ISSN 1440-771X

MONASH UNIVERSITY

AUSTRALIA

DEPARTMENT OF ECONOMETRICS AND BUSINESS STATISTICS

Estimation of Hyperbolic Diffusion Using MCMC Method

Y.K. Tse, Xibin Zhang and Jun Yu

Working Paper 18/2002

Estimation of Hyperbolic Diffusion using MCMC Method

Y.K. Tse
School of Economics and Social Sciences
Singapore Management University
Singapore 259756
yktse@smu.edu.sg

Xibin Zhang
Department of Econometrics and Business Statistics
Monash University
Clayton, Vic., 3800, Australia
bill.zhang@buseco.monash.edu.au

Jun Yu
Department of Economics
University of Auckland
Private Bag 92019, Auckland, New Zealand
j.yu@auckland.ac.nz

September 2002

Abstract: In this paper we propose a Bayesian method for estimating hyperbolic diffusion models. The approach is based on the Markov Chain Monte Carlo (MCMC) method after discretization via the Milstein scheme. Our simulation study shows that the hyperbolic diffusion exhibits many of the stylized facts about asset returns documented in the financial econometrics literature, such as a slowly declining autocorrelation function of absolute returns. We demonstrate that the MCMC method provides a useful tool to analyze hyperbolic diffusions. In particular, quantities of posterior distributions obtained from MCMC outputs can be used for statistical inferences.

Key Words: Markov Chain Monte Carlo, Hyperbolic diffusion, Milstein approximation, ARCH, Long Memory

JEL Classification: C11, C15, G15 and C63.

Acknowledgment: The third author gratefully acknowledges financial support from the Royal Society of New Zealand Marsden Fund under number 01-UOA-015.

1 Introduction

In the theoretic finance literature many diffusion models have been used to describe the movement of stock prices. As a nonlinear diffusion process, hyperbolic diffusions proposed by Bibby and Sorensen (1997) have received some attention (see, e.g., Rydberg (1999)). A variety of statistical properties possessed by hyperbolic diffusions are found to be consistent with many well known stylized features of financial time series. For example, the unconditional distribution of returns generated by hyperbolic diffusions has fatter tails than the normal distribution.

Although the stationary distribution of the hyperbolic diffusion process is known to be hyperbolic and hence has a closed-form expression, the transition density has no closed-form solution. Due to the lack of knowledge of the transition density, econometric estimation of the model using the likelihood approach is intractable. To circumvent this difficulty Bibby and Sorensen (1997) proposed to estimate the hyperbolic diffusion using the martingale estimation function method. However, although the estimator based on the martingale estimation function is consistent and asymptotically normally distributed, it is inefficient in general. Furthermore, the computation of the standard errors is difficult and requires techniques such as parametric bootstrapping.

In this paper we propose to use the Markov Chain Monte Carlo (MCMC) method to estimate the parameters of the hyperbolic diffusion after discretizing the diffusion via the Milstein scheme. The MCMC method is based on Bayesian analysis and offers a full likelihood-based inference. It provides a general mechanism to sample the parameter vector from its posterior distribution, enabling exact finite-sample inference via Monte Carlo methods, and is thus highly efficient. Empirical illustrations reported by Rydberg (1999) show that a member of the generalized hyperbolic diffusion can induce long-memory features in the squared return.[1] In this paper we report the ability of the

[1]Rydberg (1999) considered the normal inverse Gaussian diffusion, which differs from the hyperbolic

hyperbolic diffusion in reproducing other stylized facts found in the financial econometrics literature, in particular, the so-called Taylor effect.

This paper is organized as follows. Section 2 reviews the hyperbolic diffusion model and its properties. We also discuss how the Milstein scheme can be used to discretize the model. Section 3 describes the MCMC method. In Section 4 we fit the model to three stock market indices over a decade of daily data. Statistical inference is made via the posterior quantities. In Section 5 we illustrate the statistical properties of a sample path generated by the hyperbolic diffusion. We find that many of the stylized facts for stock returns in the empirical finance literature documented by Ryden, Teräsverta and Asbrink (1998) are satisfied. Section 6 provides further comments and concludes.

2 The Model

Consider the following continuous-time parametric diffusion

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t, \tag{1}$$

where X_t is a state variable, W_t is a standard Brownian motion defined on the probability space $(\Omega, \Im^B, (\Im_t^B)_{t\geq 0}, P)$, $\mu(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ are known functions, and θ is a vector of unknown parameters.

Many empirical studies have shown that asset returns are not normally distributed. Barndorff-Nielsen (1978) suggested using the hyperbolic distribution to describe unconditional asset returns. The density of the hyperbolic distribution is proportional to $1/b^2(x)$, with

$$b(x) = \exp\left\{\frac{1}{2}\left[\alpha\sqrt{\delta^2 + (x-\mu)^2} - \beta(x-\mu)\right]\right\}, \tag{2}$$

where α, β, δ and μ are the parameters of the distribution satisfying $\alpha > |\beta| \geq 0$ and $\delta > 0$. It is noted that δ is the scale parameter, μ is the location parameter, β determines

diffusion in the form of the stationary density.

the symmetry (the distribution is symmetrical about μ if $\beta = 0$) and α determines the steepness of the distribution.

We assume that the stock price S_t depends on the state variable X_t through the process

$$S_t = \exp(X_t + \kappa t) \tag{3}$$

where κ is the (constant) drift rate. Following Bibby and Sorensen (1997) we consider the following hyperbolic diffusion process to describe the movement of stock prices[2]

$$dS_t = S_t \left\{ \left[\kappa + \frac{1}{2}\sigma^2 b^2(\ln S_t - \kappa t) \right] dt + \sigma b(\ln S_t - \kappa t) dW_t \right\}. \tag{4}$$

Bibby and Sorensen (1997) obtained some interesting statistical properties of the process S_t. For instance, they showed that the marginal distribution of $\ln S_t$ is hyperbolic and hence $\ln S_t$ is approximately hyperbolically distributed after a sufficiently long time period. Also, the distribution of increments over short intervals has fat tails while an increment over a long interval follows a distribution that is close to being hyperbolic.

To derive the dynamic properties of stock returns, we apply Ito's lemma to obtain

$$dX_t = \sigma\, b(X_t) dW_t, \tag{5}$$

which represents a diffusion process with no drift. As dW_t are uncorrelated over nonoverlapping intervals, increments of the log-prices (i.e., the continuously compounded rates of return) are serially uncorrelated. Similar to the stochastic volatility models, however, the squared increments of the log-prices are generally serially correlated.

