A Bayesian Analysis of Log-Periodic Precursors to Financial Crashes*

George Chang[†] and James Feigenbaum[‡]

October 8, 2004

Abstract

A large number of papers have been written by physicists documenting an alleged signature of imminent financial crashes involving so-called log-periodic oscillations–oscillations which are periodic with respect to the logarithm of the time to the crash. In addition to the obvious practical implications of such a signature, log-periodicity has been taken as evidence that financial markets can be modeled as complex statistical-mechanics systems. However, while many log-periodic precursors have been identified, the statistical significance of these precursors and their predictive power remain controversial in part because log-periodicity is ill-suited for study with classical methods. This paper is the first effort to apply Bayesian methods in the testing of log-periodicity. Specifically, we focus on the Johansen-Ledoit-Sornette (JLS) model of log periodicity. Using data from the S&P 500 prior to the October 1987 stock market crash, we find that, if we do not consider crash probabilities, a null hypothesis model without log-periodicity outperforms the JLS model in terms of marginal likelihood. If we do account for crash probabilities, which has not been done in the previous literature, the JLS model outperforms the null hypothesis, but only if we ignore the information obtained by standard classical methods. If the JLS model is true, then parameter estimates obtained by curve fitting have small posterior probability. Furthermore, the data set contains negligible information about the oscillation parameters, such as the frequency parameter that has received the most attention in the previous literature.

JEL Classification: G13, C11.

Keywords: Financial crashes; Bayesian inference; log-periodicity.

---

*The authors would like to thank Dave DeJong, John Geweke, Anders Johansen, Jean-Francois Richard, Gene Savin, Pedro Silos, and Chuck Whiteman for suggestions and discussions on this topic.

[†]Department of Finance; Bloch School of Business; University of Missouri - Kansas City; Kansas City, MO 64110-2499. E-mail: changg@umkc.edu.

[‡]Department of Economics; University of Pittsburgh; 4S34 W. W. Posvar Hall; 230 South Bouquet St.; Pittsburgh, PA 15260. E-mail: jfeigen@pitt.edu. URL: www.pitt.edu/~jfeigen.

An analogy has often been drawn between crashes in financial markets and other disruptive events like earthquakes. Recently, a number of physicists have pursued this qualitative analogy in a more quantitative fashion, suggesting that such complex "rupture" events can be modeled like phase transitions in statistical mechanical systems. A physical system can exist in different phases when the optimal, energy-minimizing structure of the system is different for different values of the exogenous parameters. The parameter space can then be partitioned into regions with different optimal structure. A phase transition occurs when the parameters are adjusted in such a way as to cross the boundary between two such regions. If x is the distance to the boundary, then an observable M that varies with x will typically exhibit a power law relationship $M \sim x^{-\alpha}$ for some α. If the underlying scaling symmetries of the system are discrete rather than continuous, the exponent of this power law can be complex (Sornette (1998)). In that case,

$$M(x) \sim \mathrm{Re}[x^{-\alpha+i\omega}] = \mathrm{Re}[x^{-\alpha}e^{i\omega \ln x}] = x^{-\alpha}\cos(\omega \ln x),$$

and such a periodic relationship with respect to $\ln x$ has come to be known as a log-periodic relationship.

In modeling complex rupture events as phase transitions, physicists have supposed that whatever exogenous parameter x that corresponds to the distance to the phase boundary varies at a constant rate over time, and so the time remaining until the critical time t_c when the boundary will be crossed can stand as a proxy for x. In that case, an observable $M(t)$ of this complex system should exhibit a time-series relationship of the form

$$M(t) \sim (t_c - t)^{-\alpha}\cos(\omega \ln(t_c - t) + \phi),$$

where the phase ϕ is introduced to compensate for the change in units between x and $t_c - t$. This idea of viewing rupture events as phase transitions gained momentum after such a log-periodic relationship was discovered in historical data of ion concentrations within well water near Kobe, Japan prior to the 1995 earthquake there (Johansen et al (1996)). Soon afterwards, two groups independently discovered such a relationship in the S&P 500 prior to the famous October 1987 crash (Feigenbaum and Freund (1996), and Sornette, Johansen, and Bouchaud (1996)) as can be seen in Fig. 1, where the S&P 500 index $s(t)$ is fitted to the specification

$$q(t) = \ln s(t) = A - B(t_c - t)^{\beta}\left[1 + C\cos(\omega \ln(t_c - t) + \phi)\right]. \tag{1}$$

This finding sparked an intensive search of time series data on financial prices, and several examples of such log-periodic precursors to financial crashes were soon identified.[1] In addition to the scientific question of whether financial

[1] For a review of log-periodic research and other examples of how physicists have applied their methods to problems of economic interest, see Feigenbaum (2003). For a more thorough discussion of the evidence in favor of log-periodicity, see Sornette (2003).

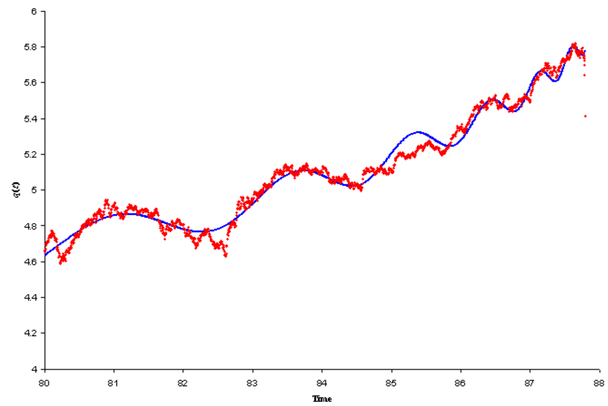

Figure 1: The S&P 500 from 1980 to 1987 and a fit according to Eq. (1) with $A = 6.74$, $B = 0.155$, $C = 2.546$, $t_c = 6/23/88$, $\beta = 0.323$, $\omega = 14.651$, and $\phi = 0.294$. The sum of the squares of the residuals is $e^T e = 4.517$.

markets can actually be modeled as complex systems, there is much practical interest in the question of whether log-periodicity can be used to forecast imminent stock market crashes.

Perhaps the most convincing evidence in favor of a correlation between log-periodic precursors and ensuing crashes comes from systematic searches over all time windows of a given length. Graf v. Bothmer and Meister (2003) fit the Dow Jones Industrial Average for every window of 750 trading days between 1912 and 2000 to a log-periodic specification. If the parameters in the best fit for a given window matched a selected profile, they found there was a 54.1% probability of a crash occurring within a year of the end of that window. Similarly, Sornette and Zhou (2003) investigated the distribution of frequencies ω obtained by fitting a log-periodic specification to every window of a given length. The conditional distribution of ω for windows that ended in crashes was significantly different from the distribution conditional on the window not ending in a crash. In particular, the distribution conditional upon a crash had a mean of 6.4 and a standard deviation of 1.6, highlighting the common observation that log-periodic spells which precede crashes usually have frequencies within a narrow range. This universality of the log-periodic frequency has been interpreted by some as further evidence of an underlying mechanism that governs the leadup to a crash.[2]

[2] Although he uses a more stringent definition of both a crash and a log-periodic spell than other authors who have done systematic searches, Johansen (2004) argues that there is very nearly a 1-1 correspondence between log-periodic precursors and crashes that cannot be linked to a clear external cause.

Nevertheless, since the existence of such log-periodic precursors would have radical implications for the Efficient Markets Hypothesis (Fama (1970)), much skepticism persists about whether they, indeed, are statistically significant and–even if, for the sake of argument, we accept that they are significant–whether existing models to explain these precursors are valid (Feigenbaum (2001a,b), Ilinski (1999), Laloux et al (1999,2002)). Most investigations into log-periodicity have focused on whether the time series of financial prices fits well to a log-periodic specification or whether the Fourier transform of the series with respect to $\ln(t_c - t)$ has large peaks. However, as Phillips (1986) has shown, nonstationary time series like financial time series can produce regressions with deceptively high measures of goodness of fit that do not reflect any properties of the underlying data generating process.

According to the leading model of log-periodicity, the Johansen-Ledoit-Sornette (JLS) model (2000), log-periodicity in the price process is a consequence of herding behavior on the part of irrational investors, which causes the probability of a crash to vary log-periodically. Rational investors perceive this log-periodic variation in the crash probability, so, if markets are efficient, stock returns must reflect this time-varying probability. Note that because markets are efficient in the JLS model, *rational investors in this model cannot exploit the information about possible future crashes conveyed by a log-periodic trend.*

If the JLS model is true, it is not just stock prices but also the daily returns on stock prices, i.e. the first differences $q_t - q_{t-1}$, that must behave log-periodically. Feigenbaum (2001) implemented a test of this more rigorous prediction about daily returns. However, classical estimation of this log-periodic model is complicated because the likelihood function has an infinity of local maxima, and no one maximum totally overshadows all the others. The small-sample properties of nonlinear least squares estimation in such a context are not well-understood, so it is difficult to assess the statistical significance of a nonzero log-periodic component in the largest peak. Disregarding any secondary peaks, Feigenbaum (2001) used Monte Carlo simulations to estimate standard errors for linear coefficients and obtained mixed results regarding the statistical significance of the amplitude C of the oscillations in the precursor to the 87 crash.[3] The null hypothesis of no log-periodicity could be rejected or not rejected, depending on the beginning and end points of the data set. Moreover, when the log-periodic coefficients are close to zero, the variance-covariance matrix of the nonlinear parameters will be the inverse of a nearly singular matrix, so precise estimates of the frequency ω, exponent β, and other nonlinear parameters could not be obtained.

In the present paper, we examine the same data with an alternative statistical paradigm that sidesteps these technical issues that plague classical estimation. Whereas classical estimation tries to construct functions of the observed data that converge in probability to the parameters, Bayesian estimation employs a more straightforward approach. In the Bayesian paradigm, the

[3]Feigenbaum (2001) did not impose the constraint $|C| \leq B$ imposed in this paper, so the paper tested for log-periodicity primarily by testing for the significance of the oscillation term.

researcher defines his prior beliefs about the parameters as a probability distribution over the parameter space and then uses Bayes' Law to update those beliefs based on the observed data, producing a so-called posterior distribution for the parameters. This procedure is not dependent on any special assumptions about the likelihood function. We can avoid the complicated problem of obtaining the true global maximum. Bayesian methods also provide intuitive and exact finite-sample inference regarding any function of the parameters without relying on asymptotic distributions.

Naturally, all of these advantages come with a price. We have to choose a prior distribution for the parameters, and the results may be sensitive to this choice of a prior. Because there is a wide variance of opinion about the validity of the log-periodic hypothesis, we consider two sets of priors. Both priors satisfy the restrictions on the parameters imposed by the JLS model. To match the preponderance of the evidence regarding the frequency ω, in both cases ω has a mean of 6.4 and a standard deviation of 1.6. Within these constraints, we consider one model that an agnostic financial economist might favor with a diffuse prior distribution spread wide over most of the parameter space. We also consider a model that a log-periodic researcher might favor with a tight prior distribution concentrated around the parameters favored by classical curve-fitting methods.

In addition to varying the priors, we consider variations of the model along two other dimensions. One of the prime advantages of the Bayesian paradigm is that we can take into account the JLS model's predictions regarding the probability of a crash on each trading day, something that cannot be incorporated into a curve-fitting analysis. To compare to the previous literature, we begin by analyzing the marginal likelihood and posteriors for a model that ignores the crash dynamics, and we then go on to analyze the full JLS model with crash probabilities. In addition, we also assess the issue of how time is measured in the model. Researchers have generally studied log-periodicity with calendar-time models that measure the time t of (1) in terms of physical days, but we also look at a market-time model that measures t in trading days.

Using data from 1983 to 1987, we find in all cases that the market-time model has a marginal likelihood orders of magnitude better than the calendar-time model. Disregarding the contribution of crash probabilities to the likelihood function, as has been done in the previous literature, then for all four combinations of timing conventions and priors, the marginal likelihood of observing the data is higher in a model where the log-periodic coefficient B is restricted to zero than in a model where it is distributed over nonzero values. While the difference is small with diffuse priors, the posterior probability for the JLS model is quite small compared to the corresponding null hypothesis model with tight priors. This implies that the tight priors are concentrated in a region with low posterior probability. Consistent with this result, we always find that the posterior distribution for the sum of squared residuals minimized by classical methods has a mode away from the minimum, so classical methods do not obtain the most likely set of parameters in the JLS model. Moreover, while the previous literature has focused much attention on the frequency ω, with diffuse priors we find that the data set contains negligible information about the oscillatory parameters of the JLS model, i.e. the frequency, amplitude, and phase. Precisely speaking, the posterior distributions for the oscillatory parameters are essentially unchanged from their prior distributions.

If we do account for the crash probabilities, the marginal likelihood comparison reverses in favor of the JLS model with diffuse priors. However, this happens because the JLS model gives a high probability for a crash to occur on 10/19/87. Our findings regarding the oscillatory parameters remain the same. Overall, Bayesian methods find no evidence that log-periodic oscillations in daily returns are responsible for log-periodic precursors to financial crashes.

We must emphasize that the Bayesian framework is only capable of testing fully specified probability models. As such, while the evidence reported here counts against the JLS model, we make no claims regarding the general log-periodic hypothesis that log-periodic spells are a signal of an imminent crash. The log-periodic phenomenon may indeed be real, but if it is it then it would appear the explanation for the phenomenon remains a mystery.

