Quantitative Finance
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rquf20

Stochastic volatility models including open, close, high and low prices

Enrique Ter Horst [a b], Abel Rodriguez [c], Henryk Gzyl [b] & German Molina [d]

[a] Euromed Management, Marseille, France

[b] Instituto de Estudios Superiores de Administración, Caracas 1010, Venezuela

[c] Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA

[d] Tudor Capital Europe LLP, London, UK

Published online: 23 Nov 2010.

To cite this article: Enrique Ter Horst, Abel Rodriguez, Henryk Gzyl & German Molina (2012) Stochastic volatility models including open, close, high and low prices, Quantitative Finance, 12:2, 199-212, DOI: 10.1080/14697688.2010.492233

To link to this article: http://dx.doi.org/10.1080/14697688.2010.492233

Quantitative Finance, Vol. 12, No. 2, February 2012, 199–212

Stochastic volatility models including open, close, high and low prices

ENRIQUE TER HORST*†‡, ABEL RODRIGUEZ§, HENRYK GZYL‡ and GERMAN MOLINA¶

†Euromed Management, Marseille, France
‡Instituto de Estudios Superiores de Administración, Caracas 1010, Venezuela
§Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA
¶Tudor Capital Europe LLP, London, UK

(*Received 8 February 2009; in final form 5 May 2010*)

Mounting empirical evidence suggests that the observed extreme prices within a trading period can provide valuable information about the volatility of the process within that period. In this paper we define a class of stochastic volatility models that uses opening and closing prices along with the minimum and maximum prices within a trading period to infer the dynamics underlying the volatility process of asset prices and compare it with similar models presented previously in the literature. The paper also discusses sequential Monte Carlo algorithms to fit this class of models and illustrates its features using both a simulation study and real data.

Keywords: Volatility modelling; Time series analysis; Stochastic volatility; Statistics; Bayesian statistics

1. Introduction

Understanding the volatility of financial assets is an extremely important problem in econometrics and finance (Abramov and Klebaner, 2007), especially in the derivatives world where volatility plays a key role in option pricing. From the early autoregressive conditional heteroskedastic (ARCH) model of Engle (1982) and its generalized version (GARCH) (Bollerslev 1986), to the stochastic volatility approaches (Shephard 2005), the literature has expanded rapidly, and applications now include not only derivatives pricing but also asset management and risk management.

Historically, most statistical and econometric models have used opening and closing prices over the frequency considered for estimation and forecasting, e.g. daily opening and closing prices to estimate daily volatility. However, the last two decades have seen a major change in terms of the quantity and quality of data available; from having access to only daily or weekly closing prices we have now moved into a world where we have access to tick-by-tick/bid–ask data for many asset prices. This increased availability of information has induced new classes of models and algorithms that combine the best possible use of the information available with an efficient way to process it, so that they remain practical and useful for practitioners. Important work has been accomplished in this direction: for example Russell and Engle (1998) use non-equally spaced observations to extract volatility features, Andersen *et al*. (2001) and Barndorff-Nielsen and Shephard (2002) use data at higher frequencies to improve estimation of the volatility and its persistence at lower frequencies, while Hofmann (2002), Rosenbaum (2007) and Gatheral and Oomen (2010) explore the effects that microstructure noise and long-range dependence can have asymptotically and in finite samples on the volatility dynamics. However, although these papers have focused on the use of larger amounts of the information available to achieve the common goal of better fitting and forecasting, they still tend to focus their efforts on closing prices at the highest frequency considered and ignore the paths between those points, even if this information is available.

Closing prices, especially at the lower frequencies (daily or less), have become less and less reliable to ascertain realized volatilities, given the patterns recently seen in volumes. For example, stocks in the S&P500 index show unusually large volumes of trades in the last 15 minutes

*Corresponding author. Email: enriqueterhorst@gmail.com

Quantitative Finance
ISSN 1469–7688 print/ISSN 1469–7696 online © 2012 Taylor & Francis
http://www.tandfonline.com
http://dx.doi.org/10.1080/14697688.2010.492233

of sessions, which have been linked to several factors such as high-frequency funds unwinding positions accumulated during the day, or exchange-traded funds and hedgers operating in the last minutes to adjust their positions. Also, volatility can be very large and still end up with closing prices not far from where they started, while big swings have happened throughout the day. This is especially the case during periods of low liquidity, where the average holding period of any position diminishes significantly, and people react more violently to moves. Stop-loss orders can exacerbate those intra-day swings, while squaring of daily positions for intra-day funds can produce the opposite effect. This is a reflection of higher realized volatility and lower liquidity, but it will not be captured by models that only use close-to-close returns.

