



Article

Bayesian estimation of the fractional Ornstein-Uhlenbeck instantaneous rate of asset return process: Evidence from high-frequency stock price data

Theodore Simos*

Department of Economics, University of Ioannina, University Campus, 45110 Ioannina, Greece, tel. +30-26510-05916, e-mail: tsimos18@gmail.com

Mike G. Tsionas

Lancaster University Management School, LA1 4YX, UK, tsionas@otenet.gr

04 July 2016

Abstract

Using recent developments in econometrics and computational statistics we consider the estimation of the instantaneous rate of asset return process when the underlying Data Generating Mechanism (DGM) is an Ornstein-Uhlenbeck process, driven by fractional noise, and sampled at fixed intervals of length h . To address the problem we adopt throughout the paper an exact discretization approach. This enable us to exploit the fact that a flow sampling scheme arises naturally when observing the DGM. For, while the instantaneous rate of return process is unobservable at points in time, its time integral over successive observations is observable since it equals the increment of log-prices. Exact discretization delivers an ARIMA(1,1,1) model for log-prices with a fractional driving noise. Building on the resulting exact discretization formulae and covariance function, a new Markov Chain Monte Carlo (MCMC) scheme is proposed and we examine the properties of both the time and frequency domain likelihoods / posteriors through Monte Carlo. For the exact discrete model we adopt a general sampling interval of length h . This allow us to determine the optimal choice of h independent of the sample size. An empirical application using high frequency stock price data is presented showing the relevance of aggregation over time issues in modelling asset prices.

Keywords: Bayesian modeling, long memory/anti-persistence; continuous time modeling; MCMC.

JEL codes: C11, C22, C58

*Corresponding author.

1. Introduction

Continuous time models are gaining wide acceptance in econometrics and finance in view of data availability at very high frequencies. Modern finance theory rely heavily on continuous time analysis, see for example the influential book by Merton (1990). However, modelling in discrete time may induce temporal aggregation problems that occur when observing a flow variable or the timing of observations is different from that of the underlying DGM. These issues are largely ignored in applied work, however proper accounting for them is important to ensure the correct specification of asset price dynamics. The seminal article of Bergstrom (1983) and subsequent work provided a rigorous treatment of the estimation problem by adopting an exact discretization procedure and deriving the Exact Discrete Model (EDM) for a range of continuous time models.

In this paper we consider the problem of estimating the instantaneous rate of asset return process when sampled at fixed time intervals of length h using exact discretization techniques. Note that we allow for a general, not necessary unit-spacing, sampling interval h . Zhu and Taqqu (2005) noted, in the case of a fBm, that although the rate of convergence of the variance parameter is \sqrt{N} , where

$N = \frac{T}{h}$ is the number of observations and T is the time span, irrespective of the sampling interval

the value of optimal spacing depends from the Hurst exponent H . Using Bayesian technics, we are

able to determine an optimal h that in general will depend from the values of the parameters and the sample size.

We assume that between successive observations the, unobservable, asset return process obeys an Ornstein-Uhlenbeck (OU) process driven by a fractional noise. A flow sampling scheme allow us to express the integral of the realization of an asset return process over successive observations as the increment of log-prices. Thus, we are able to show that for a fixed sampling interval of length h , log-prices are generated according to a fractional type ARIMA (1,1,1) model. The source of the particular ARIMA structure is traced back to the temporal aggregation properties of the flow sampling scheme. Furthermore, the derived exact discrete model does not contradict a benchmark EMH since in the stationary region, $a < 0$, log prices possess a unit root exhibiting a stochastic trending behaviour.

Using high frequency stock price data from Athens Stock Exchange (ASE) we evaluate the importance of the aggregation over time effect on asset prices. Lo and Wang (1995) as well as Perron and Vodounou (2004, p. 207) noted also, in continuous time, the ARIMA structure of log prices. However, in their work the ARIMA structure is originated in the reduced form version of the stationary AR(1) plus random walk model. In discrete time, Lo and MacKinlay (1999, pp. 79-81) noted the empirical relevance of the ARMA(1,1) model for small to medium size log-price increments. Ait-Sahalia, Mykland and Zhang (2005) also noted the MA(1) structure of innovations in the presence of microstructure noise, and Griffin and Oomen (2008) the ARMA(1,1) dependence structure of high frequency returns in tick time and transaction time.

We model persistence, through the Hurst parameter H , by introducing a correlated fractional driving noise in the OU process. Fractional noise is a flexible tool for modelling long term persistence or anti-persistence in asset price time series. The OU process is mean reverting and its consequence for portfolio and consumption decisions has been considered in Wachter (2002) while empirical evidence and tests of mean reverting stock returns are stated in Lo and MacKinlay (1999) section 2. Long range dependence and anti-persistence are empirically relevant in high frequency data coming from financial markets where microstructure noise come into play. Moreover, it has important implications both in derivative pricing as well as in testing for possible arbitrage opportunities and departures from the Efficient Market Hypothesis (EMH).

The estimation method, in this work, is based on closed-form solutions for the covariance matrix of the instantaneous rate of asset return process. Then, assuming that the instantaneous rate of return is a Gaussian fractional process the likelihood can be formulated easily and formal statistical inferences via an efficient MCMC scheme are possible. We use recently developed differential geometry samplers to cope with the likelihood / posterior and provide Monte Carlo results as well as an application to real-world data from the ASE. We examine both the time-domain and spectral-domain likelihoods / posteriors and compare their relative performance. Using proper numerical techniques we find that the likelihood / posterior is well behaved, autocorrelation of the MCMC samplers is mitigated to a great extent and inferences can be obtained in a fraction of time relative to more conventional MCMC schemes.

In the Bayesian literature there have been many studies which implement Markov Chain Monte Carlo (MCMC) sampling schemes to draw from the posterior of the model. Most studies focus on an Euler discretization and treat the missing observations on the continuous path as latent variables. This overhead is quite important because when the time interval of observations is longer relative to the actual spacing of observations, the latent variables are highly autocorrelated and the performance of MCMC schemes can deteriorate very quickly.

The difficulty is that the likelihood function is not available in closed form for most models, and so one has to resort to various approximations. Pedersen (1995) and Santa-Clara (1995) (see also Brandt and Santa-Clara, 2002) proposed a simulation-based approach, which is based on integrating out unobserved states of the process between pairs of observations. See also Gallant (2002). Ait-Sahalia (1999, 2002) proposed a method to approximate the transition density in a one-dimensional diffusion setting. The MCMC methods have been used by Eraker (2001), Jones (1999), Elerian et al. (2001), and Kim et al. (1998). Typically, MCMC methods simulate sample paths across unobserved intermediate points between two observations.

When the amount of augmentation is large MCMC methods that rely on data augmentation can break down. Papaspiliopoulos et al. (2003) find that high correlation between unknown parameters in a diffusion coefficient and the missing data results in arbitrarily slow rates of convergence of sampling schemes such as the Gibbs sampler proposed in Eraker (2001) and Golightly and Wilkinson (2005). See also Golightly and Wilkinson (2008).

Some attempts have been made to rectify this problem; Roberts and Stramer (2001) transform the stochastic differential equation to obtain constant diffusion coefficients so that the dependence problem is overcome. Beskos et al. (2006) propose alternative Monte Carlo methods that rely on non-centered parametrizations of the process. In this paper we show that although the density cannot be obtained in closed form the likelihood function can be obtained analytically for the class of continuous time autoregressive models with a fractional parameter. Since we rely on analytical results computation of the likelihood function is trivial and efficient Monte Carlo methods can be used to evaluate it. In addition, although the simple Metropolis-Hastings algorithm can be used, we implement more efficient alternatives and show in artificial and real world data that they perform well. The analytical methods of the paper in conjunction with Monte Carlo methods of inference can be used in a variety of contexts involving stochastic differential equations and can be thought of as extremely efficient alternatives to existing methods.

Tsai and Chan (2005 a, b) consider maximum likelihood estimation of continuous time ARMA (p, q) processes driven by fractional Brownian motion while Tsai and Chan (2005 c) obtain the autocorrelation properties of aggregates from a CARFIMA model. However, it should be noticed, that unlike Tsai and Chan (2005 a, b and c) we base our analysis and empirical application on the derived EDM. The EDM is particularly convenient in the current framework, since, in addition to using it in the estimation procedures, forms the basis for the Monte Carlo investigation of the estimates.

The paper is organized as follows. Section 2 sets the continuous time DGM and its assumptions. The exact discrete model, satisfying the underlying continuous time process, along with its second-order properties are stated in section 3. The likelihood functions, in time and spectral domain, are provided in section 4. The algorithms implementing Bayesian inference are described in section 5. An empirical implementation using high frequency data is discussed in section 6. Section 7 concludes. The derivation of the exact discretization formulae is provided in the Appendix.

