Bubbles, Can We Spot Them? Crashes, Can We Predict Them?

Xiao Qin
Division of Economics
School of Humanities and Social Sciences
Nanyang Technological University of Singapore
cqxiao@ntu.edu.sg

Tan Gee Kwang, Randolph
Division of Economics
School of Humanities and Social Sciences
Nanyang Technological University of Singapore
arandolph@ntu.edu.sg

Abstract:

Johansen and Sornette proposes that the crash has fundamentally an endogenous origin and exogenous shocks only serve as triggering factors. This endogenous force is shown in price as power law log-periodicity (PLLP) signature prior to a crash. We estimate the highly nonlinear model developed by them using a hybrid approach which combines scatter search, genetic adaptor and tabu search. The model is applied to two property data sets (Hong Kong Office Price Index and Seoul Hosing Price Index) and one property related stock price (Korea General Construction Stock Price Index). The fitting of the original model to these data sets was unsuccessful, due to the lack of the power law. We hence fit the data using a modified model, and the results are encouraging when crash-date prediction is the aim.

1. Introduction

The burst of fully-fledged bubbles is often cited as an important cause of large market crashes. A market crash is defined as a significant sudden decline in price over a short time span, which the world has seen a number of times in the past two decades in all major asset markets, such as stock, property and commodity markets.

A market crash is often epidemic and destructive to the economy at large. If market crashes are indeed, in many cases, caused by the bursts of speculative bubbles, understanding the behavior of the speculative bubbles and the relationship between bubbles and market crashes will help the policy makers to minimize the damage of speculative bubbles to economy at large.

In recent years, Johansen, Sornette and their co-authors (J&S et al) developed an interesting analytical framework for market crash. Their theory has a close connection with a class of bubble postulated by Blanchard and Watson (1982) which collapses periodically, specifically, with their idea that the probability that the bubble ends may depend on how long the bubble has lasted, or by how far the price is from market fundamentals.

The J&S et al framework has its deep roots in the recent findings of the behavior economics and the concept of criticality evolved in statistical physics. They propose that market crashes are large draw-downs which occur because the market has entered an unstable phase and any small disturbance or process may have triggered the instability. The collapse is fundamental due to the unstable position. The instantaneous cause of the collapse is secondary. In another word, the crash has fundamentally an endogenous origin, and exogenous shocks only serve as triggering factors. The origin of crashes is constructed progressively by the market as a whole, as a *self-organizing* process. In this sense, the true cause of a crash could be termed a systemic instability.

Using the model developed, J&S et al identified a precursory pattern, namely, the power law log periodic signature (PLLP), of asset prices prior to major crashes. In their papers, they verified the presence of PLLP for different asset markets in many developed countries and some emerging economies. J&S et al argue that this precursory pattern originates from some very fundamental and robust properties of asset markets: the herding behavior among the traders and the self-organizing markets, which lead to accelerating speculative bubbles that often end in crashes.

At the back of “Why Stock Markets Crash” (D. Sornette, 2003), Robert Shiller writes “a professor of geophysics gives a very different perspective, informed by his scientific training, on the stock market. I am sure that his view will be highly controversial, but the book is fascinating, and mind-expanding, reading.”

In this paper, we apply the methodologies of J&S et al to Seoul Housing Price (SHP) and Hong Kong Office Price (HKOP). As SHP show slow decaying throughout the 1990s rather than large crashes of the kind experienced by HKOP (Figure 5.1), we also include a third time series, namely Korea General Construction Stock Price (KGCSP) in this paper (Figure 7.1).

The remaining of this paper includes five parts. In part 2 we layout the economic rational behind the J&S et al model, which is followed by a summary of the model in part 3. The details of estimation and optimization strategies are described in part 4. Our empirical findings, the main contributions of this paper, are presented in part 5. This is followed by concluding remarks in part 6.

2. The Economic Rational

2.1. Self-organization, Market Efficiency and Speculative bubble

Macroscopic systemic self-f-organization can emerge from some simple rules of repeated actions at the microscopic level. This idea follows Adam Smith's notion that selfish, greedy individuals, if allowed to pursue their interests largely unchecked, would interact to produce a wealthier society as if guided by an "invisible hand." Smith himself never worked out a proof that this invisible hand existed. The mathematical proving of the existence of invisible hand was carried out by Arrow and Debreu (1954) under a set of very restricted assumptions.

The main tool in the analysis of Arrow and Debreu is constrained optimization. However, this is not an entirely satisfactory representation of reality, as most people are not versed in economic optimization reasoning.

But this does not mean we shall fail to function effectively in social and economic exchanges in life. The remarkable insight of Adam Smith is that people have natural intuitive mechanisms enabling them to "read" situations and the intentions and the likely reactions of others without deep, tutored cognitive analysis.

This fact has been established by experimental economists. Their experiments show that economic agents can achieve efficient outcomes that are not part of their intentions---a key principle formulated by Adam Smith. Experiments on markets with both insiders and uninformed traders show that equilibrium prices do reveal insider information after several trials of the experiments, suggesting that the markets disseminate information efficiently.

However, these results are not always present if the following conditions are not fulfilled: identical preferences, common knowledge of the dividend structure, and complete contingent claims (i.e. existence of a full spectrum of derivative instruments allowing one to probe the expectation of future risks). Under these situations, information aggregation is a more complicated process, and market efficiency, defined as full information aggregation, depends on the complexity of the market.

Thus the "emergence of self-organization" does not imply that the market will always be equivalent to an efficient and global optimization machine. Empirical economics shows that market forces may lead to plenty of imperfections, problems and paradoxes. In fact, rational behavior could lead to less-than-optimal market outcomes, such that the creation of rational speculative bubbles that give false signals on the fundamental values of assets.

2.2. The Power Law

The power law refers to the power law acceleration of market price prior to a crash.

Consider a purely speculative asset which pays no dividends. The no-arbitrage condition plus rational expectation would imply that the price of this asset should be zero at all time. Any deviation of the price from the zero value signals the presence of speculative bubble.

A speculative bubble emerges from "self-reinforcing imitation" among traders in a self-organizing process. However, fundamental forces would make bubbles transient phenomena. The fighting between a bubble and the fundamental forces leads to repeated price fluctuations around its fundamental values.

The self-reinforcing imitation process leads to the blossoming of a bubble which often (but not necessarily) ends in a market crash.

A crash is not certain but can be characterized by its hazard rate: the probability per unit time that the crash will happen in the next instant provided it has not happened yet. The crash hazard rate quantifies the probability that a large group of traders place sell orders simultaneously and creates enough imbalances in the order book of market makers to be unable to absorb the other side without lowering prices substantially. It is computed as the ratio of scenarios that give crash to all possible scenarios.

Since crash is not a certain deterministic outcome of the bubble---there is a finite probability of “landing softly”---it remains rational for the traders to stay invested provided they are compensated by higher rate of growth of the bubble for taking the risk of a crash.

This means that the *critical time,* which is defined as the time when a bubble ends, is not the time of the crash, but the most probable time for the crash to occur.

Assuming the movement of a price is driven by risks. In particular, for each period (e.g. a day), there are two components and only two compete to determine the price increment from one day to the next: A daily market return and the possibility that a crash will occur.

With the no-arbitrage condition and rational expectation, the daily return should compensate exactly the average loss due to the possibility of a crash. It implies that the total average return at any time is exactly zero. It also implies that market return is proportional to the crash hazard rate: the higher the risk of a crash, the higher the price return.

Most traders in the world are organized into a network and they influence each other locally through this network. Imitation among traders (herding behavior) creates a speculative bubble. The same imitation force also brings the end to the bubble.

On approaching the critical time, the time when the bubble ends, imitation among traders strengthens in an accelerating manner (fueled by expected large capital gains, perhaps). The market becomes more and more sensitive to news or rumors. These lead to the fueling up of the crash hazard rate, which in turn leads to the *power law* acceleration of price upon the approaching of the critical time.

The above paragraphs lay out the "risk-driven model" in the terminology of J&S et al. We may, however, assume price driving risks rather that the other way round. They we would have a "price-driven model" according to J&S et al.

In the price-driven model, the price drives the crash hazard rate. The price itself is driven up by the imitation and herding behavior of the "noisy" investors. The occurrence of a crash is again characterized by its hazard rate.

Let the price variation in an elementary time period to be the sum of two components: a certain instantaneous return and a random return. The first embodies the remuneration due to estimated risks as well as the effect of imitation and herding. The second embodies the noise component of the price dynamics with volatility. The volatility can also have a systematic component controlled by imitation as well as many other factors.

Noisy investors (as opposed to investors who base their investment strategy on fundamentals) look at the market price going up, they speak to each other, develop herding, buy more and more of the stock, thus pushing prices further up. As the price variation speeds up, the no-arbitrage condition together with rational expectations implies that there must be an underlying risk, not yet revealed in the price dynamics, which justifies this apparent free ride and free lunch.

To capture the phenomenon of speculative bubbles, we focus on the class of models with positive feedbacks. In the present context, this means that the instantaneous return as well as the volatility becomes larger and larger when past prices and/or past returns and/or past volatilities become large. Such positive feed backs with increasing growth rate may lead to singularities in a finite time[i]. Here it means that, unchecked, the price would blow up without bounds. However, two effects compete to tamper with this divergence. First, the stochastic component impacting the price variations makes the price much more erratic, and the convergence to the critical time becomes a random, uncertain event. The second effect that tampers with the possible divergence of the bubble price is the impact of the price on the crash hazard rate: as the price blows up due to imitation, herding, speculation and randomness, the crash hazard rate increases even faster, so that a crash will occur and drive the price back closer to its fundamental value. Hence this model proposes two scenarios for the end of a bubble: either a spontaneous deflation or a crash.

