfy (6.9), we suggest

$$p(x)\Big|_{x \leq 0} \approx \frac{1}{< x >_-} e^{\frac{p_- x}{< x >_-}} \Rightarrow \hat{P}(s) \approx \frac{1}{1 + s\mu}, \tag{6.13}$$

in the sense that if $P(s) - \frac{p_-}{1+p_+ s\mu} = o(s)$ then $\hat{P}(s) - \frac{1}{1+s\mu} = o(s)$, since

$$\begin{aligned}
\frac{p_+}{p_-}\frac{P(s)}{1-P(s)} - \frac{1}{1+s\mu} &= \frac{P(s)[1 + p_+ s\mu] - p_-}{(1-P(s))(1+s\mu)}, \\
&= \frac{(1 + p_+ s\mu)\left[P(s) - \frac{p_-}{1+p_+ s\mu}\right]}{(1-P(s))(1+s\mu)}, \\
&= a(s)b(s),
\end{aligned} \tag{6.14}$$

where $a(s) = \frac{(1+s\mu p_+)}{(1-P(s))(1+s\mu)}$ and $b(s) = \left[P(s) - \frac{p_-}{1+p_+s\mu}\right]$. It follows that

$$\begin{aligned}\lim_{s\to 0}\frac{1}{s}\left[\frac{p_+}{p_-}\frac{P(s)}{1-P(s)} - \frac{1}{1+s\mu}\right] &= \lim_{s\to 0} a(s) \lim_{s\to 0} b(s)/s, \\ &= 0,\end{aligned}$$

since $\lim_{s\to 0} = 1/p_+$ and $b(s) = o(s)$ by assumption. In sum, an exact exponential left tail (6.9) leads to an exact exponential distribution for d (6.6-6.7). We suggest that, for real markets, (6.6-6.7) should remain approximately true if (6.9) remains a good approximation.

In the sequel, we assume that the result in (6.7) holds exactly. In this case, it follows that the distribution of $d_{t_m,t_{min}}$ is $\exp(\lambda)$, where $\exp(\lambda)$ is the usual exponential distribution on $[0,\infty)$. In *[55]*, the authors supposed that it was the difference in the raw prices rather than in the log-prices that satisfied (6.6). Since this corresponded to a very simple model for prices, the authors went beyond this theoretically derived result to consider a more general Weibull or stretched-exponential model for drawdowns. Johansen and Sornette then classified drawdowns that violated this Weibull model as 'outliers'. However, in the exponential-Lévy market considered here, the result in (6.7) is seen to lead naturally to a generalised Pareto distribution for drawdowns, suggesting a power-law tail probability for drawdowns on the original (non-logarithmic) scale.

By construction, the drawdown $D_{t_m,t_{min}}$ is given by

$$D_{t_m,t_{min}} = 1 - e^{-d_{t,t_{min}}}. \tag{6.15}$$

From (6.15) the distribution of $D_{t_m,t_{min}}$ follows a generalized Pareto distribution, with parameters $\zeta = -1/\lambda$ and $\beta = 1/\lambda$ (*[77]*, Chapter 7), since if X is $\exp(\lambda)$ then it follows that the distribution function of $Y = 1 - e^{-X}$ is given by

$$\begin{aligned}F_Y(x) &= Pr(Y \leq x), \\ &= Pr(X \leq -\ln(1-x)), \\ &= 1-(1-x)^{\lambda},\end{aligned} \tag{6.16}$$

using $F_X(x) = 1 - e^{-\lambda x}$. The log-likelihood function can readily be calculated from (6.16) as

$$l(\lambda|\mathbf{x}) = n\log(\lambda) + (\lambda-1)\sum_{i=1}^{n}\log(1-x_i),$$

leading to the maximum likelihood estimate

$$\hat{\lambda} = \frac{-n}{\sum_{i=1}^{n} \log(1 - x_i)}. \tag{6.17}$$

As an application we looked at the sequence of drawdowns from the DJIA using daily data covering the period October 1st 1928 to 22nd May 2006. Results are shown in Figure 6.1, and show an extremely close correspondence between the empirical CDF and the theoretical CDF predicted by (6.16). Finally, we mention possible applications for

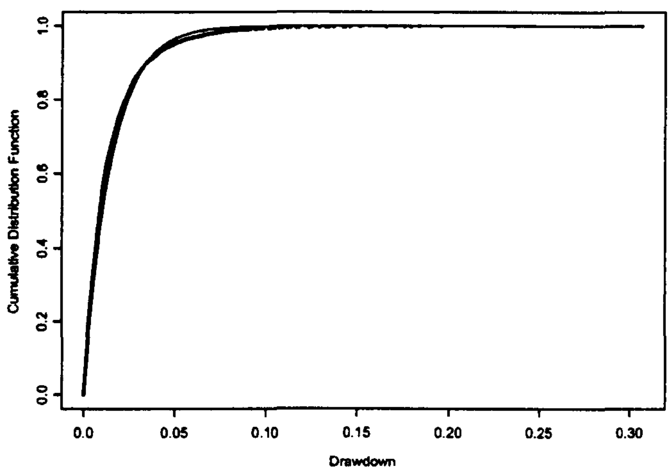


Figure 6.1: Plot of fitted GPD ($\hat{\lambda} = 66.13597$) for drawdowns on the DJIA and empirical CDF (dots)

options pricing, risk management and capital allocation for banks. In terms of generic risk management and options pricing, we may calibrate a Lévy model for daily log-returns to historically observed sequences of drawdowns by equating the maximum likelihood estimate in (6.17) to the value of λ by ensuring

$$\frac{1}{\lambda} = \frac{- < x >_-}{p_- p_+}, \tag{6.18}$$

where $< x >_-$ denotes the mean of the negative log-returns and p_+ and p_- denote the probabilities of observing positive and negative log-returns. Another interesting possibility deals with setting capital levels for banks. For some $0 < \alpha << 1$, the problem is to find a capital amount C so that the losses L relating to a stock market portfolio satisfy $Pr(L > C) = \alpha$. Suppose we have holdings $\phi_1, \ldots, \phi_n$ in a portfolio of stocks $X_1, \ldots, X_n$, for some large n. Let $Y_i(t) = \sum_{i=1}^{n} \phi_i X_i(t)$ denote a time series of historical values of the portfolio. If we assume that $Y_i(t)$ corresponds to an exponential-Lévy market (at least approximately), we can use (6.18) to estimate λ. Let I denote the level of investment in the stock market, C the level of capital reserves and LP denote a percentage loss corresponding to a drawdown sequence. For a given non-exceedance probability α, set

$$1-\alpha = 1-(1-LP)^{\lambda}.$$

The capital levels C and investment levels I can then be chosen to satisfy

$$C \geq I.LP.$$

so that with probability $1-\alpha$ the capital levels C are not exceeded by the aggregate losses. Risk management studies, as described for example in the guidelines of the Basel Commitee of Banking Supervision, have previously advocated calculating capital levels from analysis of daily risk levels. However, results here suggest that such an analysis should be possible - and may even be more natural - for runs of consecutive losses. For background on risk management issues and mathematical modelling we refer to *[77]*. Finally, methods in this section can easily be extended to incorporate Sornette and Johansen's concept of ϵ-drawdowns by setting $p_+ = Pr(x_j > \epsilon)$ and $p_- = Pr(x_j < \epsilon)$.

6.2 Sornette-Helmsetter method for complex systems

We consider a simple model of the activity level $A(t)$ of a system at time t, viewed as the noisy response to all past perturbations:

$$A(t) = \int_{-\infty}^{t} \eta(t)K(t-\tau)d\tau, \tag{6.19}$$

where $\eta(t)$ denotes standardized Gaussian white noise, and $K(\cdot)$ is referred to variously as the memory kernel, propagator, Green function, or response function *[95]*. Further, we assume that $K(t)$ is a causal function ensuring that the system is not anticipative[2]. One of the most interesting choices of $K(\cdot)$ described by Sornette and Helmsetter in *[95]* is fractional Gaussian noise which corresponds to the kernel

$$K(t-\tau) = \frac{1}{(l+t-\tau)^{3/2-H}} = \frac{1}{(l+t-\tau)^{1-\theta}}, \tag{6.20}$$

where $H = 1/2+\theta$ and l is a small constant whose inclusion avoids blow-up at the origin. In (6.20) $0 < \theta < 1/2$ corresponds to persistence where successive variations are positively correlated, $-1/2 < \theta < 0$ antipersistence, where positive variations are preferentially followed by negative ones (*[105]*). The comment is made in *[95]* that it is the case $0 < \theta < 1/2$ that seems to be of practical interest, and further that (6.20) appears able to

[2]A function $g(t)$ is called a causal function if is zero when $t < 0$. This constraint arises naturally in many physical systems, due to the fact that observed effects cannot precede cause and the restrictions that this then places on functions describing the system.

explain some of the phenomenology associated with Omori's power-law for earthquakes[3].

6.2.1 A simple model for exogenous shocks

Suppose an external shock occurs at $t = 0$, which we model as $A_0\delta(\tau)$, where $\delta(\cdot)$ denotes Dirac's delta function. From (6.19) we see that the response of the system becomes

$$A(t) = A_0 K(t) + \int_{-\infty}^{t} \eta(t)K(t-\tau)d\tau,$$

leading to the expected response

$$E(A(t)) = A_0 K(t).$$

In the case of a fractional Gaussian noise kernel, we obtain

$$E(A(t)) = A_0 t^{\theta-1}, \qquad t >> l. \tag{6.21}$$

6.2.2 A simple model for endogenous shocks

Suppose there is a large "internal" shock A_0 at time 0. From the integral representation (6.19), we see that $\{A(t) : t\in\mathbb{R}\}$ is a Gaussian process. Thus for a given $t > 0$, $(A(t), A(0))$ is a bivariate normal pair. For an arbitrary bivariate normal pair X, Y we know that

$$Y|X{\sim}N\left(\mu_Y + \frac{Cov(X,Y)(X-\mu_x)}{\sigma_X^2}, \sigma_Y^2 - \frac{(Cov(X,Y))^2}{\sigma_X^2}\right), \tag{6.22}$$

(*[46]*, Chapter 4). Also, from the integral representation (6.19) we see that

$$\mathrm{Cov}(A(t), A(0)) = \int_{-\infty}^{0} K(t-\tau)K(-\tau)d\tau. \tag{6.23}$$

Using the substitution $v = -\tau$, it follows that the expected value of the system at time t obeys

$$E(A(t)) \quad \propto \quad A_0 \int_0^{\infty} K(t+v)K(v)dv. \tag{6.24}$$

Consider the fractional Gaussian noise kernel (6.20) and take the Laplace transform of (6.24), where the Laplace transform is defined as

$$L(f(t)) = \int_0^{\infty} f(t)e^{-st}dt.$$

[3] $n(t) = \frac{K}{(c+t)^p}$, where n denotes the number of aftershocks, K is the magnitude of the initial quake, c is a time offset and $p\in(0.7, 1.5)$.

We note that the RHS of (6.24) is the convolution of $K(t)$ with itself. We note that we have we have $L(t^{\theta-1} = \Gamma(\theta)s^{-\theta}$, *[61]*, Chapter 5, where $\Gamma(\cdot)$ denotes the gamma function given by

$$\Gamma(\theta) = \int_0^\infty x^{\theta-1}e^{-x}dx. \tag{6.25}$$

Since $K(t){\sim}t^{\theta-1}$, $t >> l$, it follows that $L(K(s)){\sim}\text{Const}.s^{-\theta}$. Hence, the Laplace transform of (6.24) is of the form $\text{Const}.s^{-2\theta}$, suggesting

$$E(A(t)) \quad \propto \quad t^{2\theta-1}, \qquad t >> l. \tag{6.26}$$

The predictions (6.21) and (6.26) are intended to represent generic behaviour of complex systems, with "endogenous" shocks exhibiting a significantly slower power law decay than "exogenous" shocks, for $\theta>0$.

6.3 The multifractal random walk (MRW) model

The aim here is to demonstrate how and under what circumstances approximate power-law decay of volatility is predicted by *[100]*. The discussion in this section is intended only as a very brief guide. For further details see for example *[79]* and *[8]*. The multifractal random walk model is the continuous time limit of a stochastic volatility model with the log-volatility possessing logarithmically decaying correlations. In the version of this model considered by *[100]*, for each time scale $\Delta t{\leq}t$, the returns at scale Δt, $r_{\Delta t}(t)$, can be described as a stochastic volatility model

$$r_{\Delta t}(t) = \epsilon(t)\cdot\sigma_{\Delta t}(t) = \epsilon(t)e^{\omega_{\Delta t}(t)}, \tag{6.27}$$

where $\epsilon(t)$ is a standardized Gaussian white noise independent of $\omega_{\Delta t}$, and $\omega_{\Delta t}$ is a Gaussian process with mean and covariance:

$$\mu_{\Delta t} = 1/2\ln(\sigma^2\Delta t) - C_{\Delta t}(0),$$

$$C_{\Delta t} := \text{Cov}(\omega_{\Delta t}(t), \omega_{\Delta t}(t+\tau)) = \lambda^2 \ln\left(\frac{T}{|\tau|+l}\right), \tag{6.28}$$

where $\sigma^2\Delta t$ is the return variance at scale Δt, l is a small constant to avoid blow-up at the origin and T represents the time scale over which volatility is correlated. *[22]*, Chapter 7 Section 3, state that the MRW model is attractive as it suggests both log-normal volatility fluctuations and logarithmic decay in the correlation function of the log-volatility, both of which can be verified empirically. According to *[100]* the MRW model can be written in a more accessible form in which the log-volatility $\omega_{\Delta t}(t)$ obeys an autoregressive equation

$$\omega_{\Delta t}(t) = \int_{-\infty}^{t} \eta(\tau) K_{\Delta t}(t-\tau) d\tau, \tag{6.29}$$

where $\eta(t)$ denotes a standardized Gaussian white noise and the memory kernel $K_{\Delta t}(\cdot)$ is a causal function ensuring that the system is not anticipative. The process $\eta(t)$ can be interpreted as the information flow, with $\omega(t)$ representing the response of the market to the incoming information up to time t. Rather than multifractality, it is the kernel representation shown in (6.29), see also (6.19), which is the key mathematical feature of interest here. At time t the distribution of $\omega_{\Delta t}(t)$ is Gaussian with mean $\mu_{\Delta t}$ and variance $V_{\Delta t} = \int_0^{\infty} K_{\Delta t}^2(\tau) d\tau = \lambda^2 log(\frac{T}{l})$, consistent with (6.28). The covariance – which in conjunction with the mean completely specifies the random process – is given by

$$C_{\Delta t}(\tau) \quad = \quad \int_0^{\infty} K_{\Delta t}(t) K_{\Delta t}(t + |\tau|) \quad dt. \tag{6.30}$$