2. The data generating mechanism

To fix notation in what follows, let the instantaneous rate of return be defined by $r(t) = \frac{d}{dt} \log P(t)$, where $P(t)$ is the price of an asset or portfolio at time t . In the sequel, the DGM is formulated assuming that log prices are influenced by the ongoing effects of the continuous time unobserved process $r(t)$. Therefore, although log prices are observed at points in time, as stock variables, their movement between successive observations are unobserved. Using an exact discretization procedure, we are able to express log-returns as an integral of the instantaneous rate of return $r(t)$, over the observational interval. This allows us to examine the implications and the empirical relevance of the fact that $r(t)$ is observed as a flow variable (for the econometric implications of flow variables in macroeconomic modelling, see Bergstrom, 1984). In order to obtain the exact discrete model we shall make use of the following assumption:

Assumption 1

The instantaneous rate of return is generated according to the Ornstein-Uhlenbeck type diffusion model

$$dr(t) = ar(t)dt + cdt + \varepsilon(t)dt \quad (1)$$

under the fixed initial condition

$$r(0) = x, \quad (2)$$

where the innovation $\varepsilon(t)$ is a continuous time fractional Gaussian noise that possesses the auto-covariance function

$$R_{\varepsilon}(s) = c(H)|s|^{2H-2}, \quad (3)$$

$$c(H) = \frac{\sigma^2}{2\pi} 2 \sin\left(\frac{\pi}{2}(2H-1)\right) \Gamma(2-2H).$$

Note that H is the Hurst parameter taking values on the interval $(0,1)$. If $0 < H < \frac{1}{2}$ (anti-persistence),

$c(H) < 0$ and the spectral density tends to zero as $\omega \rightarrow 0$. If $\frac{1}{2} < H < 1$ (long-memory), $c(H) > 0$ and the spectral density function tends to infinity as $\omega \rightarrow 0$. Moreover, using a frequency domain approach, the fractional Gaussian noise possess the spectral density function:

$$f_\varepsilon(\omega) = \frac{\sigma^2}{2\pi} \frac{1}{|\omega|^{2H-1}}. \quad (4)$$

For a definition and properties of the fractional Gaussian noise see Samorodnitsky and Taqqu (1994).

3. The exact discrete model

Next, we state the exact discrete model along with its second-order properties. Following the discussion in the Introduction we assume that the investigator observes asset prices at fixed equally spaced points in time $h, 2h, 3h, \dots$ over a fixed time interval $[0, T]$.

Theorem. Define the observed sample $\bar{x}_{th} = \int_{(t-1)h}^{th} r(s) ds = \log P(th) - \log P((t-1)h)$, $t = 1, 2, \dots, \frac{T}{h}$, where h is the fixed observational interval. Then under Assumption 1 \bar{x}_{th} satisfies the exact discrete model:

$$\bar{x}_{th} = e^{ah} \bar{x}_{(t-1)h} + v h + \kappa_{th}, \quad (5)$$

where

$$v = cK_1(h) \quad (6)$$

$$\kappa_{th} = \int_0^h K_1(r) \mathcal{E}(th-r) dr + \int_0^h K_2(r) \mathcal{E}(th-h-r) dr \quad (7)$$

$$K_1(x) = \int_0^x e^{as} ds = \sum_{j=0}^{\infty} \frac{1}{(j+1)!} a^j x^{j+1} \quad (8)$$

$$K_2(x) = \int_x^h e^{as} ds = \sum_{j=0}^{\infty} \frac{1}{(j+1)!} a^j [h^{j+1} - x^{j+1}], \quad (9)$$

the auto-covariance $\gamma_k^F(\tau h) = E[\kappa_{th} \kappa_{(t+\tau)h}]$ and spectral density functions of the error term are given, respectively, by

$$\begin{aligned} \gamma_k^F(\tau h) &= \int_0^h \int_0^h K_1(r) R_\varepsilon(\tau h + r - s) K_1(s) dr ds \\ &\quad + \int_0^h \int_0^h K_2(r) R_\varepsilon(\tau h + r - s) K_2(s) dr ds \\ &\quad + \int_0^h \int_0^h K_1(r) R_\varepsilon((\tau+1)h + r - s) K_2(s) dr ds \\ &\quad + \int_0^h \int_0^h K_2(r) R_\varepsilon((\tau-1)h + r - s) K_1(s) dr ds, \\ h_k^F(\omega) &= \frac{\sigma^2}{2\pi} \sum_{j=-\infty}^{\infty} \left| K\left(\omega + \frac{2\pi j}{h}\right) \right|^2 f_\varepsilon\left(\omega + \frac{2\pi j}{h}\right), \end{aligned} \quad (10)$$

where

$$K(\omega) = \int_0^h e^{-i\omega s} K_1(s) ds + e^{-i\omega h} \int_0^h e^{-i\omega s} K_2(s) ds, \quad (11)$$

and $R_\varepsilon(s)$, $f_\varepsilon(\omega)$ are stated in (3)-(4).

The derived exact discrete model, stated in the Theorem, is particularly suitable for observable variables that are sampled at different rates since, unlike more conventional discrete-time models, it retains its dynamic structure irrespective of the sampling frequency.

Remark

Under stationarity, $a < 0$, we obtain from equations (8) and (9): $K_1(r) = \frac{e^{ar} - 1}{a}$ and $K_2(r) = \frac{e^{ah} - e^{ar}}{a}$ respectively. Then from (12) it follows, after some simple algebraic manipulations

$$|K(\omega)|^2 = \frac{2(1 - \cos(\omega h))}{\omega^2(a^2 + \omega^2)} |1 - e^{ah - i\omega h}|^2,$$

where $|1 - e^{ah - i\omega h}|^2 = 1 + e^{2ah} - 2e^{ah} \cos(\omega h)$. Thus from (11) we deduce the spectral density function of the discrete error term:

$$h_k^F(\omega) = \frac{\sigma^2}{2\pi} 2(1 - \cos(\omega h)) (1 + e^{2ah} - 2e^{ah} \cos(\omega h)) \times \sum_{j=-\infty}^{\infty} \frac{1}{a^2 \left| \omega + \frac{2\pi j}{h} \right|^{2H+1} + \left| \omega + \frac{2\pi j}{h} \right|^{2H+3}}.$$

4. Gaussian likelihood - time and frequency domain methods

To construct the likelihood function conditional on the initial value $\bar{x}_0 = 0$, we first express the exact discrete model for $t=1, \dots, T/h$ in system form as

$$F\bar{x} - \bar{v} = k \quad (13)$$

where

$$F = \begin{bmatrix} 1 & 0 & . & . & . & 0 \\ -e^{ah} & 1 & 0 & . & . & 0 \\ 0 & -e^{ah} & 1 & . & . & . \\ 0 & 0 & -e^{ah} & . & . & . \\ . & . & . & . & . & 0 \\ 0 & 0 & . & 0 & -e^{ah} & 1 \end{bmatrix}, \quad (14)$$

$$\bar{x} = [\bar{x}_h \quad \bar{x}_{2h} \quad . \quad . \quad . \quad \bar{x}_T]^T, \quad \bar{v} = [v \quad v \quad . \quad . \quad . \quad v]^T h \text{ and}$$

$$k = [k_h \quad k_{2h} \quad . \quad . \quad . \quad k_T]^T.$$

Next, define the $T \times T$ covariance matrix of the innovation vector $\Omega = E[kk^T]$ and the parameter vector $\theta = \{a, c, H, \sigma^2\}$. Then, apart from a constant, minus twice the logarithm of the Gaussian likelihood function is expressed by

$$L(\theta) = \log |\Omega| + \bar{x}' F' \Omega^{-1} F \bar{x} = \log |V| + tr[V^{-1} \bar{x} \bar{x}'], \quad (15)$$

where the matrix

$$V = F^{-1} \Omega F'^{-1} \quad (16)$$

is a function of the parameter vector θ ,

$$\Omega = \begin{bmatrix} \gamma_k^F(0) & \gamma_k^F(h) & . & . & . & \gamma_k^F(T-h) \\ \gamma_k^F(h) & \gamma_k^F(0) & . & . & . & \gamma_k^F(T-2h) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \gamma_k^F(T-h) & \gamma_k^F(T-2h) & . & . & . & \gamma_k^F(0) \end{bmatrix} \quad (17)$$

and $\gamma_k^F(\tau h)$, $\tau = 0, \dots, T-h$ are calculated using by (10) using equations. (3) and (8)-(9).