The risk-driven model and the price-driven model describe a system of two populations of traders, the ‘rational’ and the ‘noisy’ traders. Occasional imitative and herding behaviors of the noisy traders may cause global cooperation among traders, causing a crash.

In the risk-driven model, the crash hazard rate determined from herding drives the bubble price.

In the price-driven model, imitation and herding induce positive feedbacks on the price, which itself create an increasing risk for a looming yet unrealized financial crash.

Both models capture a part of reality. Studying them independently is the standard strategy of dividing-to-conquer the complexity of the world. Both models embody the notion that the market anticipates the crash in subtle, self-organized and cooperative fashion, hence releasing precursory power law ‘fingerprints’ observable in stock market prices.

2.3. The Log-Periodicity

In the previous section we learnt *a critical point* in the time domain underlies stock market crashes. A crash is not the critical point itself, but its triggering rate is strongly influenced by the proximity to the critical point: the closer to the critical time, the more probable is the crash.

The hallmark of critical behavior is a *power law acceleration* of the price, its volatility, and the crash hazard rate, upon approaching the critical time. However in practice, due to the presence of the noise and the irregularities of the trajectories of stock market prices, power law acceleration is often difficult to detect.

Luckily, the looming of a crash has a second fingerprint: *log-periodicity*, which is more robust in the presence of noise. That is, in the presence of noise, the price is not monotonously accelerating, but rather the pattern of accelerating is decorated by oscillations whose frequency itself accelerates on approaching the critical time. This oscillation is called “log-periodic oscillation” as it is seen as accelerating periodic oscillation in a logarithmic representation (Sornette, 2003, page 179, Figure 6.4).

This means crash hazard rate and price increases dramatically when the interaction between investors becomes strong enough, but this acceleration is interrupted by and mixed with an accelerating sequence of quiescent phases in which the risk and price decreases.

The power law and log-periodicity signature may be explained by the interaction between trend-following traders and fundamentalists. The trend-following traders exerts positive feed back on prices and enhances the previous price trend, leading to exponential growth of the price and resulting in price to exhibit finite time singularity. The fundamentalists exert fundamental value restoring force which generates oscillations around the fundamental value which are approximately log-periodic.

3. The Model

In this section, we summarize the derivation of the model developed by J&S et al.

3.1. The Price Dynamics

Consider a purely speculative asset with no dividends payment. Ignore the interest rate, risk aversion, information asymmetry and market-clear condition. Assume markets are efficient in the sense that all available information is reflected in current market prices. Then rational expectation implies that the price follows a martingale process:

$$E\left[p(t')\right] = p(t) \qquad \forall t' > t \qquad \textbf{Equation 7.1}$$

Where $p(t)$ is the price of the asset at time t and $E_t[\cdot]$ the expectation operator conditional on information revealed up to time t. In a market without noise,

$$p(t) = p(t_0) = 0 \qquad \forall t \qquad \textbf{Equation 7.2}$$

t_0 is some initial time. A positive value of $p(t)$ would constitute a speculative bubble[ii].

Crash is likely to occur if bubble blows up the price too high. The probability of crash is characterized by the crash hazard rate, $h(t)$, defined as the probability per unit time that the crash will happen in the next instant if it has not yet happened, and $h(t) = q(t)/(1-Q(t))$, where $Q(t)$ is the cumulative probability density of the crash. The probability density is $q(t) = \frac{dQ}{dt}$. The crash is an exogenous event. Similarly, the crash hazard rate is an exogenous variable.

Define a jump process, denoted by j, with $j = 0$ before the crash, and $j = 1$ after the crash. Assume when crash occurs, price $p(t)$ drop by a fixed fraction κ, with $\kappa \in (0,1)$. Then the dynamics of asset price before the crash are governed by

$$dp = \mu(t)p(t)dt - \kappa p(t)dj$$ **Equation 7.3**

where $\mu(t)$ is a time dependent drift satisfying the martingale condition

$$E_t[dp] = \mu(t)p(t)dt - \kappa p(t)h(t)dt = 0$$ **Equation 7.4**

which implies $\mu(t) = \kappa E_t\left[\frac{dj}{dt}\right] = \kappa h(t)$. Plugging it into equation 7.3 yield

$$dp(t) = \kappa h(t)p(t)dt - \kappa p(t)dj$$ **Equation 7.5**

Since $j = 0$ before the crash, we have $dp(t) = \kappa h(t)p(t)dt$ before the crash. Hence before the crash

$$\ln\left(\frac{p(t)}{p(t_0)}\right) = \kappa \int_{t_0}^{t} h(t')dt'$$ **Equation 7.6**

The larger the value of $h(t)$, the higher the probability of crash, the faster the price must increase---investors must be compensated by the chance of a higher return in order to be induced to hold an asset that might crash.

3.2. The Crash

Equation 7.6 states that the evolution of price is pending on that of the crash hazard rate. The macro-level probability of crash, in turn, results from micro-level agents interactions, in particular, the interplay of imitation and anti-imitations among agents.

Traders in reality are organized into a network of family, friends, colleagues, etc. They influence each other locally through this network. Consider a network of investors: each one can be named by an integer i=1,2,…,I. A typical trader i has $N(i)$ neighbors. His opinion is influenced by the opinions of these neighbors as well as an idiosyncratic signal that this trader alone receives. The first force will tend to create order while the latter create disorder. A crash happens when order wins. Thus, the macro-level coordinated sells, which causes crash, is a result of micro-level imitation among agents.

Then what determines whether order or disorder wins?

Assume that agent i can be in only one of two possible states: $s=+1$ if he buys and $s=-1$ if he sells. Based on the information of the actions $s_j(t-1), j=1,2,...$, performed at time $t-1$ by her $N(i)$ "neighbors," Anne set $s_i(t-1)$ to maximizes her return.

Assume each agent can either buy or sell only one unit of the asset. The selection of one of the two states (buy or sell) is determined from small and subtle initial biases as well as from the fluctuations during the evolutionary dynamics.

The asset price variation is thus proportional to the aggregate sum $\sum_{i-=1}^{I} s_i(t-1)$ of all traders' actions, e.g. the sum is zero if there are as many buyers as there are sellers and the price does not change. Other influences impacting the price are accounted for by adding a stochastic component.

AT time $t-1$, when the price $p(t-1)$ has been announced, trader i defines her strategy $s_i(t-1)$ based on information available to maximizes her expected profit $P_E = Ep(t) - p(t-1)$. Since the price moves with the general opinion, the best strategy is to buy if $\sum_{i=1}^{I} s_i(t-1)$ is positive and sell if it is negative. However, $\sum_{i=1}^{I} s_i(t-1)$ is unknown to a given trader. The best the trader can do is to poll the opinions of his immediate $N(i)$ neighbor. Suppose the a priori probability $\Pr+$ and $\Pr-$ for each trader to buy or sell is known to all. From all these information the trader can construct his prediction of the price drift. The best guess of trader i is that the future price change will be proportional to the sum of the actions of her neighbors who she has been able to poll. Thus the strategy that maximizes his expected profit is:

$$s_i(t-1) = sign(K \sum_{j \in N(i)} s_j(t-1)) + \sigma\varepsilon_i + G \qquad \textbf{Equation 7.7}$$

where ε_i is the noise, with $\varepsilon_i \xrightarrow{d} i.i.d.N(0,1)$, and $N(i)$ is the number of neighbors with whom trader i interacts significantly. K is a positive constant measuring the strength of imitation[iii]. It is inversely proportional to the "market depth": the larger the market, the smaller the relative impact of a given unbalance between buy and sell orders, hence the smaller is the price change. The tendency towards idiosyncratic behavior is governed by σ. Thus the relative value of K to σ determines the outcome of the battle between order and disorder, and eventually the probability of a crash. G captures the global influence which tend to favor sate $+1(-1)$ if $G > 0(G < 0)$.

Equation 7.7 only describes the state of an agent at a given time. In the next instant, new ε_i's are realized, new influences propagate themselves to neighbors, and agents can change their decision. The system is thus constantly changing and reorganizing.

In a practical implementation of a trading strategy, it is not sufficient to know or guess the overall direction of the market. A trader may want to be slightly ahead of the herd to buy at a better price, before the price is pushed up for the bullish consensus. Symmetrically, she will want to exit the market a bit before the crowd.

Real markets result from agents' behaviors, which are neither fully imitative nor fully anti-imitative. A better representation of real markets requires a combination of the two. Indeed, one should distinguish the "buy" and "sell" actions from the "holding" period. In general, a typical trader would ideally like to be in the minority when entering the market, in the majority while holding her position, and again in the minority when closing her position.

Traders will try to out guess each other on when to enter the market. If all traders use the same set of decision rules, they will end up doing the same thing at the same time and cannot therefore be in the minority. To be in the minority implies striving to be different. By adaptation, traders will learn and be forced to differentiate their entry and exit strategies based on past successes and failures.

The interaction between the forces of imitation and the forces of anti-imitation is the key to understanding market crashes.

The chance that a large group of agents find themselves in agreement is called the *susceptibility of the system*. Define the average state of the system as $M = \frac{1}{I}\sum_{i=1}^{I} s_i$. In the absence of the global influence, $E[M] = 0$: agents are evenly split between the two states. In the presence of a positive (negative) global influence, agents in the positive (negative) state will outnumber the others: $E[M] \times G > 0 \quad (E[M] \times G < 0)$. Hence the system susceptibility is defined formally as

$$\chi = \left. \frac{d(E[M])}{dG} \right|_{G=0}$$ **Equation 7.8**

That is the susceptibility measures the sensitivity of the average state of the system to a tiny global influence, hence the degree of coordination of the overall system. The susceptibility can also be interpreted as the variance of the average state M around its zero mean caused by the random idiosyncratic shocks followed by imitation. *It is precisely the emergence of this global synchronization from local imitation that can cause a crash*.