The paper is organized as follows. In Section 1, we give a brief introduction to the subject of Bayesian inference. In Section 2, we review the JLS model. In Section 3, we present the first probability model that we estimate, which assumes the stock market is in a log-periodic regime for the entire data set and which does not explicitly take into account the probability of a crash. In Section 4, we discuss our choice of priors and compute marginal likelihoods. In Section 5, we describe the resulting posterior distributions. In Section 6, we discuss the posterior distribution of the sum of squared residuals, which is the focus of most of this literature. In Section 7, we repeat the analysis for the model with crash probabilities. Finally, we conclude in Section 8.

1 Bayesian Inference

Bayesian inference proceeds by computing the likelihood of observing a set of data for a given probability model.[4] The results are summarized as a probability distribution for the parameters of the model and also for unobserved quantities such as forecasts of future observations. Thus, Bayesian statistical conclusions about a parameter θ are made in terms of probability statements conditional on the observed data Q. The so-called posterior probability distribution $p(\theta \mid Q)$ contains all current information about the parameter θ. In order to make probability statements about θ given Q, we must begin with a *model that* provides a *joint probability distribution* for both θ and Q:

$$p(\theta, Q) = p(\theta)\, p(Q \mid \theta).$$

[4]For a more complete introduction to the subject of Bayesian inference, consult Berger (1985), Bernardo and Smith (1994), or Gelman et al (2003).

The distribution $p(\theta)$ represents the modeler's prior beliefs regarding the parameters θ as they stand before he confronts the data. The sampling distribution or likelihood function $p(Q \mid \theta)$ is the probability of observing the data under the model if θ is the parameter vector.

Conditional on the known values of the data Q, **Bayes' rule** yields the *posterior* density:

$$p(\theta \mid Q) = \frac{p(\theta, Q)}{p(Q)} = \frac{p(\theta)\, p(Q \mid \theta)}{p(Q)}, \tag{2}$$

where $p(Q) = \int p(Q \mid \theta)\, p(\theta)\, d\theta$. Note that $p(Q)$, known as the *marginal likelihood*, is obtained by integrating the likelihood function $p(Q \mid \theta)$ over the whole parameter space with respect to the measure $p(\theta)\, d\theta$. The marginal likelihood is important because it is the posterior likelihood that the model is correct, whatever the unobservable parameters might be.

Since $p(Q)$ does not depend on θ and can be considered a constant for a given data set Q,

$$p(\theta \mid Q) \propto p(\theta)\, p(Q \mid \theta) \tag{3}$$

is an an *unnormalized posterior density*. The primary task of any specific application of Bayesian inference is to develop the model $p(\theta, Q)$ and perform the necessary computations to summarize $p(\theta \mid Q)$ in appropriate ways. When the posterior distribution $p(\theta \mid Q)$ does not have a closed form, various posterior simulation methods can be used to access the posterior distribution. In this paper we will use the importance sampling methodology described in Appendix C.

When a discrete set of competing models is proposed, the term *Bayes factor* is sometimes used for the ratio of the marginal likelihood $p(Q|A_i)$ under one model A_i to the marginal likelihood $p(Q|A_j)$ under a second model A_j. That is,

$$\text{Bayes factor}\,(A_i; A_j) = \frac{p(Q \mid A_i)}{p(Q \mid A_j)} = \frac{\int p(\theta_{A_i}|A_i)\, p(Q|\theta_{A_i}, A_i)\, d\theta_{A_i}}{\int p\left(\theta_{A_j}|A_j\right) p\left(Q|\theta_{A_j}, A_j\right) d\theta_{Aj}},$$

where θ_{A_i} and θ_{A_j} are the vector of parameters for the models A_i and A_j respectively, which need not be the same. Suppose the researcher has prior beliefs $p(A_i)$ and $p(A_j)$ regarding the probability that each of these models is the correct model. Then the ratio of the posterior probabilities of these models is

$$\frac{p(A_i|Q)}{p(A_j|Q)} = \frac{p(A_i)p(Q|A_i)}{p(A_j)p(Q|A_j)} = \frac{p(A_i)}{p(A_j)} \times \text{Bayes factor}\,(A_i; A_j)\,.$$

Thus, if the Bayes factor of model A_i over A_j is greater than 1, the posterior of A_i will increase more relative to its prior than the posterior of A_j relative to its prior. It is through a comparison of Bayes factors that we will judge how well the JLS model explains the data relative to an alternate hypothesis.

2 The JLS Model

The Johnasen-Ledoit-Sornette (JLS) (2000) model of log-periodic precursors describes the market for a financial asset with price $s(t)$ that pays no dividends, so any nonzero price path for the asset constitutes a bubble.[5] Two types of agents participate in the market. First, there are enough rational agents to ensure that the market behaves efficiently. These agents are identical in their preferences and any other characteristics, so they can be lumped together as one representative agent. Second, there is also a group of irrational agents whose herding behavior leads to the crashes in this model.

The irrational agents reside on a network with a discrete scaling symmetry. For example, this network could have a tree structure where every node is joined to Γ other nodes without any closed loops. Each of these irrational agents can be in one of two states: bullish or bearish. Let τ_{it} represent the state of irrational agent i at time t, where $\tau_{it} = 1$ if the agent is bullish and $\tau_{it} = -1$ if the agent is bearish. Irrational agents determine their beliefs about the future of the market based largely on the influence of their nearest neighbors. If a majority of his neighbors is bearish, i will likely be bearish also, and conversely if they are bullish. Adding stochastic noise to allow the beliefs to change over time, we model the states of the irrational agents as following the Markov process

$$\tau_{i,t+1} = \operatorname{sgn}\left(K \sum_{j \in N(i)} \tau_{s,t} + \varepsilon_{i,t+1} \right),$$

where K is a positive coupling constant, $N(i)$ is the set of i's nearest neighbors, and $\varepsilon_{i,t+1}$ is a mean-zero, i.i.d. random variable.

This model is very similar in structure to the Ising model of ferromagnetism in statistical mechanics, where the τ_i correspond to the spin or magnetization of each component atom of the ferromagnet. The Ising model exhibits two phases of behavior. When the coupling constant K is high relative to the standard deviation of the noise process σ_ε, which is analogous to the system's temperature, the system will eventually settle into an ordered phase where all the spins have the same direction. In this phase, the aggregate sum $\sum_i \tau_i$ will be large in magnitude, and the system will have a measurable magnetization on the macroscopic level. When the coupling constant K is low relative to σ_ε, the system will be disordered. There will be domains of the network where the τ_i are positive and other domains where the τ_i are negative. The aggregate sum $\sum_i \tau_i$ and the aggregate magnetization will be close to zero.

In the JLS network model, a crash can be viewed as a transition from the disordered phase to an ordered phase where the bulk of irrational agents are bearish. In the disordered phase, irrational agents are split roughly equally between bullish and bearish opinions, and each group's influence on the market

[5]See Blanchard and Fischer (1989) for a review of bubble solutions.

cancels the other out. The rational beliefs of the rational agents then determine the price of the asset, which evolves in a fashion consistent with the Efficient Markets Hypothesis (Fama (1970)). In contrast, when the market transits into an ordered phase where the irrational agents all believe the price of the asset is going to fall, they all unload their holdings in the asset, causing a precipitous decline in the price.

Let σ_c denote the critical standard deviation that divides the ordered phase from the disordered phase in the analogous Ising model. Because of the discrete scale invariance of the network, aggregate properties of the ferromagnet will have a log-periodic dependence on the distance between σ_ε and σ_c. Based on this result, JLS postulated that the probability of the trading network going from a disordered phase to an ordered phase, i.e. the probability of a crash, should also vary log-periodically with respect to time.[6] The rational agents, who know the probability of a crash, respond accordingly, and this causes the price of the asset to also exhibit a log-periodic time dependence.

In more precise terms, JLS claim that the hazard rate of a crash will vary log-periodically. Consider an event that occurs at a stochastic time $\widetilde{T} \geq 0$. Let $F(t) = \Pr[\widetilde{T} \leq t]$ be the cumulative distribution function (cdf) for $\widetilde{T}$ and $f(t) = F'(t)$ be the corresponding probability density function (pdf). Then the hazard rate,

$$h(t) = \frac{f(t)}{1 - F(t)},$$

is usually interpreted as the probability that the event occurs at t given that it has not already occurred. This is not quite correct, however, since h is a density and not a probability. The proper interpretation of h is discussed in Appendix A, where we show that for $t_2 > t_1$

$$\Pr[\widetilde{T} \leq t_2 | t_1 \leq \widetilde{T}] = 1 - \exp\left(-\int_{t_1}^{t_2} h(t')dt'\right). \tag{4}$$

The hazard rate function determines the probability that a crash will occur at time t given that a crash has not yet occurred. In the JLS model, $h(t)$ evolves log-periodically in $t_c - t$, where t_c is a critical time (and a parameter of h). If a crash occurs, then $\ln q(t)$ will fall by some random amount κ drawn from a distribution with mean $\overline{\kappa}$. The model can leave the log-periodic regime in two ways. Either there is a crash, or t gets to t_c without a crash. *Note that t_c should not be interpreted as the crash time.* It is a critical time at which the potential for a crash subsides. Indeed, according to the model the crash must occur before t_c.

Suppose that $h(t)$ has the specification

$$h(t) = B'(t_c - t)^{-\alpha}[1 + C\cos(\omega \ln(t_c - t) + \phi')]. \tag{5}$$

[6] Note that the analogy between the time to a crash and the distance in the phase space to the transition boundary is not exact, and this postulate has never been rigorously established.

As Graf v. Bothmer and Meister (2003) point out, the hazard rate must be positive, so we must have $B' > 0$ and $|C| < 1$. JLS (2000) impose the further restriction $\alpha = 1/(\Gamma - 1) \in (0, 1)$. Without loss of generality we can assume that $\omega \geq 0$ and $\phi \in [0, 2\pi)$. The latter assumption also allows us to restrict C to be positive. Finally, if t^* is the time of the crash, we must have $t_c \geq t^*$.

In the absence of a crash, the price of the asset $s(t)$ is assumed to be a martingale process, consistent with the Efficient Markets Hypothesis:

$$E[ds(t)] = 0. \tag{6}$$

Let $j(t)$ be a random variable that is 1 if the crash has occurred as of time t and is 0 if no crash has yet occurred. Suppose that the price process is

$$dq(t) = \frac{ds(t)}{s(t)} = \mu(t) + \sigma dz(t) - \kappa dj(t),$$

where $dz(t)$ is a mean-zero, unit-variance stochastic innovation and $\mu(t)$ is a deterministic drift function that will be set so that $s(t)$ satisfies the martingale condition (6). JLS say little about the distribution of z. We will assume that dz is an i.i.d. normally distributed variable. The standard deviation σ will be drawn from an appropriate prior. Since the probability density that a crash will occur at t if no crash has occurred so far is $h(t)$,[7]

$$E[ds(t)] = \mu(t)s(t) - \overline{\kappa}h(t)s(t) = 0.$$

Therefore, the drift must satisfy

$$\mu(t) = \overline{\kappa}h(t),$$

so, in the absence of a crash,

$$dq(t) = \overline{\kappa}h(t)dt + \sigma dz. \tag{7}$$

Since the price of the asset is actually measured discretely at times $t_0, \ldots, t_N$, we need to translate (7) into a conditional probability distribution for $q(t_{i+1})$ given $q(t_i)$. Let us define

$$H(t) = \overline{\kappa} \int_{t_0}^{t} h(t')dt',$$

and

$$\Delta H(t_1, t_2) = H(t_2) - H(t_1).$$

Then since we have assumed dz is normally distributed,

$$q(t_{i+1}) - q(t_i) \sim N(\Delta H(t_i, t_{i+1}), \sigma^2(t_{i+1} - t_i)).$$

[7] JLS (2000) disregard the contribution of the stochastic volatility σ on the expectation of ds that comes from Ito's Lemma.

As is shown in Appendix B, if $h(t)$ has the log-periodic specification (5), then

$$H(t) = A - B(t_c - t)^{\beta} \left[1 + \frac{C}{\sqrt{1 + \left(\frac{\omega}{\beta}\right)^2}} \cos\left(\omega \ln(t_c - t) + \phi\right) \right], \tag{8}$$

where

$$\beta = 1 - \alpha$$

$$B = \frac{\overline{\kappa} B'}{1 - \alpha}$$

$$\phi = \phi' - \tan^{-1}\left(\frac{\omega}{1 - \alpha}\right),$$

and A is an unidentifiable normalization constant that can be ignored.

3 Model without Crash Probabilities

For comparison with the previous literature, we first ignore the probability of a crash. Suppose that we have data at times $t_0, t_1, t_2, \ldots, t_N < t_c$. At t_0, when the log price is q_0, we assume we are in a log-periodic regime characterized by the parameter vector $\xi = (B, C, \beta, \omega, \phi, t_c)$. Then Eq. (8) will determine $H(t; \xi)$ for all $t \in [t_0, t_c]$. Since the crash does not occur between t_0 and t_N, for $i = 1, \ldots, N$,

$$q(t_i) \sim N\left(q(t_{i-1}) + \mu(t_i - t_{i-1}) + \Delta H(t_{i-1}, t_i; \xi), \sigma^2(t_i - t_{i-1})\right).$$

Note that we allow for a constant drift μ that is not included in the original JLS model.