Performing a Fourier analysis we obtain

$$\begin{aligned}
\hat{C}_{\Delta t}(f) &= 2\int_0^{\infty}\int_0^{\infty} K_{\Delta t}(\tau) K_{\Delta t}(\tau+t) e^{-ift} \quad d\tau dt, \\
&= 2\int_0^{\infty}\int_0^{\infty} K_{\Delta t}(\tau) K_{\Delta t}(v) e^{-ifv} e^{if\tau} \quad dv d\tau, \\
&= 2 \mid \hat{K}_{\Delta t}(f) \mid^2,
\end{aligned} \tag{6.31}$$

where $\hat{K}$ denotes the Fourier transform of K. Using (6.28), we thus see that

$$\begin{aligned}
\hat{C}_{\Delta t}(f) &= 2\lambda^2 \int_0^{T} \ln\left(\frac{T}{t+l}\right) \cos(ft) \quad dt, \\
&= \frac{2\lambda^2}{f}\left[\left[\ln\left(\frac{T}{t+l}\right)\sin(ft)\right]_0^T + \int_0^T \frac{\sin(ft)}{t+l} \quad dt\right],
\end{aligned} \tag{6.32}$$

where the second line follows from integration by parts. The first term in (6.32) is approximately zero – it corresponds to a logarithmically decaying covariance function which tends to zero as it approaches T – and is equal to

$$\ln\left(1 - \frac{l}{T+l}\right)\sin(fT). \tag{6.33}$$

Turning to the second term and using the change of variable $v = tf$, we can see that

$$\int_0^T \frac{\sin(ft)}{t+l} dt = \int_0^{Tf} \frac{\sin(v)}{v+fl} dv. \tag{6.34}$$

Further, we have that

$$\left|\int_0^{Tf}\left(\frac{\sin(v)}{v}-\frac{\sin(v)}{v+fl}\right)dv\right|\leq fl\left|\int_0^{Tf}\frac{\sin(v)}{v(v+fl)}dv\right|. \tag{6.35}$$

From (6.35) we also have

$$\begin{aligned}\left|\int_0^{Tf}\frac{\sin(v)}{v(v+fl)}dv\right| &\leq \int_0^{Tf}\frac{\mid\sin(v)\mid}{v(v+fl)}dv,\\ &\leq \int_0^{Tf}\frac{1}{v+fl},\\ &= \ln(T/l+1).\\ &= \mathcal{O}(\ln(1/fl)),\end{aligned} \tag{6.36}$$

under the assumption that $1/fl > T+fl$. (6.36) follows from the fact that $\left|\frac{\sin(x)}{x}\right|\leq 1$ for all x. From the above (6.32) reduces to

$$\hat{C}_{\Delta t}(f)=\frac{2\lambda^2}{f}\left[\int_0^{Tf}\frac{\sin(v)}{v}dv+\mathcal{O}\left(fl\ln(1/fl)\right)\right]. \tag{6.37}$$

For large Tf, the integral in (6.37) approximates $\int_0^\infty\frac{\sin(x)}{x}dx=\frac{\pi}{2}$, and then (6.37) is approximately given by

$$\hat{C}_{\Delta t}(f)\approx\frac{\lambda^2\pi}{f}. \tag{6.38}$$

Assuming that (6.38) holds with equality, we have that

$$\mid\hat{K}_{\Delta t}(f)\mid=\lambda\sqrt{\frac{\pi}{2}}f^{-1/2}. \tag{6.39}$$

By (6.28) and (6.30) we see that $K_{\Delta t}(\cdot)$ must be a decreasing function of t, and it then follows from (6.39) that

$$K_{\Delta t}(\tau)\sim K_0\sqrt{\frac{1}{\tau}},\qquad \tau>>l, \tag{6.40}$$

for $K_0=\frac{\lambda}{\sqrt{2}}$ since

$$\hat{t}^{-1/2}=\int_0^\infty t^{-1/2}e^{-ift}dt=\sqrt{\frac{\pi}{2}}(f^{-1/2}-if^{-1/2}),$$

and $\mid\hat{t}^{-1/2}\mid=\sqrt{\pi}f^{-1/2}$. According to *[100]*, this slow power law decay of the memory memory kernel in (6.40) ensures the long-range dependence and multifractality of the MRW.

6.3.1 Linear response to an external shock

The transfer function formalism introduced in (6.29) gives us an interesting insight here. Assume that a new piece of information $\eta(t) = \omega_0\delta(t)$ impacts on the market (taken without loss of generality to be $t = 0$ since the system is stationary). Using the transfer function formalism, the response to the shock is

$$\begin{aligned}\omega(t) &= \mu + \int_{-\infty}^{t} (\omega_0\delta(\tau) + \eta(\tau))K(t-\tau)d\tau \\ &= \mu + \omega_0 K(t) + \int_{-\infty}^{t} \eta(\tau))K(t-\tau)d\tau,\end{aligned}$$

where for convenience the reference to the scale Δt has been omitted. Given the causal representation of the MRW, the expected volatility conditional on this new incoming information is given by

$$\begin{aligned}E[\sigma^2(t)|\omega(t)] &= e^{2\mu+2\omega_0 K(t)}E[e^{2\int_{-\infty}^{t}\eta(\tau)K(t-\tau)d\tau}] \\ &= \overline{\sigma^2(t)}e^{2\omega_0 K(t)} \\ &\simeq \overline{\sigma^2(t)}e^{2\omega_0 K_0\sqrt{\frac{1}{t}}},\end{aligned}$$

for $\Delta t << t << T$, where $\overline{\sigma^2(t)} = e^{2\mu}E[e^{2\int_{-\infty}^{t}\eta(\tau)K(t-\tau)d\tau}]$ is the unconditional average of the stochastic process $\sigma^2(t)$. Following the shock, the volatility relaxes back down to its unconditional average $\overline{\sigma^2(t)}$. This means that under a linear approximation we are left with behaviour of the form

$$\sigma^2(t) \sim \overline{\text{Const.}} t^{-\alpha}, \tag{6.41}$$

with $\alpha = 0.5$. The prediction made by (6.41) can then be used as the focus of empirical work.

6.3.2 "Conditional response" to an endogeneous shock

We consider the evolution of the system, which despite the absence of a large external shock nonetheless exhibits a large volatility burst ω_0 at time 0. Now $\omega(t)$ is a Gaussian process, and using the fact that the moment generating function of a $N(\mu, \sigma^2)$ is $e^{\mu t+\sigma^2 t^2/2}$ and that $2\omega(t) \sim N(2\mu(t), 4\sigma^2(t))$ we see that

$$\begin{aligned}E[\sigma^2(t)|\omega_0] &= E[e^{2\omega(t)}|\omega_0] \\ &= \exp(2E[\omega(t)|\omega_0] + 2Var[\omega(t)|\omega_0]).\end{aligned}$$

Further, since $\omega(t)$ is a Gaussian process we can use the relation (6.22) to see that

$$
\begin{aligned}
E[\omega(t)|\omega_0] &= E[\omega(t)] + \frac{Cov[\omega(t), \omega_0]}{Var[\omega_0]} \cdot (\omega_0 - E[\omega_0]) \\
&= (\omega_0 - \mu) \cdot \frac{C(t)}{C(0)} + \mu,
\end{aligned}
$$

and also

$$
Var[\omega(t)|\omega_0] = Var[\omega(t)] - \frac{Cov[\omega(t), \omega_0]^2}{Var[\omega_0]}.
$$

Set $e^{2\omega_0} = e^{2s}\overline{\sigma^2(t)}$. Hence s provides a measure of the "size" of ω_0, with $\overline{\sigma^2(t)} = e^{2C(0)+2\mu}$ equal to the ("unconditional") expectation $E(\sigma^2(t)) = e^{2C(0)+2\mu}$. Equating exponentials, we can see that we have $\omega_0 - \mu = s + C(0)$. By substituting into the above formula we can see that we obtain

$$
\begin{aligned}
E[\sigma^2(t)|\omega_0] &= \overline{\sigma^2(t)} \exp\left[2(\omega_0 - \mu) \cdot \frac{C(t)}{C(0)} - 2\frac{C^2(t)}{C(0)}\right] \\
&= \overline{\sigma^2(t)} \exp\left[2(s + C(0)) \cdot \frac{C(t)}{C(0)} - 2\frac{C^2(t)}{C(0)}\right] \\
&= \overline{\sigma^2(t)} \left(\frac{T}{t+l}\right)^{\alpha(s)+\beta(t)}, \qquad (6.42)
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha(s) &= \frac{2s}{\ln\left(\frac{T}{l}\right)}, \\
\beta(t) &= 2\lambda^2 \frac{\ln(t/l + 1)}{\ln(T/l)},
\end{aligned}
$$

using the form for $C(t)$ in (6.28) and in agreement with the result stated in *[100]*. If $l << t << le^{\frac{|s|}{\lambda^2}}$ we have that $\beta(t) << \alpha(s)$ and (6.42) can be seen to lead to an approximate power-law decay in the volatility:

$$
E_{endo}[\sigma^2(t)|\omega_0] \sim \overline{\sigma^2(t)} t^{-\alpha(s)}.
$$

Thus, even in the case of an endogenous crisis we may expect to find approximate power-law decay of the volatility provided the condition $l << t << le^{\frac{|s|}{\lambda^2}}$ holds. Thus, we may test the model of *[100]* by testing for a power-law $t^{-\alpha}$ decay in the historical volatility with $\alpha = 0.5$. Values less than 0.5 suggest a slower decay and an endogenous aspect, by analogy with the fractional Gaussian noise model in Section 2.

6.4 Data analysis

Given the stochastic representation (6.27), we use the squared log-returns as the appropriate volatility proxy for σ^2. As a candidate model, we consider

$$\sigma^2(t) = A + Bt^{\alpha}. \tag{6.43}$$

For the results to be consistent with the model of *[100]*, we should have $B > 0$ and $\alpha = -1/2$ in (6.43), at least for exogenous shocks. In order to estimate the model (6.43) we employ the pragmatic estimation technique commonly employed in the physics literature (see for instance *[100]*, *[54]*) by "integrating" (6.43) and fitting the model

$$V_n(t) = At + \frac{Bt^{1+\alpha}}{1+\alpha} + \epsilon_t, \tag{6.44}$$

where $V_n(t) = \sum_{n=1}^{t} \sigma_n^2$. This is a commonly used technique to estimate power law exponents in models of the form (6.43). This method is intended to ensure that the estimates obtained are more robust with respect to the presence of background noise, model mis-specification etc. We estimate (6.44) by least squares, using 100 values of daily log-returns. The resulting regressions typically return R^2 values of around 90%. We compare and contrast market responses to the 1987 crash with responses to the terrorist attacks of September 11th 2001, the attempted coup against President Gorbachev on August 19th 1991, Black Monday October 19th 1987 and the Nasdaq crash taken to be April 10th 2000. We describe the results obtained in the next subsection.

6.4.1 Empirical results

When examining markets in the aftermath of the attempted coup against President Gorbachev, one of the surprising features that we find is that much of the volatility occurs towards the end of the sample period. As a result, the model (6.43) seems inappropriate for this data and the estimated α values obtained are positive, reflecting the enhanced volatility towards the end of the sample (see Table 6.1).

Market	α
Nasdaq	1.973 (0.600)
CAC 40	6.682 (0.020)
FTSE	6.349 (0.037)
Nikkei	7.936 (0.014)
S&P 500	3.697 (30.040)

Table 6.1: Results based on 100 trading days after August 19th 1991

In contrast, there are a number of markets in which the model (6.43) appears more reasonable. In particular, we obtain estimates of $\alpha \leq -0.5$, suggestive of an exogenous root under the MRW model of *[100]*, see Table 6.2. Note, however, that the estimated exponents are seen to vary significantly from -0.5, and the model of *[100]* systematically fails to correctly estimate the empirically observed power-law exponent of the volatility decay. A better estimate would appear to be $\alpha \approx 0.7 \pm 0.15$, although non-overlapping confidence intervals do suggest some differences between the various markets.

Market	α	95% C. I
Nasdaq Oct. 19th 1987	-0.691 (0.016)	(-0.72, -0.66)
FTSE Oct. 19th 1987	-0.742 (0.006)	(-0.75, -0.73)
Nikkei Oct. 19th 1987	-0.852 (0.009)	(-0.87, -0.84)
Nikkei Sept. 11th 2001	-0.718 (0.031)	(-0.78, -0.66)
S&P 500 Oct. 19th 1987	-0.624 (0.033)	(-0.69, -0.56)
DJIA Oct. 24th 1929	-0.535 (0.021)	(-0.58, -0.49)

Table 6.2: Selected results ($\alpha < -0.5$)

Finally, on some markets we find some suggestion of an exogenous crises, with observed volatilities decaying *slower* than the value -0.5 predicted by the model of *[100]*. The results are interesting in that proposed bubbles such as the Nasdaq 2000 and the Hang Seng in 1994 are seen to lead to a depressed volatility decay in line with the interpretation of crashes as endogenous events. However, the model of *[100]* is unable to predict the exponent of the power law corresponding to an endogenous cause. In addition, we have the suggestion of slow logarithmic volatility decay on the FTSE and the CAC 40 on April 10th 2000 in response to the crash in internet stocks. The events of September 11th are also seen to lead to a volatility decay that this simple approach would more readily associate with an exogenous cause. The results are shown in Table 6.3.

Market	α	95% C. I
Nasdaq April 10th 2000	-0.177 (0.007)	(-0.19, -0.16)
Nasdaq Sept. 11th 2001	-0.071 (0.015)	(-0.10, -0.04)
CAC 40 April 10th 2000	-0.01 (0.000)	(-0.01, -0.01)
FTSE April 10th 2000	-0.01 (0.000)	(-0.01, -0.01)
FTSE Sept. 11th 2001	-0.220 (0.022)	(-0.26, -0.17)
Nikkei April 10th 2000	-0.274 (0.054)	(-0.38, -0.17)
S&P 500 April 10th 2000	-0.298 (0.034)	(-0.36, -0.23)
S&P 500 Sept. 11th 2001	-0.391 (0.048)	(-0.49, -0.3)
Hang Seng Jan. 6th 1994	-0.174 (0.101)	(-0.38, 0.03)

Table 6.3: Selected results ($\alpha > -0.5$)

6.5 Synthesis

In this section we detail briefly how the model of *[100]* can be modified to obtain a better fit with the empirical results. The comment is made in *[95]* that the multifractal random walk model corresponds, at least approximately, to the fractional Gaussian noise kernel (6.20) with $\theta = 1/2$. Suppose that we retain the integral equation for the log-volatility (6.29), but that $K(t)$ takes the form (6.20) for $\theta \neq 1/2$ in general. By the same reasoning, the formula for the linear response to an external shock (6.41) follows through, with the $t^{-1/2}$ replaced by a more general $t^{1-\theta}$ dependence. The estimated values obtained in Table 6.2 would then be consistent with the choice $\theta = 0.3 \pm 0.15$. Similar considerations also enable one to predict the exponents of the power-law decay in volatility that should accompany an endogenous shock.