The susceptibility depends on the structure of the network and the strength of imitation.

Let t_c denote the first time the imitation strength reached its critical value. That is for the first time $K = K_c$. Notice t_c is *not* the time of the crash but the time the crash is most likely. When $K < K_c$, disorder reigns: the sensitivity of the system to a small global influence is small, the clusters of agents who are in agreement remain of small size, and imitation only propagates between close neighbors. In this case, the susceptibility χ of the system to external news is small, as many clusters of different opinions react incoherently, thus more or less canceling out their responses.

When the imitation strength K increases and approaches K_c, order starts to appear: the system becomes extremely sensitive to a small global perturbation, agents who agree with each other form large clusters, and imitation propagates over long distances. These are the characteristics of so-called *critical phenomena* in natural sciences[iv]. In this case, the susceptibility χ of the system goes to infinity.

The large susceptibility means that the system is unstable: a small external perturbation may lead to a large collective reaction of the traders who may drastically revise their decision, which may abruptly produce a sudden unbalance between supply and demand, thus triggering a crash or a rally.

For even stronger imitation strength K $> K_c$, the imitation is so strong that the idiosyncratic signals become negligible and the traders *self-organize* into strong imitative behavior.

Thought the susceptibility depends on the structure of the system, notwithstanding the large variety of topological structures of social networks (e.g. horizontal or hierarchical), the qualitative conclusion of the existence of a critical transition between a mostly disordered state and an ordered one, separated by a critical point, survives by-and-large for most possible choices of the network of interacting investors.

A solution to the Equation 7.8 is

$$\chi \approx A(K_c - K)^{-\gamma} \qquad \textbf{Equation 7.9}$$

where A is a positive constant and $\gamma > 0$ is called the *critical exponent* of the susceptibility, which can be a real or complex number, depending on the structure of the network. A more general version of Equation 7.9 is

$$\chi \approx \mathrm{Re}\left[A_0 (K_c - K)^{-\gamma} + A_1 (K_c - K)^{-\gamma + i\omega} + \ldots\right] \qquad \textbf{Equation 7.10}$$

$$\approx A_0'(K_c - K)^{-\gamma} + A_1'(K_c - K)^{-\gamma} \cos[\omega \log(K_c - K) + \psi] + \ldots \qquad \textbf{Equation 7.11}$$

Where A_0', A_1', ω, ψ are real numbers, and $\mathrm{Re}[.]$ denotes the real part of a complex number. In this expression, the power law in Equation 7.9 is corrected by oscillations whose frequency explodes as we reach the critical time. These accelerating oscillations are “log-periodic”, which have the “log-frequency” $\frac{\omega}{2\pi}$.

Assume the hazard rate of crash behaves in the same way as the susceptibility in the neighborhood of the critical point. Thus

$$h(t) \approx B_0 (t_c - t)^{-\alpha} + B_1 (t_c - t)^{-\alpha} \cos[\omega \log(t_c - t) + \psi'] + \ldots$$

where $0 < \alpha < 1$, otherwise the implied price would go to infinity when approaching the critical time. Plugging it into Equation 7.6 gives

$$\begin{aligned} \ln[p(t)] &\approx \ln[p_c] + \frac{\kappa}{\beta}\{B_0 (t_c - t)^{\beta} + B_1 (t_c - t)^{\beta} \cos[\omega \log(t_c - t) + \phi]\} \\ &\equiv A + B(t_c - t)^{\beta} + C(t_c - t)^{\beta} \cos[\omega \log(t_c - t) + \phi] \end{aligned} \qquad \textbf{Equation 7.12}$$

where $\beta = 1 - \alpha$. The key feature is that oscillations appear in the price of the asset just before the critical date, with frequency

$$\lambda = e^{\frac{\omega}{2\pi}} \qquad \textbf{Equation 7.13}$$

(Sornetter, 2003), (Johansen, Ledoit, and Sornette, 2000a, page 229-33.)

3.3. Empirical Findings of J&S et al

In Johansen, Ledoit and Sornette (JLS, 2000a), Equation 7.12 is fitted by minimizing the mean squared errors, using a combination of taboo-search and downhill simplex methods. As the function is highly nonlinear, many local minima exist and the minimization algorithm can get trapped at any of these local minima. In their estimation, when more than one minimum is produced, they select the best one according to a set of criteria in addition to the variance (JLS, 2000a, page 239). Two data sets were used in this paper: S&P 500 (July 1985 to end of 1987 with 557 trading days,) and Hang Seng Index (approximately two and half years daily data points prior to the October 1997 crash). Both power law and log-periodicity are present in the two samples and the in-sample performances of Equation 7.12 are remarkably well.

But how long prior to the crash can one identify the log-periodic signature using Equation 7.12 or its variants? To investigate this, JLS (2000a) truncated S&P 500 down to an end-date of approximately equals to 1985. Then approximately 0.16 years was added consecutively and the fitting was re-launched until the full time interval was recovered. These experiments show that a year or more before the crash, the data is not sufficient to give any conclusive results at all. Approximately a year before the crash, the fit begins to lock-in the date of the crash with increasing precision. However, if one wants to actually predict the time of the crash, a major obstacle is that the fitting procedure produces several possible dates (multiple minima) for the date of the crash even for the last data set. They apply the same procedure to Dow Jones Index prior to the crash of 1929. For this data set, the fit locks in on the date of the crash approximately four months before the crash.

Sornette and Johansen (1997) argued that, based on the renormalization group theory (also refer to Sornette 2003), validate their proposed model requires that one obtain a good fit in several data sets with approximately the same parameter values. In the past few years, J&S et al produced a series of papers examining various kinds of markets. Their investigation show that, for all the bubbles in the most liquid markets, for example, USA, Hong Kong and the foreign exchange market, the log-frequency $\omega / 2\pi$ have consistently been close to 1. In the framework of power laws with complex exponents or equivalently, discrete scale invariance (Sornette, 2003, Chapter 6)[v], this corresponds to a preferred scaling ratio $\lambda = e^{\omega / 2\pi} \approx e \approx 2.7$: the local period of the log-periodic oscillations decreases according to a geometrical series with the ratio λ. For the emergent market, the value of λ shows more fluctuations, but the statistics resulting from over twenty bubbles were quite consistent with that of the large market (Johansen and Sornette, 2000b) (Table 7.1 of this thesis). However, the "universality" of the value of the real part of the exponent β quantifying the acceleration in the price has not been established. Johansen and Sornette (J&S, 2000c) explains that the technical reason for this is that the determination of β is sensitive to finite-size-effects as well as to errors in the value of t_c, the critical point. In all cases, the model show reasonable accuracy in identifying the crash date.

In J&S (2000c), the authors mentioned two false signal issued by the model. On September 17th, 1997, the signal was issued to predict a crash of the stock market at the end of October 1997. It turned out that the market dropped by 7% but quickly recovered. On October 1999, the model again issued a false alarm. The authors argue that these two examples of bubbles landing more or less smoothly by be explained by the finite probability $1 - Q(t_c)$ that no crash occurs over the whole time including the critical time of the end of the bubble, even through price show characteristic of looming crash. This finite probability obviously reduces the accuracy of the crash prediction.

4. Optimization and Estimation

4.1. The Optimization Problem

Consider the global optimization problem: let $f : D \rightarrow R$, where D is a convex set in R^n. Find a point $x^* \in D$ such that $f\left(x^*\right) \leq f\left(x\right), \forall x \in D$. When f is highly nonlinear with many local optima, finding the global optima can be very tough.

Two types of methods have been developed and implemented in practice to solve this global optimization problem: deterministic and stochastic methods. Deterministic methods attempt to generate trajectories that eventually converge to points which satisfy the criterion of local optimality. They are beneficial only when the starting point belongs to the region of attraction of the global optimum. So any deterministic method could be attracted by the local optimum instead. The stochastic methods, on the contrary, attempt to reasonably cover the whole search space and to identify all local and global optimum. In stochastic methods, points that do not strictly improve the objective function can also be created and take part in the search process. Hence stochastic methods have better chance of reaching the global optimum.

A number of stochastic methods have been proposed and used in different types of optimization problems. Al-Harkan and Trafalis (2002) suggest a hybrid approach, which incorporates scatter search, genetic adopter and tabu search, for solving the unconstrained continuous nonlinear global optimization problem. They named this approach “hybrid scatter genetic tabu”(HSGT).

The scatter search (SS) approach was introduced by Glover (1977) as a heuristic to obtain a near optimal solution to an integer programming problem. Recently the SS approach was refined and used for both discrete and continuous optimization problems (Glover 1994a, 1994b, 1995, and Fleurent et al). The SS approach generates sequences of coordinated initializations which are performed to ensure the exploration of the various parts of the solution space. The exploration of the solution space was based on a kindred strategy which was suggested in Glover (1977). Based on the kindred strategy, the SS approach directs its explorations systematically relative to a collection of points called the reference points. Hence, the SS approach begins its procedure with a set of reference points which can be obtained by applying either heuristic procedures or random methods. Then, a weighted center of gravity of the reference points is determined using a linear combination of the reference point solutions and their weights. The linear combination allows the use of negative weights which are used to allow the weighted center of gravity to go outside the area spanned by the reference points. This process is known as the diversification process which allows a variety of new solutions. Next, subsets of initial reference points and the weighted centers of gravity are used to define new sub-regions as a foundation for generating subsequent points. Then, these points are evaluated and are used as the new set of reference points. At this stage, a complete iteration of the SS approach is performed. The procedure can be repeated until some preset stopping criteria are satisfied.