In this paper, we focus on stochastic volatility models that, in addition to opening and closing prices over the frequency chosen, incorporate information on the highest and lowest infra-period prices. Data on minimum and maximum infra-period prices is available for most series through the common data providers (Datastream, Reuters, Bloomberg), even if other high-frequency information is not available. The relevance of close, high, low and open (CHLO) data, that is the relevance of the path followed by the series rather than the start and end points, is well known. Financial newspapers include it for the frequency reported (daily, weekly, monthly or yearly), data providers add it to their data series, and practitioners in technical analysis have been using it in the past to model volatility by using estimators like ranges or average true ranges (Wilder 1978). Chartists have also been using CHLO data as a key source of information from a graphical/visual standpoint to identify price patterns, trends or reversals, such as, for example, the well-known DeMark indicators (Perl 2008). By assuming that log-prices follow a log-normal diffusion, Parkinson (1980) showed that high/low prices provide a highly efficient estimator of the variance compared with estimators based solely on open/close prices. By incorporating information from CHLO prices, Garman and Klass (1980) extended the estimator of Parkinson and gained a significant amount of efficiency compared with only including open/close prices. Ball and Tourus (1984) derived a maximum likelihood analogue of the estimator of Parkinson. In the context of time series modeling of asset return volatility, Gallant *et al.* (1999) and Alizadeh *et al.* (2002) take one very relevant step forward by using the range of observed prices (rather than the maximum and minimum directly) to estimate stochastic volatility models. Brandt and Jones (2005) extend the work of Alizadeh *et al.* (2002) by including the range and the closing price in the estimation of a stochastic volatility model, but again ignore the actual levels of the maximum and minimum prices. Full CHLO prices have been used by a number of authors, including Rogers and Satchell (1991), Rogers *et al.* (1994) and Rogers (1998). In particular, Magdon and Atiya (2003) derives a maximum likelihood estimator assuming constant volatility, obtaining better performance than existing previous methods on simulated data. Their method, however, is not integrated into a GARCH or stochastic volatility framework, something done by Lildholdt (2002) using maximum likelihood estimation.

The models we discuss in this paper combine the use of CHLO data with a Bayesian approach to estimation and prediction, all in a stochastic volatility framework. Inference is performed using particle filters, which are computationally quick, relatively simple to implement and can be applied to any frequency of data. This class of algorithms should be very appealing to practitioners handling large amounts of data at the higher frequencies or for processing large numbers of series. Section 2 provides a quick introduction to the theory behind stochastic volatility models and to the notation used throughout the paper. Sections 3 and 4 present the analytical framework for the joint density of the CHLO prices, the elicitation of prior densities and a description of the particle filtering algorithm used. Sections 5 and 6 show how to deal with problems of missing data, and apply our model to weekly CHLO data of the S&P. We also provide a comparison with the stochastic volatility models using only closing prices data and show how different the results can be once the information contained in the observed extremes is added. Section 7 concludes with a summary and a description of potential applications and extensions.

In summary, our net contribution is threefold. First, we provide a coherent model that links the traditional stochastic volatility model with the CHLO data without the need of assumptions beyond those used in that traditional stochastic volatility model. Second, we introduce the use of particle filters in the context of CHLO models, which provides a quick and simple algorithm that allows for fast on-line estimation. Finally, we demonstrate that the use of the information contained in the open, maximum, minimum and close (instead of only the range) can improve estimation in stochastic volatility models.

2. Stochastic volatility models

Because of its simplicity and tractability, geometric Brownian motion is by far the most popular model to explain the evolution of the price of financial assets, and has a long history dating back to Louis Bachelier (Courtault *et al.* 2000). A stochastic process $\{S_t : t \in \mathbb{R}^+\}$ is said to follow a geometric Brownian motion (GBM) if it is the solution of the stochastic differential equation