Suppose $F(t) = At^{\theta-1}$, for $\theta > 0, A$ constant. We have the integral formulae

$$\begin{aligned}
\int_0^\infty \cos(ft)F(t)dt &= A\frac{\Gamma(\theta)}{f^\theta}\cos\left(\frac{\theta\pi}{2}\right), \\
\int_0^\infty \sin(ft)F(t)dt &= A\frac{\Gamma(\theta)}{f^\theta}\sin\left(\frac{\theta\pi}{2}\right),
\end{aligned} \tag{6.45}$$

(*[61]*, Chapter 10), where $\Gamma(\cdot)$ denotes the gamma function given by (6.25). Thus, we have

$$\mid \hat{F}(f) \mid^2 = A^2(\Gamma(\theta))^2 f^{-2\theta}.$$

Hence, since $K(t) \sim t^{\theta-1}$, $t >> l$, and using (6.31), it follows that

$$\hat{C}(f) \sim \text{Const.} f^{-2\theta},$$

and hence, from (6.45), that

$$C(t) \sim \text{Const.} t^{2\theta-1}, \qquad t >> l, \tag{6.46}$$

for $\theta < 1/2$. Inserting (6.46) into (6.42) then gives

$$\overline{\sigma^2(t)} \exp\left[2(s + C(0)) \cdot \frac{C(t)}{C(0)} - 2\frac{C^2(t)}{C(0)}\right].$$

After a sufficiently long time, the volatility relaxes down to its unconditional average $\overline{\sigma^2(t)}$. Since

$$\overline{\sigma^2(t)} \exp\left[2(s + C(0)) \cdot \frac{C(t)}{C(0)} - 2\frac{C^2(t)}{C(0)}\right] - \overline{\sigma^2(t)} \quad \sim \quad \text{Const.} t^{2\theta-1},$$

this predicts an approximate $t^{2\theta-1}$ power-law decay associated with an endogenous crisis.

$\theta \in (0.3, 0.45)$ predicts a $t^{-\alpha}$ volatility decay with exponent α in the range $(0.1, 0.4)$, in reasonable agreement with most of the values in Table 6.3.

6.6 Conclusions

In this chapter we have explored the possible applications of power laws to a mathematical theory of financial crashes. In the first section we derived results building on original work in *[55]*. In particular, we derived a generalized Pareto distribution for drawdowns on an exponential-Lévy market. This approximate result was seen to fit well to an historical series of drawdowns corresponding to the DJIA. Potential applications include calibration of Lévy process models to financial data and capital allocation problems in banks.

In Sections 2-4 we examined the method of *[100]* for determining endogenous or exogenous causes of crises. We found at least some evidence to support the analogy between stock market crashes and simple models for endogenous and exogenous perturbations of complex systems. However, stock markets are complicated and we have two notable observations. Following the attempted coup against President Gorbachev in August 1991, much of the observed volatility is seen to occur towards the end of the sample period and not in the immediate aftermath. In addition, the volatility decay corresponding to September 11th appears more like that associated with "endogenous" events such as the crash of internet stocks in April 2000, and the bubble on the Hang Seng.

However, we do seem to have at least some evidence to support the hypothesis that endogenous crises lead to a slower volatility decay than exogenous crises. However, the model of *[100]* systematically fails to recreate the empirically observed exponents corresponding to exogenous shocks, and is unable to predict the exponent corresponding to an endogenous shock. A better match with empirical data is found when the log-volatility is assumed to obey a fractional Gaussian noise process. Under this model, we are also able to extend the original formulation and predict the exponent of the power law of the volatility associated with endogenous crises, to a reasonable degree of agreement with the empirical data in Table 6.2 and Table 6.3.

Chapter 7

Evaluating contagion in economics

According to *[6]* the economic literature on currency and banking crises can be usefully divided into insurance crisis and illiquidity crisis models. Chapters 1-5, see also Appendix C on market-value models, essentially constitute insurance crisis models and deal with how the inherent risks are priced by markets. With illiquidity crisis models the main feature of interest is the transfer of shocks between different countries and sectors. It is this second theme which motivates material in this chapter. In particular, how we might seek to define and measure contagion in a manner that can give practical answers to important real questions? The layout of this chapter is as follows. In Section 1 we motivate our discussion with a brief survey of the relevant economic literature. In Section 2 we describe the relevant mathematical and statistical background. In Section 3 we give a statistical solution to the problem of evaluating contagion in economics, with an application to the Latin American currency crisis of the 1990s. Section 4 summarises.

7.1 Contagion

According to *[57]* contagion has been defined in various different ways in the literature. To some degree of generality the term contagion relates to the transfer of economic shocks across countries or sectors, although this definition is not complete in itself. *[58]* defines contagion as the situation where the knowledge of crisis in one country increases the risk of crisis in another country. Alternatively, some authors restrict the definition of contagion to situations where the magnitude of a shock exceeds that which might be expected purely on the basis of economic fundamentals.

According to Jokipii and Lacey, much of the theoretical work on contagion and propagation of crises can be categorized into three main areas: aggregate shocks which affect the economic fundamentals of more than country; country-specific shocks which affect the economic fundamentals of other countries; and shocks which are not explained by economic fundamentals and are classified as pure contagion. Interesting empirical issues are then connected to testing for the existence of contagion and also quantifying how shocks are transmitted through the financial system.

Among the most widely used procedures used to test for contagion are simple OLS regressions *[3]*, principal component analysis *[59]*, and analysis of correlation coefficients *[36]*. Tests based on correlation assume that any changes in coefficient estimates obtained reflect material changes. However, it is important to recognise a possible bias brought about by the estimates obtained being conditional on extreme market movements over the period in question. *[43]* recognises a distinction between genuine contagion and interdependence. Markets may just be interdependent and may naturally respond similarly to a common shock. Contagion in this context then reflects genuine changes in the relationship between two different markets over and above naturally occurring interdependence. According to Jokipii and Lacey, much work in the literature fails to make an adequate distinction between the two.

A second question associated with the issue of contagion is evaluating the different channels through which shocks can be propagated across countries. According to Jokipii and Lacey, much work has focussed on a small array of simple techniques. These include regression – both OLS and logit/probit models – and principal components analysis. A few studies have also tried to use news as the identifying condition for the propagation of shocks *[37]*. Eichengreen et al. study the collapse of fixed exchange rates in the ERM in 1993, with one county's collapse taken to be the external news event.

7.2 Statistical background

This section is concerned with the question of contagion in economics and statistical methods which can be used to provide insight. In particular the focus here is on methods proposed in Chapter 6 of *[71]*, with some of the background discussion taken from *[77]*, Chapter 5. In terms of the economic applications, the salient point is as follows. Contagion relates to the transfer of economic shocks across countries or sectors. However, the major complication is the need to distinguish between genuine contagion and simple interdependence. The following two sections describe how we may proceed to answer this question statistically. This is primarily an empirical question and one should not read too much into the fact that the theoretical model we use is the simple random walk model. However, even with this choice of very simple model, there are some theoretical aspects which we need to clarify. The next three subsections provide a brief description of the relevant mathematical and statistical background.

7.2.1 Copulae

The subject of *copulae* is one of the important underlying themes here. Intuitively, the copula plays the key role in multivariate distribution theory by literally *'coupling'* (linking) the various univariate marginal densities with one unique multivariate density. Consider two continuous random variables X and Y with joint distribution function H so that

$$H(x, y) = Pr[X \leq x; Y \leq y].$$

Let F and G denote the marginal distributions of X and Y respectively. Then, a bivariate copula would be a function C satisfying

$$H(x, y) = C(F(x), G(y)), \tag{7.1}$$

so that the copula couples the univariate marginals to a multivariate distribution. If the marginals are continuous then this representation is unique (See Sklar's theorem overleaf). Thus the copula fully describes the probabilistic co-dependence of X and Y, and there is a sense whereby one can talk of a "copula property" of X and Y as a true feature of the probabilistic co-dependence of X and Y. Several concordance or correlation measures, such as Spearman's rho, can be constructed as direct functionals of the underlying copula. One is thus not restricted to the simple linear correlation coefficient, which depends on both the underlying copula and the univariate marginals, as a sometimes flawed measure of co-dependence. Several examples exist whereby variables which are completely co-dependent nonetheless have a linear correlation of zero. In addition, there are interesting theoretical considerations, such as the Invariance Theorem, which motivate the use of copulae in the study of co-dependent random variables.

We consider the situation with $n > 2$ random variables. Let G be a continuous univariate distribution function. It is shown in *[77]*, Chapter 5 Section 1 that

$$G(Y) \sim U(0, 1). \tag{7.2}$$

(7.2) is known as the probability transform. We follow the presentation in *[71]*, Chapter 3. First we need the following useful definition. A function $C : [0, 1]^2 \rightarrow \mathbb{R}$ is described as being 2-increasing if

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0, \tag{7.3}$$

for any $u_1 \leq u_2, v_1 \leq v_2$. One can go a step further if we take (7.3) to be the appropriate notion of a "C-volume" in $[0, 1]^2$. This notion of a volume measure in two-dimensions can then be extended to n-dimensions. For a function $C : [0, 1]^n \rightarrow \mathbb{R}$ this measure can be defined as

$$\sum_{1}^{2} \cdots \sum_{1}^{2} (-1)^{i_1+\cdots+i_d} C(u_{1i_1}, ..., u_{di_d}), \tag{7.4}$$

where u_{j1} and u_{j2} denote the end-points of the j^{th} box in $[0,1]^n$. Using (7.4) one also has a notion of n-positivity if this sum is positive. These considerations allow us to give a mathematical definition of copulae in n-dimensions. We also have Sklar's theorem (see theorem 8.2.3 below), which describes copulae as a multivariate generalisation of (7.2). Finally we also give the result of an *invariance theorem*. This Invariance Theorem is a very powerful result, demonstrating that the copula is an intrinsic measure of dependence as it is invariant under increasing nonlinear transformations.

Definition 7.2.1 *A function $C : [0,1]^n \rightarrow [0,1]$ is a n-copula if (i)-(iii) hold:*
(i) For all $u \epsilon [0,1], C(1, ..., 1, u, 1...1) = u$.
(ii) $C(u_1, .., u_n) = 0$ if at least one of the u_i equals zero.
(iii) C is n-increasing which means that the sum in (7.4) is positive.

Here, with econometric applications in mind, it is natural to restrict attention to the case of a random vector X with continuous marginals. In this case we can say quite a lot more:

Definition 7.2.2 (Continuous case) *If the random vector X has joint distribution function F with continuous marginal distributions $F_1, ...F_n$, then the copula of F is the distribution function C of $(F_1(X_1), ..., F_n(X_n))$.*

In this continuous case, it becomes clear that the copula is simply a multivariate distribution function with $U(0,1)$ marginals. Moreover, this representation is unique:

Theorem 7.2.3 (Sklar's Theorem) *Given a n-dimensional distribution function F with continuous marginal distributions $F_1, ..., F_n$, there exists a unique n-copula C : $[0,1]^n \rightarrow [0,1]$ such that*

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)).$$

For a proof of Sklar's theorem as stated above, see *[77]*, Chapter 5. A more general result is given in *[82]*, page 18. Here, we define a scalar function h to be *strictly increasing* if $x_1 > x_2 \Rightarrow h(x_1) > h(x_2)$. We have the following:

Theorem 7.2.4 (Invariance Theorem) *Let $X_1, ..., X_n$ be continuous random variables with copula C. Then, if $h_1(X_1), ...h_n(X_n)$ are strictly increasing on the ranges of $X_1, ...X_n$, then the random variables $Y_1 = h_1(X_1), ..., Y_n = h_n(X_n)$ have exactly the same copula, C.*

For a proof see *[77]*, Chapter 5 Section 1. As an illustration we can use the invariance theorem to provide an expression for the Gaussian copula – the copula for a multivariate normal distribution. Suppose $X \sim N_n(\mu, \Sigma)$. Since the operation of standardising by subtracting the mean and dividing by the standard deviation amounts to applying a series of strictly increasing transformations[1], we can see by applying the Invariance theorem inductively that X must have the same copula as $N_n(0, P)$, where P is the correlation matrix of X. From the definition of copulae in the continuous case, we can read off

$$\begin{aligned} C_{P,X}(u_1, \dots, u_n) &= Pr(\Phi(X_1) \leq u_1, \dots \Phi(X_n) \leq u_n), \\ &= \Phi_P(\Phi^{-1}(u_1), \dots \Phi^{-1}(u_n)), \end{aligned}$$

where Φ_P denotes the distribution function of a $N_n(0, P)$ variate.

7.2.2 Dependence measures

The prototypical measure of dependence is linear correlation, defined by

$$\rho(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sqrt{Var(X)Var(Y)}}. \tag{7.5}$$

The linear correlation does satisfy $\rho(X_1, X_2) = 0$ when X_1 and X_2 are independent, but the converse does not hold in general. Moreover, the linear correlation is invariant under linear transformations in that

$$\rho(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho(X_1, X_2). \tag{7.6}$$

Trivially (7.6) holds in the case of strictly increasing linear transformations with β_1 and β_2 positive. However, correlation is not invariant under nonlinear strictly increasing transformations, so that in general we have

$$\rho(T(X_1), T(X_2)) \neq \rho(X_1, X_2), \tag{7.7}$$

for a strictly increasing nonlinear function T. In contrast, we see from the Invariance Theorem that $(T(X_1), T(X_2))$ share the same copula as (X_1, X_2). Hence, if ρ can be expressed as a functional of the underlying copula of (X_1, X_2), we would have equality in (7.7). Moreover, we can construct simple examples where two random variables X and Y are completely co-dependent but have zero linear correlation 0. [Suppose ω is $U[0, 2\pi]$ and consider $(X, Y) = (\sin(\omega), \cos(\omega))$.] Finally, linear correlation is only defined for variables with a finite variance. According to *[71]*, Chapter 4 Section 3, there are also cases when the correlation coefficient is defined but the usual sample estimator has as its asymptotic distribution a fat-tailed Lévy stable distribution. Thus sample estimates obtained may deviate quite considerably from their true underlying values.

[1]If $x_i > \hat{x}_i$ then $\frac{x_i - \mu_i}{\sigma_i} > \frac{\hat{x}_i - \mu_i}{\sigma_i}$.