The genetic adopter (GA) approach was developed by Holland in 1975 (Holland, 1992). Since GAs are adaptive and flexible, they have attracted attentions from researchers from different fields, such as computer science, operation research, business and social science, etc. The theory and application of GAs have been reported be several researches (Davis 1991, Goldberg 1989, Holland 1992, Michalewicz 1994, and Srinivas and Patnaik 1994). In these reports, the GAs were shown to be successfully applied to several optimization problems. The GAs are stochastic search techniques whose search algorithms simulate biological evolution---the strong tend to adapt and survive while the weak tend to die. At the beginning, a population of binary or non-binary chromosomes is initialized randomly. Then, each chromosome is evaluated using the fitness function. A set of better chromosomes is selected to reproduce new chromosomes. The production process is accomplished by applying the genetic operators (crossover and mutation) on the chromosomes selected. Then, each new chromosome is evaluated. At this stage, a full iteration is performed. Repeat the procedure until some preset termination criteria are satisfied.

The tabu search (TS) approach is a heuristic to solve combinatorial optimization problems. Recently, it has been applied to solve continuous global optimization problem (Cvijovi and Klinowski 1995, Fleurent et al. 1995, Glover 1994b). In a tabu search, restrictions (tabu) are imposed to guide the search process to investigate difficult regions. It starts with an initial solution for the problem under consideration which can be constructed by using either heuristics or a random solution. Then TS constructs a neighborhood from the current solution to identify adjacent solutions, and the objective function associated with each adjacent solution is evaluated. Before determining the best move, the TS approach selects the set of admissible moves which are not tabu. For instance, recent moves will be classified as tabu to prevent the search from going back to its previous position. A recent move will be tabu for the duration of a certain number of iterations which depend on the size of the tabu list or tabu tenure. The aspiration criterion can be activated if a move that was tabu results in a solution better than any visited solution so far. In this case, the move’s tabu status is broken and it becomes the best move. Otherwise the best move is selected from the set of admissible moves. By then, a complete iteration of the TS approach was performed. Repeat the procedure until the stopping criteria are met.

Harkan and Trafalis (2002) tested their proposed approach against a simulated annealing algorithm and a modified version of a hybrid scatter genetic search approach by optimizing twenty-one well know test functions. They show that HSGT approach is quite effective in identifying the global optimum.

4.2. Estimation Strategy

We fit Equation 7.12 by minimizing one half of the sum of squared residuals.

$$\min_{\theta} e^2 = \frac{1}{2}\sum_{t=1}^{T}\left(x(t) - \hat{x}(t)\right)^2$$ **Equation 7.14**

Where $x(t)$ is the first difference of the log price, $\hat{x}(t)$ given by Equation 7.12, and $\theta = (A, B, C, t_c, \beta, \omega, \phi)$. As argued in Greene (3rd. ed., Chapter 10), the values of the parameters that minimize one half of the sum of the squared residuals will be the maximum likelihood estimators, as well as the nonlinear least squares estimators. The first-order conditions for minimization of e^2 will be a set of nonlinear equations that do not have an explicit solution. This will typically require an iterative procedure for solution. In particular, the HSGT described above will be used.

To reduce the number of free parameters to be estimated, following Johansen and Sornette (2000a), the three linear parameters A, B, C are enslaved as functions of the nonlinear parameters t_c, β, ω, ϕ. This is done by requiring the objective function to have zero derivatives with respect to A, B, C at the minimum. Optimizing Equation 7.14 with respect A, B, C we get a system of three equations

$$\sum_{t=1}^{T}\begin{pmatrix} \ln[p(t)] \\ \ln[p(t)]f(t) \\ \ln[p(t)]g(t) \end{pmatrix} = \sum_{t=1}^{T}\begin{pmatrix} T & f(t) & g(t) \\ f(t) & f(t)^2 & f(t)g(t) \\ g(t) & f(t)g(t) & g(t)^2 \end{pmatrix}\begin{pmatrix} A \\ B \\ C \end{pmatrix}$$ **Equation 7.15**

where $f(t)=(t_c-t)^\beta$, $g(t)=(t_c-t)^\beta\cos[\omega\log(t_c-t)+\phi]$. Equation 7.15 will be solved via by LU decomposition, as it is an efficient algorithm for matrix inversion (Press et al., 1992), (Lee, **http://prosys.korea.ac.kr/~tclee/lecture/numerical/node14.html**).

The nonlinear parameters are estimated using the HSGT approach proposed by Al-Harkan and Trafalis (2002). Their method is summarized below:

Step 1, generate a random starting point $X_k=\{t_c,\beta,\omega,\phi\}$ from a uniform distribution between the upper and lower bounds of each variable[vi];

Step 2.1, generate a set of m random directions from a standard normal distribution, with $m=n\times 2^3$ where n the number of parameters to be estimated. In our case, $m=4\times 2^3=32$.

Step 2.2, recalculate the m directions so that all of the m vectors of dimension $n\times 1$ have unit Euclidean norm[vii];

Step 3, generate a set of m reference points by moving from the starting point in the m directions calculated previously;

Step 4, assign weight to each of the m points according to the values of the objective function at these points. The largest weight is given to the point with the smallest value;

Step 5, generate the centre of gravity $X_k=\{t_c,\beta,\omega,\phi\}$ by taking the weighted average of the m reference points obtained in Step 4. This will be the new starting point for the next round of iteration if it passes the tabu test[viii]. Otherwise, retain the old value of X_k;

Step 6, generate new search directions using genetic operator, either whole arithmetical crossover approach or general mutation approach (Al-Harkan and Trafalis, 2002, page 9-10). Normalize the m directions the same way as in step 2.2; repeat step 3 to 5 until the pre-specified stopping rule is satisfied.

4.3. Properties of Nonlinear Least Square Estimators

Consider the nonlinear regression model

$$y = h(x,\theta) + \varepsilon, \qquad \varepsilon \sim N(0,\sigma^2) \qquad \textbf{Equation 7.16}$$

The parameters θ can be estimated by minimize one half of the sum of squared residuals

$$S(\theta) = \frac{1}{2}\sum_{t=1}^{T}\left[y_t - h(x_t,\theta)\right]^2 \qquad \textbf{Equation 7.17}$$

The asymptotic properties of the nonlinear least squares estimator of θ is derived in Greene (3rd ed., CH. 10), and summarized below.

If the pseudoregressors defined in Equation 7.17 are well behaved, then

$$\hat{\theta} \xrightarrow{a} N\left[\theta, \frac{\sigma^2}{n} Q_0^{-1}\right] \qquad \textbf{Equation 7.18}$$

where

$$Q_0 = p\lim \frac{1}{n} X_0^{'} X_0 \equiv p\lim \frac{1}{n}\sum_{i=1}^{n} \frac{\partial h(x,\theta)}{\theta}\frac{\partial h(x,\theta)}{\theta'}\Big|\theta_0 \qquad \textbf{Equation 7.19}$$

With θ_0 a particular value of θ. The sample estimate of the asymptotic covariance matrix is

$$Est.Asy.Var\left[\hat{\theta}\right] = \hat{\sigma}^2 \left(X_0^{'} X_0\right)^{-1}$$ **Equation 7.20**

where

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} \left[y_t - h(x_t, \theta)\right]^2$$ **Equation 7.21**

Under normality, this is the ML estimator of σ^2.

5. Our Empirical Application

5.1. The Data

As mentioned at the start, the prime purpose of this paper is to examine the applicability of the model developed by J&S et al to Asian property markets. In particular, we apply the model to Hong Kong office price index (HKOP) and Seoul housing price index (SHP). As SHP shows slow decays throughout the 1990s, rather than large crashes of the kind seen in HKOP, we also include, in this paper, a third series: Korea general construction stock price index (KGCSP, Jan. 1980 to Nov. 2003), a stock price index which has property as its fundamental. The rational is that, due to strict capital constraints until recent and tight government control of land supply in Korea (Kim, 1999, 2000), speculation in property may be more likely to show up in stock written on property rather than property itself in this country. The plots of these price indices are displayed in Figure 5.1 and 7.1. Table 7.2 lists accounts of the price crashes of these series.

5.2. Fitting the Model

In our application, we include the after-crash points up to the end of each sample, instead of only pre-crash observations, as is generally the case in the papers of J&S et al, to see if the model will be able to identify the critical time. Notice that the monthly observations on price rally before a crash in our data sets are less than 50 before each crash in each series, rather than the 100 strong observations obtained in the data sets of Zhou and Sornette (2003). For each data set, we found values fitted using Equation 7.12 shoot too far from the actual data. To examine the problem, we took the first difference of the log price. The plots show that the differenced sets fluctuate around a zero mean, implying that the dominant force is a linear trend rather than power law acceleration. This is confirmed later by the estimates of β, which turns out to be close to one (Comments of Sornette, December 6th, 2004).

Nevertheless the pattern of log-periodicity is very obvious in HKOP: the frequency of cycle obviously increases around the crash time. This pattern is less obvious to the naked eyes in KGCSP yet discernable (Figure 7.2).