4.1 Computational procedure to evaluate the likelihood

Let the Cholesky factorization be $\Omega = MM'$, where M is a lower triangular matrix with positive elements m_{ii} along the diagonal. Then the likelihood function takes the form

$$L = \sum_{i=1}^{T/h} (z_i^2 + 2 \log m_{ii}), \quad (18)$$

where $z = [z_1 \ z_2 \ \dots \ z_{T/h}]$ is a vector whose elements z_i are evaluated recursively from the system of equations

$$Mz = k. \quad (19)$$

4.2 Whittle likelihood

Under stationarity, $a < 0$, the spectral density of the exact discrete model is given by

$$\begin{aligned} f_{\bar{x}_{th}}(\omega) &= [1 + e^{2ah} - 2e^{ah} \cos(\omega h)]^{-1} h_k^F(\omega) \\ &= \frac{\sigma^2}{2\pi} 2(1 - \cos(\omega h)) \sum_{j=-\infty}^{\infty} \frac{1}{a^2 \left| \omega + \frac{2\pi j}{h} \right|^{2H+1} + \left| \omega + \frac{2\pi j}{h} \right|^{2H+3}} \end{aligned} \quad (20)$$

where $h_k^F(\omega)$ is the spectral density of the error term κ_{th} . The Whittle likelihood function is given by

$$L(\theta) = \frac{h}{2\pi T} \sum_{j=-\frac{T}{h}}^{\frac{T}{h}} \left[\log f_{\bar{x}_{th}}(\omega_j) + \frac{I(\omega_j)}{f_{\bar{x}_{th}}(\omega_j)} \right] \quad (21)$$

Where

$$\begin{aligned} \omega_j &= \frac{2\pi jh}{T}, \\ I(\omega) &= \left| \sqrt{\frac{h}{2\pi T}} \sum_{t=1}^{\frac{T}{h}} (\bar{x}_{th} - \mu) e^{-i\omega t} \right|, \\ \mu &= (1 - e^{ah})^{-1} v h. \end{aligned}$$

5. MCMC samples from the posterior

The likelihood function is $\mathcal{L}(\theta; \mathbf{Y})$ and $p(\theta)$ denotes the prior. Our prior has the standard reference form: $p(\theta) \propto \sigma^{-1} 1_{(0,1)}(H)$. For the likelihood function we use both the time-domain and spectral-domain expressions and compare the results both in terms of parameter estimates, posterior distributions as well as convergence through the autocorrelation functions produced by the MCMC schemes.

The most significant advantage of our numerical scheme is that it reduces the difficult problem of Bayesian inference to sampling from a tri-variate posterior¹. To our knowledge, previous treatments of the problem typically rely on sampling from high-dimensional posteriors by treating the intermediate steps of the diffusion process as latent variables. Although Gibbs sampling can be used in these cases to facilitate the computations, these approaches are always prone to convergence problems associated with high autocorrelation induced when sampling the latent states in one way or another.

Although a standard Metropolis-Hastings algorithm can be used, due to the high nonlinearity of the posterior we use the method of Girolami and Calderhead (2011) to perform Bayesian inference. As the authors write “[t]he proposed methodology exploits the Riemannian geometry of the parameter space of

¹ The matrix Ω becomes frequently non-invertible numerically. To address this problem we use Tychonov regularization of the matrix. The regularization parameter is chosen so that starting from a value 10^{-12} the matrix becomes invertible numerically. We monitor this constant so that it does not become excessively large. On the average the regularization constant is close to 10^{-5} (relative to the smallest non-zero eigenvalue) for both artificial and real data. Draws for which this does not hold are discarded. This never happened in more than 5% of all draws.

statistical models and thus automatically adapts to the local structure when simulating paths across this manifold providing highly efficient convergence and exploration of the target density".

Given a kernel posterior distribution the method of Girolami and Calderhead (2011) aims at obtaining draws by exploiting the gradient and the Hessian. Roughly speaking, parameters are updated so that they visit the area of high posterior probability mass, not unlike standard Newton-Raphson numerical optimization techniques.

Following Girolami and Calderhead (2011) we utilize Metropolis adjusted Langevin and Hamiltonian Monte Carlo sampling methods defined on the Riemann manifold, since we are sampling from target densities with high dimensions that exhibit strong degrees of correlation. Consider the Langevin diffusion:

$$d\boldsymbol{\theta}(t) = \frac{1}{2} \nabla \log p(\boldsymbol{\theta}(t); \mathbf{Y}) dt + d\mathbf{B}(t),$$

where \mathbf{B} denotes the d-dimensional Brownian motion. The first-order Euler discretization provides the following candidate generation mechanism:

$$\boldsymbol{\theta}^* = \boldsymbol{\theta}^o + \frac{1}{2} \varepsilon^2 \nabla \log p(\boldsymbol{\theta}^o; \mathbf{Y}) + \varepsilon \mathbf{z},$$

where $\mathbf{z} \sim N_d(\mathbf{0}, \mathbf{I})$, and $\varepsilon > 0$ is the integration step size. Since the discretization induces an unavoidable error in approximation of the posterior, a *Metropolis step* is used, where the proposal density is

$$q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^o) = N_d\left(\boldsymbol{\theta}^o + \frac{1}{2} \varepsilon^2 \nabla \log p(\boldsymbol{\theta}^o; \mathbf{Y}), \varepsilon^2 \mathbf{I}\right),$$

$$\text{with acceptance probability } a(\boldsymbol{\theta}^o, \boldsymbol{\theta}^*) = \min\left\{1, \frac{p(\boldsymbol{\theta}^* | \mathbf{Y}) q(\boldsymbol{\theta}^o | \boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}^o | \mathbf{Y}) q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^o)}\right\}.$$

The Brownian motion of the Riemann manifold is given by:

$$d\mathbf{B}_i^{\mathbf{G}}(t) = \left| \mathbf{G}(\boldsymbol{\theta}(t)) \right|^{-1/2} \sum_{j=1}^d \frac{\partial}{\partial \boldsymbol{\theta}_j} \left[\mathbf{G}^{-1}(\boldsymbol{\theta}(t))_{ij} \left| \mathbf{G}(\boldsymbol{\theta}(t)) \right|^{1/2} \right] dt + \left[\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta}(t))} d\mathbf{B}(t) \right]_i,$$

for $i = 1, \dots, d$.

The discrete form of the above stochastic differential equations is:

$$\begin{aligned} \boldsymbol{\theta}_i^* &= \boldsymbol{\theta}_i^o + \frac{1}{2} \varepsilon^2 \left[\mathbf{G}^{-1}(\boldsymbol{\theta}^o) \nabla \log p(\boldsymbol{\theta}^o; \mathbf{Y}) \right]_i - \varepsilon^2 \sum_{j=1}^d \left[\mathbf{G}^{-1}(\boldsymbol{\theta}^o) \frac{\partial \mathbf{G}(\boldsymbol{\theta}^o)}{\partial \boldsymbol{\theta}_j} \mathbf{G}^{-1}(\boldsymbol{\theta}^o) \right]_{ij} \\ &\quad + \varepsilon^2 \sum_{j=1}^d \mathbf{G}^{-1}(\boldsymbol{\theta}^o)_{ij} \text{tr} \left[\mathbf{G}^{-1}(\boldsymbol{\theta}^o) \frac{\partial \mathbf{G}(\boldsymbol{\theta}^o)}{\partial \boldsymbol{\theta}_j} \right] + \varepsilon \left[\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta}(t))} \mathbf{z} \right]_i @ \\ &\quad \boldsymbol{\mu}(\boldsymbol{\theta}^o, \varepsilon)_i + \varepsilon \left[\sqrt{\mathbf{G}^{-1}(\boldsymbol{\theta}(t))} \mathbf{z} \right]_i. \end{aligned}$$

The proposal density is $\boldsymbol{\theta}^* | \boldsymbol{\theta}^o \sim N_d(\boldsymbol{\mu}(\boldsymbol{\theta}^o, \varepsilon), \varepsilon^2 \mathbf{G}^{-1}(\boldsymbol{\theta}^o))$ and the acceptance probability has the

$$\text{standard Metropolis form: } a(\boldsymbol{\theta}^o, \boldsymbol{\theta}^*) = \min\left\{1, \frac{p(\boldsymbol{\theta}^* | \mathbf{Y}) q(\boldsymbol{\theta}^o | \boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}^o | \mathbf{Y}) q(\boldsymbol{\theta}^* | \boldsymbol{\theta}^o)}\right\}.$$

Both algorithms are started at the maximum likelihood parameter estimates and convergence is tested using Geweke's (1992) diagnostics. Our MCMC implementation of Girolami and Calderhead (2011) uses 120,000 iterations the first 20,000 of which are discarded to mitigate the possible impact of start-up effects. Convergence was found to occur during the transient or burn-in phase. For the Metropolis-Hastings algorithm convergence was typically more difficult to obtain so we used 500,000 iterations discarding between 100,000 and 250,000 iterations, depending on the application. Every other tenth draw is omitted to approximate marginal posteriors through kernel densities to mitigate the impact of autocorrelation induced by MCMC. Posterior moments are obtained using the full set of draws along an autocorrelation-robust Newey and West covariance matrix (lags=10).