As argued in Bibby and Sorensen (1997), although the marginal distribution of the hyperbolic diffusion process is hyperbolic, the transition density is unknown. Therefore, the maximum likelihood method is difficult to implement. Bibby and Sorensen (1997)

[2]Note that μ in equation (2) and σ in equation (4) are parameters of the diffusion. They should not be confused with $\mu(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$, which are known functions of the drift and diffusion terms, respectively.

suggested using the martingale estimation function approach (see Bibby and Sorensen (1995)) to estimate the diffusion models. Their approach, however, requires knowledge of the conditional expectation and conditional variance of the diffusion. These conditional moments are usually known only for very simple models, such as when the drift term is linear in the state variable. Hence, although Bibby and Sorensen (1997) showed that their estimates are consistent and asymptotically normal, application of their method is difficult in practice.

In this paper we use the MCMC method to estimate the hyperbolic diffusion after discretizing the model via the Milstein scheme. In the next section we shall discuss the MCMC method. To conclude this section, we outline the Milstein scheme for the discretization of the hyperbolic diffusion model. As shown by Elerian (1998), the Milstein scheme provides an approximation with improved accuracy over the Euler scheme in estimating the parameters of a nonlinear diffusion. It can also be used to approximate transition densities of diffusion processes as in Pedersen (1995), Eraker (2001), and Elerian, Chib and Shephard (2001).

The Milstein (1978) approach approximates a general diffusion process such as equation (1) by the following expansion

$$X_{t+\Delta t} = X_t + \mu(X_t,\theta)\Delta t + \sigma(X_t,\theta)\Delta W_t + \frac{1}{2}\sigma(X_t,\theta)\frac{\partial\sigma(X_t,\theta)}{\partial X_t}\left[(\Delta W_t)^2 - \Delta t\right], \quad (6)$$

where $\Delta W_t = \varepsilon\sqrt{\Delta t}$ with $\varepsilon \sim N(0,1)$. This equation can be rewritten as

$$X_{t+\Delta t} - X_t - \mu(X_t,\theta)\Delta t + g(X_t,\theta)\Delta t = \sigma(X_t,\theta)\sqrt{\Delta t}\,\varepsilon + g(X_t,\theta)\,\Delta t\,\varepsilon^2, \quad (7)$$

where $g(X_t,\theta) = \frac{1}{2}\sigma(X_t,\theta)(\partial\sigma(X_t,\theta)/\partial X_t)$. Let

$$a = \sigma(X_t,\theta)\sqrt{\Delta t}, \qquad b = g(X_t,\theta)\Delta t, \quad (8)$$

then equation (7) can be represented by

$$Y = a\varepsilon + b\varepsilon^2 = b\left[\left(\varepsilon + \frac{a}{2b}\right)^2 - \frac{a^2}{4b^2}\right], \quad (9)$$

where $Y = X_{t+\Delta t} - X_t - \mu(X_t, \theta)\Delta t + g(X_t, \theta)\Delta t$.

The normality assumption implies that $\left(\varepsilon + \frac{a}{2b}\right)^2$ (denoted by, say, Z) follows a noncentral χ^2 distribution with 1 degree of freedom and noncentrality parameter $\lambda = a^2/(4b^2)$. Elerian (1998) showed that the density of Z is given by

$$f(z) = \frac{1}{2} \exp\left\{-\frac{\lambda + z}{2}\right\} \left(\frac{z}{\lambda}\right)^{-1/4} I_{-1/2}\left(\sqrt{\lambda z}\right), \tag{10}$$

where

$$I_{-1/2}(w) = \sqrt{\frac{2}{w}} \sum_{j=0}^{\infty} \frac{(w/2)^{2j}}{j!\Gamma(j + 1/2)} = \sqrt{\frac{2}{\pi w}} \cosh(w),$$

with $\cosh(w) = (1/2)\{\exp(w) + \exp(-w)\}$ being the hyperbolic cosine function. Hence the density of Y is

$$f^*(y) = \frac{1}{b} f\left(\frac{y}{b} + \frac{a^2}{4b^2}\right). \tag{11}$$

Assuming constant priors for all the parameters, and given n observations of $\mathbf{x} = \{x_t\}$, $t = 1, ..., T$, the posterior for θ (the vector of all parameters) is

$$\pi(\theta|\mathbf{x}) = \log(c) + \sum_{t=1}^{T} \left[\log\left\{f\left(\frac{y_t}{b} + \frac{a^2}{4b^2}\right)\right\} - \log(b)\right], \tag{12}$$

where $y_t = x_{t+\Delta t} - x_t - \mu(x_t, \theta)\Delta t + g(x_t, \theta)\Delta t$. Based on the above equation as the approximate posterior, we shall see in the next section how the Metropolis-Hastings algorithm can be adopted to sample the parameter vector θ.

3 Markov Chain Monte Carlo Simulation

Bayesian inference concerning a parameter vector θ conditional on data y is made via the posterior density $\pi(\theta|\mathbf{x})$. By the Bayes theorem, the posterior takes the form

$$\pi(\theta|\mathbf{x}) = c\, L(\mathbf{x}|\theta)\, \pi(\theta), \tag{13}$$

where c is a normalizing constant, $L(\mathbf{x}|\theta)$ is the likelihood of the data $\mathbf{x}$ conditional upon θ and $\pi(\theta)$ is the prior density of θ. The Bayesian approach requires that statistical inference be based on the posterior. Dealing with the posterior, however, is often analytically intractable. Nonetheless, if we can sample the parameter vector from the posterior, statistical inference about the parameter vector can be made using the usual Monte Carlo approach. The MCMC method aims to provide a general mechanism to sample the parameter vector from its posterior density. While simulating directly from the posterior distribution is typically very difficult, the MCMC method sets up a Markov chain so that its stationary distribution is the same as the posterior density. When the Markov chain converges, the simulated values may be regarded as a sample obtained from the posterior.

There are two broad categories of algorithms for implementing MCMC. Firstly, the Gibbs sampling algorithm generates the observations in which each draw is obtained by sampling sub-components of a random vector from a sequence of full conditional distributions. Secondly, the Metropolis-Hastings algorithm generates each value of the Markov chain from a proposal density, and the acceptance and rejection of the simulated value is according to the density at the candidate point relative to the density at the current point. Robert and Casella (1999) presented detailed discussions on the use of the Metropolis-Hastings algorithm and the Gibbs sampler. As the full conditional density is often difficult to derive, the Metropolis-Hastings algorithm is generally adopted in complex problems. In this paper we use the Metropolis-Hastings algorithm for its simplicity. In what follows we briefly describe the procedure of the algorithm.

The Metropolis-Hastings algorithm is based on proposing a new point according to an arbitrary proposal density (also called the candidate generating density). The acceptance (or otherwise) of the proposed value is determined by a process with an acceptance probability that depends on the current point, the new point and the proposal density from which the new point is proposed. Suppose we wish to simulate from the posterior $\pi(\theta|\mathbf{x})$. Let $q(\theta, \theta')$ be an arbitrary proposal density which describes the probability of proposing θ' given current value θ. The proposal density should be chosen so that simulation from it is straightforward. A sequence of draws from the Metropolis-Hastings algorithm is obtained as follows.