Inspired by these observations, we fit instead the following model using HGST algorithm described before

$$d\ln[p(t)] \approx C(t_c - t)^{\beta} \cos[\omega \log(t_c - t) + \phi] \qquad \textbf{Equation 7.22}$$

where $d\ln(p(t))$ is the first difference of the log price. Equation 7.22 is the essentially the same as Equation 7.12 except for the intercept and exponential trend terms. The standard errors of parameter estimates are given by

$$SE(\theta_i)=\sqrt{\frac{\hat{\sigma}^2}{\sum_{t=1}^{T}\frac{\partial\hat{x}}{\partial\theta_i}\frac{\partial\hat{x}}{\partial\theta_i}'}\Bigg|\hat{\theta}}$$

Where $\hat{x}$ given by Equation 7.22, $\hat{\theta}$ the ML estimates of θ, and

$$\frac{\partial\hat{x}}{\partial t_c}=\frac{\left[\left(c(t_c-t)^{\beta}\beta\cos(\omega\ln(t_c-t)+\phi)\right)-\left(c(t_c-t)^{\beta}\sin(\omega\ln(t_c-t)+\phi)\omega\right)\right]}{(t_c-t)}$$

$$\frac{\partial\hat{x}}{\partial\beta}=\left(c(t_c-t)^{\beta}\ln(t_c-t)\cos(\omega\ln(t_c-t)+\phi)\right)$$

$$\frac{\partial\hat{x}}{\partial\omega}=\left(-c(t_c-t)^{\beta}\sin(\omega\ln(t_c-t)+\phi)\right)\ln(t_c-t)$$

$$\frac{\partial\hat{x}}{\partial\phi}=\left(-c(t_c-t)^{\beta}\sin(\omega\ln(t_c-t)+\phi)\right)$$

5.3. Estimation Results

We initiated randomly the nonlinear parameters t_c,β,ω,ϕ on a uniform distribution, with $\beta\in\left[0.3,\ \ 0.98\right]$, $\omega\in\left[1,\ \ 10\right]$, $\phi\in\left[0,\ \ 20\right]$.These values are set with reference to the empirical results found by J&S et al. The critical time t_c is set to be $t_c\in\left[tcl,\ \ T\right]$, where T is the size of the sample, tcl an random date before a naked-eye-identified crash date.

The estimates of β, ω, ϕ are very consistent for all three data sets, which are {0.98, 10, 20}, or {0.9516, 4.5984, 5.1880}, or {0.9509, 4.5981, 5.1881} (Table 7.3-7.5). This fact is consistent with the renormalization theory which requires consistent parameter estimates from different data sets for the validation of the suggested model.

These estimates suggest that there exist two log-periodic harmonics ω, one is 10 the other near 10/2, suggesting the theoretical formula ideally should have two cosine terms with two harmonics (courtesy of Sornette). A careful examination of Figure 7.2 reveals that these plots indeed resemble that of some theoretical function which has two log-periodic oscillations of different frequencies, with one superimposed on the other (Figure 7.4). But there are cons as well as pros of fitting this second term. The obvious one is the loss of degree of freedom.

In viewing the records of the experiments, whenever we obtain the second set of estimates, the search process is trapped at the starting point. Whenever the move becomes possible, we would obtain the last set of estimates. The first set of estimates is the pre-set boundaries. Henceforth, we will refer to the set of estimates neither hit the boundary nor is trapped at the starting values as “preferred estimates”.

The estimates of the critical time however depend on the lower boundary we imposed on the starting value of t_c (Table 7.3-7.5). In the estimation process, we move the lower boundary up by 0.083 decimal year (one monthly observation) at a time until the last sample observation. As a result, we obtain 18 dates for HKOP, 11 for SHP. However, the fit for KGCSP is remarkably well: only four dates are obtained---the best of which is within one month of the actual crash.

We tried to alter the boundaries for the nonlinear parameters. Different estimates for β, ω, ϕ are obtained as a result. However, we always get consistent estimates for the three data sets for each pair of boundaries set for a parameter. As for t_c, when the lower boundaries (not only for the starting value but also for the entire estimation process) are brought backwards (observation further before the crash are admitted), we obtain more boundary estimates for the four nonlinear parameters, suggesting information is not clear for obtaining accurate estimates at such early dates. But raising the lower boundaries has no impact on the results displayed in the above tables.

Table 7.6 summarizes the best fitting for the three data sets, alongside with their standard errors. The best estimate for HKOP is within 3 months of the true crash date; That for KGCST within one month of the true date, but β is insignificant. Viewing the data shows that the price drop in December 1994 was slow and with lots of reversals until May 1996, when the price crashes all the way down. We missed the true date by near four years in SHP, however. The reason is perhaps that the decay in SHP since its peak in 1991 was very slow in pace.

5.4. Forecast

To investigate the predictive power of the model, we truncate the data sets such that they end at March 1991 (38 months ahead of the crash), September 1992 (27 months ahead) and May 1989 (23 months ahead)[1] for HKOP, KGCSP and SHP respectively. Table 7.7 reports the preferred estimates, which neither hit the boundaries nor are trapped at the starting point. The best fits are also shown is Figure 7.3.

The best forecast for HKOP is 1993.9997, which misses the true crash date (May 1994) by four months. Increasing the size of sample such that the last observation is November 1991, we obtain one more predicted critical date, 1994.9997, which overshoots the actual date by eight months. When the last observation is Jan. 1993 (16 months ahead of the crash), we obtain only one preferred estimate, 1993.9997.

KGCSP obtains three preferred estimates, the best of which is 1994.9997 (the actual crash date is 1994.92 decimal year. β is insignificant however.) These results are robust with respect to changes of the tcl. Something peculiar is that increasing the sample size in this case actually worsens the results. For instance, when the last observation is January 1994, no preferred estimate is obtained.

Both HKOP and KGCSP miss the crashes which closely follow their predecessors, namely, October 1997 crash of HKOP and May 1996 crash of KSCSP.

The unique forecast obtained for SHP (1989.9997) undershoots the true date by some 15 months. Again, increasing the sample size actually worsens the result of forecast.

[1] The points of truncation are chosen at where certain pattern has developed.

6. Conclusive Remarks

Our experiments show that the original model of J&S in the form of Equation 7.12 does not comply completely with the three data sets under our consideration. The power low behavior is missing from these data sets and replaced in stead by a linear trend. We thus estimate Equation 7.22 instead. The results show that this model is far more successful for KGCSP than for the other two data sets, where too many prediction of the critical time are obtained making it difficult to tell which one of these is the true alarm. With KGCSP we obtain only two preferred estimates, defined as estimates neither hit the preset boundary values nor are trapped at the starting point of the iteration, both are within reasonable range of the crash dates(refer to Table 7.6. One possible explanation is that the price swings in KGCSP are far spectacular than in the two other data sets. Excluding the set of estimates with insignificant β, we however captures only one out of the three crashes identified a prior.

The forecast using HKOP gives the best prediction capturing one of the two crashes after the data truncation point, with reasonable accuracy. Forecasting using both KGCSP and SHP missed out the crash dates completely when estimates with insignificant β are excluded.

The estimates of the nonlinear parameters β, ω, ϕ are nearly identical in all three samples, which is consistent with renormalization theory.

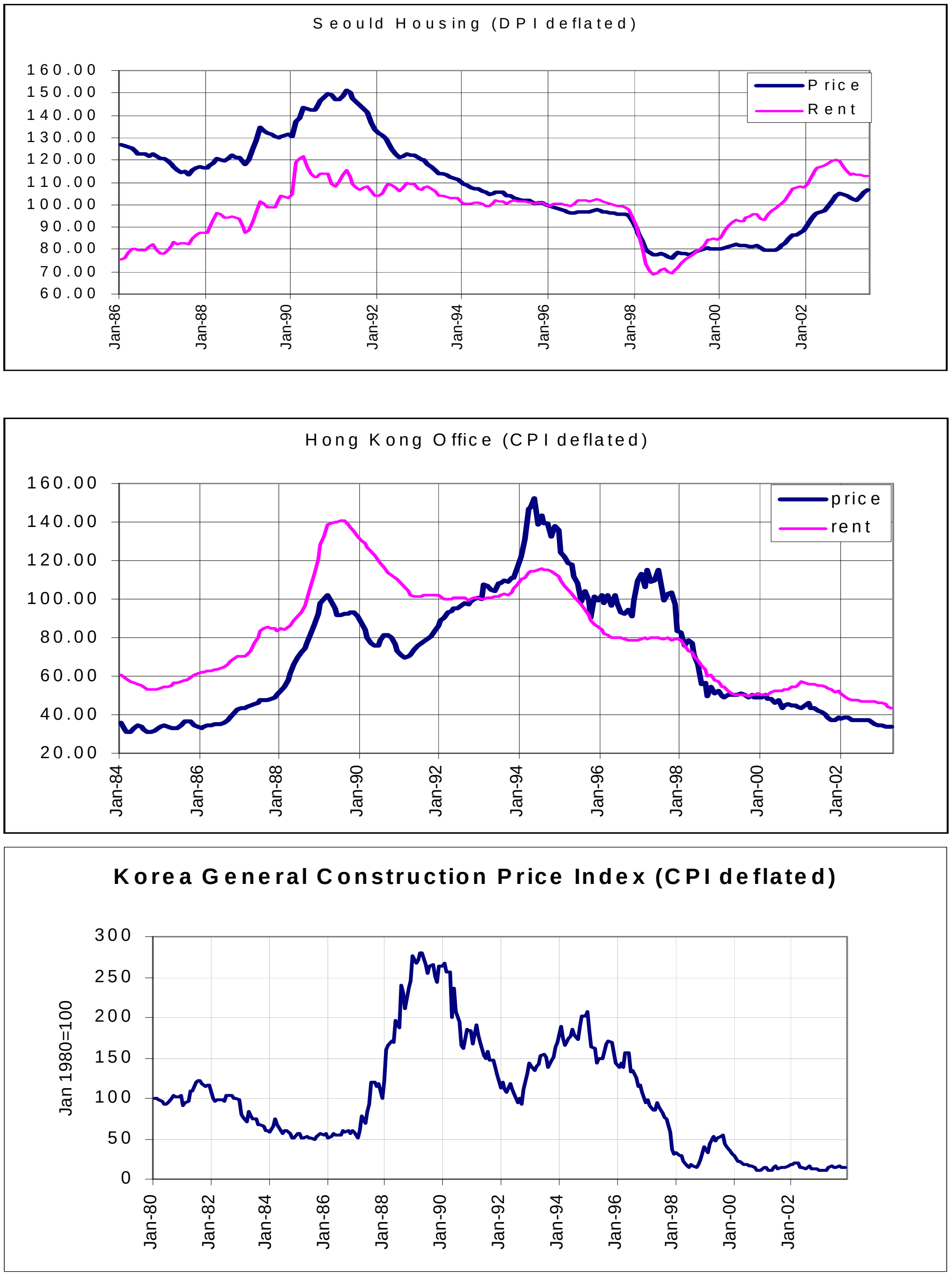
The model we estimated, however, fails to capture the fact that variances tend to increase around the crash points. We have tried a few ways to remedy it but the improvement is insignificant in terms of increasing the accuracy of predicted critical time. One of these remedies is to allow β to exceed one, which is nevertheless ruled out theoretically at the start.