We assume the following parameterizations for the priors. The drift is drawn from

$$\mu \sim N\left(\underline{\mu}, \underline{\sigma}^2\right).$$

It is convenient to characterize the distribution for the variance of the daily returns in terms of its inverse, which is known as the precision. The more precisely a random variable is known, the smaller its variance will be. The precision is then drawn from[8]

$$\tau = \frac{1}{\sigma^2} \sim \Gamma\left(\underline{\alpha_p}, \underline{\beta_p}\right).$$

The log-periodic parameters of ξ are drawn independently from distributions appropriate to the support of each parameter:

$$\begin{aligned}
B &\sim \Gamma\left(\underline{\alpha_B}, \underline{\beta_B}\right) \\
C &\sim B\left(\underline{\alpha_C}, \underline{\beta_C}\right) \\
\beta &\sim B\left(\underline{\alpha_\beta}, \underline{\beta_\beta}\right) \\
\omega &\sim \Gamma\left(\underline{\alpha_\omega}, \underline{\beta_\omega}\right) \\
\phi &\sim U\left(0, 2\pi\right) \\
t_c - t_N &\sim \Gamma(\underline{\alpha_t}, \underline{\beta_t}).
\end{aligned}$$

Let $\theta = (\mu, \tau, \xi)$ be the parameter vector. Then the prior density will be

$$\begin{aligned}
p(\theta) \propto\; & \frac{1}{\underline{\sigma_n}} \exp\left[\frac{-\left(\mu_n - \underline{\mu_n}\right)^2}{2\underline{\sigma_n}^2}\right] \tau^{\underline{\alpha_{\sigma^2_{lp}}}-1} \exp\left(-\underline{\beta_{\sigma^2_{lp}}}\tau\right) \\
& \times B^{\underline{\alpha_B}-1} \exp\left(-B\underline{\beta_B}\right) C^{\underline{\alpha_C}-1}(1-C)^{\underline{\beta_C}-1} \beta^{\underline{\alpha_\alpha}-1}(1-\beta)^{\underline{\beta_\alpha}-1} \\
& \times \omega^{\underline{\alpha_\omega}-1} \exp\left(-\omega\underline{\beta_\omega}\right)(t_c - t_N)^{\underline{\alpha_t}-1} \exp(-(t_c - t_N)\underline{\beta_t})
\end{aligned}$$

for $\theta \in \Theta = \mathbf{R} \times \mathbf{R}_+^2 \times [0,1]^2 \times \mathbf{R}_+ \times [0, 2\pi) \times [t_N, \infty)$.

Given θ and q_{t_i}, the probability density for $q_{t_{i+1}}$ will be

$$\begin{aligned}
p(q_{t_{i+1}}|q_{t_i}, \theta) =\; & \sqrt{\frac{\tau}{2\pi(t_{t+1} - t_i)}} \\
& \times \exp\left[-\frac{\tau\left(q_{t_{i+1}} - q_{t_i} - \mu(t_{i+1} - t_i) - \Delta H(t_i, t_{i+1}; \xi)\right)^2}{2(t_{i+1} - t_i)}\right].
\end{aligned}$$

Then the posterior density will be

$$p^{nc}(\theta|q_{t_0}, q_{t_1}, \ldots, q_{t_N}) \propto p(\theta) \prod_{i=0}^{N-1} p(q_{t_{i+1}}|q_{t_i}, \theta).$$

[8] The beta distribution $B(\alpha, \beta)$ has support [0,1] with density proportional to $x^{\alpha-1}(1-x)^{\beta-1}$. The mean is $\alpha/(\alpha+\beta)$ and the variance is $\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$. The gamma distribution $\Gamma(\alpha, \beta)$ has support $[0, \infty)$ with density proportional to $x^{\alpha-1}\exp(-\beta x)$. The mean is α/β and the variance is α/β^2.

4 Marginal Likelihood

We compute the log marginal likelihood as

$$\mathcal{L} = \ln(\mathcal{M}) = \ln\left(\int_{\Theta} p\,(\theta)\, p\,(Q|\theta)\, d\theta\right)$$

where

$$p(Q|\theta) = \prod_{i=0}^{N-1} p(q_{t_{i+1}}|q_{t_i}, \theta),$$

and $Q = (q_{t_0}, q_{t_1}, \ldots, q_{t_N})$. The marginal likelihood of a model can be interpreted as the likelihood that the model is true given the observed data.

Initially, we considered the same data set as Feigenbaum (2001a), the S&P 500 from January 2, 1980 to October 19, 1987. Actually, in Feigenbaum (2001a), the data set was cut off at September 30, 1987, following the common practice in the curve-fitting literature to not include the last days before the crash. This end-cutting practice has never been given much justification other than that it leads to improved fits and is not appropriate when comparing the likelihoods of different models, although we will discuss what happens when we end-cut the data later in the paper.

We started with the following relatively diffuse priors with large variances that would hopefully encompass the "true" values of the parameters:

$$\begin{aligned} \mu &\sim N(0.0003, (0.01)^2) \\ \tau &\sim \Gamma\left(1.0, 10^{-5}\right) \\ B &\sim \Gamma(1.0, 100) \\ C, \beta &\sim U(0,1) \\ \omega &\sim \Gamma(16.0, 2.5) \\ \phi &\sim U(0, 2\pi) \\ t_c - t_N &\sim \Gamma(1.0, 0.01). \end{aligned} \tag{9}$$

We chose $E[\mu] = 0.0003$ and $E[\tau] = 10^5$ to roughly match the behavior of daily returns for the 80s. Since B is the critical coefficient that determines whether there is any log-periodicity or not, we chose $E[B] = V[B]^{1/2} = 0.01$ to encompass a wide range of possible values. Since fits typically obtain values of t_c within a few months after the crash, we chose $E[t_c - t_n] = V[t_c - t_n]^{1/2} = 100$. Since β, C, and ϕ are bounded both above and below, we chose uniform priors that give equal weighting to all points in their domains. Finally, we chose $E[\omega] = 6.4$ and $V[\omega]^{1/2} = 1.6$ to match the observations reported by Sornette and Zhou (2003) based on log-periodic fits over several data sets.

Let $A_{lp,c}^{nc}$ be the model with these priors and calendar (real) time. Let $A_{lp,m}^{nc}$ be the model with these priors and market time (so t_i is replaced by i). Let $A_{n,c}^{nc}$ be a null hypothesis model with H set to zero and calendar time. Let $A^{nc}_{n,m}$ be the same null hypothesis model in market time. Finally, let $Q^{t_l}_{t_k} = (q_{t_k}, \ldots, q_{t_l})$. With 1,000,000 simulations we find that

$$\begin{aligned}
\mathcal{L}\left(A^{nc}_{n,c}|Q^{10/16/87}_{1/2/80}\right) &= 6373.9034 \pm 0.0580 \\
\mathcal{L}\left(A^{nc}_{lp,c}|Q^{10/16/87}_{1/2/80}\right) &= 6373.3688 \pm 0.0707 \\
\mathcal{L}\left(A^{nc}_{n,m}|Q^{10/16/87}_{1/2/80}\right) &= 6421.5131 \pm 0.0531 \\
\mathcal{L}\left(A^{nc}_{lp,m}|Q^{10/16/87}_{1/2/80}\right) &= 6421.0307 \pm 0.0681
\end{aligned}$$

Notice that in both market and calendar time, the marginal likelihood is essentially the same whether we include log-periodicity or not. This suggests that the likelihood function does not significantly depend on the log-periodic parameters ξ, which it will not for B close to 0. Nevertheless, with both timing conventions, the marginal likelihood is higher for the null hypothesis model.

As we have discussed, one issue that comes up in Bayesian inference is that the results may be dependent on the choice of a prior distribution. Although the priors (9) may seem reasonable to an economist who is skeptical at best regarding the log-periodic hypothesis, the hypothesis has been around for almost a decade, so there are researchers who have strong priors that the hypothesis is probably valid. Since the hypothesis is based on the results of nonlinear least squares (NLLS) curve fitting, and the JLS model in particular is based on the hypothesis that oscillations in $q(t)$ reflect oscillations in $E[q(t)] = H(t)$, the formation of a log-periodic researcher's priors would presumably be guided by such curve fitting. However, for the 80-87 data set the best fit to the specification (1) has $C = 2.546$, which is outside the range $0 \leq C \leq 1$ required for the hazard rate to be positive in the JLS model (Graf v. Bothmer and Meister (2003)). Other researchers (Graf v. Bothmer and Meister (2003), and Sornette and Zhou (2003)) have found that the S&P data do fit this restriction if a smaller data set is used, so we next tried the same fit using data only from January 1983 to September 30, 1987. This is presented in Fig. 2 and has parameters $A = 5.918$, $B = 0.013$, $C = 0.966 \in [0, 1]$, $\beta = 0.580$, $\omega = 5.711$, $t_c = 10/20/87$, and $\phi = 4.844$. The parameters that have been given the most attention in the literature are B, β, and ω. Our prior for ω was already chosen to match the observations of log-periodic researchers, and we do not change this. However, we do tighten the parameters of B and β and also t_c, since it originally had a very diffuse prior. For these three parameters, we set their prior means to match the values in the fit. We then set the prior standard deviation for each of these variables to 10% of the corresponding prior mean. Finally, since there was no drift in the JLS model, we center the prior for μ around 0 with a standard deviation smaller than the diffuse standard deviation by a factor of

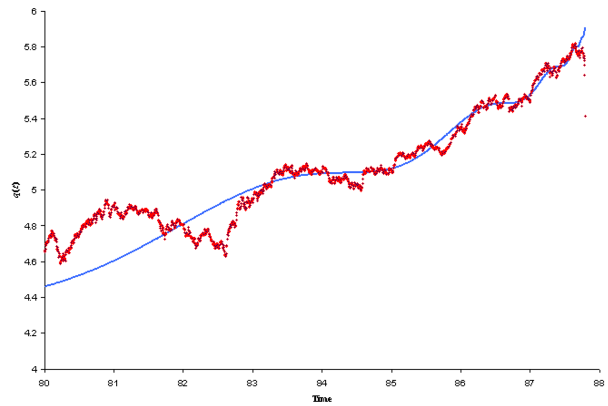

Figure 2: The S&P 500 from 1983 to 1987 and a fit according to Eq. (1) with $A = 5.92$, $B = 0.013$, $C = 0.966$, $t_c = 10/23/87$, $\beta = 0.580$, $\omega = 5.711$, and $\phi = 4.845$. The sum of the squares of the residuals is $e^T e = 1.762$.

ten. This gives the priors

$$
\begin{aligned}
\mu &\sim N(0.0, 10^{-6}) \\
\tau &\sim \Gamma\left(1.0, 10^{-5}\right) \\
B &\sim \Gamma(100.0, 7613.8) \\
C &\sim U(0,1) \\
\beta &\sim B(41.3834, 29.9228) \\
\omega &\sim \Gamma(16.0, 2.5) \\
\phi &\sim U(0, 2\pi) \\
t_c - t_N &\sim \Gamma(100.0, 25.0).
\end{aligned}
$$

We will denote models with this tight prior by B. With 1,000,000 simulations we find that

$$
\begin{aligned}
\mathcal{L}_{nc}\left(B^{nc}_{n,c} | Q^{10/16/87}_{1/3/83}\right) &= 3982.8349 \pm 0.0145 \\
\mathcal{L}_{nc}\left(B^{nc}_{lp,c} | Q^{10/16/87}_{1/3/83}\right) &= 3977.8087 \pm 0.0975 \\
\mathcal{L}_{nc}\left(B^{nc}_{n,m} | Q^{10/16/87}_{1/3/83}\right) &= 4031.1120 \pm 0.0143 \\
\mathcal{L}_{nc}\left(B^{nc}_{lp,m} | Q^{10/16/87}_{1/3/83}\right) &= 4026.8213 \pm 0.0724.
\end{aligned}
$$

If we use the original diffuse priors with this smaller 83-87 data set, we get

$$\begin{aligned}
\mathcal{L}_{nc}\left(A^{nc}_{n,c}|Q^{10/16/87}_{1/3/83}\right) &= 3980.6134 \pm 0.0437 \\
\mathcal{L}_{nc}\left(A^{nc}_{lp,c}|Q^{10/16/87}_{1/3/83}\right) &= 3979.9157 \pm 0.0600 \\
\mathcal{L}_{nc}\left(A^{nc}_{n,m}|Q^{10/16/87}_{1/3/83}\right) &= 4028.9448 \pm 0.0420 \\
\mathcal{L}_{nc}\left(A^{nc}_{lp,m}|Q^{10/16/87}_{1/3/83}\right) &= 4028.6634 \pm 0.0640.
\end{aligned}$$

The log-periodic model actually does better with the diffuse priors. Indeed, under both timing conventions, tightening the prior actually decreases the marginal likelihood of both log-periodic models, which suggests that the posterior mode is not in the region favored by NLLS fitting to $q(t)$. In contrast, tightening the prior for the two null hypothesis models increases the marginal likelihood. Presumably this happens because we are tightening the prior around the posterior mode of the null hypothesis models.