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + S_t \sigma \,\mathrm{d}B_t, \tag{1}$$

where B_t is a standard Wiener process and μ and σ are, respectively, the instantaneous drift and instantaneous volatility of the process. GBM implies that the increments of $Y_t = \log S_t$ over intervals of constant length Δ are independent, stationary and identically distributed, i.e. $Y_{t+\Delta} - Y_t = \log S_{t+\Delta} - \log S_t \sim \mathbb{N}(\Delta\mu, \Delta\sigma^2)$, or, equivalently,

$$S_{t+\Delta} = S_t \exp\{\Delta\mu + \sigma(B_{t+\Delta} - B_t)\}.$$

By construction, GMB models assume that the volatility of returns is constant. However, empirical evidence going back at least to Mandelbrot (1963), Fama (1965) and Officer (1973) demonstrates that the price volatility of financial assets tends to change over time and therefore the simple model in (1) is generally too restrictive. The GBM model can be generalized by assuming that both the price and volatility processes follow general diffusions (Hull and White 1987),

$$\begin{aligned} \mathrm{d}S_t &= \mu(S_t, \sigma_t)\mathrm{d}t + \nu(S_t, \sigma_t)\mathrm{d}B_{t,1}, \\ \mathrm{d}\sigma_t &= \alpha(S_t, \sigma_t)\mathrm{d}t + \beta(S_t, \sigma_t)\mathrm{d}B_{t,2}. \end{aligned}$$

One particularly simple (and popular) version of this approach assumes that, just as before, the price follows a GBM with time-varying drift, and that the log-volatility follows an Ornstein–Uhlenbeck (OU) process,

$$\mathrm{d}Y_t = \mu\,\mathrm{d}t + \sigma_t\,\mathrm{d}B_{t,1}, \tag{2}$$

$$\mathrm{d}\log(\sigma_t) = \kappa(\theta - \log(\sigma_t)) + \tau\,\mathrm{d}B_{t,2}. \tag{3}$$

Practical implementation of these models typically relies on a discretization of the continuous time model in (2) and (3). For the remainder of the paper, we assume that the drift and volatility are scaled so that $\Delta = 1$ corresponds to one trading period (e.g., day or week) and we focus attention on the state-space model

$$Y_t = \mu + Y_{t-1} + \epsilon_t^1, \quad \epsilon_t^1 \sim \mathsf{N}(0, \sigma_t^2), \tag{4}$$

$$\log(\sigma_t) = \alpha + \phi[\log(\sigma_{t-1}) - \alpha] + \epsilon_t^2, \quad \epsilon_t^2 \sim \mathsf{N}(0, \tau^2), \tag{5}$$

where $\alpha = \kappa\theta$ and $\phi = 1 - \kappa$. It is common to assume that $0 \leq \phi < 1$ and $\log(\sigma_0) \sim \mathsf{N}(\alpha, \tau^2/(1-\phi^2))$, so that volatilities are positively correlated and the volatility process is stationary (which ensures that the process for the prices is a martingale) with α determining the median of the long-term volatility, $\nu = \exp\{\alpha\}$. Therefore, we expect the volatility of returns to take values greater than ν half of the time, and *vice versa*.

Unlike ARCH and GARCH models (Engle 1982, Bollerslev 1986), where a single stochastic process controls both the evolution of the volatility and the observed returns, stochastic volatility models use two coupled processes to explain the variability of the returns. By incorporating dependence between ϵ_t^1 and ϵ_t^2, the model can accommodate leverage effects, while additional flexibility can be obtained by considering more general processes for the volatility (for example, jump process, and linear or nonlinear regression terms).

Although theoretical work on stochastic volatility models goes back at least to the early 80s, practical application was limited by computational issues. Bayesian fitting of stochastic volatility models has been discussed by various authors, including Jacquier *et al.* (1994) and Kim *et al.* (1998). Popular approaches include Gibbs sampling schemes that directly sample from the full conditional distribution of each parameter in the model, algorithms based on offset mixture representations that allow for joint sampling of the sequence of volatility parameters, and particle filter algorithms. In the sequel, we concentrate on sequential Monte Carlo algorithms for the implementation of stochastic volatility models.

3. Incorporating extreme values in stochastic volatility models

3.1. Joint density for the closing price and the observed extremes of geometric Brownian motion

The goal of this section is to extend the stochastic volatility model described in the previous section to incorporate information on the full CHLO prices. To do so, note that the Euler approximation in (4) implies that, conditionally on the volatility σ_t, the distribution of the increments of the asset price follows a geometric Brownian motion with constant volatility σ_t during the tth trading period. Therefore, the discretization allows us to interpret the process generating the asset prices as a sequence of conditionally independent processes defined over disjoint and adjacent time periods; within each of these periods the price process behaves like a GBM with constant volatility, but the volatility is allowed to change from period to period. This interpretation of the discretized process is extremely useful, as it allows us to derive the joint distribution for the closing, maximum and minimum price of the asset within a trading period using standard stochastic calculus tools.