In conclusion, we consider the PLLP theory reasonably successful in our data sets. Though the prediction of the critical time is not as neat as we would prefer, the model nevertheless provides a very useful signature of price behavior: a power-law accelerating of prices, and/or an accelerating price oscillation is a highly reliable signal of looming market crash. Furthermore, the model relates the price trajectory to the crash hazard rate, and provides another useful model for extrapolating the price variations.

We also obtained some observations on the HSGT approach to global optimization. The algorithm reaches convergence, in general, in less than three runs, when applied to Equation 7.22. However, when applied to Equation 7.12, it takes too much CPU time and is difficult to implement given the typical state of computer a researcher has at hand[ix]. This is due to the fact that the LU decomposition module needs to be called $2 + n \times 2^3$ times, where n the number of nonlinear parameters to be estimated, in each iteration of the grid search. Thus, some simpler method, such as downhill simplex, may be more practical.

CPI Deflated Price Index

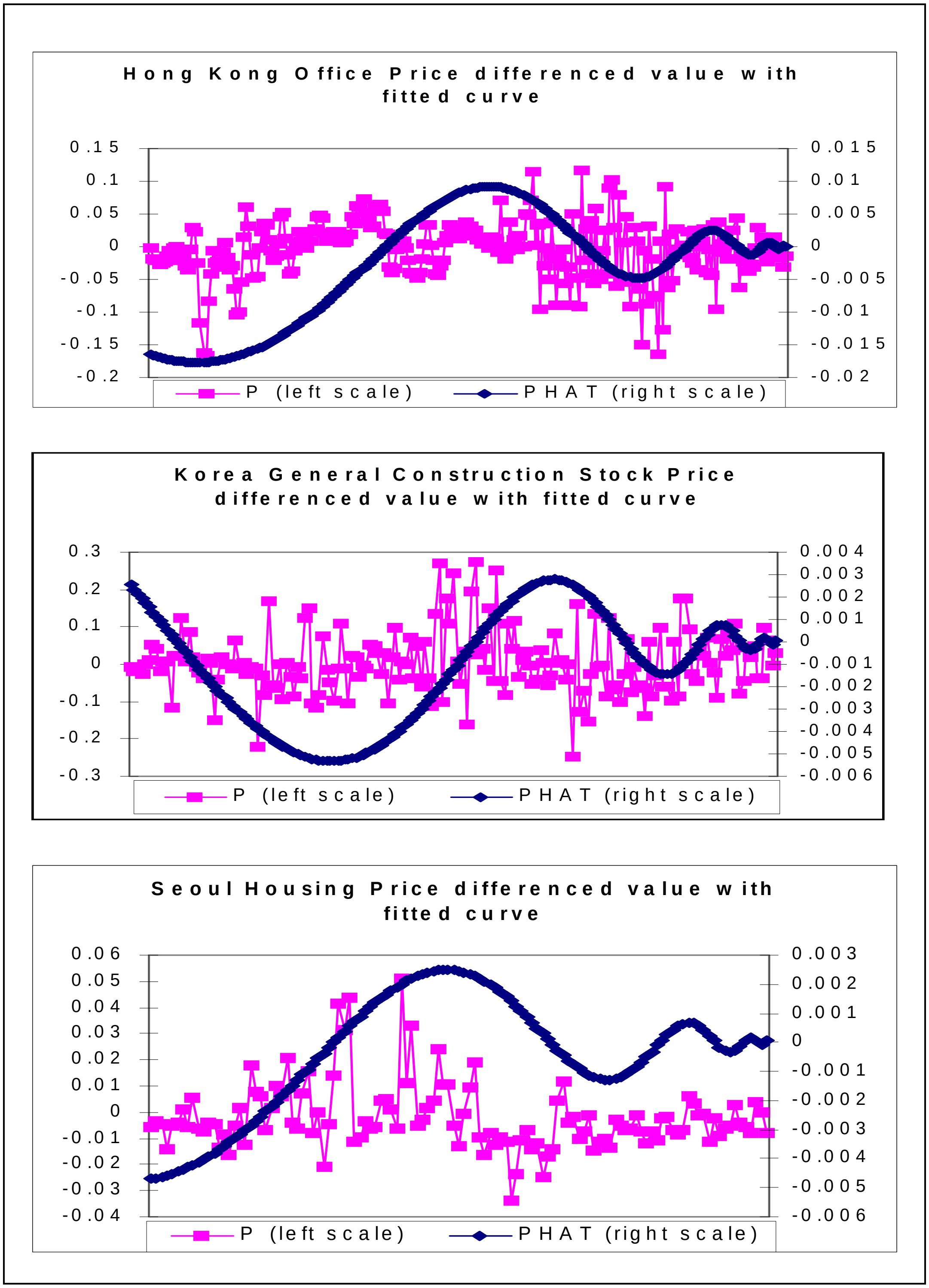
Figure 7.1 *Korea General Construction Price Index (CPI deflated)*