As a side point, if we truncate the data set to end at 9/30/87 as we did when obtaining the curve fitting estimates that guided the prior, the likelihood results for the log-periodic model get closer to the null hypothesis but remain smaller. For the two sets of priors, we get the following:

$$\begin{aligned}
\mathcal{L}_{nc}\left(A^{nc}_{n,c}|Q^{9/30/87}_{1/3/83}\right) &= 3983.5710 \pm 0.0427 \\
\mathcal{L}_{nc}\left(A^{nc}_{lp,c}|Q^{9/30/87}_{1/3/83}\right) &= 3983.4500 \pm 0.0679 \\
\mathcal{L}_{nc}\left(A^{nc}_{n,m}|Q^{9/30/87}_{1/3/83}\right) &= 4024.8567 \pm 0.0409 \\
\mathcal{L}_{nc}\left(A^{nc}_{lp,m}|Q^{9/30/87}_{1/3/83}\right) &= 4024.7732 \pm 0.0600.
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_{nc}\left(B^{nc}_{n,c}|Q^{9/30/87}_{1/3/83}\right) &= 3985.7188 \pm 0.0145 \\
\mathcal{L}_{nc}\left(B^{nc}_{lp,c}|Q^{9/30/87}_{1/3/83}\right) &= 3984.8788 \pm 0.0299 \\
\mathcal{L}_{nc}\left(B^{nc}_{n,m}|Q^{9/30/87}_{1/3/83}\right) &= 4026.9231 \pm 0.0146 \\
\mathcal{L}_{nc}\left(B^{nc}_{lp,m}|Q^{9/30/87}_{1/3/83}\right) &= 4026.5267 \pm 0.0225.
\end{aligned}$$

In sum, for both sets of priors and both sets of timing conventions, the null hypothesis always outperforms the JLS model in terms of marginal likelihood. With diffuse priors, the difference is small. For example, consider the market-time model (and the whole 83-87 data set). The Bayes factor for the JLS model relative to the matching null hypothesis model is $\exp\left(\mathcal{L}_{nc}\left(A^{nc}_{lp,m}|Q^{10/16/87}_{1/3/83}\right) - \mathcal{L}_{nc}\left(A^{nc}_{n,m}|Q^{10/16/87}_{1/3/83}\right)\right) = 0.75$. The data cause an agnostic economist to put less weight on the JLS model, although they do not compel him to quash this weight. In contrast, if we consider the calendar-time model with tight priors, the corresponding Bayes factor is 0.0066. If a researcher has strong priors about the parameters of the JLS model based on the results of NLLS curve fitting, then the JLS model fares quite poorly.

Note that the marginal likelihood assesses the overall likelihood that a model is true without saying anything about its most likely parameter values. Thus, even if the JLS model fares well or not so poorly in terms of marginal likelihood, the parameters favored by the data may not correspond to values that translate to significant log-periodic behavior. To determine the favored parameters, we must compute their posterior distribution, which is the subject of the next section.

5 Posterior Estimation

We use importance sampling, as described in Appendix C, to estimate the posterior distribution for the parameters. For the drift μ and the precision τ, we used a multivariate t distribution with eight degrees of freedom. After using a diffuse source to estimate the posterior moments, the source was fine-tuned so the source means matched the first-stage posterior means and the source variances were chosen to be twice the first-stage posterior variances. For the log-periodic parameters in ξ, we generally used the same distribution as the prior for the source except we doubled the variance.[9] In the special case of a parameter with a uniform prior, we maintained a uniform source.

The means of the posteriors for the eight parameters of the JLS model are given for three versions of the model in Table 1. For the case of the parameters in ξ, the estimates obtained by nonlinear least squares curve fitting to Eq. (1) are also given. Note that for both the market- and calendar-time models, the posterior means are generally far from the NLLS estimates. With the tight priors, the distributions for B, β, ω, and t_c are confined to the neighborhood of the NLLS estimates, so they cannot stray too far from those estimates. Nevertheless, excepting t_c, the means are still beyond two standard errors of the NLLS estimates.

However, the posterior means are not necessarily a good measure of the center of the posterior distributions because some of these distributions are skewed. The posterior marginal densities for each parameter give us a better picture of what values of the parameters are favored by each model.

5.1 Posterior Graphs for Market-Time Model with Diffuse Priors

For the remainder of the paper, we will primarily focus on the market-time model with diffuse priors ($A^{nc}_{lp,m}$) since this has the highest marginal likelihood

[9] If a t_c was drawn that was too close to t_N, this could cause problems so we truncated the source distribution so $t_c > t_N + 0.5$. To be consistent, we must truncate the prior distribution in the same way. This was not done in the marginal likelihood estimates reported above (and below). Nevertheless, truncating the priors had a negligible effect on marginal likelihoods.

	NLLS	$A^{nc}_{lp,c}$	$B^{nc}_{lp,c}$	$A^{nc}_{lp,m}$
μ	-	0.000294	0.000032	0.000420
		(6.16×10^{-6})	(4.59×10^{-6})	(4.78×10^{-6})
τ	-	15757	15698	13437
		(2.39)	(11.70)	(2.43)
B	0.0130	0.007195	0.012553	0.007359
		(3.36×10^{-5})	(2.77×10^{-5})	(3.44×10^{-5})
C	0.966	0.498202	0.721244	0.500030
		(0.001705)	(0.005045)	(0.001618)
β	0.580	0.361958	0.530319	0.384268
		(0.002057)	(0.000998)	(0.001726)
ω	5.711	6.429387	5.940680	6.425584
		(0.005232)	(0.010669)	(0.006496)
ϕ	4.845	3.124951	3.818973	3.149531
		(0.012147)	(0.048272)	(0.009522)
t_c	10/20/87	2/9/88	10/20/87	3/31/98
		(0.504492)	(0.008439)	(0.436634)
$e^T e$	1.762	53.482	64.473	44.897
		(0.233)	(1.379)	(0.189)

Table 1: Estimates of parameters using nonlinear least squares and from posterior means for the 1983 to October 16, 1987 data set without crash probabilities. Standard errors for the posterior means are given underneath. Results are based on one million draws from an importance sampler.

of the four versions of the JLS model that we have considered thus far. In Figs. 3-10, we plot the marginal posterior density for each of the eight parameters. Fig. 3 gives the posterior for the precision τ of the stochastic innovations to the price process. Although JLS (2000) say very little regarding these stochastic innovations, it is necessary to specify a distribution for these distributions to complete the model. Given our choice of a normal distribution for the innovations, the precision is actually the parameter that the likelihood function is most sensitive to. If the precision is too large, the likelihood function will penalize a parameter vector if it gives an expected return series that varies far from the observed return series. On the other hand, if the precision is too small, the daily returns will be allowed to vary much more. Probability will be spread over a larger space of possible return series, putting less density on any return series that is covered, including the observed return series. However, while the likelihood function is quite sensitive to τ, JLS make no claims about it, so it is of less interest.

The drift μ is another parameter that we have added to the model since most financial models would allow for a small drift. Its posterior is given in Fig. 4. A constant drift term would produce exponential variation in the price series. A term of the form $\exp(\mu t)$ does not appear in the log-periodic specification (1) and has generally not been considered by log-periodic researchers since an exponential term would tend to diminish the significance of the nonoscillatory power term. However, the peak of the posterior for μ is clearly away from zero in Fig. 4.

The posterior for the log-periodic coefficient B is plotted in Fig. 5. A

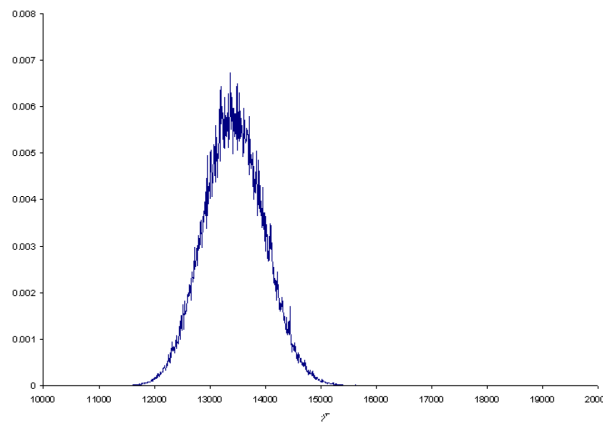

Figure 3: Posterior density for the precision τ in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

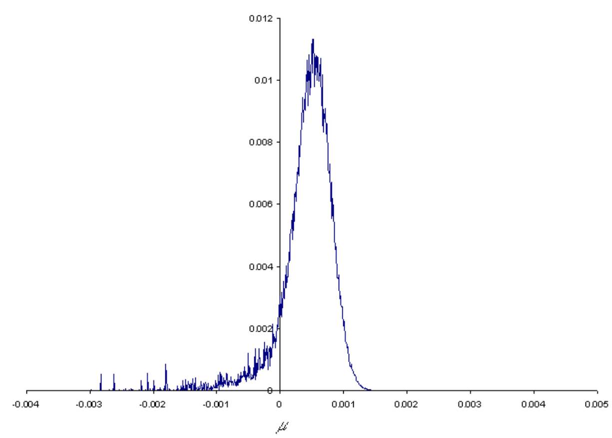

Figure 4: Posterior density for the drift μ in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

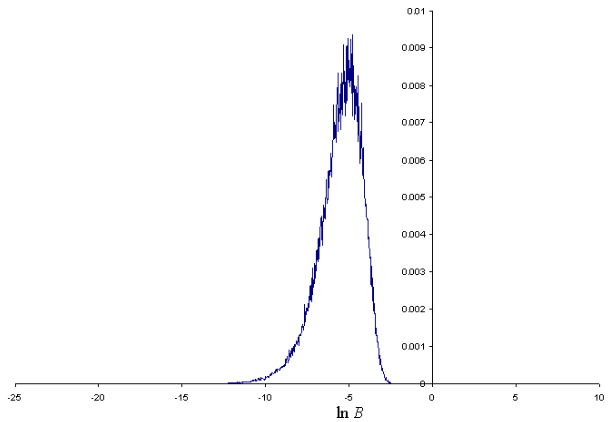

Figure 5: Posterior density for the log of the log-periodic coefficient $\ln B$ in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

classical approach to testing the log-periodic hypothesis would be to test the null hypothesis that $B = 0$. One might think that a natural analog to this in the Bayesian paradigm would be to look at whether the posterior for B peaks at zero. However, the prior density for B at $B = 0$ is zero. That is why we have to consider the null hypothesis where $B = 0$ as the separate model $A^{nc}_{n,m}$. The graph does show though that $\ln B$ peaks about -5 with $B \sim 0.007$, which is roughly the posterior mean. This is smaller than the NLLS estimate of 0.13.[10]

We did not plot the corresponding priors in the above graphs of the precision, drift, and log-periodic coefficient because the data must necessarily be informative about the values of these three parameters. In the remaining graphs, we also give the prior for comparison. If $B = 0$, the likelihood function will not depend on β, t_c, ω, C, or ϕ. In that case, the remaining five parameters would not be identified by the model, so estimating them would be pointless. Since the posterior for B is concentrated around small values, it may still be the case that the likelihood function is practically independent of these parameters, in which case the posterior and prior densities will be the same, and we can view these parameters as "spurious".

Both the posterior and prior densities for the exponent β are given in Fig. 6. Clearly, the data are informative about β since the posterior is monotonically decreasing whereas the uniform prior is flat. In this case, the posterior mean of 0.38 understates how bad the NLLS estimate of $\beta = 0.58$ is

[10]Strictly speaking, we should be comparing the market-time model to NLLS estimates from the market-time model. However, there is not much difference between the posteriors for the market and calendar-time models with diffuse priors.

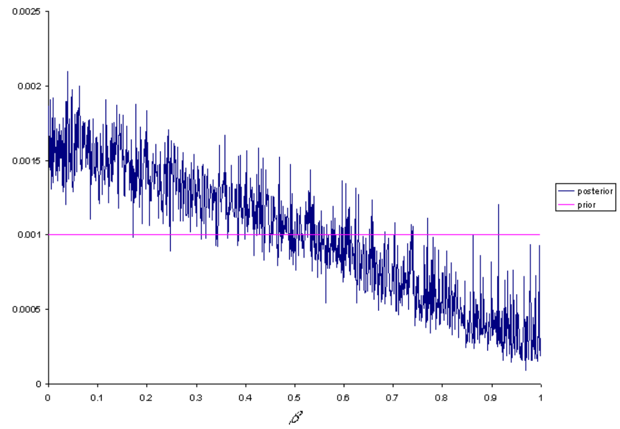

Figure 6: Posterior and prior densities for the exponent β in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

since the posterior mode is near zero.

The posterior and prior densities for the critical time t_c are plotted in Fig. 7. For most of the parameter space the posterior and prior are fairly similar, but the posterior density falls off at small values close to t_N while the exponential prior peaks at zero. The posterior peaks at about 20 trading days after t_N, or the week of 11/9/87. The NLLS estimate, in contrast, is at 10/20/87, or a day after the crash.

Finally, the posterior and priors for the frequency ω, amplitude C, and phase ϕ are given in Figs. 8-10. These three parameters appear only in the oscillatory term of Eq. (1) and will be referred to as the oscillatory parameters. For all three of these parameters, there is no significant difference between the posterior and prior. Because the mean is tightly estimated, the posterior mean of 6.42 for ω is more than three standard errors away from the prior mean of 6.40.[11] However, this 0.3% difference is not economically significant, particularly since there is no diminishment of the variance from the prior to the posterior. Therefore, we have to conclude that the data contains negligible information about the oscillatory parameters of the JLS model. This finding presents a problem for the JLS model since it was these oscillations that the model was invented to explain, yet the model actually has nothing to say about them.

[11]For ϕ and C, the difference between the posterior and prior means is not statistically significant.