Step 1: Let the current value be $\theta^{[i]}$. Generate a proposal, θ' from the proposal density $q(\theta^{[i]}, \theta')$.

Step 2: Calculate the acceptance probability

$$T(\theta^{[i]}, \theta') = \begin{cases} \min\ \{\frac{\pi(\theta'|\mathbf{x})q(\theta',\theta^{[i]})}{\pi(\theta^{[i]}|\mathbf{x})q(\theta^{[i]},\theta')}, 1\} & \text{if } \pi(\theta^{[i]}|\mathbf{x})q(\theta^{[i]}, \theta') > 0 \\ 1 & \text{if } \pi(\theta^{[i]}|\mathbf{x})q(\theta^{[i]}, \theta') = 0 \end{cases} \tag{14}$$

Step 3: With probability $T(\theta^{[i]}, \theta')$ accept the proposal and set $\theta^{[i+1]} = \theta'$. Otherwise, reject the proposal and set $\theta^{[i+1]} = \theta^{[i]}$.

Step 4: Repeat the previous steps to obtain a sequence $\{\theta^{[0]}, \theta^{[1]}, \theta^{[2]}, \cdots\}$, where $\theta^{[0]}$ denotes the initial value for θ. Discard the burn-in values (up to $\theta^{[D]}$ say) obtained whilst the algorithm converges. Then the remaining values, $\{\theta^{[D+1]}, \theta^{[D+2]}, \cdots\}$ are a correlated sequence simulated from $\pi(\theta|\mathbf{x})$, and have the same stationary posterior density as $\pi(\theta|\mathbf{x})$.

Two important points should be noted. First, the calculation of $T(\theta^{[i]}, \theta')$ does not require knowledge of the normalizing constant for the posterior function. Second, if the proposal density is symmetric, that is, $q(x, y) = q(y, x)$, then the acceptance probability reduces to $\pi(\theta'|\mathbf{x})/\pi(\theta^{[i]}|\mathbf{x})$. This is the reason why the standard normal density and the uniform density on the interval $[-0.5, 0.5]$ are often used as proposal densities.

The MCMC strategy has proved to be extremely useful in many statistical applications, and has many advantages compared to traditional independent sampling methods. For example, MCMC methods can be applied without knowing the normalizing constant of the posterior density. This point is very important in the Bayesian context, where the normalizing constant of the posterior density is almost never known. Geweke (1999) provided a survey of the fundamental principles of subjective Bayesian inference in econometrics and the implementation of these principles using posterior simulation methods, emphasizing the importance of simulation methods and describing the implementation of MCMC simulation for Bayesian inference. Due to the effective simulation methods, Bayesian inference is easier to implement than asymptotic classical inference for complex nonlinear models with latent variables and in dynamic models that are nearly nonstationary and nonidentified. The MCMC strategy has also been widely used in econometrics, such as the Bayesian analysis of stochastic volatility models by Jacquier, Polson and Rossi (1994), the likelihood inference on stochastic volatility models in Kim, Shephard and Chib (1998), Bayesian inference for GARCH models in Vrontos, Dellaportas and Politis (2000), the likelihood inference for diffusions in connection with the Euler approximation by Elerian, Chib and Shephard (2001) and the recent survey by Chib (2001), among many others.

4 Empirical Applications

In this section we apply the Metropolis-Hastings algorithm to the discretized diffusion processes and present empirical results based on some real data sets.. The data series considered are the MSCI World Index, the MSCI Europe Index and the NYSE Index. The series consist of daily observations from January 1, 1990 to December 31, 2000.

In our empirical study the joint prior for the parameter vector θ is assumed to be a constant. According to the Metropolis-Hastings rule, the normalizing constant in the posterior, as well as the constant prior, does not affect acceptance probability. In the implementation of the Metropolis-Hastings algorithm, the proposal density is the uniform density on $[-0.5, 0.5]$, and then the parameter vector θ is updated in the following way:

$$\theta' = \theta + \tau\varepsilon, \tag{15}$$

where θ' is the proposal for θ, ε is a vector of uniform random numbers on $[-0.5, 0.5]$, and τ is a tuning parameter which is chosen so that the acceptance rate is between 20% and

30%. In addition, τ might be either a scalar- or vector-constant. Generally speaking, if the parameters are of weak correlation and their values are of the same scale, τ can be a scalar constant. Otherwise, τ should be a constant vector, so that each parameter is assigned a specific tuning parameter.

In the implementation of the MCMC algorithm, the sampled path, denoted by $\{\theta^{[i]} : i = 1, 2, \cdots, N\}$, forms a Markov chain whose stationary density is the posterior $\pi(\theta|\mathbf{x})$, and the output is summarized in terms of the ergodic averages in the form of:

$$\overline{f}_N = \frac{1}{N}\sum_{i=1}^{N} f(\theta^{[i]}), \tag{16}$$

where $f(\cdot)$ is a real-valued function to be estimated. Roberts (1996) pointed out that most of the Markov chains produced in MCMC converge geometrically to the stationary distribution $\pi(\theta|\mathbf{x})$, and one of the most important consequences of the geometric convergence is that it follows the central limit theorem of ergodic averages, i.e.,

$$\sqrt{N}\left(\overline{f}_N - E_\pi[f(\theta)]\right) \xrightarrow{D} N(0, \sigma_f^2), \tag{17}$$

where $E_\pi[\cdot]$ denotes the expectation operator under $\pi(\theta|\mathbf{x})$, and the convergence is in distribution. To assess the accuracy of the ergodic average as an estimate of $E_\pi[f(\theta)]$, it is essential to estimate σ_f^2. One of the most commonly used methods is to estimate σ_f^2 using the batch means, which is discussed extensively in Geyer (1992) and Roberts (1996).

To estimate σ_f^2 using the batch means, the MCMC algorithm is run for $N = m \times n$ iterations, where n is sufficiently large so that

$$y_k = \frac{1}{n}\sum_{i=(k-1)n+1}^{kn} f(\theta^{[i]}), \tag{18}$$

for $k = 1, 2, \cdots, m$, are approximately independently distributed as $N(E_\pi[f(\theta)], \sigma_f^2/n)$. Therefore σ_f^2 can be estimated by

$$\hat{\sigma}_f^2 = \frac{n}{m-1}\sum_{k=1}^{m}(y_k - \overline{f}_N)^2, \tag{19}$$

where $\overline{f}_N$ is defined in equation (16). Thus, the standard error of $\overline{f}_N$ can be estimated by $\sqrt{\hat{\sigma}_f^2/N}$, which is called the batch-mean standard error or the Monte Carlo standard error, and is commonly used for checking the mixing performance.