As a remedy to the various deficiencies of linear correlation, *[77]* mention *rank correlations.* These provide simple scalar measures of dependence which depend solely on the copula of a bivariate distribution – in contrast to the linear correlation coefficient which depends on both. There are two practical reasons for looking at rank correlations. Theoretically, these statistics are attractive as they can be calculated as direct functionals of the copula. In practical terms, they enable calibration of copulae to empirical data, and since they are calculated by looking solely at the ranks of data – rather than the actual data values – they are also more robust with respect to outliers and extreme values. One example of the rank correlation statistics discussed by McNeil et al. is Spearman's rho. Spearman's rho is given by the linear correlation of the probability-transformed random variables:

Definition 7.2.5 (Spearman's rho) *For random variables X_1 and X_2 with marginal distribution functions F_1 and F_2. Spearman's rho is given by $\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2))$.*

For $n > 2$, we may speak of a Spearman's rho matrix S for the random vector X, where $S_{ij} = \rho_s(X_i, X_j)$. The following lemma motivates our discussion, and enables us to show that Spearman's rho can be constructed as a functional of the underlying copula *only.*

Lemma 7.2.6 (Höffding formula) *If (X_1, X_2) has joint distribution function F and marginal distribution functions F_1 and F_2, then the covariance of X_1 and X_2 when finite is given by*

$$Cov(X_1, X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_1, x_2) - F_1(x_1)F_2(x_2)) \quad dx_1 dx_2 \tag{7.8}$$

A proof can be found in *[77]*, Chapter 5 Section 2. We have the following.

Proposition 7.2.7 *Suppose X_1 and X_2 have continuous marginal distributions and unique copula C. Then Spearman's rho is given by*

$$\begin{aligned} \rho_S(X_1, X_2) &= 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) \quad du_1 du_2, \\ &= 12 \int_0^1 \int_0^1 C(u_1, u_2) \quad du_1 du_2 - 3. \end{aligned} \tag{7.9}$$

Proof. The factor of 12 appears because $u_i = F_i(X_i)$ is $U(0, 1)$ with variance $1/12$. The full formula follows by straightforward application of Höffding's formula (7.8). (7.9) shows that Spearman's rho can be constructed as a functional which depends directly on the underlying copula.

7.2.3 Conditional dependence measures

We now motivate a statistical treatment of contagion more directly, by introducing the concepts of conditional dependence measures and conditional correlation. *[71]*, Chapter 6, introduces the concept of conditional correlation. Let X, and Y denote two real random variables. If A is a subset of $\mathbb{R}$, the conditional correlation coefficient ρ_A is defined by

$$\rho_A = \frac{Cov(X, Y|Y \in A)}{\sqrt{Var(X|Y \in A) Var(Y|Y \in A)}}. \tag{7.10}$$

There are two versions of (7.10) which are of key interest to us here. These are the two cases where returns are conditioned upon exceedance of given positive and negative thresholds. We define ρ_v^+ to correspond with the case that the conditioning set A is $[v, +\infty)$, and ρ_v^- when the conditioning set A is $(-\infty, v]$. Though the conditional correlation is a slightly more sophisticated notion of linear dependence than the usual linear correlation coefficient, it remains a purely linear notion of dependence. In particular, there remains a two-fold dependence both on the marginal distributions and the joint distributions. To counter this, Malevergne and Sornette (Chapter 6, Section 4) introduce the notion of a conditional Spearman's rho statistic, which is intended to better reflect properties of the underlying copula and may be a more robust method to consider in practice.

Recall that if $U = F_X(X)$ and $V = F_Y(Y)$ then Spearman's rho was simply the linear correlation between U and V. It is this insight which is used to motivate the notion of a conditional Spearman's rho statistic. There are two conditional Spearman statistics that are of interest here:

$$\begin{aligned} \rho_s^+(v) &= \frac{Cov(U, V|V \geq v)}{\sqrt{Var(U|V \geq v) Var(V|V \geq v)}}, \\ \rho_s^-(v) &= \frac{Cov(U, V|V \leq v)}{\sqrt{Var(U|V \leq v) Var(V|V \leq v)}}. \end{aligned}$$

As shown in Malevergne and Sornette Chapter 6, it is possible to obtain analytical expressions for the conditional correlation coefficient in the case of the bivariate Normal and bivariate t-distributions. However, the conditional Spearman's rho statistic is much harder to deal with theoretically and usually has to be calculated numerically.

7.3 A statistical approach to evaluating contagion in economics

The methodology introduced in *[71]* directly uses the conditional correlation coefficient and conditional Spearman's rho to evaluate contagion in economics. One can conclude in favour of a contagion from X to Y if the empirical conditional correlation measure of

$Y|X$ exceeds the confidence interval obtained under a suitable null hypothesis. We assume that the log-price follows a symmetric random walk. Thus we restrict attention under the null hypothesis to a constant correlation structure. This may be criticised as unrealistic, but the empirical convenience of the method is attractive. As an empirical question we may reasonably seek to define contagion as that observed over and above what we would expect if we were sampling from a simple symmetric random walk. *[71]* consider both bivariate normal and bivariate Student-t models, concluding in favour of the Student-t distribution as this copula allows for asymptotic dependence and is not restricted to the case of asymptotic independence as is the Gauss copula (*[77]*, Chapter 5, Section 3). Implicit here, is financial modelling using heavy-tailed non-Gaussian *elliptically-contoured distributions* *[18]*. Elliptically contoured distributions form a natural generalisation of the family of multivariate normal distributions, sharing many of the familiar properties of the normal distribution such as stability under conditional expectations and linear transformations. Where Malevergne and Sornette consider the bivariate t distribution, we consider the symmetric *generalised hyperbolic* distribution. Let X_{Lr} and Y_{Lr} denote log-returns corresponding to X and Y respectively. As a solution to the economic question of evaluating contagion from X and Y, our proposed method follows the heuristic:

1. Look at daily log-returns, and assume a simple symmetric random walk.
2. For each $X_{Lr} = x_{Lr}$ calculate the conditional distribution of $Y_{Lr}|x_{Lr}$.
3. For each $X_{Lr} = x_{Lr}$ simulate y_{Lr} from $Y_{Lr}|x_{Lr}$, and calculate the conditional correlation measure for this pseudosample. Repeat this process.
4. There is evidence in favour of contagion if the sample conditional correlation measures lie outside the Monte Carlo confidence regions constructed in this way.

Note that as presented here the method does not require numerical integration to calculate the exact conditional concordance measures for the assumed null distribution. The main point of interest is whether or not the observed sample values lie outside the Monte Carlo confidence intervals generated using steps 1-4.

7.3.1 The generalized hyperbolic distribution

Both the generalized hyperbolic distribution and Student's t distribution are regularly used in finance as more realistic and heavier tailed alternatives to the normal distribution. A comparison of the two is made in *[77]*, Chapter 3 Section 2. The generalized hyperbolic distribution was introduced in *[10]* and popularised as a financial model by *[34]* and *[35]*. The theoretical background helps to provide context. So-called *normal variance-mean mixtures* *[18]* are among the simplest ways of generating multivariate distributions with both heavy tails and asymmetry. If U is a random variable with law F on $[0, \infty)$ and

X is an $d \times 1$ random vector, then X is a normal variance-mean mixture if

$$X|(U = u) \sim N_d(\mu + u\beta, u\Delta), \tag{7.11}$$

where Δ is a symmetric positive definite ('covariance') $d \times d$ matrix satisfying a determinant one identifiability constraint, μ and β are d vectors. μ is called the position and β is called the drift and models asymmetry. Now financial time series – such as log-returns calculated over intervals of one day or longer – tend to be only very slightly asymmetric. Thus it is common practice to model financial time series using distributions which retain symmetry but have heavy non-Gaussian tails. This can be achieved by considering *normal variance mixtures*, corresponding to the case where the drift parameter β in (7.11) is zero. Both the generalized hyperbolic and Student t distributions can be considered as normal variance-mean mixture models. In this case, the generalized hyperbolic distribution corresponds to the case where U in (7.11) is generalized inverse Gaussian ($GIG(\lambda, \chi, \psi)$).The Student t_v distribution is the case where U is inverse gamma ($IG(v/2, v/2)$).

There are further theoretical reasons for considering the generalized hyperbolic distribution and the Student t distribution as suitable candidate models in finance. Firstly, if we restrict attention to the symmetric normal variance mixture case, then both of these distributions belong to the class of *elliptically contoured distributions.* This family constitutes a natural semi-parametric generalisation of the multivariate normal family. In the regular case where the density exists and the covariance matrix Σ is of full rank, the density of a elliptically contoured random variable can be written as

$$f(x) = |\Sigma|^{-1/2} g\left((x-\mu)'\Sigma^{-1}(x-\mu)\right),$$

(*[19]*, equation (EC)). The multivariate normal distribution is the special case $g(x) = (2\pi)^{-d/2}e^{-x/2}$. Further, in a dynamic setting, the generalized hyperbolic and Student t distributions are also of interest in building heavy-tailed Lévy process generalisations of the geometric Brownian motion model of the stock market. Note that both the generalized hyperbolic and Student t distributions are infinitely divisible. So that if for example we were looking at daily prices, there would be a Lévy process whose daily increments were generalized hyperbolic (with the same parameters) and also a Lévy process whose daily increments were Student t.

In the general asymmetric normal variance-mean mixture case, the density of the generalized hyperbolic distribution is given by

$$\begin{aligned} f(x) &= \frac{(\psi/\chi)^{\lambda/2}}{(2\pi)^{d/2}\alpha^{\lambda-d/2}K_\lambda(\sqrt{\chi\psi})}\left(\chi + (x-\mu)'\Delta^{-1}(x-\mu)\psi\right)^{\lambda/2-d/4} \\ &\times \; K_{\lambda-d/2}\left(\alpha\sqrt{(\chi + (x-\mu)'\Delta^{-1}(x-\mu))}\right)e^{\beta'(x-\mu)}, \end{aligned}$$

where $\alpha^2 = \psi + \beta'\Delta^{-1}\beta$. If X has a generalized hyperbolic distribution, we use the shorthand $gh_d(\lambda, \alpha, \beta, \chi, \mu, \Delta)$, where d denotes the dimension of the random vector X. We restrict attention here to the symmetric normal variance mixture case. In this case the density of a generalized hyperbolic random variable simplifies considerably, and is given by

$$f(x) = \frac{\chi^{-\lambda/2}\psi^{d/2-\lambda/2}}{(2\pi)^{d/2}K_\lambda(\sqrt{\chi\psi})} \frac{K_{\lambda-d/2}\left(\sqrt{(\chi + (x-\mu)'\Delta^{-1}(x-\mu))\psi}\right)}{(\chi + (x-\mu)'\Delta^{-1}(x-\mu)\psi)^{d/4-\lambda/2}}, \tag{7.12}$$

where K_λ is a modified Bessel function of the third kind satisfying the integral representation

$$K_\lambda(x) = \frac{1}{2}\int_{-\infty}^{\infty} u^{\lambda-1} \exp\left\{\frac{1}{2}\left(u + \frac{1}{u}\right)\right\} du \qquad (x > 0).$$

We employ the pragmatic estimation method used by *[14]* – which is applicable for the symmetric elliptically contoured case considered here. We estimate μ by the empirical mean vector and Δ by the empirical variance scaled to have determinant one. Then we base estimation of the other parameters upon maximization of the likelihood suggested by the formula given in (7.12). We perform the optimisation using the Nelder-Mead method (*[81]*).

There are further reasons why the generalized hyperbolic distribution might be of interest for use in financial modelling. Namely tractability, as this class of distributions remains closed under linear transformations and forming conditional distributions. The full results are given in *[20]*. Quoting the result for conditional distributions we have

Theorem 7.3.1 *Suppose that X is a d-dimensional variate distributed according to the generalized hyperbolic distribution $gh_d(\lambda, \alpha, \beta, \chi, \mu, \Delta)$. Let (X_1, X_2) be a partitioning of X, and let r and k denote the dimensions of X_1 and X_2 respectively. Let (β_1, β_2) and (μ_1, μ_2) be the corresponding partitions of β and μ. Let*

$$\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}$$

be a partition of Δ such that Δ_{11} is a $r \times r$ matrix. Then the conditional distribution of X_2 given $X_1 = x_1$ is $gh_k(\tilde{\lambda}, \tilde{\alpha}, \tilde{\beta}, \tilde{\chi}, \tilde{\mu}, \tilde{\Delta})$ where

$$\begin{aligned}
\tilde{\lambda} &= \lambda - r/2, \\
\tilde{\alpha} &= \alpha|\Delta_{11}|^{\frac{1}{2k}}, \\
\tilde{\beta} &= \beta_2, \\
\tilde{\chi} &= |\Delta_{11}|^{\frac{-1}{k}}(\chi + (x_1 - \mu_1)'\Delta_{11}^{-1}(x_1 - \mu_1)),
\end{aligned}$$

$$\tilde{\mu} = \mu_2 + (x_1 - \mu_1)\Delta_{11}^{-1}\Delta_{12},$$
$$\tilde{\Delta} = |\Delta_{11}|^{\frac{1}{k}}(\Delta_{22} - \Delta_{21}\Delta_{11}^{-1}\Delta_{12}).$$

Here we are simply interested in the symmetric 2-dimensional case, which allows for considerable simplification. As a simple corollary we have:

Corollary 7.3.2 *Let (X_1, X_2) be generalized hyperbolic $gh_2(\lambda, \mu, \Delta, \chi, \psi)$. Then the conditional distribution of X_2 given X_1 is $gh_1(\lambda - 1/2, \tilde{\mu}, 1, \tilde{\chi}, \tilde{\psi})$ where*

$$\tilde{\mu} = \mu_2 + (x_1 - \mu_1)\frac{\Delta_{12}}{\Delta_{11}},$$
$$\tilde{\chi} = \frac{\chi}{\Delta_{11}} + \frac{(x_1 - \mu_1)^2}{\Delta_{11}^2},$$
$$\tilde{\psi} = \psi\Delta_{11}.$$

Finally, we remark that the formulation of the hyperbolic distribution as a normal variance mixture allows for an easy way to simulate from $gh_d(\lambda, \mu, \Delta, \chi, \psi)$. Let $\Sigma = u\Delta$. The simulation algorithm is as follows:

1. Generate u from $GIG(\lambda, \chi, \psi)$.
2. Return $Y = \mu + A'v$,

where $A'A = \Sigma$ via Cholesky decomposition and v is $N_d(0, I_d)$ independent of u. Thus the problem effectively reduces to a one-dimensional problem, namely that of simulating from a generalized inverse Gaussian random variable. Here this is achieved by using the R function *rgig* downloaded from Professor David Scott's website (http://www.stat.auckland.ac.nz/~dscott/).

7.3.2 Financial application

As a financial application we consider the contagion problem associated with Latin America. In particular, the period encapsulates the Mexican crisis in 1994 and a crisis in Argentina that started in 2001. *[71]*, Chapter 6, look at daily log-returns from four national stock markets: Argentina (Merval index), Brazil (IBOV index), Chile (IPSA index) and Mexico (Mexbol index). The sample period covers January 15th 1992 to 15th June 2002, and thus includes both crisis periods. Malevergne and Sornette conclude in favour of an asymmetric contagion effect, whereby Mexico and Chile can be potential sources of contagion towards Argentina and Brazil, but the reverse does not hold. According to Malevergne and Sornette this offers an attractive economic interpretation in terms of market-oriented and state-intervention oriented countries. The hypothesis is that currency floating regimes are able to adapt to important manufacturing sectors and deliver more competitive real exchange rates (Chile and Mexico) than fixed exchange rates (Argentina before the 2001 crisis, and Brazil). In short, a more flexible exchange rate might be seen to provide a safety net and allow a decoupling between the national stock market and external influences.