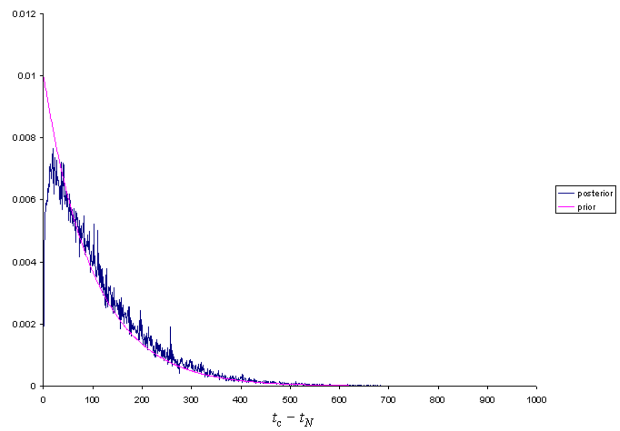

Figure 7: Posterior and prior densities for the critical time t_c in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

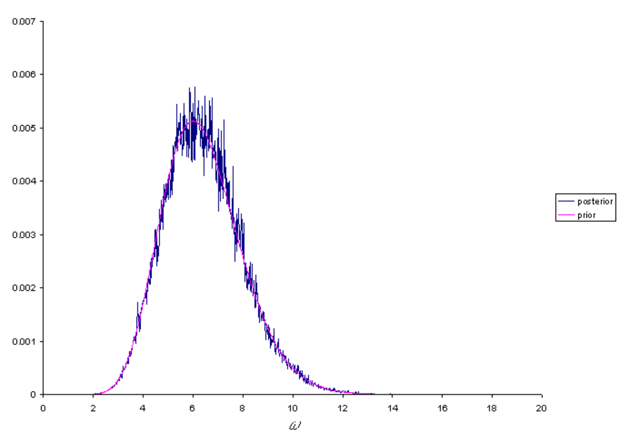

Figure 8: Posterior and prior densities for the frequency ω in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

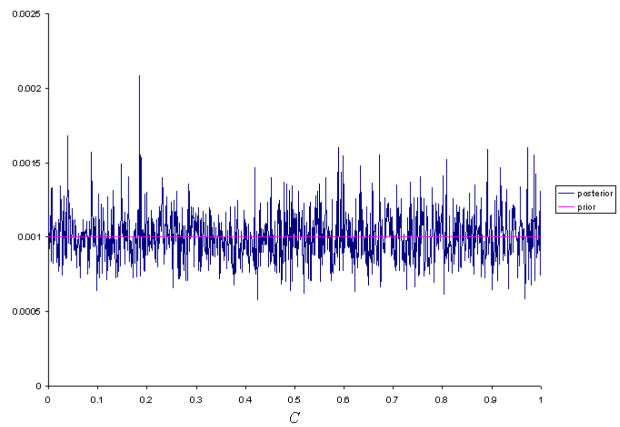

Figure 9: Posterior and prior densities for the amplitude C in the market-time model with diffuse priors and no crash probabilities ($A_{lp,m}^{nc}$).

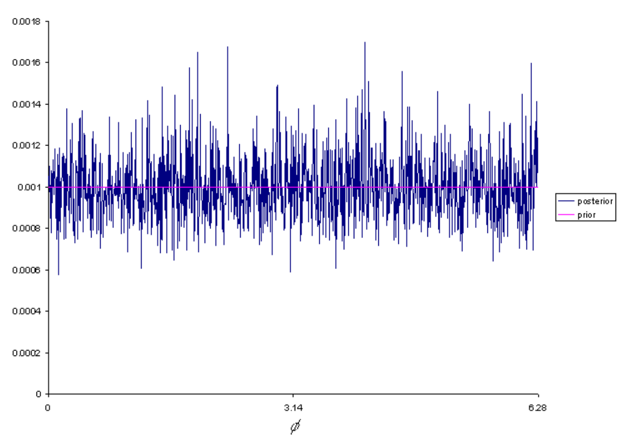

Figure 10: Posterior and prior densities for the phase ϕ in the market-time model with diffuse priors and no crash probabilities ($A_{lp,m}^{nc}$).

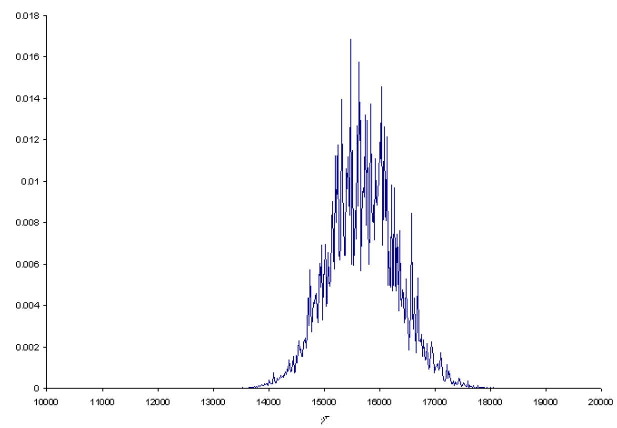

Figure 11: Posterior density for the precision τ in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

5.2 Posterior Graphs for Calendar-Time Model with Tight Priors

The calendar-time model with tight priors is the model that would most resemble what has been discussed in the previous literature, so we also plot the marginal posterior densities for its parameters. These are shown in Figs. 11-18. Unlike with diffuse priors, the posteriors for all the log-periodic parameters, including the oscillatory parameters, differ significantly from the priors.

The tight priors for the exponent β and the critical time t_c both have a single, roughly symmetric peak away from zero, so they are markedly different from the corresponding diffuse priors. Consequently, the posteriors for these parameters are also quite different from the corresponding posteriors with diffuse priors. In Figs. 14-15, both parameters exhibit posteriors with roughly symmetric peaks like the prior. Consistent with the difference between the prior and posterior means for β reported in Table 1, the posterior mode for β is lower than the prior mode, confirming that the model favors lower values of β than would be obtained with NLLS curve fitting. This is also true for t_c although the difference in the prior and posterior modes is less than a day. Note that in both cases the posterior mode cannot move too far away from the prior mode since if the prior puts zero weight on a parameter vector the posterior must do the same.

Fig. 16 demonstrates that the posterior for the frequency ω has at least three peaks. As is shown in Fig. 17, the posterior for the amplitude C is not uniform like for the diffuse prior, but is weighted toward the maximum value of 1. Keeping in mind that the posterior for the phase ϕ has to wrap around

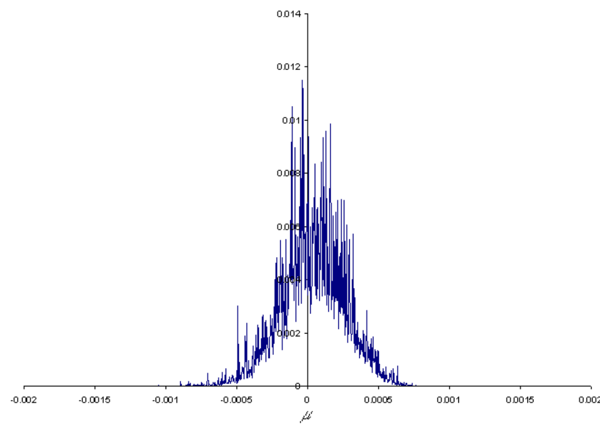

Figure 12: Posterior density for the drift μ in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

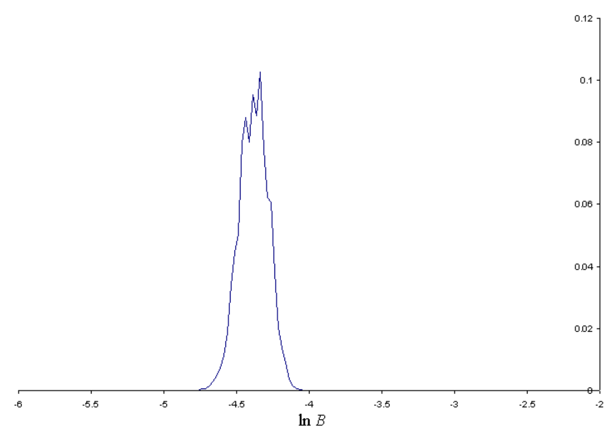

Figure 13: Posterior density for the log of the log-periodic coefficient $\ln B$ in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

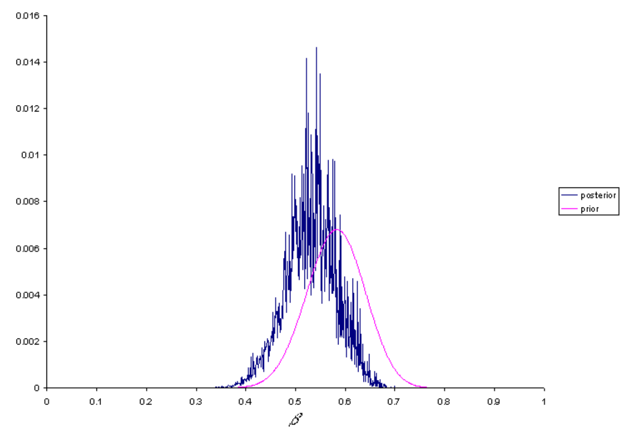

Figure 14: Posterior and prior densities for the exponent β in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

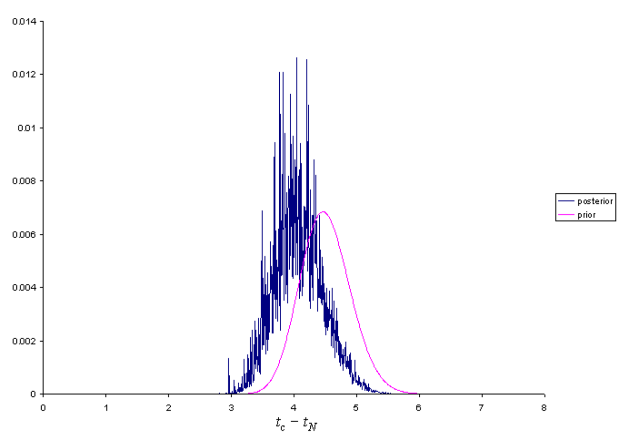

Figure 15: Posterior and prior densities for the critical time t_c in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

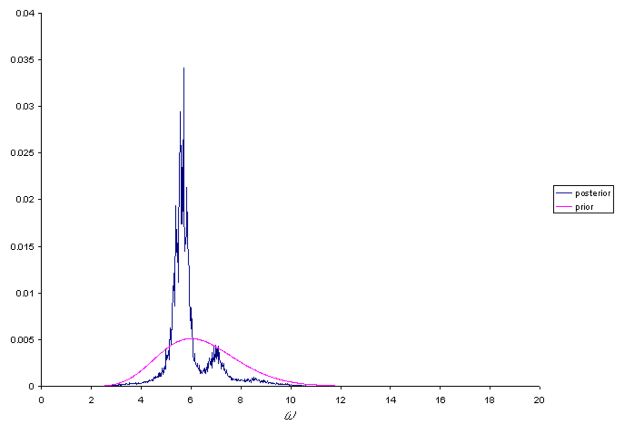

Figure 16: Posterior and prior densities for the frequency ω in the calendar-time model with tight priors and no crash probabilities ($B^{nc}_{lp,c}$).

continuously from 2π back to 0, we see in Fig. 18 that the ϕ posterior has a single peak between $3\pi/2$ and 2π. These results suggest that the likelihood function is not, in fact, completely insensitive to the oscillatory parameters. Tightening the prior allows us to probe the likelihood function at a smaller scale, and this reveals that there is structure to the likelihood function along the oscillatory dimensions.

However, if we do not have a reason to believe that the parameters of a model are most likely to fall in a given region, it does not make sense to concentrate the probability density of our prior within that region. We chose our tight prior based on NLLS curve fitting, but, as we have seen, tightening the prior in this way actually diminishes the marginal likelihood, suggesting that the most likely parameter values are not close to their NLLS estimates. This interpretation is also supported by the above findings regarding t_c and β. In the next section, we will see further evidence of this in the posterior distribution of the residuals sum minimized by the NLLS procedure.

6 Nonlinear Least Squares

In addition to computing the posterior distribution for the parameters of a model, the Bayesian paradigm allows us to compute posterior distributions for any function of the parameters, including the sum of the squared residuals between the raw financial time series and the model's prediction for the time

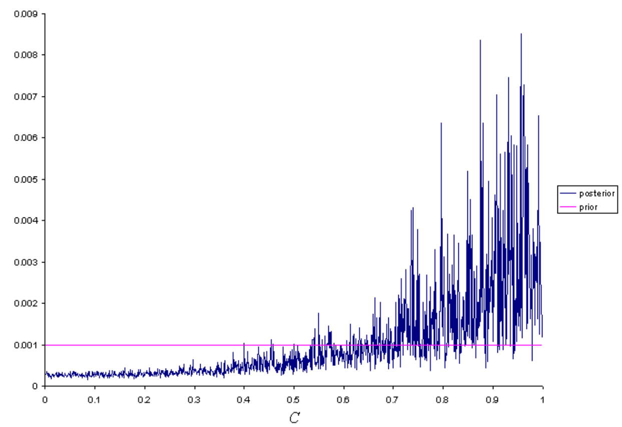

Figure 17: Posterior and prior densities for the log-periodic amplitude C in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

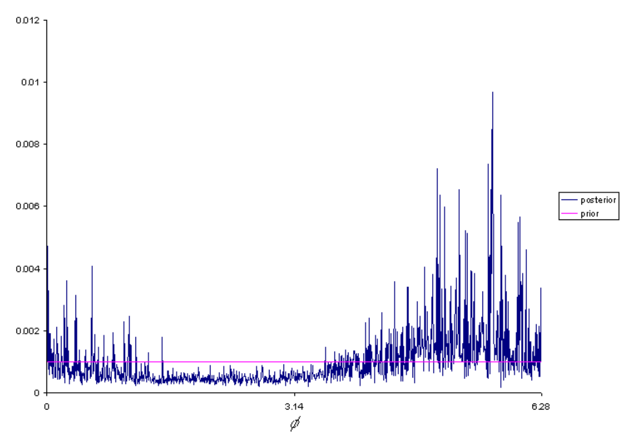

Figure 18: Posterior and prior densities for the phase ϕ in the calendar-time model with tight priors and no crash probabilities ($B_{lp,c}^{nc}$).

series. According to the JLS model, the expectation of the log price at time t_i is

$$E[q_{t_i}] = q_{t_0} + \mu(t_i - t_0) + \Delta H(t_i, t_0). \tag{10}$$

If we define

$$e_i = q_{t_i} - E[q_{t_i}] \quad i = 0, \dots, m,$$

then the NLLS procedure, which has generally been followed in the existing literature, is to choose the parameters ξ that minimize

$$e^T e = \sum_{i=0}^{N} e_i^2.$$

Note that the specification of Eq. (10) differs from the standard log-periodic specification of (1) in two respects. First, as we have noted before, we have included a linear drift term in (10), which generally is not done. Second, and more important, the constant term in (10) is the log of the price at the initial time t_0. In contrast, the constant term in (1) is a free parameter that is also determined by the fitting procedure. Since the linear term in (10) is partially correlated with the non-oscillatory component of the ΔH term, adding the linear term cannot help in minimizing $e^T e$ as much as varying the constant will. As a result, even if we augment the JLS model by adding the linear term to $E[q_{t_i}]$, it is not possible to get values of $e^T e$ as low as one would get by using NLLS to estimate A and ξ in the standard specification of (1).