In addition to the batch-mean standard error, one may also compute the standard deviation $\tilde{\sigma}_f$ directly based on the sampled paths using the formula

$$\tilde{\sigma}_f = \left\{ \frac{1}{N-1} \sum_{i=1}^{N} \left[f(\theta^{[i]}) - \overline{f}_N \right]^2 \right\}^{1/2}. \tag{20}$$

Kim, Shephard and Chib (1998) indicated that the mixing performance of the sampled paths can be measured using the simulation inefficiency factor (SIF), also called the integrated autocorrelation time by Sokal (1996), which is estimated as the variance of the sample mean divided by the variance of the sample mean from a hypothetical sampler that draws independent random observations from the posterior distribution. Meyer and Yu (2000) showed that SIF is given by

$$\text{SIF} = \frac{\hat{\sigma}_f^2}{\tilde{\sigma}_f^2}. \tag{21}$$

We apply the Metropolis-Hastings algorithm to the MSCI World Index, the MSCI Europe Index and the NYSE Index. The burn-in period is taken as $D = 10,000$ iterations and the number of total recorded iterations after the burn-in period is $N = 50,000$. The sampled paths for the parameters of the MSCI World Index are plotted in Figures 1.1 to 1.6, which show that the sampled paths are reasonably well mixed. Based on the sampled path for each data set, we calculate the ergodic average (or mean) and standard deviations. The MC standard errors are obtained using the batch-mean approach described in equations (16) through (19) with $f(x) = x$. The number of batches is $m = 50$, and there are $n = 1000$ points in each batch. Table 1 summarizes the ergodic averages, standard deviations, 95% Bayes confidence intervals, Monte Carlo standard errors, and the simulation inefficiency factors for each data set. The Bayes confidence interval can be used to test the significance of a parameter. For example, for the World Index, all the parameters are significantly different from zero. Although our MCMC algorithm is not very efficient judged by the simulation inefficiency factors, the inefficiency is tolerable in view of the number of runs we generated.

Table 1. Empirical Results

data	para.	mean	95% confidence interval	s.d.	MC s.e.	SIF
MSCI World	κ	–0.07950	(–0.12554, –0.03552)	0.02263	0.00083	67.41
	α	1.53933	(1.26136, 1.85925)	0.14764	0.00442	44.30
	δ^2	0.02636	(0.00037, 0.12875)	0.03635	0.00153	87.95
	μ	6.48224	(6.30377, 6.66300)	0.09125	0.00292	52.16
	β	0.35768	(0.06210, 0.58873)	0.13177	0.00382	41.24
	σ^2	0.00741	(0.00513, 0.00953)	0.00109	0.00004	62.65
MSCI Europe	κ	–0.01422	(–0.06831, 0.03407)	0.02553	0.00100	76.17
	α	1.54634	(0.99561, 2.12376)	0.27217	0.01247	102.84
	δ^2	0.05139	(0.00068, 0.24200)	0.06502	0.00350	146.10
	μ	6.29925	(6.06792, 6.54124)	0.12231	0.00391	51.06
	β	0.27821	(–0.23627, 0.73844)	0.24608	0.01044	89.79
	σ^2	0.01006	(0.00584, 0.01371)	0.00199	0.00010	134.66
NYSE	κ	–0.03117	(–0.07709, 0.01361)	0.02332	0.00094	81.94
	α	1.65771	(1.33921, 2.04863)	0.17835	0.00744	86.90
	δ^2	0.01999	(0.00024, 0.11207)	0.03067	0.00152	119.81
	μ	5.73078	(5.54249, 5.92304)	0.09845	0.00370	72.31
	β	0.23127	(–0.11584, 0.48954)	0.15436	0.00511	54.95
	σ^2	0.00766	(0.00522, 0.01003)	0.00118	0.00005	87.56

Note: MC s.e. refers to the Monte Carlo standard error computed through the batch-mean approach. SIF refers to the simulation inefficiency factor and is given in equation (21).

It is noted that the estimates of α, namely, the steepness parameter, are quite similar across the three indices. While the Europe index and the NYSE index are symmetrical (the sampled posterior β is not significantly different from zero), the World index is asymmetric. For the scale and volatility parameters (i.e., δ and σ, respectively), the values are similar between the World and the NYSE indices, but different for the Europe index.

5 Hyperbolic Diffusions and Some Stylized Facts of Stock Returns

Several "stylized facts" regarding the return of stocks have been well documented in the financial econometrics literature. Let r_t denote the return of a stock. Ryden, Teräsverta and Asbrink (1998) summarized the following properties of r_t found in many empirical studies:

1. r_t are not autocorrelated, except possibly at lag one.
2. The autocorrelation functions of $|r_t|$ and r_t^2 decay slowly, and $\text{corr}(|r_t|, |r_{t-k}|) > \text{corr}(r_t^2, r_{t-k}^2)$. The decay is much slower than the exponential rate of the autocorrelation function of a stationary ARMA process r_t.
3. $\text{corr}(|r_t|, |r_{t-k}|) > \text{corr}(|r_t|^\theta, |r_{t-k}|^\theta)$, $\theta \neq 1$. The autocorrelations of powers of absolute return are highest at power one. This is called the Taylor effect.

Rydberg (1999) reported simulation results of the normal inverse Gaussian diffusion in which the autocorrelation function of r_t^2 declines very slowly, thus satisfying partly property 2 above. In this section we examine in more detail whether the hyperbolic diffusion would give rise to the statistical properties described above.

Using the fitted values of the hyperbolic diffusion based on the MSCI World Index, we generate a path of daily price series with 2000 observations and plot it in Figure 2. A time interval of 15 minutes is used (i.e., we use $\Delta t = 1/7000$ year, assuming 7 hours of trading per day and 250 trading days per year). The data are generated using the Milstein scheme. The hourly returns and the daily returns are plotted in Figures 3 and 4. Figure 4 is typical of many return series, exhibiting clustering of volatility.

Figure 5 presents the autocorrelation function of the hourly, daily and weekly return series up to 100 lags. It shows that returns are uncorrelated over the different intervals.

The autocorrelation functions of the absolute return and the squared return are plotted in Figure 6. It can be seen that the autocorrelation functions decline *very* slowly. Figure 7 provides further results on the autocorrelations of $|r_t|^\theta$ for $\theta = 1.2, 1.5$ and 1.8. From Figures 6 and 7, we can see that the sample path exhibits property 3 of the stylized facts above.

In sum, the hyperbolic diffusion appears to satisfy the *temporal* stylized properties as summarized by Ryden, Teräsverta and Asbrink (1998).