Here, we look at daily log-returns from the Mexican IPC index and the Brazilian Bovespa index, from May 3rd 1993 to April 30th 2003. Whilst Malevergne and Sornette use a bivariate student-t distribution, we use the bivariate generalized hyperbolic distribution. For our results to be consistent with the analysis of Malevergne and Sornette we should see that the IPC affects the Bovespa but the Bovespa does not affect the IPC. Here we measure "affect" in terms of deviation from the sampling distribution of the correlation measures under the null hypothesis that the log price follows a symmetric 2-d generalized hyperbolic random walk. The results are shown in Figures 7.1 and 7.2. In the sequel the graphs have the following structure. The X axis denotes x-values, standardised by subtracting the mean and dividing by the standard deviation. The Y-values denote conditional correlation or conditional Spearman's ρ values. For X-values less than zero we calculate $\rho^-(v)$ or $\rho_s^-(v)$. Thus we are conditioning on values of y when the X takes values *less than or equal to* x. For X-values greater than zero we calculate $\rho^+(v)$ or $\rho_s^+(v)$. Thus we are conditioning on values of y when the X takes values *greater than or equal to* x. The dashed lines indicate 95% confidence regions – obtained by Monte Carlo simulation – under the null hypothesis that the log-prices follow a symmetric bivariate generalized hyperbolic random walk.

Further details of the Monte Carlo algorithm are as follows. Note that under the null hypothesis of a random walk, the log-returns constitute a random sample from a bivariate generalized hyperbolic distribution. Let X and Y refer to the two time series of log-returns. First we consider the observed values x_i of X, and simulate $Y_{i,s}$ from the distribution of $Y_i|X_i = x_i$ using Corollary 7.3.2 for the conditional distributions of the 2-d symmetric generalized hyperbolic distribution. From this $(x_i, Y_{i,s})$pseudosample, we calculate $\rho^-(x_i)$ and $\rho_s^-(x_i)$ or $\rho^+(x_i)$ and $\rho_s^+(x_i)$ for each x_i, depending on whether or not $x_i > \mu_X$. We repeat this procedure 10,000 times to generate pointwise 95% confidence intervals.

7.3.3 Empirical results

The results for the Mexican IPC conditional on the Brazilian Bovespa are shown in Figure 7.1. The results for the Brazilian Bovespa conditional on the Mexican IPC are shown in Figure 7.2. In Figure 7.1 the observed sample conditional correlation coefficient lies within confidence limits. The conditional Spearman's rho shows some minor deviation from this null hypothesis, although it seems rather grand to label this contagion. Thus, we seem to have some suggestion of really quite minor contamination from Brazil to Mexico.

In contrast, we seem to have stronger evidence for a contagion effect from Mexico affecting Brazil. This is shown most strongly by the conditional Spearman's rho statistic, although this is also highlighted by the conditional linear correlation coefficient in the left panel. We note from the scale of the x-axis in the plots that we appear to have roughly twice as much contagion from Mexico to Brazil using the conditional correlation measure as from Brazil to Mexico using the conditional Spearman's rho statistic. The conditional Spearman's rho statistic then shows an added degree of contamination from Mexico to Brazil than that measured by the conditional correlation measure. The results thus appear to be in agreement with Malevergne and Sornette's analysis. There seems to be an asymmetric contagion effect whereby Mexico affects Brazil but there is a comparatively minor influence in the opposite direction.

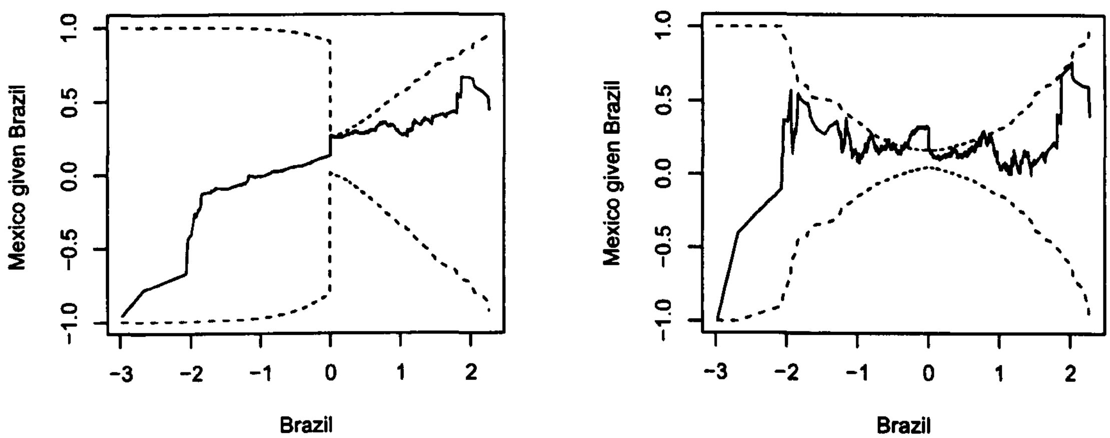
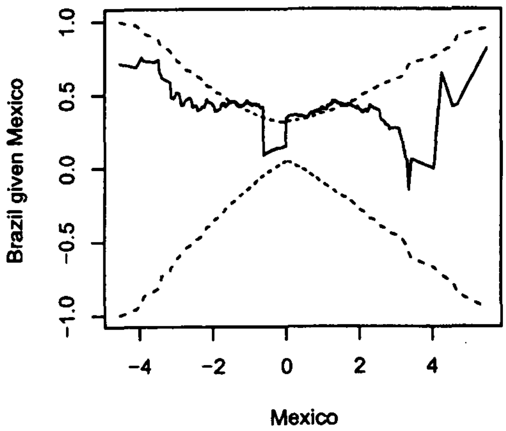


Figure 7.1: Left panel: Values of $\rho^{\pm}(v)$ Right panel: Values of $\rho_s^{\pm}(v)$

7.4 Conclusions

This chapter has discussed a statistical approach to evaluating contagion in economics based on conditional concordance measures and the null hypothesis of a heavy-tailed random walk. Here we used a 2-d symmetric generalized hyperbolic random walk. Appropriate mathematical and statistical background was discussed and, in particular, copulae were introduced as the appropriate way to describe the probabilistic co-evolution of multivariate random variables. Further, as an alternative to the simple linear correlation coefficient, Spearman's rho was introduced and, as presented in [77] and [71], was shown to be a direct functional of the underlying copula. Two conditional concordance measures were discussed, the conditional correlation coefficient and the conditional Spearman's rho. The conditional Spearman's rho appears to be the more sensitive measure of co-



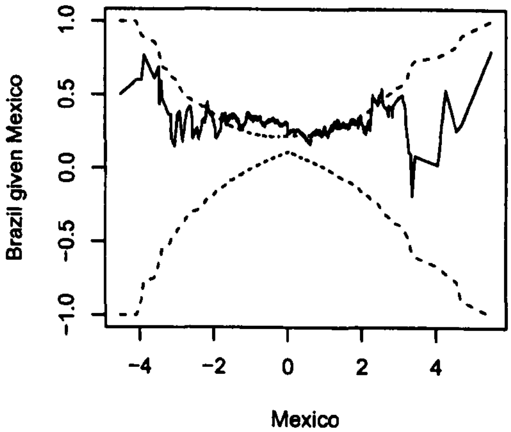


Figure 7.2: Left panel: Values of $\rho^{\pm}(v)$ Right panel: Values of $\rho_s^{\pm}(v)$

dependence, with results using this measure showing greater deviation from the null hypothesis of a random walk. However, results using both measures showed some similarities. In terms of the financial application, we have evidence for contagion from Mexico to Brazil but a relatively minor influence in the opposite direction. These results coincide with the interpretation of *[71]* and may admit the economic interpretation that currency floating regimes provide greater protection from external influences than fixed exchange rate regimes.

Chapter 8

Conclusions and further work

8.1 Conclusions

We can summarise the main conclusions of this thesis as follows. In Chapters 1-3 we analysed financial data corresponding to the S&P 500, the Nasdaq and the Hang Seng prior to observed crashes. In Chapter 1 we reviewed the log-periodic method of Sornette and Johansen. In Chapter 2 we were able to show that simple random variation can generate an apparent log-periodic signal. Further, rather than log-periodicity, it appears sufficient to model super-exponential growth. A definite precursor of crashes appears to be super-exponential growth, with the price continuing to accelerate before the observed crashes. In Chapter 2 we combined these simple observations with the interpretation of *[84]*, namely that bubble episodes may involve rational investors appropriately compensated for added levels of risk. Mathematically, we use a slightly simplified version of the JLS power-law with interest rate *[52]*. The resulting formulae are intuitive and lead to bubbles exhibiting super-exponential growth. We call this model the SEG model. We then use this model to obtain interesting backward predictions for these historically observed crashes.

In Chapter 3 we studied SDE models for bubbles. First we examined a proposed hierarchy, *[2]*, of no bubbles vs. fearless bubbles vs. fearful bubbles. However, the proposed fearless bubble model, an SDE formulation of the JLS power-law model, is not seen to offer a significant improvement over geometric Brownian motion. In contrast, the proposed fearful bubble model is seen to offer a significant improvement. We made two identifiability constraints and derived maximum likelihood equations for the Sornette-Andersen model *[94]* to refine the statistical analysis of *[2]* with a likelihood-ratio test for the presence of bubbles. These results are also interesting as they demonstrate that SDEs a flexible and natural tool to use in this context and form the natural and meaningful generalisation of the super-exponential growth ODE models considered previously.

Further, the original formulation of the Sornette-Andersen model can be improved by using a heavy-tailed hyperbolic Lévy process, [34], in order to take into account non-Gaussian behaviour in financial markets ([31], Chapter 7 page 210). The likelihood ratio test still applies and gives us a more robust test for bubbles in financial markets. The results using the hyperbolic re-formulation are seen to be generally less significant than when the background driving noise process is assumed to be Brownian motion. We use this test to repeat the analysis of [2] and also in out-of-sample analysis in Chapter 5 as we compare our model with some of the analysis in [93].

In Chapter 4 we discuss volatility and liquidity precursors. Liquidity is a general concept from economics, it refers to the unrestricted flow of money in "healthy" markets and generalises the purely statistical notion of volatility. We developed a simple test, based on the Sornette-Andersen model, to gauge which factors were significant once we account for super-exponential growth. Results are seen to offer an improvement over a simple nonlinear regression approach using ordinary least squares, since in this case the error terms appear to be very highly correlated. We have some evidence to support the interpretation of [22] that crashes occur on volatile, illiquid markets. Further, the economic interpretation of these results seems interesting, with different features seen to be significant on each of the three markets. The JLS power-law model makes explicit the analogy between stock market crashes and complex systems. In particular, this model assumes that economic variables display power-laws and approximate phase-transition behaviour prior to crashes. Our statistical analysis in Chapter 4 and Chapter 5 gives some evidence in support of this and liquidity measures closely linked to the trading volume seem particularly important. Hence, we also find some additional evidence to support the analogy between stock market crashes and complex systems.

In Chapter 5 we synthesise material in Chapters 2-4. In particular, our method follows the heuristic given by Figure 5.1, using the likelihood ratio test of the hyperbolic Sornette-Andersen model as a formal statistical test for super-exponential growth. We test for bubbles using four-years of data as suggested by the results of our backwards predictions using the SEG model in Chapter 2. In Section 1 we find evidence of a bubble in the US$:DEM series from February 1981- February 1985 but reject the hypothesis of a bubble in the US$:CAD series July 1994-July 1998. Both are cited as bubbles in [93]. In Section 2 we discuss log-periodic false predictions as discussed in [93]. In both cases we find strong evidence for bubbles, but find that the original method seems to over-state the extent to which a crash is imminent. The second prediction is of a crash on the Nasdaq in Oct. 1999 based on data up to the end of Sept. 1999. Using both our SEG model of Chapter 2 and theregime-switching regression model of [103] we conclude that in actual fact, the level of risk seems to peak around 2000, the actual time of the crash and around six months later than the original prediction. Section 4 illustrates the prediction heuristic Figure 5.1, and we are able to obtain interesting results corresponding to historical crashes using data for the Dow Jones Industrial Average Sept. 1983-Sept. 1987 and for the Nasdaq 100 index from March 1996-March 2000.

In Chapter 6 we extended the approach taken in *[55]* to derive an approximate generalised Pareto distribution for drawdowns on an exponential-Lévy market. This result refutes the previous suggestion in *[55]* that large drawdowns are outliers. The result obtained appears to have possible applications for calibrating Lévy process models of financial markets and capital allocation problems in banks. Also in Chapter 6 we discussed the method of *[100]* for determining the origin of crises. The theory states that crises with an exogenous origin should involve a faster volatility decay than crises with a more involved endogenous origin, with an approximate power-law decay in volatility in both cases. This is based on an analogy between stock markets and simple models for exogenous and endogenous shocks in complex systems *[95]*. From the results obtained we observed at least some empirical results to support the intuition of *[100]*. However, it appears that this method is flawed in that it systematically fails to estimate the empirically observed exponents for "endogenous" shocks, and also is not able to predict values for the exponent corresponding to endogenous shocks. It appears that rather than the multifractal random walk, a better match with empirical data can be found with a stochastic volatility model where the log-volatility obeys a fractional Gaussian noise process. Further, this simple model is able to predict the exponent of the power law decay corresponding to an "endogenous" shock, with a reasonable correspondence with the empirical data.

According to the economic literature on currency and banking crises there is a subtle distinction between insurance crisis and illiquidity crisis models *[6]*. The class of illiquidity crises models motivates the question of evaluating contagion in economics. Following the approach of *[71]*, we formulate this question as "Is there any evidence of contagion over and above that which we might expect under the null hypothesis of a symmetric generalized hyperbolic random walk?". The financial application was with regard to the Latin American currency crises of the 1990s and Brazil and Mexico in particular. The results obtained were consistent with the conclusions of *[71]*, Chapter 6, and appear to show an asymmetric contagion effect whereby Mexico infects Brazil but not vice-versa. It has been suggested that this is consistent with the economic interpretation that currency floating regimes provide greater protection from external influences than fixed exchange rate regimes.

8.2 Further work

There are rich potential avenues for empirical application of ideas expressed in this thesis. This includes use of the SEG bubble models in Chapters 2-5, the empirical power-laws in Chapters 4 and 5, and the work in evaluating contagion in economics in Chapter 7. In particular, the empirical power-laws in Chapter 4 and 5 are of interest as they highlight, quantitatively, the analogy between stock markets crashes and complex systems. The implicit assumption is that this feature should be very general and not a specific artefact of the limited number of markets examined in this thesis. Possible avenue for further investigation include models for time-varying expectations and models for change-points in financial series. Change-point models are discussed in *[60]*.