For example, in the context of the calendar-time model, the smallest value of $e^T e$ obtained by our importance sampler was 2.72. This was obtained with the parameters $\mu = -0.0003$, $B = 0.00964$, $C = 0.769$, $\beta = 0.675$, $\omega = 11.589$, $\phi = 6.259$, and $t_c = 10/20/87$. The corresponding fit used to determine the tight priors in Section 4 had a much smaller $e^T e$ of 1.76. However, if we adjust the constant A in that fit from 5.918 to 5.711 so that $E[q_{t_0}] = q_{t_0}$, then the residual sum increases to 9.12.

The posterior densities for $e^T e$ in the market-time model with diffuse priors and in the calendar-time model with tight priors are respectively given in Figs. 19 and 20. In both cases, most of the density is concentrated around small values of $e^T e$. However, the density is not maximized at the minimum possible value of $e^T e$. For the calendar-time model, the posterior is maximized at $e^T e = 6.25$, which is small but twice as large as the minimum observed value of 2.72. A similar story can be told regarding the market-time model.

Thus, the most likely value of $e^T e$ in the JLS model is not the minimum possible value, and it is easy to see why. The justification for least-squares minimiziation, when it is appropriate, is that the likelihood function is a decreasing function of $e^T e$, as would happen for example if we were estimating a model of the form

$$y_j = \sum_{i=1}^{m} a_i x_{ij} + \varepsilon_j, \tag{11}$$

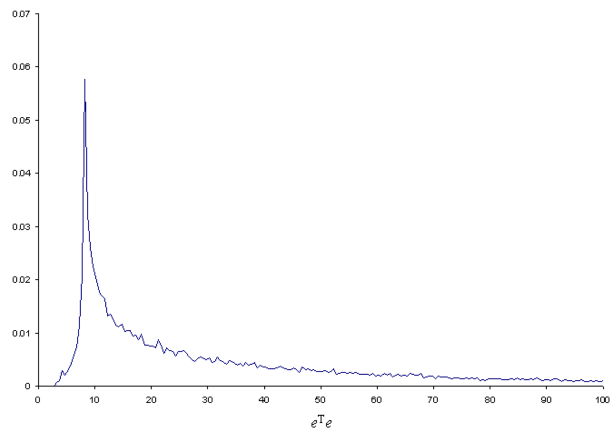

Figure 19: Posterior density for the nonlinear least squares residuals sum $e^T e$ in the market-time model with diffuse priors and no crash probabilities ($A^{nc}_{lp,m}$).

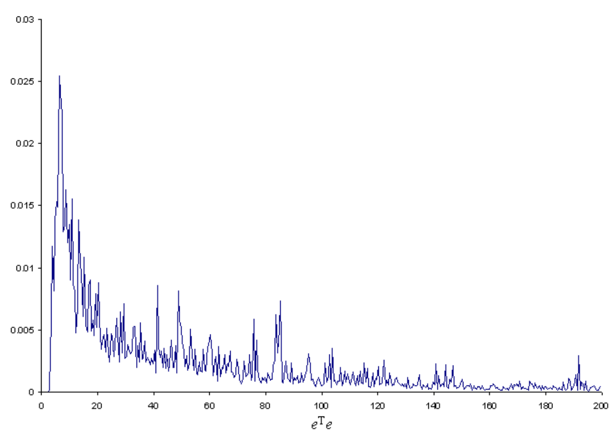

Figure 20: Posterior density for the nonlinear least squares residuals sum $e^T e$ in the calendar-time model with tight priors and no crash probabilities ($B^{nc}_{lp,c}$).

based on J observations of combinations of a dependent variable y and m independent variables $x_1, \ldots, x_m$. If the disturbance term ε_j is normally distributed, then the likelihood function will depend only on the sum of the residuals for this model and the variance of the ε_j.

As was done in Feigenbaum (2001a), one can estimate the JLS model in this matter, but it is the daily returns $q_{t_i} - q_{t_{i-1}}$ that behave as a (nonlinear) model similar to (11), not the q_{t_i} themselves. Consequently, the likelihood function is, modulo an irrelevant constant,

$$\mathcal{L} = \frac{N}{2} \ln \tau - \frac{\tau}{2} S^2,$$

where

$$\begin{aligned} S^2 &= \sum_{i=1}^{N} \frac{\left(q_{t_i} - q_{t_{i-1}} - E[q_{t_i} - q_{t_{i-1}}]\right)^2}{t_i - t_{i-1}} \\ &= \sum_{i=1}^{N} \frac{\left(q_{t_i} - q_{t_{i-1}} - \mu(t_i - t_{i-1}) - \Delta H(t_i, t_{i-1})\right)^2}{t_i - t_{i-1}}. \end{aligned}$$

Thus the likelihood function will be maximized with respect to μ and ξ if S^2 is minimized. While it is true that $e^T e$ and S^2 are highly correlated, it is not true that the global minimum of S^2 corresponds to the global minimum of $e^T e$. Indeed, for the calendar-time model with tight priors, the draw with the highest observed posterior density[12] had $S^2 = 0.0769$ and $e^T e = 5.25$ while the draw with the minimum value of $e^T e = 2.72$ had $S^2 = 0.0778$.

In the language of econometrics, NLLS estimation of the specification (1) does not consistently estimate the parameters of the JLS model. As Feigenbaum (2001a) argued, in order to determine the parameters of the JLS model one needs to focus on the daily returns and not the raw price time series. The JLS model defines a stochastic process for the daily returns, and the prices are obtained by integrating those returns. The literature has generally proceeded in the reverse fashion, by regressing a model of the prices and obtaining returns by differencing the prices, and this backwards approach is not appropriate for model estimation.

7 Model with Crash Probabilities

Thus far, we have considered a model in which daily returns vary log-periodically, but we have ignored the reason why they vary log-periodically in the JLS model, which is that the probability of a crash also varies log-periodically. Classical curve-fitting and spectral methods are unable to take into account this

[12] The parameters of this highest posterior draw were $\mu = 9.87 \times 10^{-5}$, $\tau = 15323$, $B = 0.013$, $C = 0.94$, $\beta = 0.55$, $\omega = 5.70$, $\phi = 4.96$, and $t_c = 10/20/87$.

time-varying crash probability, so the existing literature has been dominated by discussion of an incomplete model.[13] This is of concern because, if the model predicts that the probability of a crash on 10/19/87 is on the order of 10^{-5}, this probability is smaller than the empirical frequency of crashes and the model would not do a very good job of explaining crashes. On the other hand, if the model predicts that the probability of a crash on any day between 1983 and 1987 was on the order of 10%, the model would run into the opposite problem because the crash should have occurred much sooner than 10/19/87.

With Bayesian methods, it is straightforward to complete the model by adding in the crash probabilities, and this is what we will do in this section. To do this we must first specify the distribution of crash sizes. We only have the one crash event in our 1980 to 1987 data set, so we will assume a dogmatic, degenerate distribution for the crash size κ. The S&P 500 fell from 282.70 to 224.84 on 10/19/87, which translates to a fall of 0.22900 in the logarithm q. Thus, we will assume that $\kappa = 0.22900$ with probability 1.

From Eq. (4), conditional upon a crash not occurring at the prior times $t_1, \ldots, t_{i-1}$ a crash will occur at $t_i \geq t_1$ with the probability

$$1 - \exp\left(-\frac{\Delta H(t_{i-1}, t_i)}{\kappa}\right).$$

Thus, the posterior density inclusive of the crash probabilities is

$$\begin{aligned} p^c_{lp}(\theta | Q^{t_{N+1}}_{t_0}) &= p^{nc}(\theta | Q^{t_{N+1}}_{t_0}) \exp\left(-\sum_{i=1}^{N} \frac{\Delta H(t_{i-1}, t_i; \xi)}{\kappa}\right) \\ &\quad \times \left(1 - \exp\left(-\frac{\Delta H(t_N, \min\{t_{N+1}, t_c\}; \xi)}{\kappa}\right)\right). \end{aligned} \tag{12}$$

Note that in the last factor, the probability that a crash happens between t_N and $t_{N+1} = 10/19/87$, we integrate the hazard rate from t_N to either t_{N+1} or t_c, whichever is smaller. Once t_c is reached, if the crash has not occurred already, it will not occur. What happens after t_c is not specified by the JLS model.

As a null hypothesis, we will construct an alternative model with a constant probability of a crash of magnitude κ on any given trading day.[14] From 1962 to 1998, there were 9190 trading days, and there were no other days when the market dropped as much as it did on 10/19/87. Thus, we will assume a constant probability of $p_{cr} = 1/9190$ for the market to drop by κ on any given trading day. Given that a crash did not occur on the initial day in our data set, 1/3/83, there were $N = 1211$ days between 1/3/83 and 10/16/87 on which

[13] Note that there are classical methods that can take into account the crash probabilities. Maximum likelihood estimation (MLE) of the JLS model could incorporate the crash probabilities into the likelihood function as we have done, but, to our knowledge, no researcher has used MLE to estimate the model.

[14] If daily returns are normally distributed, the standard deviation would have to be on the order of 1% to fit most price data, in which case the probability of the market dropping by as much as 20% on any given day would be astronomically small. So we must introduce another mechanism to plausibly account for crashes in the null hypothesis.

a crash again did not occur, and then there was one day in which a crash did occur. Thus, under the null hypothesis the posterior density inclusive of the crash probabilities is

$$p_n^c(\theta|Q_{t_0}^{t_{N+1}}) = p^{nc}(\theta|Q_{t_0}^{t_{N+1}})(1-p_{cr})^N p_{cr}.$$

Taking into account the crash probabilities, we find the following marginal likelihoods with 1,000,000 draws:

$$\begin{aligned}
\mathcal{L}_c\left(A_{n,c}^c|Q_{1/3/83}^{10/16/87}\right) &= 3971.3871 \pm 0.0428 \\
\mathcal{L}_c\left(A_{lp,c}^c|Q_{1/3/83}^{10/16/87}\right) &= 3971.9977 \pm 0.1369 \\
\mathcal{L}_c\left(A_{n,m}^c|Q_{1/3/83}^{10/16/87}\right) &= 4019.7889 \pm 0.0405 \\
\mathcal{L}_c\left(A_{lp,m}^c|Q_{1/3/83}^{10/16/87}\right) &= 4020.4599 \pm 0.0654
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_c\left(B_{n,c}^c|Q_{1/3/83}^{10/16/87}\right) &= 3973.6069 \pm 0.0144 \\
\mathcal{L}_c\left(B_{lp,c}^c|Q_{1/3/83}^{10/16/87}\right) &= 3970.6058 \pm 0.0814 \\
\mathcal{L}_c\left(B_{n,m}^c|Q_{1/3/83}^{10/16/87}\right) &= 4021.8448 \pm 0.0145 \\
\mathcal{L}_c\left(B_{lp,m}^c|Q_{1/3/83}^{10/16/87}\right) &= 4019.8423 \pm 0.0439.
\end{aligned}$$

In the complete model, with tight priors the null hypothesis still wins out over the JLS model in terms of marginal likelihood. On the other hand, with diffuse priors the JLS model now outperforms the null hypothesis. So, once again, we find that the parameter estimates obtained by least-squares estimation are not very probable under the JLS model.

The reversal of marginal likelihoods for the diffuse-prior models is not entirely surprising since, without crash probabilities, the marginal likelihood of the null hypothesis models was only slightly higher than for the JLS models. The JLS model has an explicit explanation why a crash should happen on or around 10/19/87 while the null hypothesis does not. Accounting for the crash probabilities should benefit the JLS model relative to the null hypothesis, and with the diffuse priors it does so enough to push the balance in favor of the JLS model. With market timing, adding crashes increases the Bayes factor for the JLS model relative to the null hypothesis from 0.75 to 1.96.