6 Discussions and Conclusions

To understand why a hyperbolic diffusion generates the ARCH and long memory properties, we apply the Euler approximation to the diffusion model for the log-price (i.e., equation (5)) and obtain

$$Y_t \approx \sigma \exp\left[\frac{1}{2}\left\{\alpha\sqrt{\delta^2 + (\sum_{i=1}^{\infty} Y_{t-i} - \mu)^2} - \beta(\sum_{i=1}^{\infty} Y_{t-i} - \mu)\right\}\right] e_t, \tag{22}$$

where $Y_t = \ln S_{t+\triangle t} - \ln S_t$ denotes the return and $e_t \sim iidN(0, \triangle t)$. Equivalently this equation can be re-written as

$$\begin{aligned} Y_t &\approx \sigma \exp\left\{\frac{1}{2}h_t\right\} e_t \\ h_t &= \alpha\sqrt{\delta^2 + (\sum_{i=1}^{\infty} Y_{t-i} - \mu)^2} - \beta(\sum_{i=1}^{\infty} Y_{t-i} - \mu). \end{aligned}$$

Comparing the above specification with the well-known ARCH(∞) model (Engle (1982)),

$$\begin{aligned} Y_t &= \sigma_t e_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{\infty} \alpha_i Y_{t-i}^2, \end{aligned}$$

it can be seen that the hyperbolic diffusion model can be regarded as a special case of the following nonlinear ARCH(∞) model

$$Y_t = \sigma \exp\left\{\frac{1}{2}h_t\right\} e_t$$

$$h_t = f(Y_{t-1}, Y_{t-2}, \cdots).$$

Not surprisingly, the hyperbolic diffusion model generates the ARCH effect. Furthermore, the nonlinear relationship between h_t and Y_{t-i} seems to have caused the long-memory properties in the absolute return as well as the squared return. Our simulation study has shown that the hyperbolic diffusion is able to exhibit many of the stylized facts about asset returns documented in the financial econometrics literature. The MCMC method provides a useful tool to analyze hyperbolic diffusions. Empirical quantities of the posterior distributions can be calculated from the MCMC method for statistical inference.

References

1. Barndorff-Nielsen, O. (1978), Hyperbolic Distributions and Distributions on Hyperbolae, *Scandinavia Journal of Statistics*, **5**, 151-157.

2. Bibby, B.M. and M. Sorensen (1995), Martingale Estimation Functions for Discretely Observed Diffusion Processes, *Bernoulli*, **1**, 17-39.

3. Bibby, B.M. and M. Sorensen (1997), A Hyperbolic Diffusion Model for Stock Prices, *Finance and Stochastics*, **1**, 25-41.

4. Chib, S. (2001), Markov Chain Monte Carlo Methods: Computation and Inference, in *Handbook of Econometrics*, eds by J.J. Heckman, and E. Leamer, North-Holland, Amsterdam, **5**, 3569-3649.

5. Elerian, O. (1998), A Note on the Existence of a Closed-form Conditional Transition Density for the Milstein Scheme, Economics discussion paper 1998-W18, Nuffield College, Oxford.

6. Elerian, O., S. Chib and N. Shephard (2001), Likelihood Inference for Discretely Observed Non-Linear Diffusions, *Econometrica*, **69**, 959-993.

7. Engle, R.F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, **50,** 987-1007.

8. Eraker, B. (2001), MCMC Analysis of Diffusion Models with Application to Finance, *Journal of Business and Economic Statistics*, **19**, 177-191.

9. Geyer, C.J. (1992), Practical Markov Chain Monte Carlo (with discussion), *Statistical Science*, **7**, 473-511.

10. Geweke, J. (1999), Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication, *Econometric Reviews*, **18**, 1-73.

11. Kim, S., N. Shephard, and S. Chib (1998), Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models, *Review of Economic Studies* **65**, 361-393.

12. Meyer, R. and J. Yu (2000), BUGS for a Bayesian Analysis of Stochastic Volatility Models, *Econometrics Journal*, **3,** 198-215.

13. Milstein, G.N. (1978), A Method of Second-order Accuracy Integration of Stochastic Differential Equations, *Theory of Probability and its Applications*, **23**, 396-401.

14. Pedersen, A. (1995), A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observation, *Scandinavian Journal of Statistics*, **22**, 55-71.

15. Robert, C.P. and G. Casella (1999), Monte Carlo Statistical Methods, Springer.

16. Roberts, G.O. (1996), Markov Chain Concepts Related to Sampling Algorithms, in *Markov chain Monte Carlo in Practice* (eds. by W.R. Gilks, S. Richardson and D.J. Spiegelhalter), 45-57, London: Chapman & Hall.

17. Rydberg, T.H. (1999), Generalized Hyperbolic Diffusion Processes with Applications in Finance, *Mathematical Finance*, **9**, 183-201.

18. Ryden, T., T. Teräsverta and S. Asbrink (1998), Stylized Facts of Daily Return Series and the Hidden Markov Model, *Journal of Applied Econometrics*, **13**, 217-244.

19. Sokal, A.D. (1996), Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms, Lectures at the Cargese Summer School on "Functional Integration: Basics and Applications".

20. Vrontos, I. D., P. Dellaportas, and D. N. Politis (2000), Full Bayesian Inference for GARCH and EGARCH Models, *Journal of Business and Economic Statistics*, **18**, 187-198.