There are also a number of possible ways in which one might consider improvements to models considered here. One might consider extending the SEG model of Chapter 2 using a regime-switching framework or by a Bayesian formulation of the problem. *[28]* consider a Bayesian formulation of log-periodic models using the exact definition of the hazard rate (see Appendix A).

Chapter 3 suggests ways in which theoretical models may be developed. For the Sornette-Andersen model we may not only wish to consider option pricing problems but also extensions to the model itself. Here, we replaced the background noise process with a non-Gaussian hyperbolic Lévy process. However, the Sornette-Andersen model was initially constructed as a nonlinear generalisation of the Black-Scholes model and one might equally consider nonlinear generalisations of simple stochastic volatility models.

The derived formula for drawdowns suggest possible applications for calibration of Lévy models and risk management. It might be of interest to develop a more rigorous mathematical analysis and investigate ways in which similar power-law behaviour might emerge in complex systems. Tail behaviour under random stopping is briefly discussed in *[17]*, Chapter 8. The method of *[100]* is interesting and probably has a very broad potential application to a wide range of areas. However, in Chapter 6 it appears that rather than the multifractal random walk a better description of the empirical data is given by a stochastic volatility model where the log-volaility obeys a fractional Gaussian noise process *[95]*. Analysis and estimation of this model would be interesting. Further, the causal integral representation (6.29) is a key feature of interest here. In particular, one might examine the subject of volatility precursors of financial crashes if it is assumed, as in the JLS model, that markets price risks of the form $B(t_c - t)^{-\alpha}$ prior to crashes.

It may also be of interest to consider valuation models and ways of constructing fundamental prices. Simple valuation models include the constant fundamental price-dividend ratio model described in the appendix of *[103]* and vector autoregressive techniques, for example *[25]*. One particular problem is that in the first case the price-dividend ratio is assumed constant, and in the second case is assumed constant subject to short term variations regarding future dividend payments. Empirical results using both these methods are seen to be broadly similar *[23]*, *[103]*. However, are these modelling assumptions realistic when the price dividend ratio is seen to increase over time, particularly in recent years? Some simple present value models are discussed in Chapter 7 of *[26]*. One suggestion is that markets might be adjusting to high anticipated earnings associated with new technologies which are yet to deliver. A paper discussing some of these issues is *[91]*. Another possibility might be a Bayesian formulation of the pricing problem, where the price-dividend ratio ρ and other model parameters are considered to be unknown random variables which have to be estimated by a representative market agent. For an application of these ideas to a slightly different context see *[51]*.

There are also a number of less obvious ways in which one might consider extending this thesis. Given the close analogies between stock markets and complex systems described in Chapters 1-5 one might consider other ways in which this could be explicitly accounted for in the modelling. In Chapter 7, it may be of interest to extend our purely statistical approach by considering stochastic epidemic models. However, it is not obvious how this may be achieved in a manner which retains the practical nature of the existing method.

Appendix A

Exact definition of the hazard rate

Chapter 2 discussed the hazard-rate in relation to a first-order martingale approximation in the JLS power-law model. Here, for the sake of completeness we list the exact interpretation of the hazard rate as discussed in the appendix of *[28]*. Suppose that a crash has not occurred by time t_1. The conditional probability that X will occur at or before time t_2, $(t_2 > t_1)$, is given by

$$Pr(X \le t_2 | X > t_1) = \frac{F(t_2) - F(t_1)}{1 - F(t_1)}, \tag{A.1}$$

where F denotes the cdf of X. The hazard rate is obtained by taking

$$h(t_1) = \left(\frac{d}{dt_2} Pr(X \le t_2 | X > t_1) \right) \Bigg|_{t_2 = t_1}.$$

Let $S(t) = 1 - F(t)$. Then $S'(t) = -f(t)$ and we have that

$$h(t) = \frac{-S'(t)}{S(t)} = \frac{-d}{dt} \ln(S(t)),$$

and $S(0) = 1$. Continuing, we see that $S(t) = \exp\{-\int_0^t h(u)du\}$ and (A.1) reduces to

$$\begin{aligned} Pr(X \le t_2 | X > t_1) &= \frac{S(t_1) - S(t_2)}{S(t_1)}, \\ &= 1 - \frac{S(t_2)}{S(t_1)}, \\ &= 1 - \exp\{-\int_{t_1}^{t_2} h(u)du\}. \end{aligned} \tag{A.2}$$

Thus (A.2) is the exact conditional probability statement corresponding to the hazard rate $h(t)$.

Appendix B

Derivation of the van Norden-Schaller model

This model requires a decomposition of a stock price into a bubble component and a fundamental price component. We now describe a very simple way how this approximate decomposition might be achieved before we give a full derivation of the model.

B.0.1 Constructing fundamental values

This subsection describes a simple model to estimate fundamental values. This approach is considered by both *[103]* and *[23]*, and is the simplest way to construct fundamental values. Here, we interpret fundamental values as price levels that would appear reasonable given observed dividend payments. Both papers also consider a more complicated vector autoregressive method, described in *[25]*, to achieve this fundamental value-bubble decomposition. However, both these papers report that results seem to be similar for these two methods. According to the appendix in *[103]*, the starting point is the equilibrium condition in the *[69]* model for economy-wide market prices and quantities:

$$P_tU'(D_t) = \beta E_tU'(D_{t+1})[P_{t+1} + D_{t+1}], \tag{B.1}$$

where $U'(\cdot)$ denotes the derivative of a utility function relating to consumption of dividends and P_t, D_t are the price at time t and E_t corresponds to the conditional expectation with respect to some filtration $\mathcal{F}_t$. β denotes a subjective discount factor $0 < \beta < 1$. Next, this simple model assumes, for sake of tractability, a constant relative risk aversion utility:

$$U(D_t) = (1+\gamma)^{-1}D_t^{1+\gamma}, \tag{B.2}$$

where γ is the coefficient of relative risk aversion. Substituting (B.2) into (B.1) gives us

$$P_tD_t^{\gamma} = \beta E_tD_{t+1}^{\gamma}(P_{t+1} + D_{t+1}). \tag{B.3}$$

This model then assumes, for sake of tractability, that log dividends are a random walk with drift, so that letting $d_t = \log(D_t)$ we have

$$d_t = a_0 + d_{t-1} + \epsilon_t, \tag{B.4}$$

where a_0 is a drift parameter, and ϵ_t is an i.i.d. sequence of $N(0, \sigma^2)$ random variables. To solve this model, we propose a solution of the form

$$P_t = \rho D_t. \tag{B.5}$$

Plugging (B.5) into (B.3), we obtain

$$\rho D_t^{1+\gamma} = \beta E_t[(\rho + 1) D_{t+1}^{\gamma+1}]. \tag{B.6}$$

From (B.4) and using the fact that the moment generating function of $N(\mu, \sigma^2)$ is $e^{\mu t + \frac{\sigma^2 t^2}{2}}$ we have that

$$E_t[D_{t+1}^{1+\gamma} | D_t] = D_t^{1+\gamma} e^{a_0(1+\gamma) + \frac{(1+\gamma)^2\sigma^2}{2}}, \tag{B.7}$$

which follows from the fact that $d_{t+1}|d_t$ is $N(a_0 + d_t, \sigma^2)$. Plugging (B.7) into (B.6) we are left with

$$\begin{aligned}
\rho D_t^{1+\gamma} &= \beta(\rho+1) D_t^{1+\gamma} e^{a_0(1+\gamma) + \frac{(1+\gamma)^2\sigma^2}{2}}, \\
\text{so} \quad \rho \left[1 - \beta e^{a_0(1+\gamma) + \frac{(1+\gamma)^2\sigma^2}{2}}\right] &= \beta e^{a_0(1+\gamma) + \frac{(1+\gamma)^2\sigma^2}{2}}, \\
\text{i.e.} \quad \rho &= \frac{\beta e^{a_0(1+\gamma) + \frac{(1+\gamma)^2\sigma^2}{2}}}{1 - \beta e^{a_0(1+\gamma) + \frac{(1+\gamma)^2\sigma^2}{2}}}.
\end{aligned} \tag{B.8}$$

Hence, it follows that (B.5) holds with ρ given by (B.8). The implications are that under this simple model, the fundamental value P_t^* is simply tied to the dividend as:

$$P_t^* = \rho D_t.$$

In empirical work, ρ is estimated by the mean price-dividend ratio.

B.0.2 A Regime-Switching Regression model

We use the simple model of the previous subsection to construct fundamental prices. Let P_t denote the stock market price at time t, and B_t be the deviation from fundamental price defined as $B_t = P_t - P_t^*$, where P_t^* is the fundamental price $P_t^* = \rho D_t$. We have two states S a speculative regime with a bubble and C a collapsing bubble regime. The transition between the state S and the state C corresponds to a crash. Assuming that B_t

is statistically independent of D_t, we have that if the equilibrium condition is satisfied

$$B_t D_t^{\gamma} = \beta E_t \left[D_{t+1}^{\gamma}\right] E_t[B_{t+1}]. \tag{B.9}$$

This can be seen as follows. The equilibrium condition in (B.1) is

$$P_t D_t^{\gamma} = \beta E_t(D_{t+1}^{\gamma}(P_{t+1} + D_{t+1})). \tag{B.10}$$

The LHS in (B.10) becomes

$$(B_t + \rho D_t)D_t^{\gamma} = B_t D_t^{\gamma} + \rho D_t^{\gamma+1},$$

using $B_t = P_t - P_t^*$. Similarly, the RHS in (B.10) becomes

$$\begin{aligned}\beta E_t(D_{t+1}^{\gamma}(B_{t+1} + \rho D_{t+1} + D_{t+1})) &= \beta E_t(D_{t+1}^{\gamma} B_{t+1}) + \beta E_t((1+\rho)D_{t+1}^{\gamma+1}),\\ &= \beta E_t(D_{t+1}^{\gamma} B_{t+1}) + \rho D_t^{1+\gamma},\end{aligned}$$

from (B.6). Finally, (B.9) follows from the assumption that the B_t and the D_t are independent. If we retain the assumption that $d_t = \log(D_t) = \alpha_0 + d_{t-1} + \epsilon_t$, where the ϵ_t are i.i.d. $N(0, \sigma^2)$, then (B.9) can be re-written as

$$\begin{aligned}\frac{E_t[B_{t+1}]}{B_t} &= \frac{D_t^{\gamma}}{\beta E_t[D_{t+1}^{\gamma}]},\\ &= \beta^{-1} e^{-\gamma(\alpha_0 + \gamma\sigma^2/2)},\\ &= M,\end{aligned} \tag{B.11}$$

for some constant M say, using the fact that the moment generating function of $N(\mu, \sigma^2)$ is $e^{\mu t + \sigma^2 t^2/2}$. Let $b_t = B_t/P_t$ be the relative bubble size at time t and consider the state S. If we are in state S at time t, the probability of being in S at time $t+1$ is given by $q = q(b_t)$, so that the probability of remaining in state S at time $t+1$ depends on the relative bubble size. Further $q(\cdot)$ is assumed to satisfy

$$\frac{dq(b_t)}{d|b_t|} < 0, \tag{B.12}$$

so that as the relative bubble size increases, a collapse becomes more likely. Consider the state C. In the collapsing regime C, b_{t+1} is expected to increase less than proportionately with b_t. To incorporate this van Norden and Schaller define

$$E_t[B_{t+1}|C] = u(b_t)P(t), \tag{B.13}$$

where $u(\cdot)$ is a continuous and everywhere differentiable function satisfying $u(0) = 0$ and $0 \leq u' \leq 1$. This means that

$$\begin{aligned}\frac{dE_t[b_{t+1}|C]}{db_t} &= u'(b_t), \\ &< 1.\end{aligned}$$

In order to retain (B.11), we must have

$$E_t[B_{t+1}|S] = \frac{MB_t}{q(b_t)} - \frac{1-q(b_t)}{q(b_t)}u(b_t)P_t. \tag{B.14}$$

Let R_{t+1} denote the returns given by $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t} = \frac{P^*_{t+1}+D_{t+1}+(P_{t+1}-P^*_{t+1})}{P_t}$. It follows that

$$E_t[R_{t+1}] = \frac{E_t[P^*_{t+1} + D_{t+1}]}{P_t} + \frac{E_t[B_{t+1}]}{P_t}.$$

Using the definition of fundamental price, it follows that

$$\begin{aligned}\frac{E_t[P^*_{t+1} + D_{t+1}]}{P_t} &= \frac{E_t[(1+\rho)D_{t+1}]}{P_t}, \\ &= \frac{D_t(1+\rho)e^{\alpha_0+\sigma^2/2}}{P_t}.\end{aligned} \tag{B.15}$$

From (B.15), it follows that

$$\begin{aligned}E_t[R_{t+1}|S] &= \frac{D_t(1+\rho)e^{\alpha_0+\sigma^2/2}}{P_t} + \frac{E_t[B_{t+1}|S]}{P_t}, \\ &= \frac{D_t(1+\rho)e^{\alpha_0+\sigma^2/2}}{P_t} + \frac{Mb_t}{q(b_t)} - \frac{(1-q(b_t))u(b_t)}{q(b_t)},\end{aligned} \tag{B.16}$$

since $b_t = \frac{B_t}{P_t}$. Now $P_t - B_t = P^*_t = \rho D_t$. So $\frac{D_t}{P_t} = \frac{1-b_t}{\rho}$. Inserting this into (B.16) gives

$$E_t[R_{t+1}|S] = \frac{(1+\rho)}{\rho}e^{\alpha_0+\sigma^2/2}(1-b_t) + \frac{Mb_t}{q(b_t)} - \frac{(1-q(b_t))u(b_t)}{q(b_t)}. \tag{B.17}$$

Similarly, $E[R_{t+1}|C]$ now becomes

$$\begin{aligned}E_[R_{t+1}|C] &= \frac{(1+\rho)}{\rho}e^{\alpha_0+\sigma^2/2}(1-b_t) + \frac{E[B_{t+1}|C]}{P_t}, \\ &= \frac{(1+\rho)}{\rho}e^{\alpha_0+\sigma^2/2}(1-b_t) + u(b_t).\end{aligned} \tag{B.18}$$

Since the previous expressions are all differentiable functions of b_t, a first-order linear approximation[1] enables us to write

$$\begin{aligned} E(R_{t+1}|S) &= \beta_{S0} + \beta_{Sb} b_t, \\ E(R_{t+1}|C) &= \beta_{C0} + \beta_{Cb} b_t, \\ q_{t+1} &= \beta_{q0} + \beta_{qb}|b_t|. \end{aligned} \quad \text{(B.19)}$$

This simple linear approximation enables us to say something about the coefficients in the model (B.19). We have

$$\frac{dE_t[R_{t+1}|C]}{db_t} = \frac{-(1+\rho)}{\rho} e^{\alpha_0+\sigma^2/2} + u'(b_t). \quad \text{(B.20)}$$