Numerical singularity of covariance matrices typically occurs in the time-domain likelihood. In that case, we always used a “sweep inverse” and the results have been checked against Tykhonov regularization, see footnote 1. The spectral likelihood has been found much better behaved, which is probably responsible for the excellent behavior of the Girolami and Calderhead (2011) scheme in this case. Even with the time-domain likelihood, the Girolami and Calderhead (2011) MCMC scheme had no problems converging relatively fast and provide autocorrelations which were significantly lower compared to those of the Metropolis-Hastings update. Our implementation of the Metropolis-Hastings update relies on using a trivariate Student-t distribution with 5 degrees of freedom. The location and scale parameters were set to those obtained through ML estimation. The covariance matrix is scaled so that the algorithm provides an acceptance rate between 20% and 30%. The Girolami and Calderhead (2011) scheme was configured to produce acceptance rates in the same range by setting ε appropriately after experimentation.

Table 1 reports total MSE for parameters a , H and σ estimated by Whittle and Gaussian likelihood methods on the basis of the exact discrete model stated in the Theorem. The results in Table 1 reveal that the Whittle likelihood method provides better parameter estimates, in terms of MSE, for both the anti-persistent and long memory cases.

Table 1

Table 1. Simulations ($a = -0.8$, $\sigma = 0.1$).

	Whittle Likelihood					Gaussian likelihood				
$H = 0.75$	$h = 1/2$	$h = 1/5$	$h = 1/20$	$h = 1/50$	$h = 1/100$	$h = 1/2$	$h = 1/5$	$h = 1/20$	$h = 1/50$	$h = 1/100$
$T = 50$	0.141	0.111	0.087	0.052	0.044	0.171	0.135	0.101	0.061	0.047
$T = 100$	0.101	0.071	0.042	0.025	0.017	0.154	0.093	0.065	0.029	0.018
$T = 500$	0.035	0.022	0.017	0.0082	0.0065	0.044	0.028	0.018	0.0085	0.0067
$H = 0.25$										
$T = 50$	0.121	0.089	0.044	0.037	0.028	0.138	0.094	0.047	0.038	0.028
$T = 100$	0.087	0.055	0.032	0.011	0.007	0.095	0.061	0.035	0.015	0.008
$T = 500$	0.013	0.0085	0.0064	0.0032	0.0017	0.017	0.0091	0.0068	0.0034	0.0017

The table reports total MSE for parameters a , H and σ estimated by Whittle and Gaussian likelihood

methods. For each experiment 1,000 different samples from the true DGM have been used. Starting

values are obtained by matching autocorrelations up to lag 10.

5.1 Unknown sampling interval

As long as h is not too close to zero (to avoid singularity of Ω) the sampling parameter h can be treated as unknown. The likelihood function is $\mathcal{L}(\boldsymbol{\theta}; \mathbf{Y})$ and $p(\boldsymbol{\theta})$ denotes the prior, where the parameter vector is $\boldsymbol{\theta} = [a, \sigma, H, h]'$. Our prior has the form: $p(a, \sigma, H) \propto \sigma^{-1} 1_{(0,1)}(H) h^{-1}$, where $p(h) \propto h^{-1}$. Other

choices of the prior are $p(h) \propto 1_{(\varepsilon,1)}(h)$ and Jeffreys' prior $p(h | a, \sigma, H) \propto \left\| -E_{\mathbf{Y}|\boldsymbol{\theta}} \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}; \mathbf{Y})}{\partial h^2} \right\|^{1/2}$, where

$E_{\mathbf{Y}|\boldsymbol{\theta}}$ denotes sampling expectation. This can be considered as a reference prior on the sampling parameter. The full prior is $p(\boldsymbol{\theta}) = p(a, \sigma, H) p(h | a, \sigma, H)$.

Although the derivative can be computed numerically, the sampling expectation cannot. For this reason, approximation of Jeffreys' prior requires a simulation of the DGM for given a, σ, H and a given h , where the prior is to be evaluated².

An optimal h can be obtained via the log marginal likelihood. Given a model whose likelihood function is $L(\theta; Y)$ for the parameter $\theta \in \Theta \subseteq R^k$ and the prior is $p(\theta)$ the posterior distribution is

$$p(\theta | Y) = \frac{L(\theta; Y)p(\theta)}{\int_{\Theta} L(\theta; Y)p(\theta)d\theta}$$

From this expression we have the identity (known as "candidate's formula"):

$$M(Y) = \frac{L(\theta; Y)p(\theta)}{p(\theta; Y)}, \forall \theta \in \Theta,$$

where the marginal likelihood is $M(Y) = \int_{\Theta} L(\theta; Y)p(\theta)d\theta$. Since this holds for every θ we can choose $\bar{\theta}$, the posterior mean, and assume that the posterior is nearly normal with mean $\bar{\theta}$ and covariance V which can be computed from the MCMC draws. Therefore, the log marginal likelihood can be computed approximately as:

$$\log M(Y) \approx \log L(\bar{\theta}; Y) + \log p(\bar{\theta}) + \frac{K}{2} \log(2\pi) + \frac{1}{2} \log |V|.$$

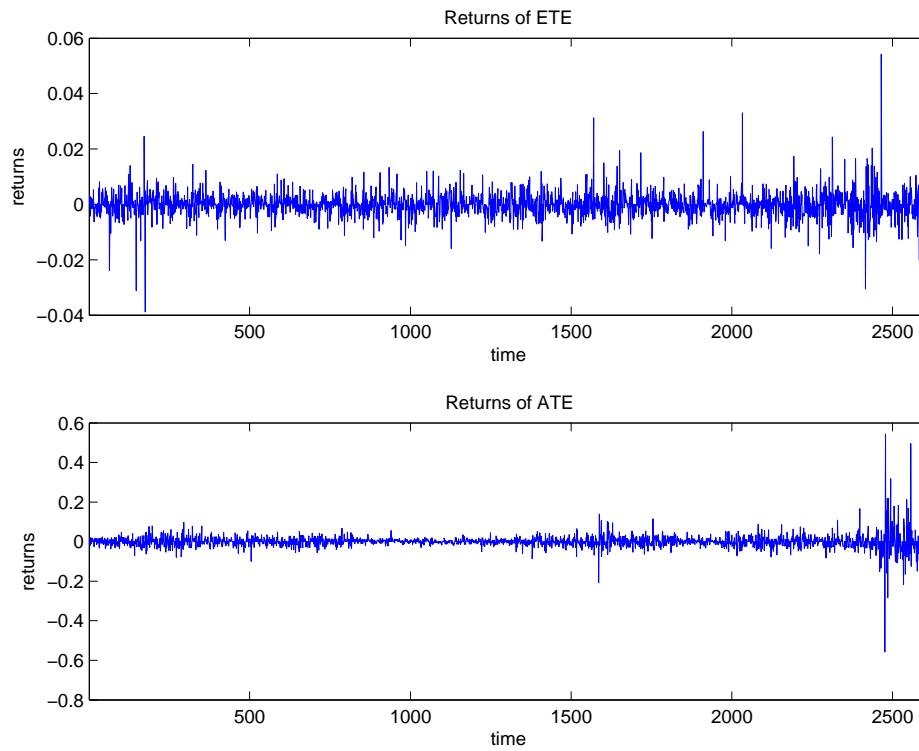
There is a related work in several papers that investigate the impact of the sampling rate on the properties of parameter estimates as well as optimality criteria of sampling schemes. See Zhu and Taqqu (2005) that examine the effects of sampling rate in the case of the fBm by utilizing the asymptotic covariance matrix of the estimates. Chambers (2004) considers the consistency properties of tests in unit root cases in flow data when the time span and/or frequency of observations increases.

6. Empirical results

The exact discretization (5) induces an ARIMA(1,1,1) model for log prices driven by fractional noise. Therefore, a negative a results in mean reversion in asset returns and a stochastic trending behavior for log prices, a well-known stylized feature of asset prices generated by competitive markets. When $0 < H < 0.5$ the driving noise is anti-persistent meaning that it reverts too frequently, while for $0.5 < H < 1$ it exhibits non-periodic long cycles. Econometric implications and empirical evidence of long memory/anti-persistence form a vast area in the relevant literature. Evidence for the fractal behavior of stock prices are provided for example in Lo and MacKinlay (1999, pp 147-184) and in Mulligan (2004). Moreover, Lillo and Farmer (2004) and Bouchaud, Kockelkoren & Potters (2006) investigate the fractal structure of market orders using high frequency data. In this work we compare three priors for the ETE (National Bank of Greece) and ATE (Agricultural Bank of Greece) price data, listed in the ASE, using $M=1,000$ simulations of the process. We implement the algorithms using two high frequency data sets. The first data set is obtained setting the sampling interval at 48 minutes and the second data set at a denser 5-minute interval. The data are from the period June 2009 to May 2010 (10:00 am to 17:00 pm) comprising 2,611 observations in the first case and 65,534 observations in the second case. The returns are plotted in Figure 1.