Figure 1.1: Sampled Path for kappa

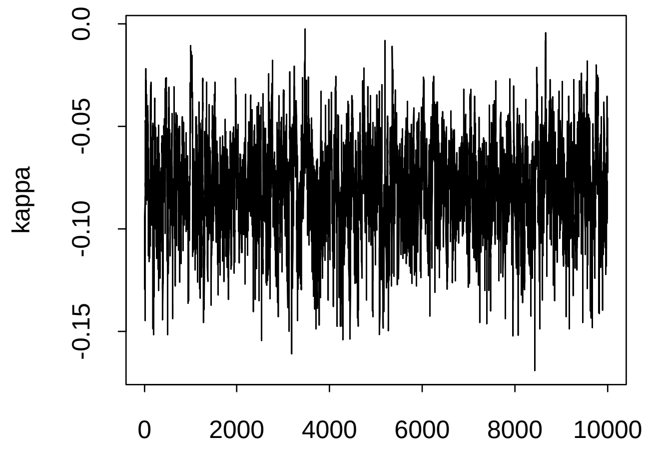

Figure 1.2: Sampled Path for alpha

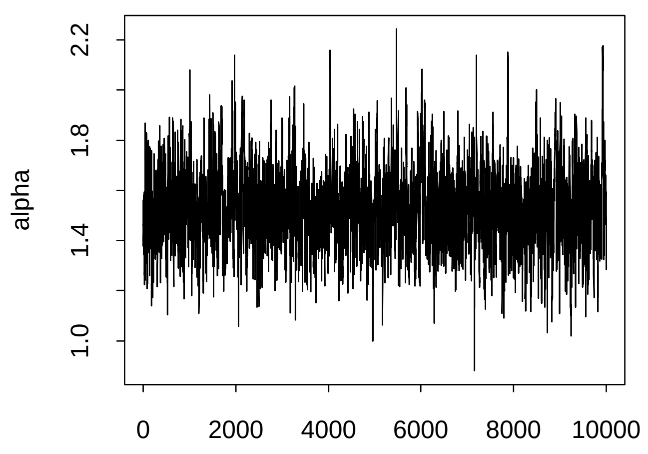

Figure 1.3: Sampled Path for delta^2

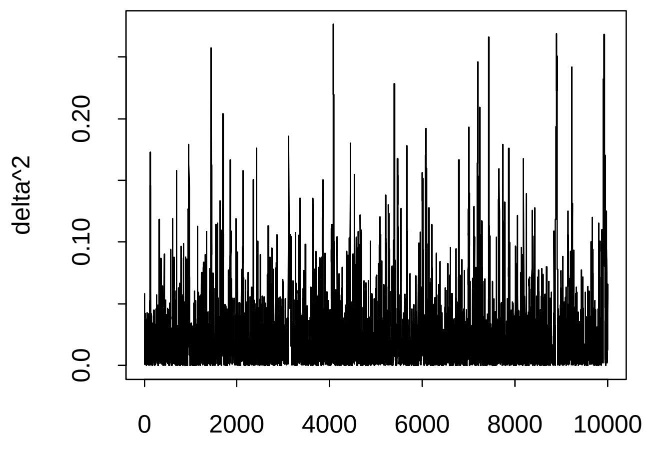

Figure 1.4: Sampled Path for mu

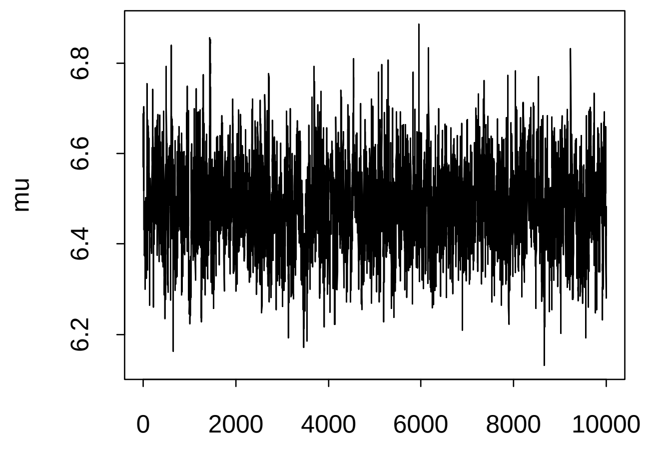