Source: CEIC database

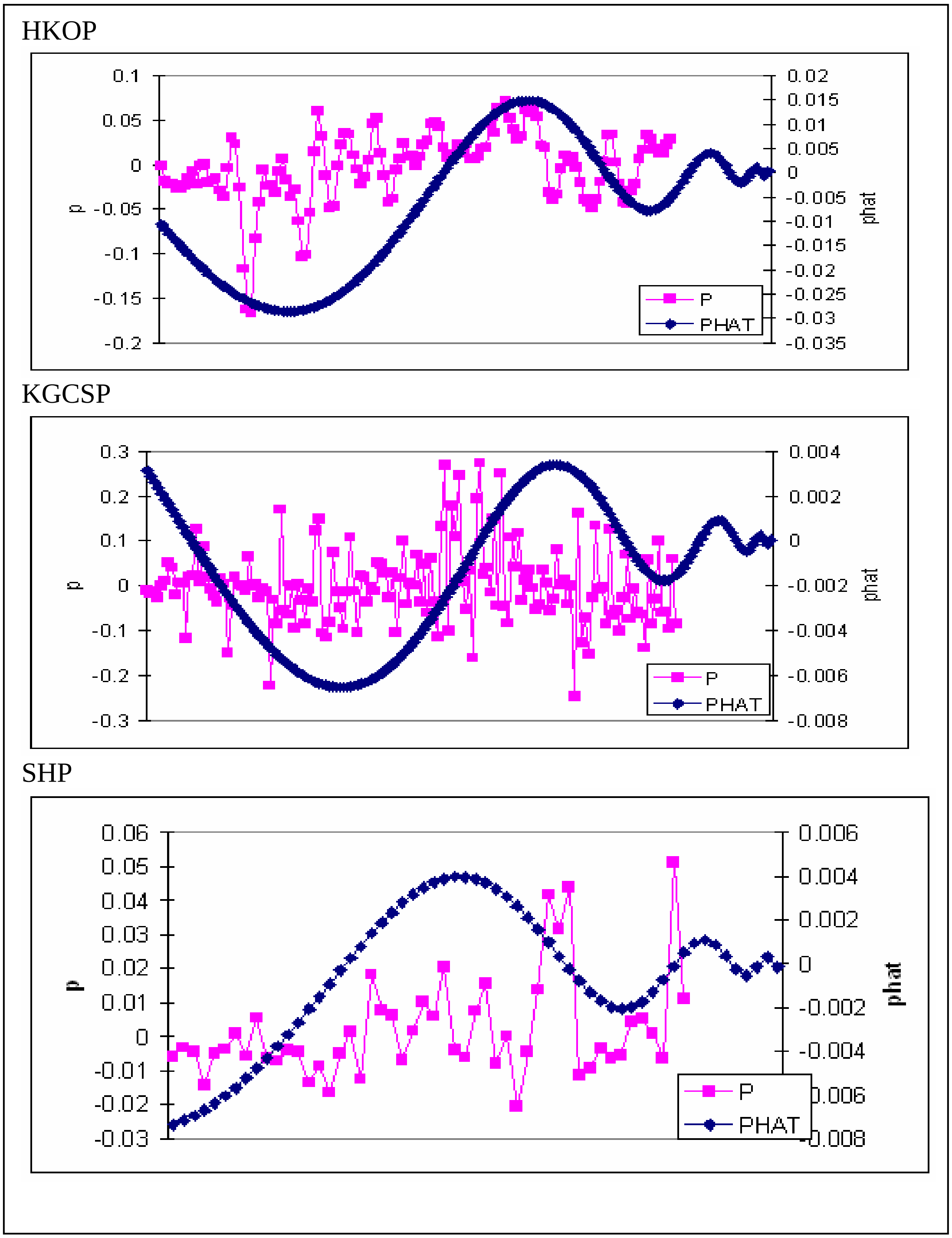
Differenced Log Series

Figure 7.2 *Differenced Log Price and fitted value*



Out-Of-Sample Forecast

Figure 7.3 *Forecast*



Plot of Theoretical Model with Two Cosine Terms

Figure 7.4 *Function Plot*

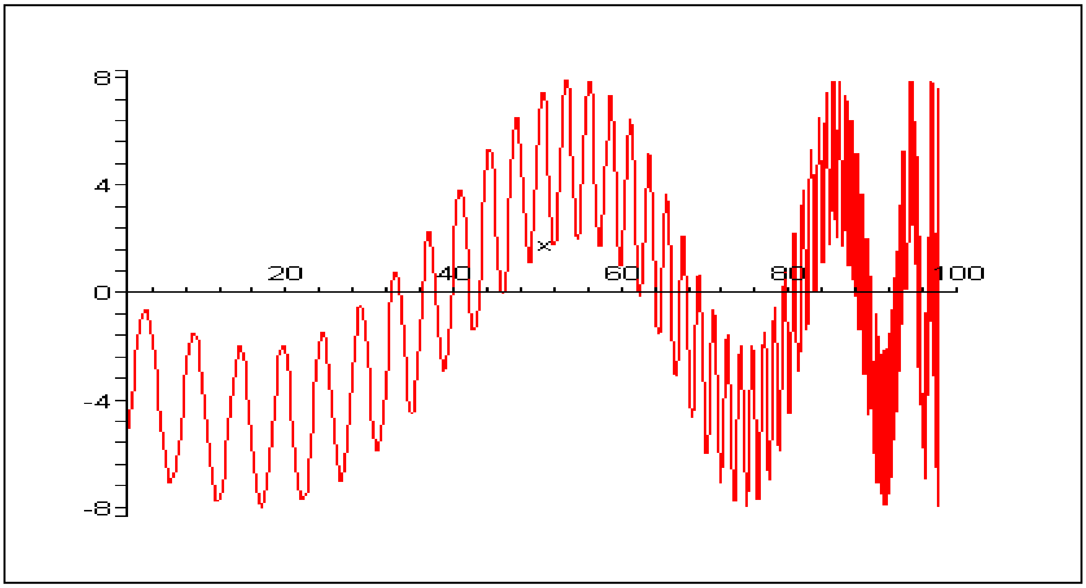


Table 7.1 Empirics from Johansen and Sornette (2000c)

Crash data (time series)	t_c	$t_{\max}$	$t_{\min}$	***%drop***	β	ω	λ
1929 (DJ)	30.22	29.65	29.87	47%	0.45	7.9	2.2
1985(DM)	85.20	85.15	85.30	14%	0.28	6.0	2.8
1985(CHF)	85.19	85.18	85.30	15%	0.36	5.2	3.4
1987(S&P)	87.74	87.65	87.80	30%	0.33	7.4	2.3
1987(HK)	87.84	87.75	87.85	50%	0.29	5.6	3.1
1994(HK)	94.02	94.01	94.04	17%	0.12	6.3	2.7
1997(HK)	97.74	97.60	97.82	42%	0.34	7.5	2.3
1998(S&P)	98.72	98.55	98.67	19.4%	0.60	6.4	2.7
1999(IBM)	99.56	99.53	99.81	34%	0.24	5.2	3.4
2000(P&G)	00.04	00.04	00.19	54%	0.35	6.6	2.6
2000(Nasdaq)	00.34	00.22	00.29	37%	0.27	7.0	2.4

Note: This table is reproduced from J&S 2000c. t_c *is the critical time predicted from equation 14. The fit is performed up to the time* $t_{\max}$ *at which the market index achieved its highest maximum before the crash.* $t_{\min}$ *is the time of the lowest point of the market after the maximum. The percentage loss is calculated from the total loss from* $t_{\max}$ *to* $t_{\min}$ *.*

Table 7.2 Facts on Crashes

	Pre-identified crash dates	*Percentage change and time taken*
HKOP	Mar. 89	↓31.697% in 23 months
	May 94	↓41.427% in 17 months
	Oct. 97	↓51.867% in 11 months
SHP	Apr. 91	↓20.075% in 15 months
KGCSP	Jan. 90	↓36.097% in 8 months
	Dec. 94	↓28.308% in 4 month
	May 96	↓89.494% in 25 months

Table 7.3 HKOP Parameter Estimates

tc l_01[2]	t_c (s.e.)	β (s.e)	ω (s.e.)	ϕ (s.e)	*C*	*Variance*
Apr. 83	1982.92[3] (0.2388)	0.98#[4] (4.5570)	10 (4.2163)	20 (2.0603)	0.0027	0.0023
May 83-Jul. 84	1987 (0.4102)	0.9516 (0.3080)	4.5984 (0.4391)	5.1880 (0.6200)	-0.0042	0.0019
Aug. 84-Nov. 86	1988 (0.1576)	0.9516 (0.1232)	4.5984 (0.1025)	5.1880 (0.1742)	-0.0082	0.0014
Dec. 86	**1988.9997**[5] **(0.2090)**	**0.9509 (0.1638)**	**4.5981 (0.1071)**	**5.1881 (0.1971)**	**-0.0062**	**0.0017**
Jan. 87-Mar. 88	1990 (2.5067)	0.9516# (1.3844)	4.5984 (1.1869)	5.1880 (2.2544)	0.0005	0.0018
Apr. 88-Jun. 89	1991 (0.4817)	0.9516 (0.1835)	4.5984 (0.2272)	5.1880 (0.4335)	0.0024	0.0016
Jul. 89-Aug. 90	1992 (0.2847)	0.9516 (0.0846)	4.5984 (0.1269)	5.1880 (0.2471)	0.0037	0.0013
Sep. 90-Nov. 91	1992.9997 (0.2236)	0.9509 (0.0593)	4.5981 (0.0813)	5.1881 (0.1701)	0.0042	0.0011
Dec. 91-Feb. 93	**1993.9997 (0.2990)**	**0.9509 (0.0791)**	**4.5981 (0.0848)**	**5.1881 (0.1916)**	**0.0032**	**0.0012**
Mar. 93-May 94	1994.9997 (0.7083)	0.9509 (0.1985)	4.5981 (0.1654)	5.1881 (0.3939)	0.0014	0.0015

[2] tcl_01: the lower boundary set for the initial value of tc.
[3] In decimal years. For non-leap year,
12 months =1.00 year; 1 month=0.083 years; e.g. December 1998=1998.92.
[4] #: estimates insignificant at conventional levels.
[5] Rows in bold show best fits.

Jun. 94-Jul. 95	1996	0.9516#	4.5984	5.1880	-0.0004	0.0017
	(2.2859)	(0.6682)	(0.4681)	(1.1541)		
Aug. 95-Oct. 96	1997	0.9516	4.5984	5.1880	-0.0017	0.0017
	(0.5540)	(0.1546)	(0.1046)	(0.2638)		
Nov. 96-Jan. 98	**1998**	**0.9516**	**4.5984**	**5.1880**	**-0.0019**	**0.0019**
	(0.4944)	**(0.1197)**	**(0.0892)**	**(0.2282)**		
Feb. 98-Mar. 99	1999	0.9516	4.5984	5.1880	-0.0021	0.0021
	(0.4759)	(0.0956)	(0.0842)	(0.2171)		
Apr. 99-Jun. 00	1999.9997	0.9509	4.5981	5.1881	-0.0019	0.0020
	(0.4972)	(0.0831)	(0.0878)	(0.2268)		
Jul. 00-Sep. 01	2000.9997	0.9509	4.5981	5.1881	-0.0015	0.0020
	(0.6461)	(0.0919)	(0.1142)	(0.2951)		
Oct. 01-Dec. 02	2001.9997	0.9509	4.5981	5.1881	-0.0013	0.0020
	(0.7511)	(0.0937)	(0.1309)	(0.3396)		
Jan. 03-May 03.	2002.9997	0.9509	4.5981	5.1881	-0.0010	0.00187
	(0.9003)	(0.1019)	(0.1496)	(0.3936)		

Note: the linear parameter C is enslaved as function of the nonlinear parameters, hence only the standard errors of nonlinear parameters are provided.

Table 7.4 KGCSP Parameter Estimates

tcl_01	t_c (s.e)	β (s.e)	ω (s.e)	ϕ (s.e)	*C*	*Variance*
May 88-Dec 92	1988.33 (1.3244)	0.98# (1.4441)	10 (1.6693)	20 (2.9468)	-0.00093	0.0071667
Jan. 93-Feb. 94	**1994.9997 (3.4343)**	**0.9509# (1.0032)**	**4.5981 (0.7099)**	**5.1881 (1.7461)**	**0.0006**	**0.0071**
Mar. 94-May. 95	1996 (1.0980)	0.9516 (0.3088)	4.5984 (0.2084)	5.1880 (0.5247)	-0.0017	0.0069
Jun.95-Aug. 96	**1996.9997 (0.7960**	**0.9509 (0.1956)**	**4.5981 (0.1440)**	**5.1881 (0.3681)**	**-0.0023**	**0.0069**
Sep 96-Nov. 03	1988.33 (1.3244)	0.98# (1.4441)	10 (1.6693)	20 (2.9468)	-0.0009	0.0072