Figure 1

² The Jeffreys prior is approximated using 1,000 simulations of the data generating process. We opted for this choice after conducting experiments with 500, 1,000, and 5,000 simulations in a preliminary investigation for a wide range of parameter values.



Marginal posterior densities of parameters α and H are provided in Figure 2 for both the time domain (straight line) and the spectral domain likelihood functions for ETE (left panel) and ATE (right panel).

Figure 2

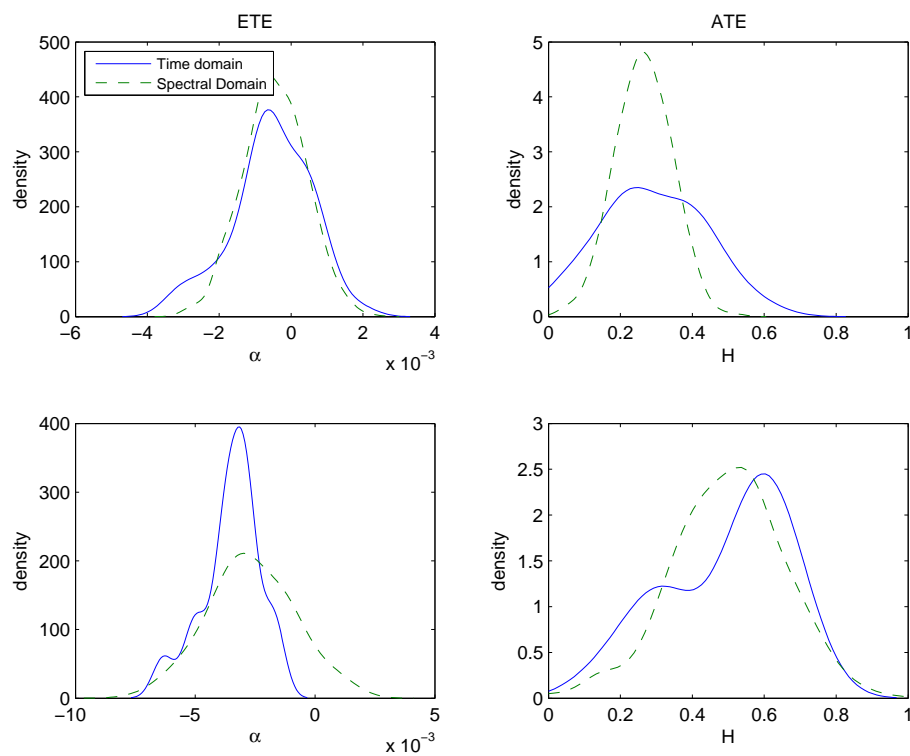


Table 2: Effects of varying spacing interval h on the parameter estimates

	ETE			ATE		
	α	σ	H	A	σ	H
h=1/14	-0.0005 0.0011	0.0131 0.0093	0.2857 0.1454	-0.0035 0.0012	0.0125 0.0097	0.4944 0.1658
	-0.0005 0.0009	0.0122 0.0095	0.2644 0.0788	-0.0027 0.0018	0.0128 0.0112	0.4992 0.1600
h=1/20	-0.0005 0.0009	0.0132 0.0095	0.2841 0.1602	-0.0026 0.0019	0.0125 0.0116	0.5111 0.1583
	-0.0005 0.0009	0.0123 0.0096	0.2572 0.0824	-0.0026 0.0019	0.0125 0.0114	0.4786 0.1630
h=1/10	-0.0005 0.0009	0.0120 0.0089	0.2870 0.1507	-0.0027 0.0018	0.0125 0.0097	0.5139 0.1489
	-0.0005 0.0009	0.0121 0.0091	0.2604 0.0821	-0.0027 0.0017	0.0126 0.0117	0.4848 0.1623
h=1/2	-0.0005 0.0010	0.0135 0.0095	0.3061 0.2513	-0.0027 0.0019	0.0125 0.0097	0.5327 0.2484
	-0.0004 0.0009	0.0125 0.0087	0.2618 0.1076	-0.0026 0.0019	0.0132 0.0077	0.4877 0.2127
h=7/10	- 0.0004 0.0011	0.0148 0.0071	0.2361 0.3026	-0.0026 0.0022	0.0155 0.0065	0.4636 0.2990
	-0.0005 0.0011	0.0165 0.0065	0.2590 0.1331	-0.0027 0.0022	0.0177 0.0053	0.4822 0.2631

— time domain likelihood

— spectral domain likelihood

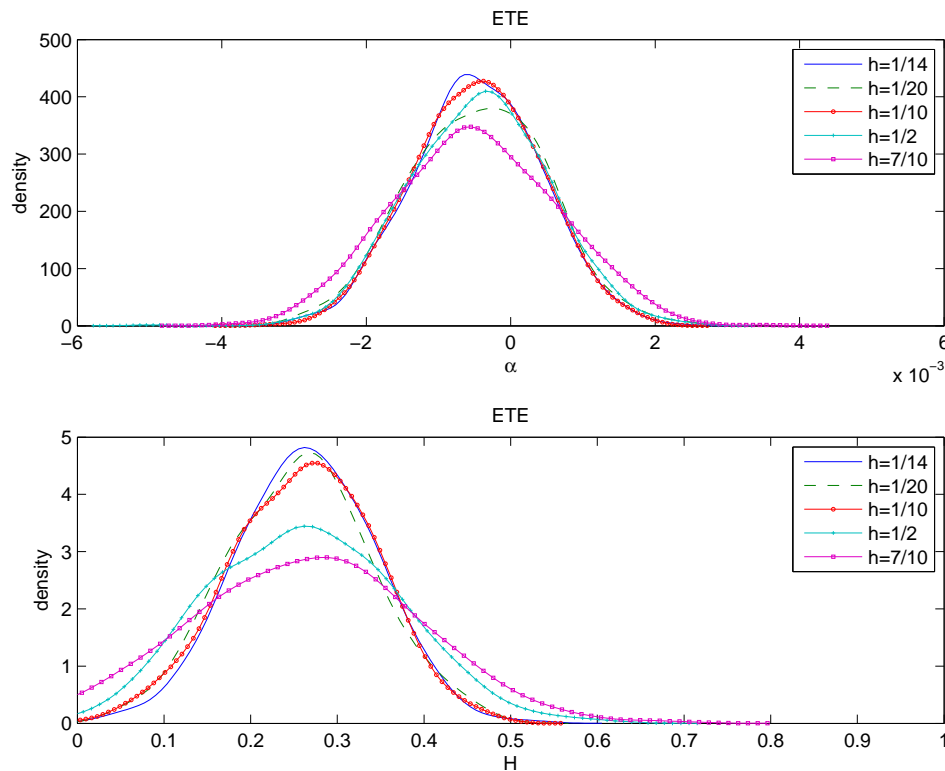
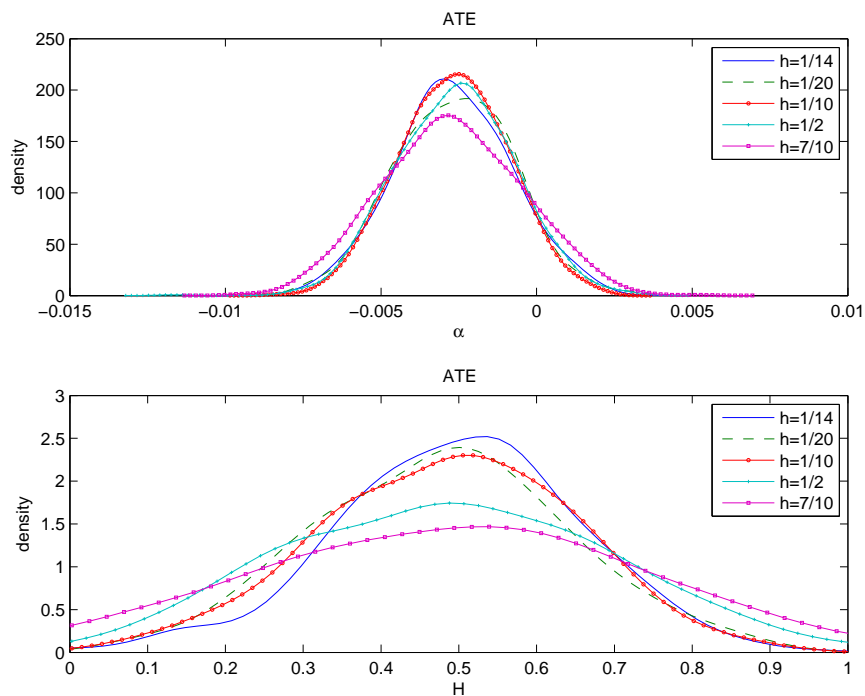
Each number is posterior mean followed below by posterior standard deviations. The MCMC algorithm is the Girolami-Calderhead (2012) Riemannian Manifold MCMC using 120,000 iterations the first 20,000 of which are discarded to mitigate start up effects.