However, the better performance of the JLS model with diffuse priors at explaining crashes does not translate to better performance at explaining log-periodic oscillations. The posterior means for the diffuse-prior models and for the tight-prior model with calendar timing are reported in Table 2. The posterior densities for the parameters, as well as the residuals sum $e^T e$, of the market-timing model with diffuse priors are plotted in Figs. 21-29. The only parameters that exhibit a major change are the log-periodic coefficient B,

	NLLS	$A^c_{lp,c}$	$B^c_{lp,c}$	$A^c_{lp,m}$
μ	-	0.000345 (7.74×10^{-10})	0.000122 (4.14×10^{-6})	0.000484 (1.22×10^{-6})
τ	-	15747 (2.49)	15687 (13.07)	13434 (2.36)
B	0.0130	0.009917 (4.80×10^{-5})	0.012525 (2.48×10^{-5})	0.009632 (5.11×10^{-5})
C	0.966	0.497774 (0.001420)	0.703691 (0.005681)	0.490394 (0.001662)
β	0.580	0.336835 (0.001031)	0.493638 (0.000957)	0.386423 (0.001132)
ω	5.711	6.409146 (0.005892)	5.988231 (0.014062)	6.426282 (0.006706)
ϕ	4.845	3.124951 (0.012147)	3.662840 (0.057786)	3.138914 (0.010267)
t_c	10/20/87	12/17/87 (0.504492)	10/20/87 (0.010463)	1/27/88 (0.364815)
e^Te	1.762	53.178 (0.202)	60.357 (1.378)	45.19711 (0.190)

Table 2: Estimates of parameters using nonlinear least squares and from posterior means for the 1983 to October 16, 1987 data set with crash probabilities. Standard errors for the posterior means are given underneath. Results are based on one million draws from an importance sampler.

the power β, and the critical time t_c. There was no reason to expect that adding crash probabilities would significantly change the posterior probabilities for the oscillatory parameters, and Figs. 26-28 confirm that the posteriors for these parameters remain essentially unchanged from their priors in the complete model.[15]

What distinguishes B, β, and t_c is that these are the three parameters that play the largest role in determining how the crash probability behaves over time. Since the crash probability at t_i is approximately proportional to $\Delta H(t_i, t_{i-1})$, and ΔH is proportional to B, a higher B means a higher crash probability, and the posterior mean for B increases from .0073 to .0096 when we add crash probabilities. We can also see the shift to higher values of B by comparing Fig. 5 to Fig. 23.

Meanwhile, if we ignore the oscillatory term, the crash probability will go roughly as

$$(t_c - t_i)^\beta - (t_c - t_{i-1})^\beta \sim \beta(t_c - t_i)^{\beta-1}(t_i - t_{i-1}).$$

Since a crash did not occur between 1/1/83 and 10/16/87 and a crash did occur on 10/19/87, the likelihood function should favor a choice of β and t_c that confers a low probability of a crash for most of the 1983 to 1987 and then has the probability shoot up just prior to 10/19/87. This can be achieved by having $\beta \approx 0$ and $t_c \approx 10/19/87$. Comparing Figs. 7 and 25, we find in the complete model that the posterior for low t_c is higher than the prior, in contrast to the

[15] The posterior for ω in Fig. 26 does exhibit one abnormally high point, but this is still within two standard errors of the prior.

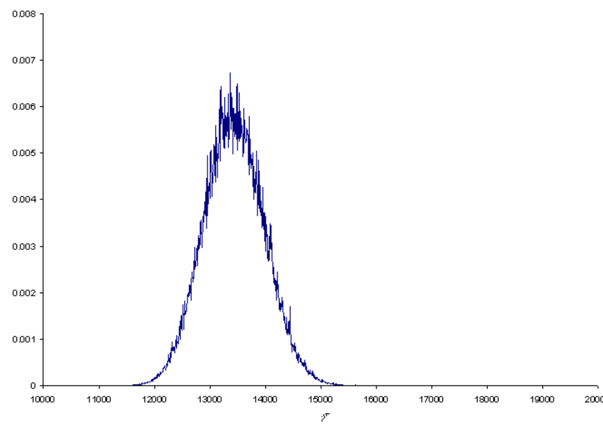

Figure 21: Posterior density for the precision τ in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

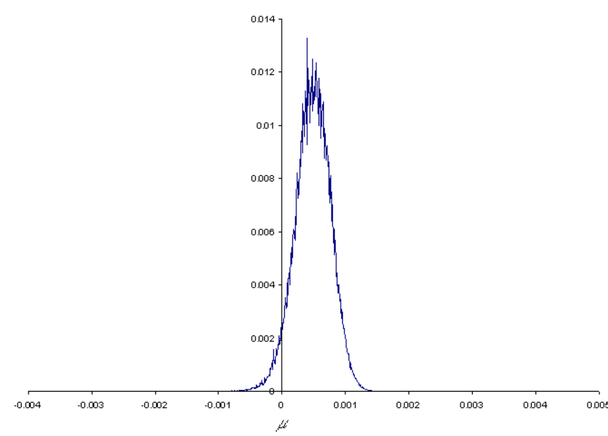

Figure 22: Posterior density for the drift μ in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

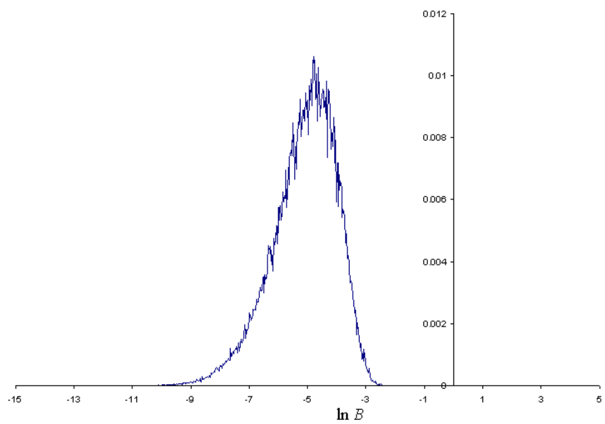

Figure 23: Posterior for the log of the log-periodic coefficient $\ln B$ in the market-time model with diffuse priors and crash probabilities $(A^c_{lp,m})$.

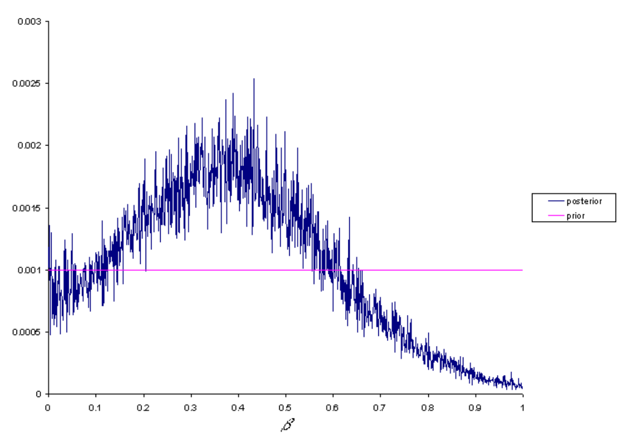

Figure 24: Posterior and prior density for the exponent β in the market-time model with diffuse priors and crash probabilities $(A^c_{lp,m})$.

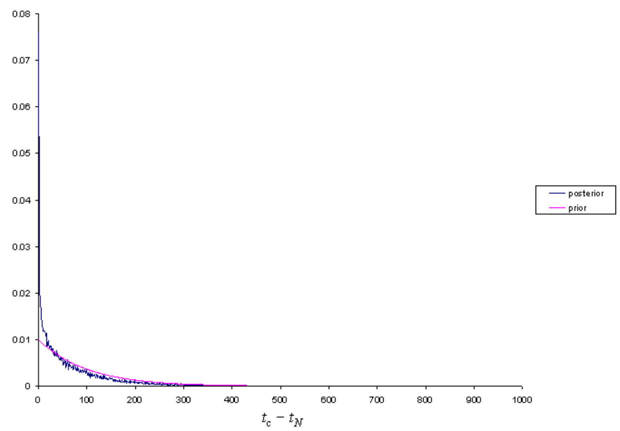

Figure 25: The posterior and prior densities for the critical time t_c in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

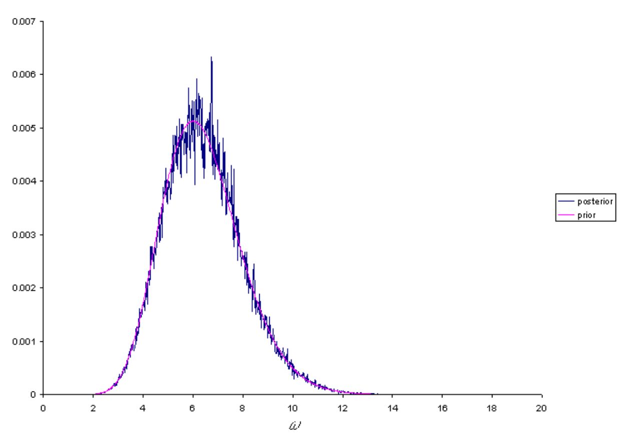

Figure 26: Posterior and prior densities for the frequency ω in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

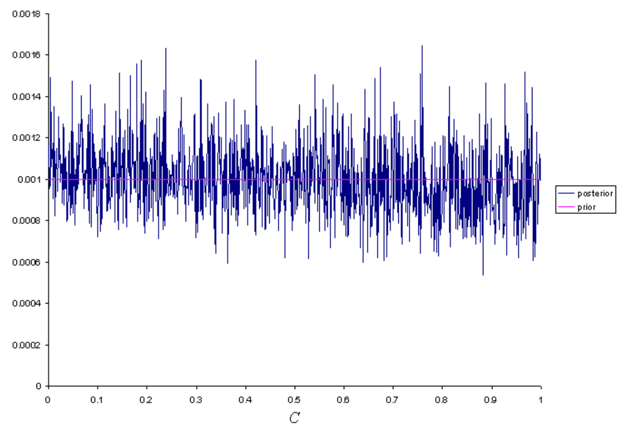

Figure 27: Posterior and prior densities for the amplitude C in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

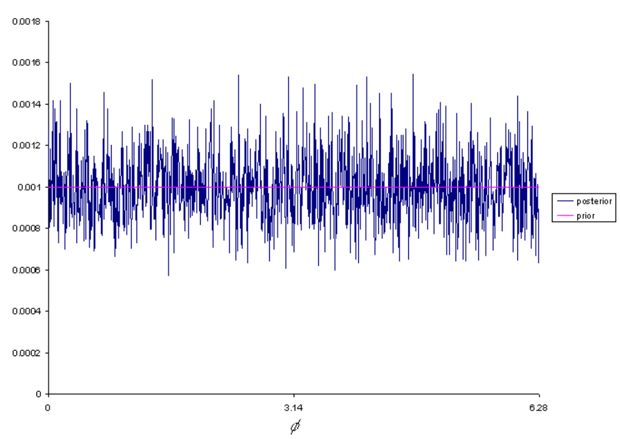

Figure 28: Posterior and prior density for the phase ϕ in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

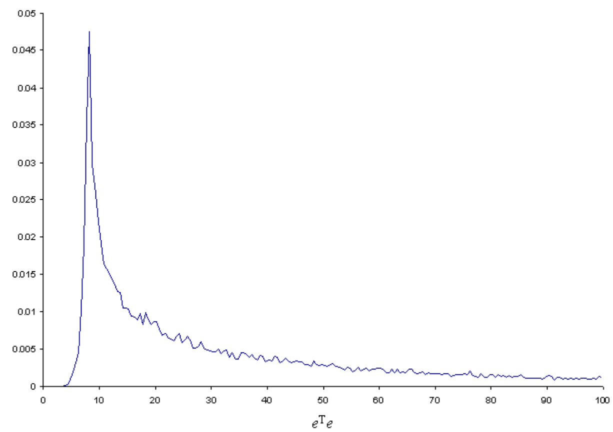

Figure 29: Posterior density for the nonlinear least squares residuals sum $e^T e$ in the market-time model with diffuse priors and crash probabilities ($A^c_{lp,m}$).

model without crash probabilities. In Fig. 24, meanwhile, we see that there is now a small secondary peak at low values of β. Counterintuitively, the posterior mode and mean for β are pushed to higher values in the complete model, which probably happens because, for high values of t_c, a large β is needed to get the crash probability to increase significantly as t approaches 10/19/87.

On a side note, since the likelihood function will put even less priority on fitting the raw price data in the complete model than it did without crash probabilities, one would expect to get larger values of the residuals sum $e^T e$. However, we do not find a statistically significant difference between the means, and the posterior density of Fig. 29 is little changed from the corresponding graph of Fig. 19 for the model without crash probabilities.

Incorporating crash probabilities of the JLS model into our analysis does not do much, good or bad, to the model's ability to explain log-periodic oscillations, although it does show that the model can outperform a model with a constant crash probability in terms of explaining when crashes occur.

8 Conclusion

Let us suppose that the log-periodic hypothesis is, indeed, correct. Estimation of the parameters of Eq. (1) by direct NLLS curve fitting of a raw price series to the log-periodic specification produces descriptive statistics that characterize any log-periodic oscillations that occur in a given time window. As we have seen, though, such descriptive statistics may not translate to the parameters of the underlying model responsible for those oscillations. For researchers interested in modeling financial markets as complex systems, it is ultimately those fundamental model parameters that are of interest, not the descriptive statistics.