Since, by assumption, $u'(b_t) \leq 1$, in the financially reasonable case that $\alpha_0 > 0$ (corresponding to dividends that tend to grow over time), it follows that (B.20) must be negative, and hence that the coefficient $\beta_{Cb} < 0$ in (B.19). By construction, we must also have $\beta_{qb} < 0$. We also have

$$\begin{aligned} \left.\frac{dE_t[R_{t+1}|S]}{db_t}\right|_{b_t=b_0} &= \frac{-(1+\rho)}{\rho} e^{\alpha_0+\sigma^2/2} + \frac{M}{q(b_0)} - \frac{M b_t q'(b_0)}{q(b_0)^2} - u'(b_0)\left(\frac{1-q(b_0)}{q(b_0)}\right) + \frac{q'(b_0)u(b_0)}{q^2(b_0)}, \\ &= \left.\frac{dE_t[R_{t+1}|C]}{db_t}\right|_{b_t=b_0} + \frac{1}{q(b_0)}\left[M - u'(b_0)\right] + \frac{q'(b_0)}{q^2(b_0)}\left[u(b_0) - M b_0\right], \\ &= \beta_{Cb} + \frac{1}{q(b_0)}\left[M - u'(b_0)\right] + \frac{q'(b_0)}{q(b_0)^2}\left[u(b_0) - M b_0\right]. \end{aligned} \quad \text{(B.21)}$$

Under the financially reasonable assumption that $M > 1$, so that the rate of return on the fundamental is positive, the second term in (B.21) must be positive since $u'(b_0) < 1$ by assumption. Next we consider the third term in (B.21). By assumption, $q'(b_t) < 0$ for $b_t > 0$ and $q'(b_t) > 0$ for $b_t < 0$. We also have that $[u(b_t) - M b_t]' = [u'(b_t) - M]$, which is negative by the above, and in this case means that $[u(b_t) - M b_t]$ is a strictly decreasing function of b_t. We also have $u(0) = 0$ (by assumption), meaning that $[u(b_t) - M b_t]$ is positive for $b_t < 0$, and negative for $b_t > 0$. These considerations imply that the final term in (B.21) is also non-negative, and we are left with $\beta_{Sb} > \beta_{Cb}$, which is intuitive as this suggests higher returns when in the speculative phase. As a purely statistical formulation of the model, *[103]* use

$$\begin{aligned} R_{t+1}|S &= \beta_{S0} + \beta_{Sb} b_t + \epsilon_{S,t+1}, \\ R_{t+1}|C &= \beta_{C0} + \beta_{Cb} b_t + \epsilon_{C,t+1}, \\ q_{t+1} &= \Phi(\beta_{q0} + \beta_{qb}|b_t|), \end{aligned} \quad \text{(B.22)}$$

[1]corresponding to $f(b_t) \approx f(b_0) + f'(b_0)(b_t - b_0)$

where $\Phi(\cdot)$ denotes the standard normal CDF. This specification then ensures that the estimated q-values lie in $[0, 1]$. The $\epsilon_{S,t+1}$ and $\epsilon_{C,t+1}$ are assumed to be independent sequences of normal white-noise errors with variances σ_S^2 and σ_C^2 respectively.

Appendix C

Market-value models for banking crises

Models in this section make the fundamental assumption that the health of banking sectors (and of companies more generally) can be distilled from publicly available market prices. Typically, fairly general tools from mathematical finance are used with the aim of producing estimates of either ***distance to default*** or ***probability of default.*** These are used as suitable proxies which may be used to measure or predict crises. For illustration, we describe two approaches here. The first model used by *[27]* uses an option pricing approach, close in spirit to the seminal paper of *[78]*, to estimate distance to default. A second method, based on the Capital Asset Pricing Model, is used by *[24]* to construct estimates of the probability of default.

C.1 Merton's model

The asset value of the firm V is assumed to follow a geometric Brownian motion with drift equal to the risk-free interest rate r and volatility σ:

$$dV_t = V_t(rdt + \sigma dW_t), \tag{C.1}$$

where W is a standard Brownian motion. The firm defaults when its asset value at maturity V_T is less than or equal to its debt at maturity D. We introduce the distance to default d as a measure of default risk. From Itô's rule, we have

$$d = log(V_T) - log(D) = log(V_0) + (r - \frac{\sigma^2}{2}T + \sigma W_T - log(D)),$$

using the fact that in the Black-Scholes model $log(V_T) \sim N(log(V_0) + (r - \frac{\sigma^2}{2})T, \sigma^2 T)$ with $W_T \sim N(0, T)$. The normalized distance-to-default DD – which is more commonly used in practice – is defined as

$$DD = \frac{d}{\sigma\sqrt{T}} - \frac{W_T}{\sqrt{T}} = \frac{log(V_0/D) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}.$$

This renormalised distance to default DD is given by $\frac{\mu_d}{\sigma_d}$ (where μ_d and σ_d denote the mean and standard deviation of d) and is thus a measure of the number of standard deviations a firm is from default. Calculating the distance-to-default requires knowledge of the asset value and the asset volatility of the firm. However, if the face value of debt D and the maturity T are known, then the two unobservable variables can be calculated from the firm's equity value E and its volatility σ_E. These latter two variables can be measured empirically as functions of the asset value of the firm. Under Merton's model, the distance to default DD can be estimated from a series of nonlinear equations.

The value of the firm's equity at maturity T is given by

$$E_T = max(V_T - D, 0).$$

From this we see that the equity is equivalent to a European call option on the price with maturity T and strike D. Thus, starting at time t and maturing T periods later, the European call option formula gives

$$E_t = V_t\Phi(d_1) - D_t e^{-rT}\Phi(d_2). \tag{C.2}$$

V_t and D_t denote the company value and the value of debt at time t, $\Phi(\cdot)$ is the $N(0,1)$ c.d.f. and d_1 and d_2 are defined as

$$\begin{aligned} d_1 &= \frac{log\left(\frac{V_t}{D_t}\right) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, && \text{(C.3)} \\ d_2 &= d_1 - \sigma\sqrt{T}. && \text{(C.4)} \end{aligned}$$

Now equity is a function of asset value. Thus we can use Itô's lemma to calculate the instantaneous equity volatility $\sigma_{E,t}$. Suppose that X_t follows an Itô process of the form

$$dX_t = \mu(X_t, t)dt + X_t\sigma(X_t, t)dW_t,$$

with instantaneous volatility given by $\sigma(X_t, t)$. It follows from Itô's formula that the process $Y_t = G(X_t, t)$ has an instantaneous volatility defined by $Y_t\sigma(Y_t, t) = \frac{\partial Y_t}{\partial X_t}X_t\sigma(X_t, t)$. From this result we can see from (C.2) that we must have

$$E_t\sigma_{E,t} = \frac{\partial E_t}{\partial V_t}\sigma V_t = \Phi(d_1)\sigma V_t, \tag{C.5}$$

where $\sigma_{E,t}$ denotes the instantaneous equity volatility. This result can be seen as follows.

From (C.2) we have that

$$\frac{\partial E_t}{\partial V_t} = \Phi(d_1) + \frac{1}{\sigma\sqrt{T}}\left[\Phi'(d_1) - \frac{D_t e^{-rT}}{V_t}\Phi'(d_2)\right]$$

Completing the square, we can see that

$$\frac{\Phi'(d_1)}{\Phi'(d_2)} = e^{-1/2(d_1^2-d_2^2)} = \exp\{-1/2(d_1+d_2)(d_1-d_2)\}.$$

We have from (C.3-C.4) that

$$\begin{aligned} d_1 + d_2 &= \frac{2\left(log\left(\frac{V_t}{D_t}\right) + rT\right)}{\sigma\sqrt{T}}, \\ d_1 - d_2 &= \sigma\sqrt{T}. \end{aligned}$$

Combining, we see that

$$\frac{\Phi'(d_1)}{\Phi'(d_2)} = \frac{D_t e^{-rT}}{V_t},$$

from which it follows that

$$\Phi'(d_1)V_t - D_t e^{-rT}\Phi'(d_2) = 0.$$

Chan-Lau et al. use $T = 1$ year, the 1-year U.S. treasury yield as the proxy for the risk-free interest rate, and calculate equity volatility from the three-month moving average from daily equity data. The value of debt D_t at time t is interpolated from annual stock debt estimates. The set of nonlinear equations (C.2-C.5) can then be used to estimate the distance to default DD.

C.2 CAPM-based modelling approach

The model here is that given in used originally in *[49]* for individual banks, and by *[24]* to represent a particular country's banking sector. A representative bank has both assets and liabilities, and if we assume that these claims are all priced efficiently by the market, then the stock price S_t of the bank is given by

$$S_t = \frac{1}{N}\sum_{i=1}^{N} P_{it}X_{it},$$

where N is the number of stocks, P_{it} is the price of the bank's asset or liability i at time t, and X_{it} is the amount of the asset/liability at time t. We assume that the expected value of the stock in the future in conjunction with the variability of the underlying value around this expectation is informative about the probability of the bank actually defaulting. Conditional upon information available up to and including time $t-1$, the CAPM expresses the expected return of a stock, $E_{t-1}(R_t)$, as the the sum of a risk-free return RF_t (for example a treasury bill) plus a time-varying risk premium RP_t:

$$E_{t-1}(R_t) = \frac{E_{t-1}(S_t - S_{t-1})}{S_{t-1}} = RF_t + RP_t. \tag{C.6}$$

The risk premium can be thought of as representing the amount of risk that an investor has to be compensated for multiplied by the market price of this risk λ_t. According to the CAPM not all risk can be compensated for, and in equilibrium only non-diversifiable risk is priced. Thus, only non-diversifiable risk should be compensated for in the market by a higher return than the risk free return. If we denote the amount of expected non-diversifiable risk by $E_{t-1}(ND_t)$ we can write

$$R_t = RF_t + \lambda_t E_{t-1}(ND_t) + \epsilon_t, \tag{C.7}$$

where the ϵ_t is a zero-mean stochastic error term. From (C.6) and (C.7) we can see that the value of the bank capital at time t can be calculated as

$$S_t N = S_{t-1} N(1 + RF_t + \lambda_t E_{t-1}(ND_t) + \epsilon_t). \tag{C.8}$$

(C.8) can be decomposed into the sum of a deterministic average and a stochastic error term:

$$S_t N = E_{t-1}(S_t N) + S_{t-1} N \epsilon_t, \tag{C.9}$$

where the conditional mean is given by the sum of the first three terms in (C.8). The conditional variance is given by $(S_{t-1}N)^2 \sigma_{\epsilon_t}^2$, where σ_{ϵ_t} denotes the time-varying standard deviation of the innovations ϵ_t. Under the assumption of market efficiency, so that the time series $S_t N$ is a martingale, we can divide the the value of the bank $S_{t-1}N$ by the conditional standard deviation $S_{t-1}N\sigma_{\epsilon_t}$ to obtain

$$\frac{S_{t-1}N}{S_{t-1}N\sigma_{\epsilon_t}} = \frac{1}{\sigma_{\epsilon_t}}. \tag{C.10}$$

We see from (C.9) that (C.10) as a ratio of a conditional mean and a conditional standard deviation is a distance to default measure, giving the number of standard deviations that the bank is away from default (if we assume that default occurs when the value of the bank becomes negative). (C.10) can also be used to directly calculate probabilities of default under the assumption that ϵ_t is normal. In this case the probability of default at time t can be calculated as

$$Pr(\epsilon_t \leq -1) = 1 - \Phi\left(\frac{1}{\sigma_{\epsilon_t}}\right), \tag{C.11}$$

which is the probability that (C.9) becomes negative. We return to the definition of the CAPM model in order to show how we may obtain estimates of σ_{ϵ_t}. According to CAPM, see for example *[24]*, the expected returns can be written as

$$E_{t-1}(R_t) = RF_t + \beta_t E_{t-1}(RM_t - RF_t),$$

where RM_t is the return on the market portfolio and β_t is the expected conditional CAPM coefficient defined as

$$\beta_t = \frac{E_{t-1}(\sigma_{R_t,RM_t})}{E_{t-1}(\sigma^2_{RM_t})},$$

where $\sigma_{R_t,RM_t} = Cov(R_t, RM_t)$ and $\sigma^2_{RM_t} = Var(RM_t)$. Since the variance of the market portfolio corresponds to the amount of non-diversifiable risk, it follows from (C.7) that

$$E_{t-1}(RM_t) = RF_t + \lambda_t E_{t-1}(\sigma^2_{RM_t})).$$

We can rearrange this formula to provide an expression for the market price of risk:

$$\lambda_t = \frac{E_{t-1}(RM_t - RF_t)}{E_{t-1}(\sigma^2_{RM_t}))},$$

since the risk free return RF_t is regarded as non-stochastic. Rewriting (C.7) for the market portfolio we see that

$$RM_t = RF_t + \lambda_t E_{t-1}(\sigma^2_{RM_t}) + v_t = RF_t + \lambda_t E_{t-1}(\sigma^2_{v_t}) + v_t, \tag{C.12}$$

where v_t is a zero-mean random error term. Finally using the definition of the CAPM coefficient β_t we can see that the returns of the representative bank are given by

$$\begin{aligned} R_t &= RF_t + \frac{E_{t-1}(\sigma_{R_t,RM_t})E_{t-1}(RM_t - RF_t)}{E_{t-1}(\sigma^2_{RM_t})} + \epsilon_t \\ &= RF_t + \lambda_t E_{t-1}(\sigma_{R_t,RM_t}) + \epsilon_t, \\ &= RF_t + \lambda_t E_{t-1}(\sigma_{\epsilon_t,v_t}) + \epsilon_t. \end{aligned} \tag{C.13}$$

In sum, (C.12) and (C.13) give equations for time-varying expectations of variances and covariances which depend the time series behaviour of ϵ_t and v_t. For parsimony Byström uses a bivariate Garch-M (1, 1) process which allows for changes in the conditional mean to occur alongside changes in conditional variances and co-variances. *[24]* uses FTSE-All-World banking sector indices and FTSE-ALL World country indices to represent R_t and RM_t respectively, before using (C.11) to estimate default probabilities. We refer to *[24]* for full details.

Bibliography

[1] A. Admati and P. Pfleiderer. A theory of intraday patterns; volume and price variability. *Review of Financial Studies*, 1:3–40, 1988.

[2] J. V. Andersen and D. Sornette. Fearless versus fearful speculative financial bubbles. *Physica A*, 337:565–585, 2004.

[3] T. G. Andersen, T. Bollerslev, F. X. Diebold, and C. Vega. Micro effects of macro announcements. Real-time price discovery in foreign exchange. *American Economic Review*, 93:38–62, 2003.

[4] J. C. Anifrani, C. Le Flo'ch, D. Sornette, and B. Souillard. Universal log-periodic correction to renormalization group scaling for rupture stress prediction from accoustic emissions. *J. Phys. I France*, 5:631–638, 1995.