Table 3. Posterior moments (5 min data)

	A		H	
ETE	-0.1695 * (0.2451) *	-0.1837 * (0.2424) *	0.4073 (0.1166)	0.4109 (0.2193)
ATE	-0.2398 ** (0.0751) **	-0.2400 ** (0.0959) **	0.4124 (0.1490)	0.4445 (0.2437)

(*) times 10^{-4} .(**) times 10^{-3} .

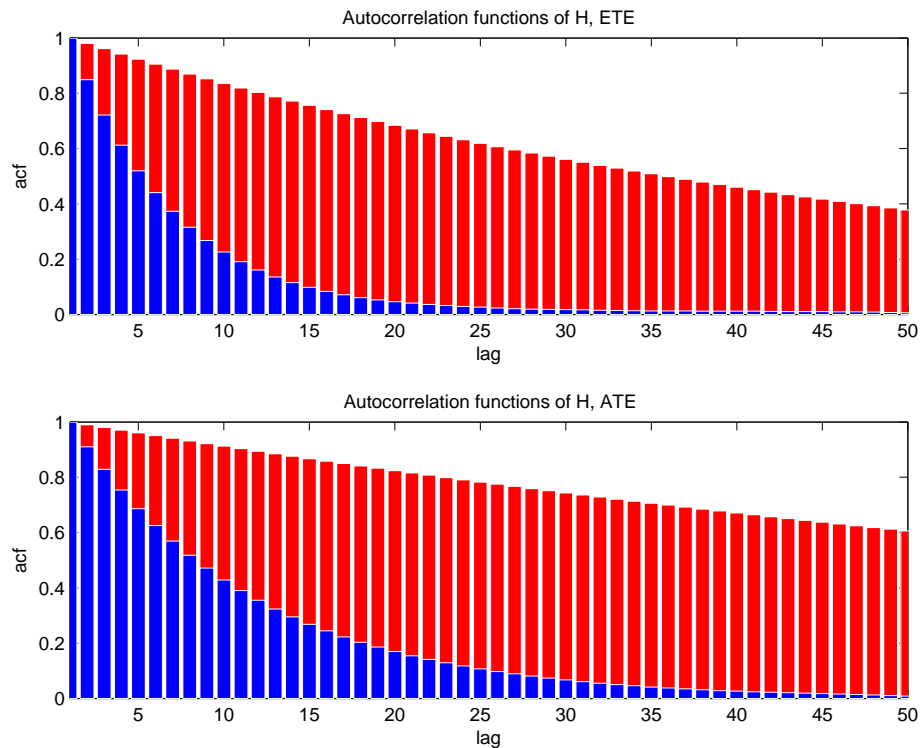
The parameter α , reported in Tables 2 and 3, is always negative but not significantly far from 0. In order to check the robustness of the estimates we report empirical results for different values of the parameter h in Table 2. Parameter H is concentrated around 0.3, in the anti-persistent region, for ETE and 0.5 for ATE. A Hurst exponent $H=0.5$ corresponds to an uncorrelated driving noise. The spectral domain posteriors seem to be more concentrated around the mean. The effect of changing h for parameters α and H are reported in Figure 3 for ETE and Figure 4 for ATE for the case of 48-min sampling data set.

Figure 3**Figure 4**

It is apparent that for $h=1/14$ marginal posteriors become broadly more concentrated around the posterior mean, therefore this value is close to an optimal one for our model. This effect is stronger in the case of the adjustment parameter α compared to the long memory parameter H . Zhu and Taqqu (2005), from a theoretical standpoint, investigate the determination of an optimal sampling interval h and its dependence from the parameter values. The Monte Carlo results concerning the existence of an optimal sampling deserve further investigation.

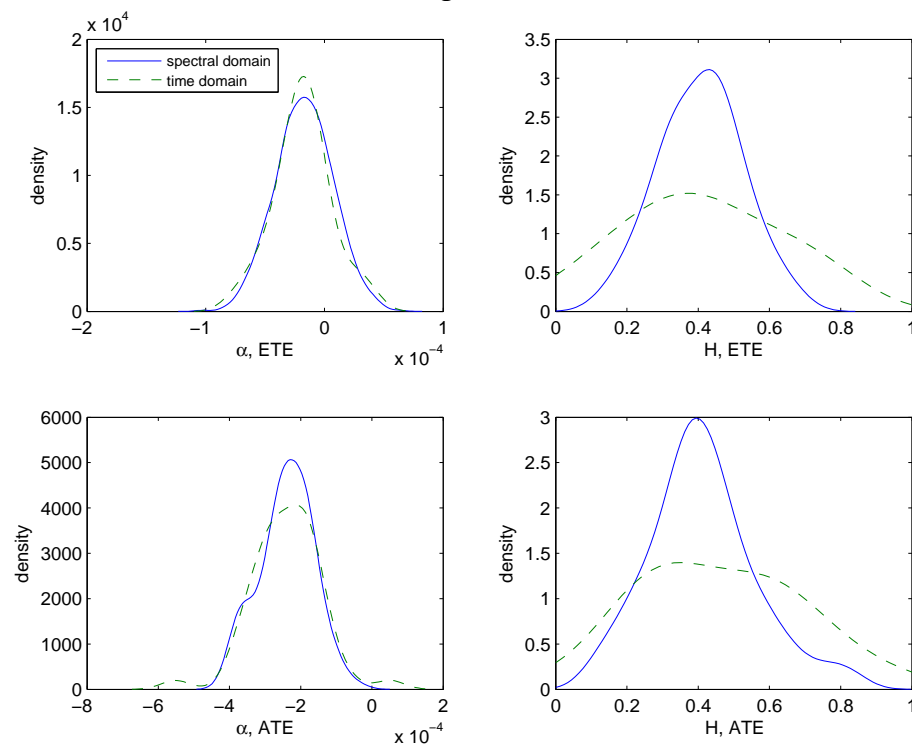
In Figure 5, we report autocorrelation functions corresponding to the MCMC draws for parameter H for the Girolami-Calderhead sampler (blue bars) and a simple Metropolis-Hastings sampler (red bars).

Figure 5



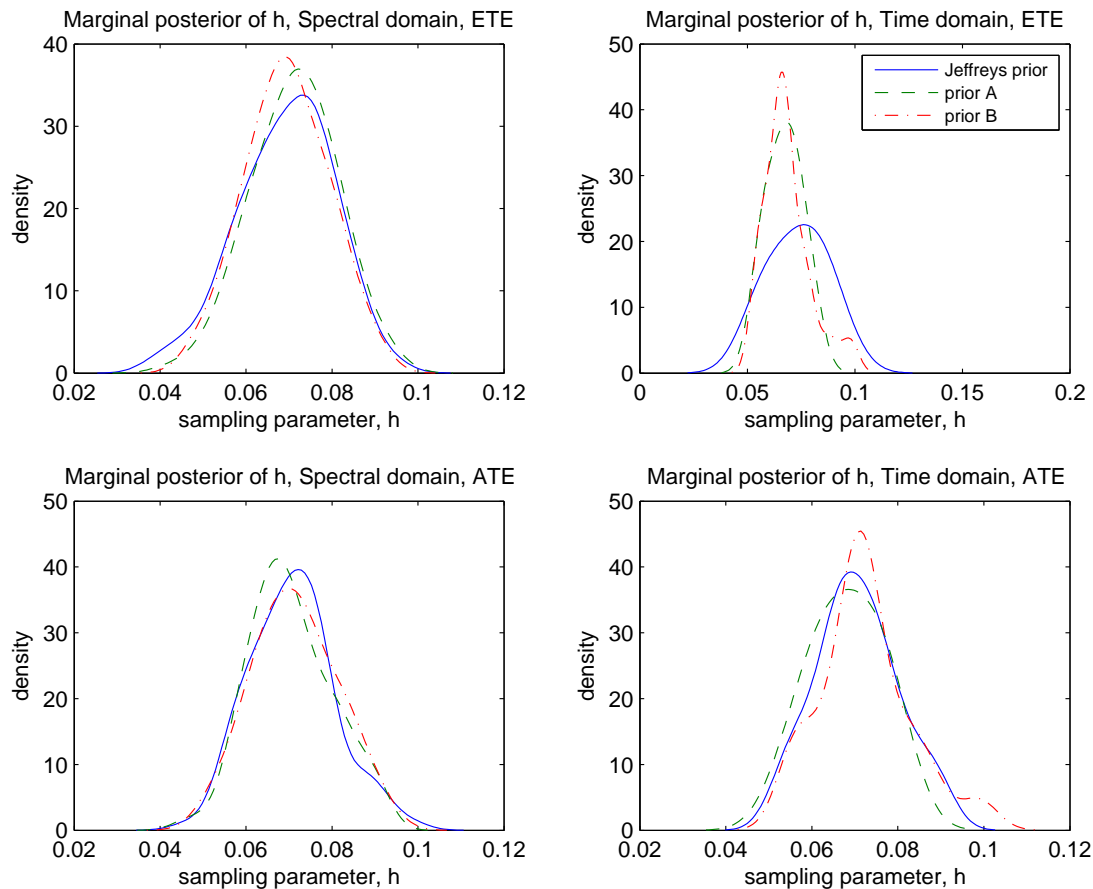
The behavior of the Girolami-Calderhead sampler is better by far as autocorrelations die out more rapidly and are practically zero after lag 30, which justified our choice to thin every other 50th draw. In Figure 6, we report marginal posteriors of α and H for the denser 5-min spacing data set of ETE and ATE from which it can be seen that the spectral domain posteriors behave much better at least for parameter H .