Figure 1.5: Sampled Path for beta

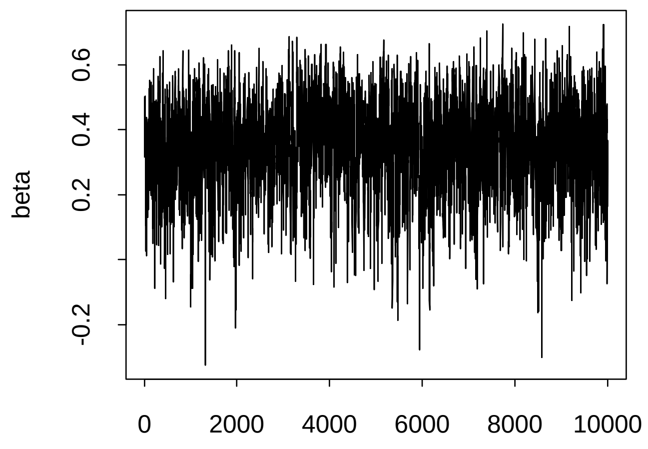

Figure 1.6: Sampled Path for sigma^2

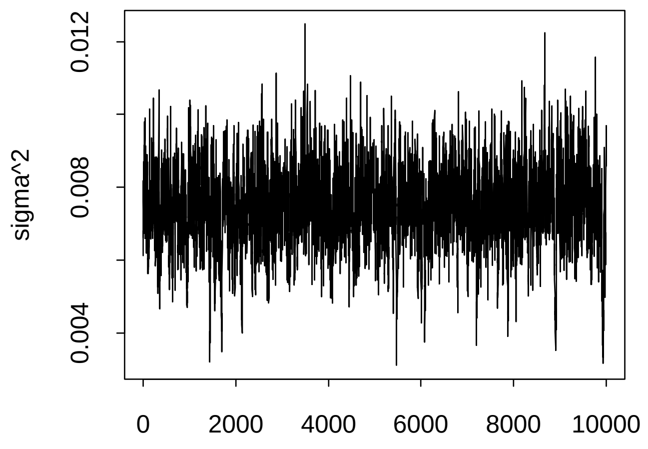

Figure 2: Generated Time Series with Dt = 1/7000

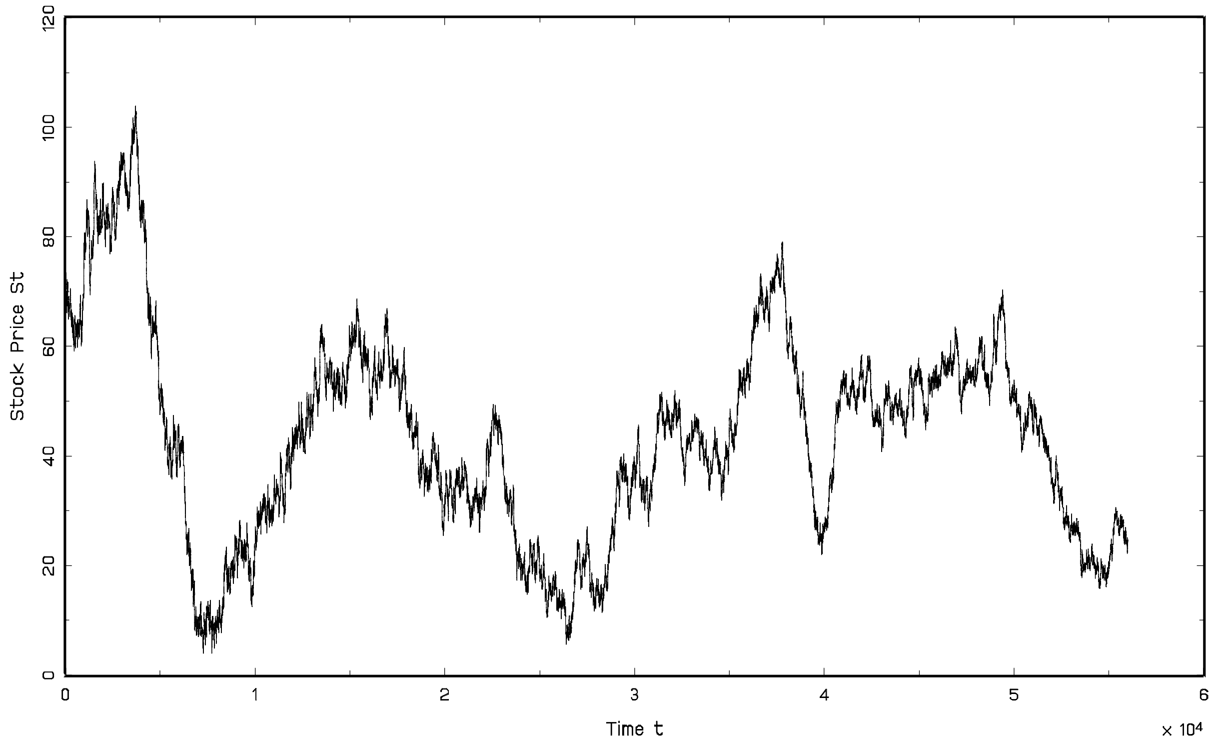

Figure 3: Hourly Return of the Generated Series

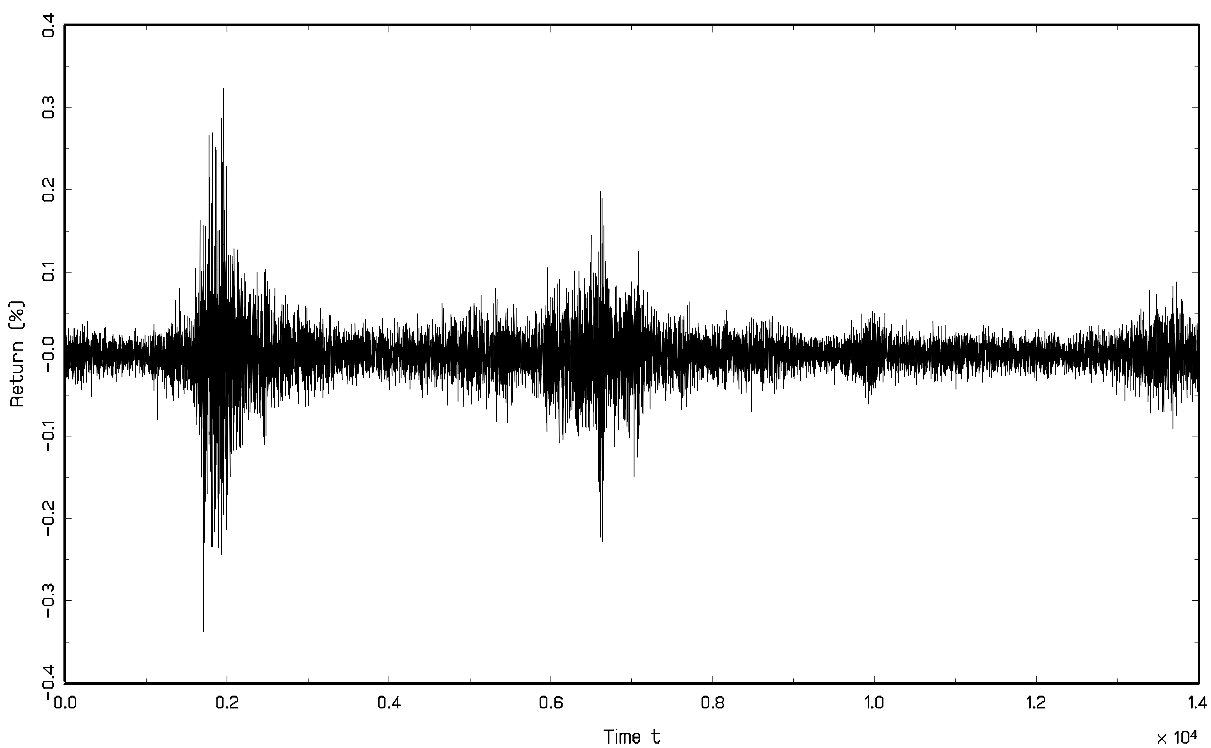