[5] D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge, 2004.

[6] G. Arias. *Currency crises: what we know and what we still need to know.* C. E. F. I. Working Paper No. 2003/13, 2003.

[7] S. Asmussen. *Applied probability and queues.* Wiley, Chichester, 1987.

[8] E. Bacry, J-. F. Muzy, and J. Delour. Multifractal random walk. *Physical Review E*, 64:026103, 2001.

[9] P. Ball. *Critical mass: how one thing leads to another.* Arrow Books, London, 2005.

[10] O. E. Barndorff-Nielsen. Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics*, pages 151–157, 1978.

[11] O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 1:41–68, 1998.

[12] O. E. Barndorff-Nielsen and K. Prause. Apparent scaling. *Finance and stochastics*, pages 199–207, 2001.

[13] D. M. Bates and D. G. Watts. *Nonlinear regression analysis and its applications.* Wiley, New York Chichester Brisbane Toronto Singapore, 1988.

[14] C. Bauer. Value at risk using hyperbolic distributions. *Journal of Economics and Business*, 52:455–467, 2000.

[15] B. J. Baxter. *Exactly solved models in statistical mechanics.* Academic Press, 1989.

[16] N. H. Bingham. On the limit of a supercritical branching process. *Journal of Applied Probability*, pages 215–228, 1988.

[17] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation.* Cambridge University Press, Cambridge, 1987.

[18] N. H. Bingham and R. Kiesel. Semi-parametric modelling in finance: theoretical foundations. *Quantative Finance*, 2:241–250, 2002.

[19] N. H. Bingham, R. Kiesel, and R. Schmidt. A semi-parametric approach to risk management. *Quantative Finance*, 3:426–441, 2003.

[20] P. Blæsild. The two-dimensional hyperbolic distribution and related distributions, with an application to Johannsen's bean data. *Biometrika*, 68:251–263, 1981.

[21] M. Boldrin and D. K. Levine. Growth cycles and market crashes. *Journal of Economic Theory*, 96:13–39, 2001.

[22] J-P. Bouchaud and M. Potters. *Theory of financial risk and derivative pricing. From statistical physics to risk management. Second edition.* Cambridge University Press, Cambridge, 2003.

[23] C. Brooks and A. Katsaris. A three-regime model of speculative behaviour: Modelling the evolution of the S&P 500 Composite Index . *The Economic Journal*, 115:767–797, 2005.

[24] H. N. E. Byström. *The market's view on the probability of banking sector failure: Cross-country comparisons.* Working Papers 2003:2, Lund University, Department of Economics, 2003.

[25] J. Campbell and R. Shiller. Cointegration and tests of present value models. *Journal of Political Economy*, 61:1062–1088, 1987.

[26] J. Y. Campbell, A. W. Lo, and A. C. Mackinlay. *The econometrics of financial markets.* Princeton University Press, Princeton, 1997.

[27] J. A. Chan-Lau, A. Jobert, and J. Kong. *An option-based approach to Bank vulnerabilities in emerging markets.* IMF Working Paper No. 04/33, 2004.

[28] C. Chang and J. Feigenbaum. A Bayesian analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 6:15–36, 2006.

[29] D. Cochrane and G. H. Orcutt. Application of least-squares regression to relationships containing autocorrelated error terms. *Journal of the American Statistical Association*, 44:32–61, 1949.

[30] R. Cont and J-P. Bouchaud. Herd behaviour and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4:170–196, 2000.

[31] R. Cont and P. Tankov. *Financial modelling with jump processes.* Chapman and Hall/CRC, Boca Raton London Hiedelberg London Washington D.C., 2004.

[32] B. Derrida, L. De Seze, and C. Itzykon. Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics*, 33:559–569, 1983.

[33] E. Eberlein. Applications of generalized hyperbolic Lévy motions to finance. *Lévy Processes: Theory and Applications ed. O. E. Barndorff-Nielsen, T. Mikosch and S. Resnick (Birkhäuser, Boston, MA)*, pages 319–337, 1999.

[34] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, pages 281–299, 1995.

[35] E. Eberlein, U. Keller, and K. Prause. New insights into smile, mispricing, and value at risk: the hyperbolic model. *Journal of Business*, pages 371–405, 1998.

[36] S. Edwards. *Interest rate volatility, capital controls and contagion.* NBER Working Paper No. 6756, 1998.

[37] B. Eichengreen, A. Rose, and C. Wyplosz. *Contagious currency crises.* NBER Working Paper No. 5681, 1997.

[38] J. Feigenbaum. Financial physics. *Reports on Progress in Physics*, 66:1611–1649, 2003.

[39] J.A. Feigenbaum. A statistical analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 1:346–360, 2001a.

[40] J.A. Feigenbaum. More on a statistical analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 1:527–538, 2001b.

[41] J.A. Feigenbaum and P. G. O. Freund. Discrete scale invariance in stock markets before crashes. *International Journal of Modern Physics B*, 10:3737–3745, 1996.

[42] W. Feller. *An introduction to probabaility theory and its applications. Volume II. Second Edition.* Wiley, New York London Sydney Toronto, 1970.

[43] K. J. Forbes and R. Rigobon. No contagion, only interdependence: Measuring stock market co-movements. *Journal of Finance*, 57:2223–2261, 2002.

[44] M. Glazer and J. Wark. *Statistical mechanics. A survival guide.* Oxford University Press, Oxford, 2001.

[45] N. Goldenfeld. *Lectures on phase transitions and the renormalization group .* Addison-Wlesey, 1991.

[46] G Grimmett and D. Stirzaker. *Probability and Random Processes. Third Edition.* Oxford University Press, Oxford, 2001.

[47] O. Häggström and J. Jonasson. Phase transition in the random triangle model. *Journal of Applied Probability*, 36:1101–1115, 1999.

[48] H. Haken. *Synergetics.* Springer-Verlag, Berlin Hiedelberg New York, 1978.

[49] S. G. Hall and D. K. Miles. *Measuring the risk of financial institution's portfolios: Some suggestions for alternative techniques using share prices.* In Henry, S. G. B. and Pattersen, K. D. (eds.) *Economic modelling and policymaking in banking and finance.* Chapman and Hall, 1990.

[50] H. J. Jensen. *Self-organized criticality.* Cambridge Lecture Notes in Physics, Cambridge University Press, Cambridge, 1998.

[51] A. Jobert, A. Platania, and L. C. G. Rogers. *A Bayesian solution to the equity premium puzzle.* Preprint. *http://www.statslab.cam.ac.uk/~chris/*, 2006.

[52] A. Johansen, O. Ledoit, and D. Sornette. Crashes as critical points. *International Journal of Theoretical and Applied Finance*, 3:219–255, 2000.

[53] A. Johansen and D. Sornette. Modelling the stock market prior to large crashes. *European Physical Journal B*, pages 167–174, 1999.

[54] A. Johansen and D. Sornette. Download relaxation dynamics on the WWW following newspaper publication of URL. *Physica A*, 276:338–345, 2000.

[55] A. Johansen and D. Sornette. Large stock market drawdowns are outliers. *Journal of Risk*, 4:69–110, 2001.

[56] A. Johansen and D. Sornette. Endogeneous versus exogeneous crashes in financial markets. in press in "*Contemporary Issues in International Finance*" (Nova Science Publishers, 2004) (http://arXiv.org/abs/cond-mat/0210509), 2004.

[57] T. Jokipii and B. Lucey. *CEE banking sector co-movement: contagion or interdependence?* IIIS Discussion Paper No. 77, 2005.

[58] G. Kaminsky and C. Reinhart. The twin crises: The cause of banking and balance-of-payments problems. *American Economic Association*, 89:473–500, 1999.

[59] G. Kaminsky and C. Reinhart. On crises, contagion and confusion. *Journal of International Economics*, 51:145–168, 2000.

[60] A. Kirman and Teyssière. Testing for bubbles and change-points. *Journal of Economic Dynamics and Control*, 29:765–799, 2005.

[61] E. Kreyszig. *Advanced Engineering Mathematics. Eighth Edition.* Wiley, New York Chichester Brisbane Toronto Singapore, 1999.

[62] P. Krugman. *Crises: the next generation?* In Helpman, E. and Sadka, E. (eds.) *Economic Policy in the International Economy: Essays in Honor of Assaf Razin pp 15-32.* Cambridge, University Press, Cambridge, 2003.

[63] A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53:1315–1355, 1985.

[64] L. Laloux, M. Potters, R. Cont, J-P. Aguilar, and J-P. Bouchard. *Are financial crashes predictable?* Available at http://arXiv:cond-mat/9804111v1, 1998.

[65] L. Laloux, M. Potters, R. Cont, J.-P. Aguilar, and J.-P Bouchaud. Are Financial Crashes Predictable? *Europhysics Letters*, 45:1–5, 1999.

[66] L. D. Landau and E. M. Lifshitz. *Statistical Physics, Part 1 Third Edition.* Pergammon, London, 1980.

[67] I. H. Lee. Market crashes and informational avalanches. *Review of Economic Studies*, 65:741–759, 1998.

[68] O. Linton. *Financial applications of tests for ellipsoidal symmetry.* Work in progress, Economics Department, LSE, 2004.

[69] R. E. Jr. Lucas. Asset prices in an exchange economy. *Econometrica*, 66:429–445, 1978.

[70] M. Maggiore. *A modern introduction to quantum field theory.* Oxford University Press, Oxford, 2004.

[71] Y. Malevergne and D. Sornette. *Extreme financial risks.* Springer, Berlin Hiedelberg New York, 2005.

[72] B. B. Mandelbrot. *Fractals and scaling in finance: Discontinuity, concentration and risk. Selected Works Volume E.* Springer, 1997.

[73] B. B. Mandelbrot. The variation of certain speculative assets. *Journal of Business*, XXXVI:392–417, 2000.

[74] M. Mandjes and J. H. Kim. An analysis of the phase transition phenomenon in packet networks. *Advances in Applied Probability*, 33:260–280, 2001.

[75] R. Marschinski and L. Matassini. *Financial markets as a complex system: A short time scale perspective.* Deutsche Bank Research Notes in Economics and Statistics, 2001.

[76] B. McCoy and T. S. Wu. *The Two-dimensional Ising model.* Harvard University Press, Cambridge, Massachussets, 1973.

[77] A. J. McNeil, R. Frey, and P. Embrechts. *Quantative risk management.* Princeton University Press, Princeton, 2005.

[78] R. C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.

[79] J-. F. Muzy, J. Delour, and E. Bacry. Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *The European Physical Journal B*, 17:537–548, 2000.

[80] K. Nakagawa. Application of Tauberian theorem to the exponential decay of the tail probability of a random variable. *IEEE transactions on information theory*, 53:3239–3249, 2007.

[81] J. A. Nelder and R. Mead. A simplex algorithm for function minimization. *Computer Journal*, 7:308–313, 1965.

[82] R. B. Nelsen. *An introduction to copulas.* Springer, 1999.

[83] B. Øksendal. *Stochastic differential equations: An introduction with applications. Sixth edition.* Springer, Berlin Hiedelberg New York, 2003.

[84] E. Osborne. Financial crashes in the globalization era. *The Independent Review*, 6:165–184, 2001.

[85] W. Paul and J. Baschnagel. *Stochastic processes: From physics to finance.* Springer-Verlag, Berlin Hiedelberg New York, 2000.

[86] Peralta-Ata. Comment on "Growth cycles and market crashes". *Journal of Economic Theory*, 111:147–148, 2003.

[87] W.H. Press, B.P. Flannery, S.A. Teukovsky, and W.T. Vetterling. *Numerical recipes.* Cambridege University Press, Cambridge, 1986.

[88] G. A. F. Seber and C. J. Wild. *Nonlinear regression.* Wiley, 1989.

[89] R. J. Shiller. *Irrational Exuberance.* Princeton University Press, Princeton, 2000.

[90] D. Sornette. Discrete scale invariance and complex dimensions. *Physics Reports*, 297:239–270, 1998.

[91] D. Sornette. Stock market speculation: Spontaneous symmetry breaking of economic valuation. *Physica A*, 284:355–375, 2000.

[92] D. Sornette. Critical market crashes. *Physics Reports*, 378:1–98, 2003.

[93] D. Sornette. *Why stock markets crash. Critical events in complex financial systems.* Princeton University Press, Princeton, 2004.

[94] D. Sornette and J. V. Andersen. A nonlinear super-exponential rational model of speculative financial bubbles. *International Journal of Modern Physics C*, 17:171–188, 2002.

[95] D. Sornette and A. Helmsetter. Endogeneous versus exogeneous shocks in systems with memory. *Physica A*, 318:577–591, 2003.

[96] D. Sornette and A. Johansen. Large financial crshes. *Physica A*, 245:411–422, 1997.

[97] D. Sornette and A. Johansen. Significance of log-periodic precursors to financial crashes. *Quantative Finance*, 1:452–471, 2001.

[98] D. Sornette, A. Johansen, and J-P. Bouchaud. Stock market crashes, precursors and eeplicas. *J. Phys. I France*, 6:167–175, 1996.

[99] D. Sornette and Y. Malevergne. From rational bubbles to crashes. *Physica A*, 299:40–59, 2001.

[100] D. Sornette, Y. Malevergne, and J-F. Muzy. Volatility fingerprints of large shocks: endogeneous versus exogeneous. *Applications of Econophysics, Proceedings of the second Nikkei symposium on econophysics*, H. Takayasu, ed., Springer Verlag, ISBN 4-431-14028-X, 2004.

[101] D. Stauffer and A. Aharony. *Introduction to percolation.* Taylor-Francis, London, 1994.

[102] C. Tsuji. Is volatility the best predictor of market crashes? *Asia-Pacific Financial Markets*, 10:163–185, 2003.

[103] S. van Norden and S. Schaller. *'Speculative behaviour, regime-switching, and stock market crashes' in (P. Rothman, ed.),* Nonlinear time series analysis of economic and financial data. *pp321-356* . Dordtrecht Kluwer, 1999.

[104] W.N. Venables and B.D. Ripley. *Modern applied statistics with S-Plus. Second edition.* Springer-Verlag, New York, 1998.

[105] J. Voit. *The statistical mechanics of financial markets. Third edition.* Springer, Berlin Hiedelberg New York, 2006.

[106] G. N. Watson. *A treatise on the theory of Bessel functions. Second Edition.* Cambridge University Press, Cambridge, 1944.

[107] J. M. Yeomans. *Statistical mechanics of phase transitions.* Oxford University Press, Oxford, 1992.

[108] J. Zeira. Informational overshooting, booms, and crashes. *Journal of Monetary Economics*, 43:237–257, 1999.

[109] W-X. Zhou and D. Sornette. Fundamental factors versus herding in the 2000-2005 US stock market and prediction. *Physica A*, 360:459–482, 2006.