Figure 6



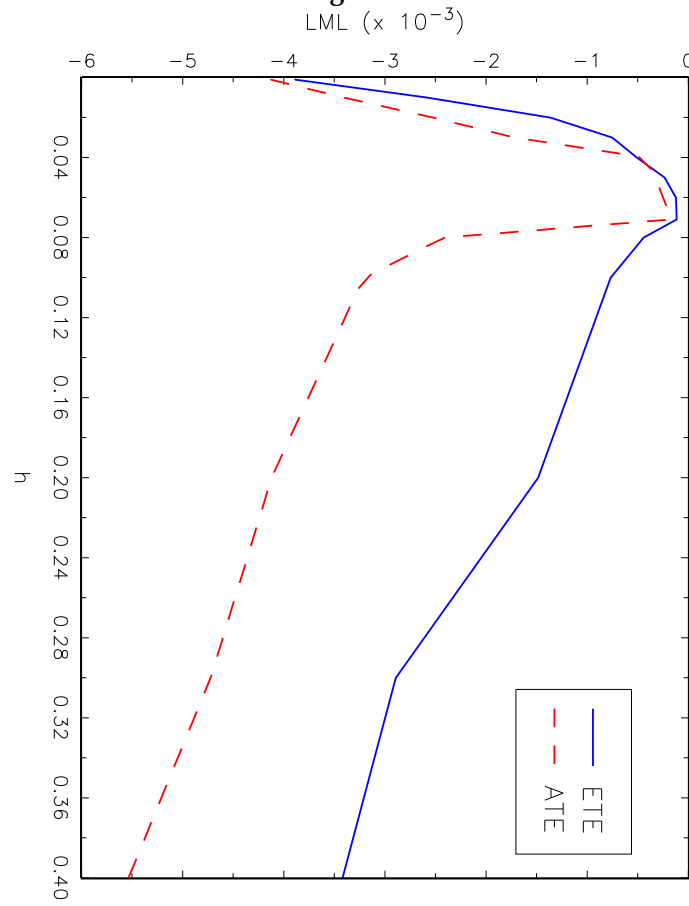
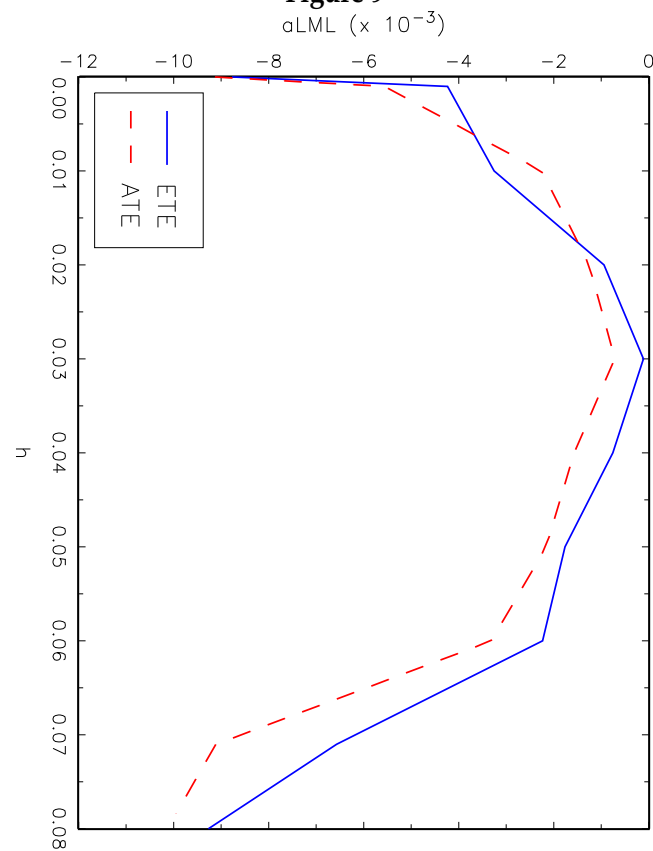
Marginal posterior densities of the parameter h for the 48-min sampling data set are reported in Figure 7.

Figure 7



The marginal posteriors are obtained using the Jeffreys prior and the priors A and B that we described earlier and plotted are also the Jeffreys prior itself for comparison. The marginal posteriors are not very different showing that posterior results about the sampling interval parameter h are not excessively sensitive to the choice of the prior. The ETE returns in the time domain show some sensitivity, which is not, however, replicated by the spectral domain posterior, so it must be attributed to the different, perhaps inferior, behavior of the former. From the spectral domain marginal posteriors of h it turns out that the posterior mean is close to 0.07 and between 0.04 and 0.10 with high posterior probability mass.

In figures 8 and 9 we plot the log marginal likelihood for h , considered as a parameter, for the 48-min and 5-min data sets respectively.

Figure 8**Figure 9**

The results provide an explanation of Figure 3 , since an optimal h for the 48-min sampled data is close to 0.07 ($\approx 1/14$) while for the 5-min sampled data an optimal h is close to 0.03. In view of the actual sampling rate of the original data this is quite reasonable and leads us to believe that the new methods perform quite well.

In Table 4, we report posterior moments from MCMC using conventional discretization methods by treating the missing data as parameters to be estimated following the method of Eraker (2001); see also Elerian, Chib and Shephard (2001). The reported posterior standard deviations were computed using a Newey-West HAC covariance matrix with 10 lags.

Table 4

Table 4. Behavior of conventional MCMC

	N=150,000	N=250,000	N=500,000	N=1,000,000
ETE α	-0.2557 (0.3545)*	-0.1991 (0.3211)*	-0.1743 (0.2782)*	-0.1782 (0.2433)*
H	0.4671 (0.1632)*	0.4578 (0.1742)*	0.4128 (0.1552)*	0.4081 (0.1245)*
ACF	0.977	0.985	0.977	0.970
RNE α	0.0171	0.0196	0.0199	0.0212
RNE H	0.0155	0.0173	0.0255	0.0344
ATE α	-0.2399 (0.051)**	-0.2913 (0.062)**	-0.2619 (0.095)**	-0.2399 (0.082)**
H	0.4301 (0.1747)**	0.3128 (0.1820)**	0.4572 (0.2233)**	0.4255 (0.1942)**
ACF	0.981	0.972	0.931	0.927
RNE α	0.0382	0.0417	0.0530	0.0683
RNE H	0.0172	0.0252	0.0302	0.0704

(*) times 10^{-4} .

(**) times 10^{-3} .

N: is the number of draws after determining the length of burn-in phase using Geweke's (1992) convergence diagnostic. This ranged from 500,000 to 1,000,000 draws. ACF is maximum absolute autocorrelation at lag 50.

RNE is Geweke's (1992) relative numerical efficiency. If IID sampling from the posterior were possibly, RNE should be close to one.

The autocorrelations even at lag 50 are still quite large, and relative numerical efficiency is quite low. Conventional MCMC was run to obtain one million draws which we used to obtain the results in Table 4. This indicates the vast improvement in performance of MCMC which can result from exact discretization and use of the time-domain likelihood function. Moreover, conventional MCMC is reasonably close to the results from time-domain or spectral-domain likelihoods showing that the new methods perform quite well.

Discussion

In this paper, we proposed a methodology for the parametric estimation of the instantaneous rate of asset returns process and assessed the empirical implementation using high frequency stock price data. The assumed data generating process is a mean reverting Ornstein-Uhlenbeck process with fractional noise. The exact discretization delivers an ARIMA(1,1,1) fractional model for log-prices and allows us to use closed-form formulae for the covariance matrix which facilitates greatly the formulation of a Gaussian likelihood / posterior in the time-domain or the frequency-domain. In addition, we derive the exact discrete model using a general non-unit sampling interval h . This permit us to optimize h along with the rest of the parameters. We have introduced a new MCMC scheme and have examined the properties of both time-domain and frequency-domain likelihoods / posteriors through a Monte Carlo application. The new techniques applied to two datasets from the Athens Stock Exchange. Relative to existing MCMC samplers, which use approximate discretization and treat the missing observations of the time path as latent variables, we found that our new

procedures result in vast improvements in terms of the autocorrelation properties of the MCMC scheme and inferences about the Hurst (and other) parameters can be performed reliably and efficiently at a fraction of the time relative to more conventional MCMC schemes. The methods presented in this work can be extended to more general continuous time ARMA(p,q) stochastic processes for the instantaneous rate of return, in which case the resulting exact discretization will yield an ARMA(p, p) model for the observed log-price differences or equivalently an ARIMA(p, 1, p) representation for log prices. Moreover, the exact discretization and estimation methods can be extended to unequal spacing observational intervals. This is interesting from an applications viewpoint when the observations are tick-data.