Table 7.5 SHP Parameter Estimates

tcl_01	t_c ***(s.e)***	β ***(s.e)***	ω ***(s.e)***	ϕ ***(s.e)***	***C***	***Variance***
May 89-Nov. 91	1989.33 (1.7243)	0.98# (8.9809)	10# (11.3343)	20 (9.7965)	-0.0002	0.0002
Dec. 91-Feb 93	1994 (0.2854)	0.9516 (0.2270)	4.5984 (0.1479)	5.1880 (0.2711)	0.0015	0.0002
Mar. 93-Apr. 94	1994.9997 (0.2302)	0.9509 (0.1317)	4.5981 (0.1093)	5.1881 (0.2074)	0.0016	0.0002
May 94-Jul. 95	1995.9997 (0.7059)	0.9509 (0.2764)	4.5981 (0.3331)	5.1881 (0.6354)	0.0005	0.0002
Aug. 95- Oct. 96	1997 (2.3786)	0.9516# (0.7175)	4.5984 (1.0718)	5.1880 (2.0780)	-0.0002	0.0002
Nov. 96-Dec. 97	1998 (0.7814)	0.9516 (0.2081)	4.5984 (0.2897)	5.1880 (0.6022)	-0.0004	0.0001
Jan. 98-Mar. 99	1999 (0.7156)	0.9516 (0.1887)	4.5984 (0.2066)	5.1880 (0.4645)	-0.0005	0.0002
Apr. 99-Jun. 00	1999.9997 (0.8442)	0.9509 (0.2354)	4.5981 (0.1998)	5.1881 (0.47429)	-0.0004	0.0002
Jul. 00-Sep. 01	2000.9997 (6.3115)	0.9509# (1.8436)	4.5981251 (1.3046)	5.1881# (3.2089)	-0.0001	0.0002
Oct. 01-Nov. 02	2002 (12.371)	0.9516# (3.4797)	4.5984 # (2.3477)	5.1880# (5.9117)	0.0000	0.0002
Dec. 02-Jun. 03	2003	0.9516#	4.5984	5.1880	0.0001	0.0002

	(2.0486)	(0.5031)	(0.3706)	(0.9471)	

Table7. 6 The Best Fits

	True crash date	t_c	β	ω	ϕ
HKOP	1989.17	1988.9997	0.9509	4.5981	5.1881
		(0.2090)	(0.1638)	(0.1071)	(0.1971)
	1994.33	1993.9997	0.9509	4.5981	5.1881
		(0.2990)	(0.0791)	(0.0848)	(0.1916)
	1997.75	missed			
KGCSP	1990.00	missed			
	1994.92	1994.9997	0.9509#	4.5981	5.1880982
		(3.4343)	(1.0032)	(0.7099)	(1.7461)
	1996.33	1996.9997	0.9509	4.5981	5.1880982
		(0.7960)	(0.1956)	(0.1440502)	(0.3681423)
SHP	1991.25	1994.9997	0.9509	4.5981	5.1881
		(0.2302)	(0.1317)	(0.1093)	(0.2074)

Table 7.7 Forecast Summery

	tcl_01	*True crash date*	t_c *(s.e.)*	β *(s.e.)*	ω *(s.e.)*	ϕ *(s.e.)*	*C*	*Variance*
HKOP	Nov. 91-Dec. 92	1989.17	1992.9997 (0.2365)	0.9509 (0.0627)	4.5981 (0.0860)	5.1881 (0.1799)	0.0041	0.0012
	Jan. 93 – Mar. 94	**1994.33**	**1993.9997**[6] **(0.3023)**	**0.9509 (0.0800)**	**4.5981 (0.0857)**	**5.1881 (0.1937)**	**0.0033**	**0.0013**
		1997.75	missed					
KGCSP	Sep. 92-May 93	1990.00	1992.9997 (0.7327)	0.9509 (0.1933)	4.5981 (0.2118)	5.1881 (0.4759)	0.0032	0.0071
	Jun. 93 – Aug. 94	False alarm	1993.9997 (0.8231)	0.9509 (0.2295)	4.5981 (0.1948)	5.1881 (0.4624)	0.0027	0.0072
	Sep. 94 – Nov. 95	**1994.92**	**1994.9997 (2.8551)**	**0.9509# (0.8340)**	**4.5981 (0.5901)**	**5.1881 (1.4516)**	**0.0007**	**0.0074**
		1996.33	missed					
SHP	Sep. 89 - Oct. 90	1991.25	1989.9997 (1.0090)	0.9509# (2.7059)	4.5981 (1.6484)	5.1881 (1.9031)	0.0006	0.0002

[6] Highlights are best fits.

Bibliography

1. Cvijovi, D. and Klinowski, J. “taboo search: an approach to the multiple minima problem”, Science 267 (1995) 664-666;
2. Davis, L. “handbook of genetic algorithms”, ed. Van Nostrand Reinhold, New York, 1991;Goldberg, D. E. “genetic algorithms in search”, Optimization and Machine Learning, Addison-Wesley, Reading, Mass., 1989;
3. Dennis, J. E. Jr. and Schnabel, R. B. “numerical methods for unconstrained optimization and nonlinear equations” *SIAM*, 1996;
4. Glover, F., Michelon., P. and Valli, Z. “a scatter search approach for unconstrained continuous optimization” Working Paper, University of Colorado, Boulder, 1995;
5. Glover, F. “heuristics for integer programming using surrogate constraints”, Decision Sciences 8/1 (1977) 156-166;
6. Glover, F. and Laguna, M. “tabu search” C. R. Reeves (ed.), Modern Heuristic Techniques for Combinatorial Problems, Blackwell Scientific Publications, Oxford, 1993, 70-141;
7. Glover, F. “genetic algorithms and scatter search: unsuspected potentials”, Statistics and Computing 4 (1994a) 131-140;
8. Glover, F. “tabu search nonlinear and parametric optimization (with links to genetic algorithms)” Discrete Applied Mathematics 49 (1994b) 231-255;
9. Glover, F. “scatter search and start-paths: beyond the genetic metaphor”, OR Spectrum 17 (1995) 125-137;

10. Glover, F. and Laguna, M., “tabu search”, Kluwer Academic Publishers, 1997;
11. Greene, W. H. “econometric analysis”, Prentice Hall, 3rd, ed. Chapter 10;
12. Holland, J. H., “adaptation in natural and artificial systems”, MIT press, 1992;
13. Jensen, M. H., Johansen, A. and Simonsen, I., “inverse fractal statistics in turbulence and finance”, *International Journal of Modern Physics B*, Vol. 17, Nos. 22, 23 &24 (2003) 4003-4012;
14. Jensen, M. H., Johansen, A. and Simonsen, I, “inverse statistics in economics: the gain-loss asymmetry”, *Physica A*, 324 (2003) 338-343;
15. Johansen, A. and Sornette, D., “stock market crashes are outliers”, *the European Physical Journal B*, 1, 141-143 (1998);
16. Johansen, A. and Sornette, D., “modeling the stock market prior to large crashes”, *The European Physical Journal B*, 9, 167-174 (1999);
17. Johansen, A., Ledoit, O. and Sornette, D., “crashes as critical points” International *Journal of Theoretical and Applied Finance*, Vol. 3, No. 2 (2000a) 219-255 ;
18. Johansen, A. and Sornette, D. “log-periodic power law bubbles in Lating-American and Asian markets and correlated anti-bubbles in Western stock markets: an empirical study”, Inte. J. Theo. Appl. Finance (2000b);
19. Johansen, A. and Sornette, D., “the Nasdaq crash of April 2000: yet another example of log-periodicity in a speculative bubble ending in a crash”, *The Europeand Physical Journal B*, 17, 319-328 (2000c);
20. Johansen, A. and Sornette, D., “finite-time singularity in the dynamics of the world population, economic and financial indices”, *Physica A*, 294 (2001) 465-502;

21. Johansen, A., “characterization of large price variations in financial markets”, Physica A, 324 (2003) 157-166;
22. Michalewicz, Z. “genetic algorithms + data structure = evolution programs”, 2nd ed. Springerverlag, New York, 1994;
23. Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. “numerical recipes”, Cambridge University Press, 1988-1992.
24. Simonsen, I., Jensen, M. H. and Johansen, A., “optimal investment horizons” *The European Physical Journal B*, 27, 583-586 (2002);
25. Sornette, D. and Johansen, A., “large financial crashes”, *Physica A*, 245 (1997) 411-422;
26. Sornette, D. and Johansen, A., “a hierarchical model of financial crashes” *Physica A,* 261 (1998) 581-598;
27. Sornette D., “why stock market crash” *Princeton University Press*, 2003;
28. Srinivas, M. and Patnaik, L. M. “genetic algorithms: a survey”, Computer 27/1 (1994) 17-26;
29. Trafalis, T.B. and I. Al-Harkan, “A hybrid scatter genetic tabu approach for continuous global optimization”, Combinatorial and Global Optimization, (P.M. Pardalos, A. Migdalas and R.E. Burkard, eds.), *Series on Applied* Mathematics, 14:11-29, 2002, World Scientific Publishing Co.

---

[i] Finite-time singularity refers to the appearance of an infinite slope or infinite value in a finite time.

[ii] More generally, the price $p(t)$ can be interpreted as the price in excess of the fundamental value of the asset.

[iii] More generally, K could be heterogeneous across pairs of neighbors. Some of the K_{ij} 's could even be negative, as long as the average of all K_{ij} 's was strictly positive.

[iv] In physics, *critical points* are widely considered to be one of the most interesting properties of *complex systems*. A system goes critical when local influences propagate over long distances and the average state of they system becomes exquisitely sensitive to a small perturbation. That is different parts of the system become highly correlated. A critical system is self-similar across scales. Critical self-similarity is why local imitation cascades through the scales into global coordination.

[v] Scale invariance means reproducing oneself on different time or space scales. The hall mark of it is power law. Scale invariance holds exactly at the critical point. When not exactly at the critical point, only a weaker kind of scale invariance, discrete scale invariance, holds. With discrete scale invariance, the system obeys scale invariance only for specific choices of scaling ratio λ. "The signature of discrete scale invariance is the presence of a power law with complex exponent, which manifests itself in data by log-periodic oscillations providing corrections to the simple power law scaling." (Sornette, 2003, page 207).

[vi] The upper and lower bound of parameters are set initially in reference to the empirical results in the series of papers written by Johansen and Sornette.

[vii] The Euclidean norm is the square root of the sum of the absolute squares of the vector elements.

[viii] A move is tabu, if it is both closely located to the previous point and the change in its objective function is very small. The tabu status of a move can be overridden if the aspiration criterion is met.

[ix] In fact, the RAM of my computer had to be upgraded from 512MB to 1GB in order to run the complete program.