Figure 4: Daily Return of the Generated Series

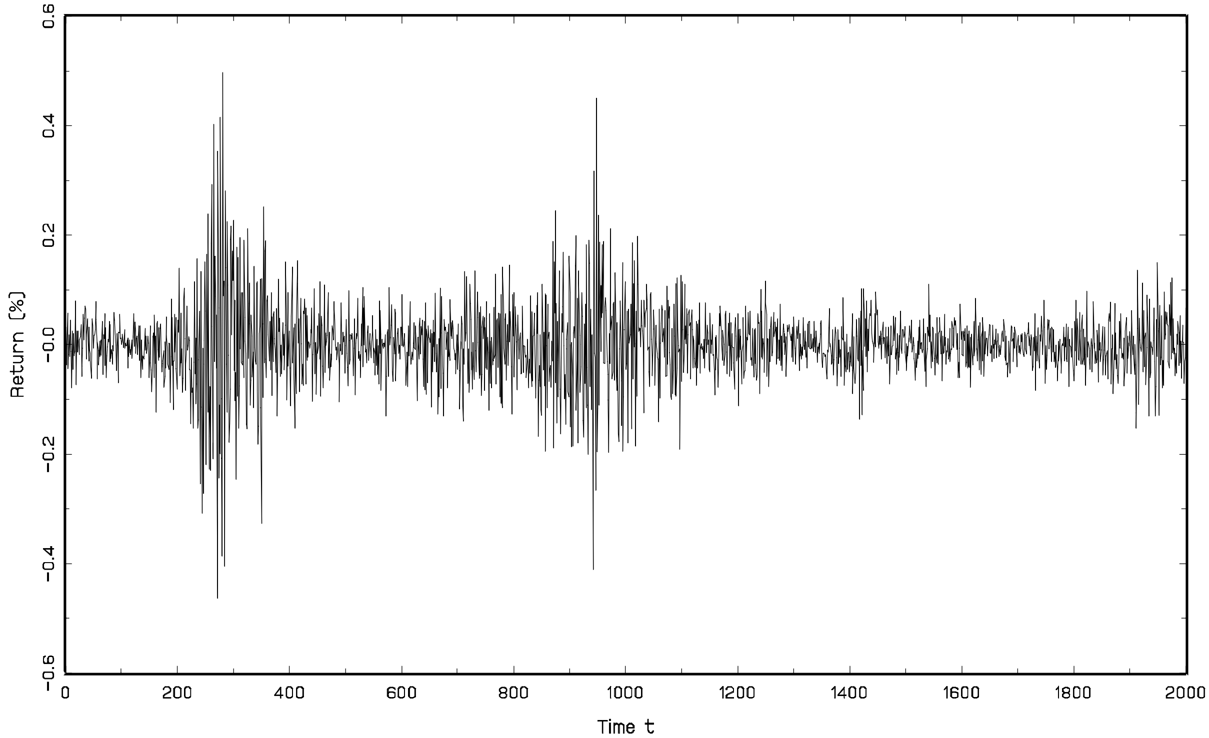

Figure 5: ACF of Hourly, Daily and Weekly Returns

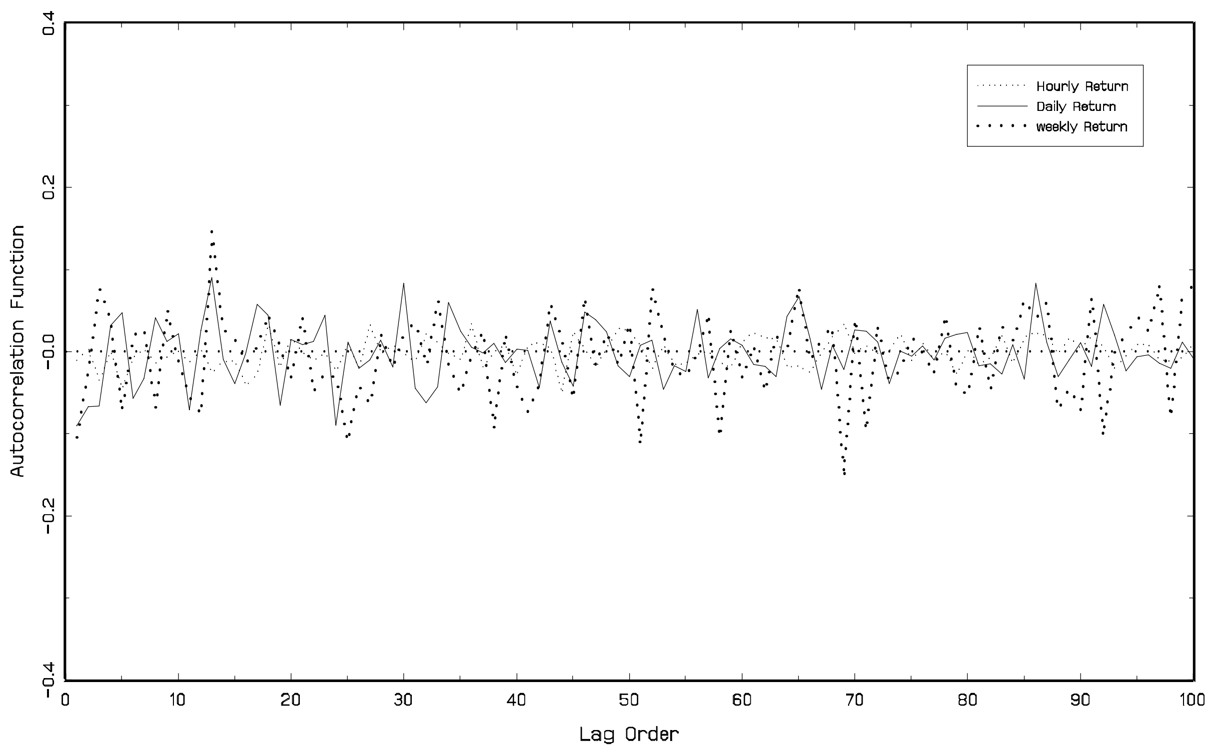

Figure 6: ACF of Absolute and Squared Daily Returns

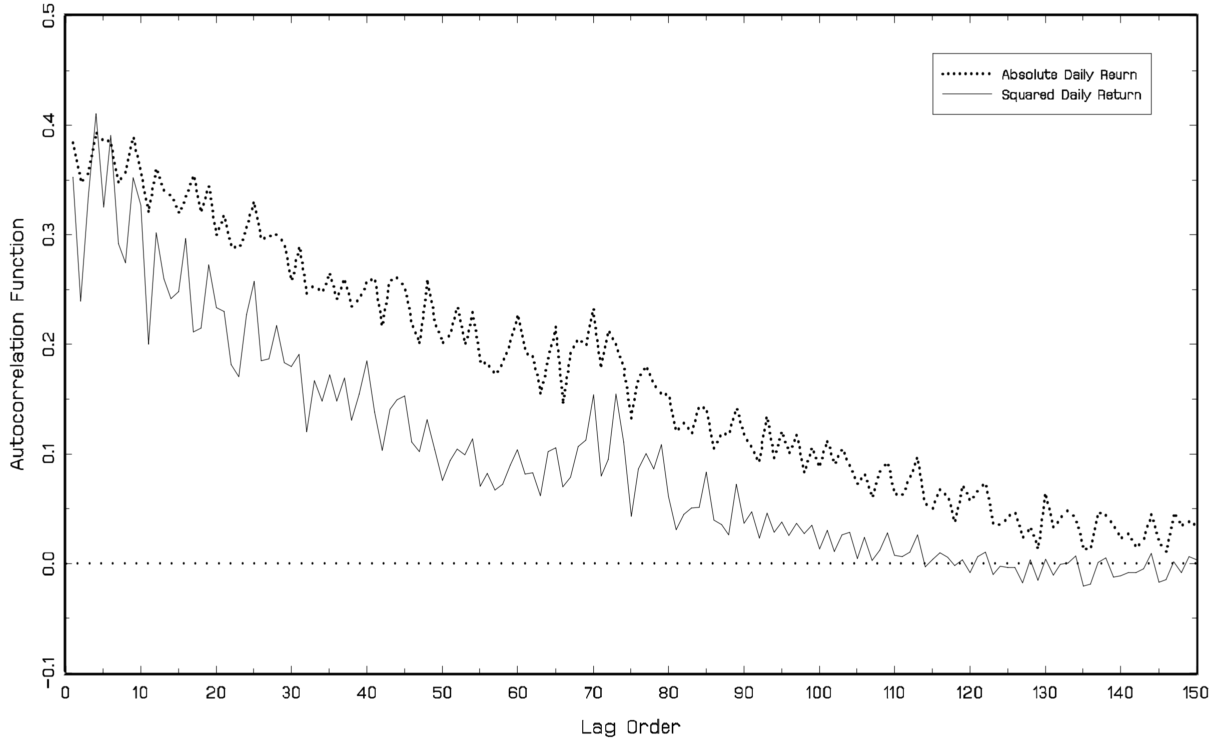

Figure 7: ACF of Abs(Rt)^theta: Daily Return