APPENDIX

Proof of Theorem. Under Assumption 1 the solution of (1) in the interval $((t-1)h, th)$ satisfy

$$r_{th} = e^{ah} r_{(t-1)h} + v + \xi_{th}, \quad (\text{A-1})$$

where $r_{th} = r(th)$,

$$v = c \int_0^h e^{as} ds \quad (\text{A-2})$$

and

$$\xi_{th} = \int_{(t-1)h}^{th} e^{a(th-s)} \varepsilon(s) ds = \int_0^h e^{as} \varepsilon(th-s) ds. \quad (\text{A-3})$$

To obtain the exact discrete model for the observed h-period log-price returns $\int_{(t-1)h}^{th} r(s) ds = \log P(th) - \log P(th-h)$, $t=1,2,\dots, T/h$, we integrate (A-1) over $((t-1)h, th)$ to obtain

$$\int_{(t-1)h}^{th} r(s) ds = e^{ah} \int_{(t-2)h}^{(t-1)h} r(s) ds + vh + \kappa_{th}, \quad (\text{A-4})$$

where κ_{th} is stated in (7) and

$$K_1(r) = \int_0^r e^{as} ds \quad (\text{A-5})$$

$$K_2(r) = \int_r^h e^{as} ds \quad (\text{A-6})$$

The stated formulae (8)-(9) are deduced by (A.5)-(A.6). In view of eq. (7) and the definition

$$\begin{aligned} \gamma_k^F(\tau h) &= E \left[\int_0^h K_1(r) \varepsilon(th-r) dr + \int_0^h K_2(r) \varepsilon(th-h-r) dr \right] \\ &\times \left[\int_0^h K_1(s) \varepsilon((t+\tau)h-s) ds + \int_0^h K_2(s) \varepsilon((t+\tau)h-h-s) ds \right] \end{aligned}$$

we obtain the auto-covariance function (10). Finally, using the spectral transfer function $K(\omega)$, eq. (12), we deduce the continuous time spectral representation of the error term κ_{th}

$$\gamma_k^F(\tau h) = \int_{-\infty}^{\infty} e^{i\omega\tau h} |K(\omega)|^2 f_{\varepsilon}(\omega) d\omega, \quad (\text{A-7})$$

from where (11) follows.

References

- Ait-Sahalia, Y., 1999. Transition densities for interest rate and other nonlinear diffusions. *Journal of Finance* 54,1361–1395.
- Ait-Sahalia, Y., 2002. Maximum likelihood estimation of discretely sampled diffusion: a closed-form approximation approach. *Econometrica* 70, 223–262.
- Ait-Sahalia, Y., Mykland P. A. and L. Zhang, 2005. How often to sample a continuous time process in the presence of market microstructure noise. *Review of Financial Studies* 18, 351–416.
- Brandt, M.W., Santa-Clara, P., 2002. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. *Journal of Financial Economics* 63, 161–210.

- Bergstrom A.R., 1983. Gaussian estimation of structural parameters in higher order continuous time dynamic models, *Econometrica*, 51, 117-152.
- Bergstrom, A.R., 1984, Continuous time stochastic models and issues of aggregation over time, *Handbook of Econometrics*, Volume II, Edited by Z. Griliches and M.D. Intriligator, Elsevier Science Publishers BV.
- Beskos, A., Papaspiliopoulos, O., Roberts, G.O., Fearnhead, P., 2006. Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes. *J. Roy. Statist. Soc. Ser. B. Methodology* 68, 1–29
- Chambers, M.J. 1996. The estimation of continuous parameter long-memory time series models. *Econometric Theory*, 12, 374-390.
- Chambers, M.J., 2004. Testing for unit roots with flow data and varying sampling frequency, *Journal of Econometrics* 119, 1-18.
- Dahlhaus, R. 1989. Efficient parameter estimation for self-similar processes. *Annals of Statistics*, 17, 1749-1766.
- Dahlhaus, R., 2006, Correction note. Efficient parameter estimation for self-similar processes. *Annals of Statistics* 34, 1045-1047.
- Elerian, O., Chib, S., Shephard, N., 2001. Likelihood inference for discretely observed non-linear diffusions. *Econometrica* 69, 959–993
- Eraker, B., 2001. MCMC analysis of diffusion models with application to finance. *Journal of Business and Economic Statistics* 19 (2), 177–191.
- Fox, R., and M.S. Taqqu, 1986. Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics* 14, 157-532.
- Girolami, M. and Calderhead B., 2011. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B*, Volume 73, Issue 2, pages 123-214.
- Golightly, A., Wilkinson, D.J., 2005. Bayesian inference for stochastic kinetic models using a diffusion approximation. *Biometrics* 61 (3), 781–788
- Golightly, A., Wilkinson, D.J., 2008. Bayesian inference for nonlinear multivariate diffusion models observed with error. *Computational Statistics & Data Analysis* 52, 1674-1693.
- Griffin, J.E, and R.C.A. Oomen, 2008. Sampling returns for realized variance calculations: Tick time or transaction time?. *Econometric Reviews* 27(1-3), 230-253.
- Jones, C.S., 1999. Bayesian estimation of continuous-time finance models. Unpublished working paper. University of Rochester.
- Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65, 361–393.
- Lieberman, O., R. Rosemarin and J. Rousseau, 2012. Asymptotic Theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes
- Lillo, F. and Farmer, J.D., 2004. The long memory of efficient markets. *Studies in Nonlinear Dynamics & Econometrics*, vol. 8, 1.
- Lo, A.L. and J. Wang, 1995. Implementing option pricing models when asset returns are predictable, *The Journal of Finance*, Vol. 50, Issue 1, pp. 87-129.
- Lo, A. L and A.C. Mackinlay, 1999, *A non- Random Walk Down Wall Street*. Princeton University Press.
- Bouchaud J. P. , Kockelkoren J. & M. Potters, 2006. Random walks, liquidity molasses and critical response in financial markets, *Quantitative Finance*, 6:02, 115-123.
- Merton, R. C., 1990. Continuous time finance. Blackwell Publishing Ltd.
- Mulligan R., 2004. Fractal analysis of highly volatile markets: an application to technology equities. *The Quarterly Review of Economics and Finance* 44, 155–179.
- Papaspiliopoulos, O., Roberts, G.O., Skold, M., 2003. Non-centered parameterizations for hierarchical models and data augmentation. In: Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A.F.M., West, M. (Eds.), *Bayesian Statistics 7*. Oxford University Press, pp. 307–326.
- Pedersen, A.R., 1995. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics* 22, 55–71.

- Perron, P. Vodoumou, C., 2004. Tests of return predictability: An analysis of their properties based on a continuous time asymptotic framework. *Journal of Empirical Finance* 11, 203-230.
- Roberts, G.O., Stramer, O., 2001. On inference for partially observed nonlinear diffusion models using the Metropolis–Hastings algorithm. *Biometrika* 88 (4), 603–621.
- Samorodnitsky, G., Taqqu, M.S., 1994. *Stable non-Gaussian random processes*. Chapman and Hall.
- Tsai, H. and Chan, K. S., 2005a. Maximum likelihood estimation of linear continuous time long memory processes with discrete time data. *Journal of the Royal Statistical Society, Series B*, 67, Part 5, pp 703-716.
- Tsai, H. and Chan, K. S., 2005b. Quasi-maximum likelihood estimation for a class of continuous-time long memory processes. *Journal of Time Series Analysis*, Vol. 26, No. 5, pp 691-713.
- Tsai, H. and Chan, K. S., 2005c. Temporal aggregation of stationary and nonstationary continuous-time processes. *Scandinavian Journal of Statistic*, 32, pp. 583-597.
- Wachter, J. A., 2002. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, Volume 37, No. 1.
- Zhang L., Mykland P. A. and Ait-Sahalia, 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, Volume 100, No. 472, 1394-1411.
- Zhu, Z. and Taqqu, M.S., 2005. Impact of the sampling rate on the estimation of the parameters of fractional Brownian motion, *Journal of Time Series Analysis*, p. 367-380, Vol.27, No3.



© 2016 by the authors; licensee Preprints.org, MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).