Previous attempts to estimate the parameters of the Johansen-Ledoit-Sornette (2001) model of log-periodicity have been hampered by technical issues. Here, we have skirted those issues by employing Bayesian methods, which are better suited for the analysis of complicated time-series models like the JLS model. However, our findings do not provide support for the claim that the JLS model can explain log-periodic oscillations. In examining the period preceding the stock market crash of October 1987, we find that the complete model outperforms a null hypothesis model in terms of explaining why a crash would occur on 10/19/87, yet the data set is uninformative about the oscillatory parameters of the model. This suggests that those oscillatory parameters play no actual role.

The scope of our analysis was limited because we only considered data from a single log-periodic spell. One argument in favor of the log-periodic hypothesis is that similar values of the frequency ω are observed throughout the set of known log-periodic precursors. If one generalized the JLS model so that it described the behavior of financial markets both during periods of log-periodicity and during more quiescent periods, one could use the entire history of financial prices to obtain a posterior for the distribution from which ω is drawn for each log-periodic spell. If the value of ω is truly a universal property of log-periodic spells, the set of all log-periodic spells should be more informative about the oscillatory parameters of the JLS model than one log-periodic spell in isolation. Such a global analysis might then provide more support for the JLS model than could be found in the short data set considered here.

Even so, a negative result for the JLS model should not be interpreted as a negative result for the log-periodic hypothesis as a whole. The head and shoulders phenomenon, long recognized by technical traders, is another pattern in stock prices that has recently been found to predict excess returns over time scales of a month or longer (Savin, Weller, and Zvingelis (2003)). So it is not unreasonable to think that a pattern like log-periodicity might also have predictive power, but a different approach may be needed to explain how this happens.

In the JLS model, log-periodicity in financial prices reflects log-periodicity in traders' expectations regarding the future path of those prices. Another model of this class is pursued by Sornette and Ide (2003). They construct a deterministic model in which log-periodicity results from the interaction of fundamental and technical traders. Since the real world is not deterministic, Sornette and Ide presumably have in mind that actual prices dance around an expected price path that behaves as in their model. For both models, if the expected price path is log-periodic, then expected daily returns should also behave log-periodically, yet we find no evidence of such log-periodicity in returns.

This suggests that researchers should look to other explanations of log-periodicity that do not involve expectations. For example, Stauffer and Sornette (1998) constructed a model where prices are governed by a biased diffusion process that produces log-periodic behavior. They offered no microfoundations for why the market should behave in this way, but one could presumably explain this behavior in terms of, possibly psychological, barriers. If the market has a tendency to reverse direction when prices hit a floor or ceiling, log-periodic oscillations would occur if prices go back and forth between the two barriers at an increasing rate. One could then imagine that an acceleration in the rate of hitting barriers might make the market more susceptible to a crash. This approach would not share the JLS model's consistency with the Efficient Markets Hypothesis, but the mainstream finance literature is starting to become more receptive to behavioral models that deviate from rationality (Barberis and Thaler (2002)).

A The Hazard Rate

One complication of this model is that while the log-periodic form of the hazard rate was "derived" in the context of continuous time, the data is discrete. How do we interpret the hazard rate in our discrete-time context?

Let us suppose that at t_1 the event has not occurred. What is the probability that it will occur at or before $t_2 > t_1$? This probability is

$$\begin{aligned}\Pr[\widetilde{T} &\leq t_2 | t_1 \leq \widetilde{T}] = \frac{\Pr[t_1 \leq \widetilde{T} \leq t_2]}{\Pr[t_1 \leq \widetilde{T}]} \\ &= \frac{F(t_2) - F(t_1)}{1 - F(t_1)}.\end{aligned}$$

Then we get the hazard rate by taking

$$h(t_1) = \left(\frac{d}{dt_2} \Pr[\widetilde{T} \leq t_2 | t_1 \leq \widetilde{T}] \right) \Bigg|_{t_2 = t_1}.$$

Let $S(t) = 1 - F(t)$ be the survivor function. Then $S'(t) = -F'(t) = -f(t)$. Thus,

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

$$\ln S(t) - \ln S(0) = -\int_0^t h(t')dt'.$$

Since $S(0) = 1$,

$$S(t) = \exp\left(-\int_0^t h(t')dt'\right).$$

Then

$$\Pr[\widetilde{T} \le t_2 | t_1 \le \widetilde{T}] = \frac{S(t_1) - S(t_2)}{S(t_1)} = \frac{\exp\left(-\int_0^{t_1} h(t')dt'\right) - \exp\left(-\int_0^{t_2} h(t')dt'\right)}{\exp\left(-\int_0^{t_1} h(t')dt'\right)}$$

Thus,

$$\Pr[\widetilde{T} \le t_2 | t_1 \le \widetilde{T}] = 1 - \exp\left(-\int_{t_1}^{t_2} h(t')dt'\right). \tag{13}$$

In the limit of small $\int_{t_1}^{t_2} h(t')dt'$,

$$\Pr[\widetilde{T} \le t_2 | t_1 \le \widetilde{T}] \approx \int_{t_1}^{t_2} h(t')dt', \tag{14}$$

which is what people usually say the conditional probability will be. However, (13) is the exact probability, and it will always be between 0 and 1, unlike the approximate result (14).

B Properties of the Hazard Rate

An important property of the hazard rate function is that its integral will have the same form as itself, a power law times a periodic function of $\ln(t_c - t)$ with frequency ω. This is because

$$h(t) = \mathrm{Re}\left[B(t_c - t)^{1-\alpha}\left\{1 + Ce^{i\phi'}(t_c - t)^{i\omega}\right\}\right],$$

and the derivative or integral of a power law is also a power law (assuming the exponent is not -1). Focusing on a strictly real representation of h, consider a function

$$f(t) = F_1(t_c - t)^{1-\gamma}[1 + F_2\cos(\omega\ln(t_c - t) + \phi) + F_3\sin(\omega\ln(t_c - t) + \phi)]. \tag{15}$$

This has derivative

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

which can be rewritten

$$f'(t) = -F_1'(t_c - t)^{-\gamma}[1 + F_2' \cos(\omega \ln(t_c - t) + \phi) + F_3' \sin(\omega \ln(t_c - t) + \phi)], \tag{16}$$

where

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

Inverting this matrix

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

From (5), the hazard rate is

$$h(t) = B'(t_c - t)^{-\alpha}[1 + C \cos(\omega \ln(t_c - t) + \phi')].$$

If we set $\gamma = \alpha$, $F_1' = -B'$, $F_2' = C$, and $F_3' = 0$, then $h(t)$ has the form of Eq. (16) Therefore, the integral of the hazard rate has the form of (15) up to a constant, so

$$\begin{aligned} H(t) &= \overline{\kappa} \int_{t_0}^{t} h(t')dt' = A - \frac{\overline{\kappa} B'}{1-\alpha}(t_c - t)^{1-\alpha} \\ &\times \left[1 + C \frac{\cos(\omega \ln(t_c - t) + \phi') + \frac{\omega}{1-\alpha} \sin(\omega \ln(t_c - t) + \phi')}{1 + \left(\frac{\omega}{1-\alpha}\right)^2} \right]. \end{aligned}$$

This can be further simplified as follows. Let

$$\theta = \tan^{-1} \frac{\omega}{1-\alpha}.$$

Then

$$\begin{aligned} H(t) &= A - \frac{\overline{\kappa} B'}{1-\alpha}(t_c - t)^{1-\alpha} \\ &\quad \times \left[1 + \frac{C\cos\theta\cos(\omega\ln(t_c - t) + \phi') + \sin\theta\sin(\omega\ln(t_c - t) + \phi')}{\sqrt{1 + \left(\frac{\omega}{1-\alpha}\right)^2}}\right] \\ &= A - B(t_c - t)^{1-\alpha}\left[1 + \frac{C}{\sqrt{1 + \left(\frac{\omega}{1-\alpha}\right)^2}}\cos(\omega\ln(t_c - t) + \phi)\right], \end{aligned}$$

where $B = \overline{\kappa} B'/(1-\alpha)$ and $\phi = \phi' - \theta$.

C Importance Sampling

Denote the target density by $p(\theta)$. Denote the source density by $j(\theta)$, and an arbitrary kernel of the source density $k_j(\theta) = c_j \cdot j(\theta)$ for any $c_j \neq 0$. Denote an arbitrary kernel of the target density by $k_p(\theta) = c_p \cdot p(\theta)$ for any $c_p \neq 0$. The following result is due to Geweke (1989). Suppose that the sequence $\left\{\theta^{(m)}\right\}_{m=1}^{M}$ is independent and identically distributed, with $\theta^{(m)} \sim j(\theta)$. Define the weighting function $w(\theta) = \frac{k_p(\theta)}{k_j(\theta)}$. Suppose $E[g(\theta)]$ exists, $E[w(\theta)]$ exists, and the support of $j(\theta)$ include Θ. Then

$$\overline{g}^{(M)} = \frac{\sum\limits_{m=1}^{M} g(\theta^{(m)})w(\theta^{(m)})}{\sum\limits_{m=1}^{M} w(\theta^{(m)})} \rightarrow E[g(\theta)]$$

Assuming $V[g(\theta)]$ exists, then

$$M^{1/2}(\overline{g}^{(M)} - E[g(\theta)]) \xrightarrow{d} N(0, \tau^2),$$

where

$$\widehat{\tau}^{2(M)} = \frac{M\sum_{m=1}^{M}\left[g(\theta^{(m)}) - \overline{g}^{(M)}\right]^2 w(\theta^{(m)})^2}{\left(\sum\limits_{m=1}^{M} w(\theta^{(m)})\right)^2} \xrightarrow{a.s.} \tau^2.$$

References

[1] Barberis, Nicholas and Richard Thaler, (2002), "A Survey of Behavioral Finance," forthcoming in the *Handbook of the Economics of Finance*.

[2] Berger, J. O., (1985), *Statistical Decision Theory and Bayesian Analysis* (Springer-Verlag: New York).

[3] Bernardo, J. M. and A. F. M. Smith, (1994), *Bayesian Theory* (Wiley: New York).

[4] Blanchard, Olivier Jean and Stanley Fischer, (1989), *Lectures on Macroeconomics* (MIT Press, Cambridge).

[5] Fama, Eugene F., (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* 25: 383-417.

[6] Feigenbaum, James, (2001), "A Statistical Analysis of Log-Periodic Precursors to Financial Crashes," *Quantitative Finance* 1: 346.

[7] Feigenbaum, James, (2001), "More on A Statistical Analysis of Log-Periodic Precursors to Financial Crashes," *Quantitative Finance* 1: 527-532.

[8] Feigenbaum, James (2003), "Financial Physics," *Reports on Progress in Physics* 66: 1611-1649.

[9] Feigenbaum, James and Peter G. O. Freund, (1996), "Discrete Scale Invariance in Stock Markets Before Crashes," *International Journal of Modern Physics B* 10: 3737-3745.

[10] Gelman, A., J. Carlin, H. Stern, and D. Rubin, (2003), *Bayesian Data Analysis* (Chapman & Hall: Boca Raton).

[11] Geweke, John, (1989), "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica* 57: 1317-1340.

[12] Graf v. Bothmer, Hans-Christian and Christian Meister, (2003), "Predicting Critical Crashes? A New Restriction for Free Variables," *Physica A* 320: 539-547.

[13] Ilinski, Kirill, (1999), "Critical Crashes?," *International Journal of Modern Physics C* 10: 741-746.

[14] Johansen, Anders, (2004), "Origin of Crashes in Three US Stock Markets: Shocks and Bubbles," *Physica A* 338: 135-142.

[15] Johansen, Anders, Olivier Ledoit, and Didier Sornette, (2000), "Crashes as Critical Points," *International Journal of Theoretical and Applied Finance* 3: 219-255.

[16] Johansen, A., D. Sornette, G. Wakita, U. Tsunogai, W. I. Newman, and H. Saleur, (1996), "Discrete Scaling in Earthquake Precursory Phenomena: Evidence in the Kobe Eartquake, Japan," *J. Phys. I.* (France) 6: 1391-1402.

[17] Laloux, Laurent, Marc Potters, Rama Cont, Jean-Pierre Aguilar, and Jean-Philippe Bouchaud, (1999), "Are Financial Crashes Predictable?" *Europhysics Letters* 45: 1-5.

[18] Laloux, L., M. Potters, J. P. Aguilar, and J. P. Bouchaud, (2002), "Reply to Johansen's Comment," *Europhysics Letters* 45: 1.

[19] Phillips, P. C. B., (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics* 33: 311-340.

[20] Savin, Gene, Paul Weller, and Janis Zvingelis, (2003), "The Predictive Power of Head-and-Shoulders Price Patterns in the U.S. Stock Market," Working Paper.

[21] Sornette, Didier, (1998), "Discrete-Scale Invariance and Complex Dimensions," *Physics Reports* 297: 239-270.

[22] Sornette, Didier, (2003), *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (Princeton University Press: Princeton, NJ).

[23] Sornette, Didier and Kayo Ide, (2003), "Theory of Self-Similar Oscillatory Finite-Time Singularities in Finance, Population and Rupture," *International Journal of Modern Physics C* 14: 267-275.

[24] Sornette, Didier, Anders Johansen, and Jean-Phillipe Bouchaud, (1996), "Stock Market Crashes, Precursors and Replicas," *J. Phys. I.* (France) 6: 167-175.

[25] Stauffer, Dietrich and Didier Sornette, (1998), "Log-Periodic Oscillations for Biased Diffusion on Random Lattice," *Physica A* 252: 271-277.

[26] Sornette, Didier and Wei-Xing Zhou, (2003), "Predictability of Large Future Changes in Complex Systems," cond-mat/0304601.