Theorem 3.1: *Let Y_t be a Brownian motion with drift and consider the evolution of the process over a time interval of unit length where Y_{t-1} and Y_t are the values of the process at the beginning and end of the period, and let $M_t = \sup_{t-1 \leq s \leq t}\{Y_s\}$ and $m_t = \inf_{t-1 \leq s \leq t}\{Y_s\}$ be, respectively, the supremum and the infimum values of the process over the period. If we denote by μ and σ_t the drift and volatility of the process between $t-1$ and t (which are assumed to be fixed within this period), the joint distribution of M_t, m_t and Y_t conditional on $Y_{t-1} = y_{t-1}$ is given by*

$$\begin{aligned} &\Pr(m_t \geq a_t, M_t \leq b_t, Y_t \leq y_t \mid Y_{t-1} = y_{t-1}) \\ &\quad = \int_{-\infty}^{y_t} q(z, a_t, b_t \mid y_{t-1})\mathrm{d}z, \end{aligned}$$

for $m_t \leq \min\{y_t, y_{t-1}\}$, $M_t \geq \max\{y_t, y_{t-1}\}$ and $m_t \leq M_t$, where

$$\begin{aligned} q(z, a_t, b_t \mid y_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left\{-\frac{[\mu^2 - 2\mu(z - y_{t-1})]}{2\sigma_t^2}\right\} \\ \times \sum_{n=-\infty}^{\infty} (\exp\{-d_{1,n}(z)\} - \exp\{-d_{2,n}(z)\}) \end{aligned} \tag{6}$$

and

$$\begin{aligned} d_{1,n}(z) &= \frac{[z - y_{t-1} - 2n(b_t - a_t)]^2}{2\sigma_t^2}, \\ d_{2,n}(z) &= \frac{[z + y_{t-1} - 2a_t + 2n(b_t - a_t)]^2}{2\sigma_t^2}. \end{aligned}$$

The proof of this theorem, which is a simple extension of the results of Freedman (1971), Dym *et al.* (1985) and Klebaner (2005), can be seen in appendix A. To understand why the summation arises in expression (6), it is useful to go back to the original proof of Freedman (1971). The summation arises by considering a series of events corresponding to intersections of increasingly larger number events by the use of the reflexion and the inclusion/exclusion principles. An equivalent but more involved expression was obtained by Lildholdt (2002) by solving the Fokker–Planck differential equation associated with (1) subject to the appropriate boundary conditions, and use it to construct GARCH models that incorporate information on CHLO prices. Note that if we take both $a_t \to -\infty$ and $b_t \to \infty$, the cumulative distribution in (6) reduces to the integral of a Gaussian density with mean $y_{t-1}+\mu$ and variance σ_t^2, which agrees with (4). To simplify the expressions that follow, from now on we slightly abuse notation by employing $p(y \mid x)$ to denote a generic density of random variable Y conditional on X; the specific meaning is provided by the arguments x and y on each specific instance.

Corollary 3.2: *The joint density for* m_t, M_t, $Y_t \mid Y_{t-1} = y_{t-1}$ *is given by*

$$\begin{aligned} p(a_t,b_t,y_t \mid y_{t-1}) &= -\frac{\partial^2 q}{\partial a_t\,\partial b_t} q(y_t,a_t,b_t \mid y_{t-1}) \\ &= \frac{1}{\sqrt{2\pi}\sigma_t^3}\exp\left\{-\frac{\mu^2-2\mu(y_t-y_{t-1})}{2\sigma_t^2}\right\} \\ &\quad \times \sum_{n=-\infty}^{\infty}[4n^2(2d_{1,n}(y_t)-1)\exp\{-d_{1,n}(y_t)\} \\ &\quad -4n(n+1)(2d_{2,n}(y_t)-1))\exp\{-d_{2,n}(y_t)\}], \end{aligned} \tag{7}$$

for $m_t \le \min\{y_t, y_{t-1}\}$, $M_t \ge \max\{y_t, y_{t-1}\}$ *and* $m_t \le M_t$, *and zero otherwise.*

Equation (7) provides the likelihood function for the closing, maximum and minimum prices given the volatility and drift, under the first-order Euler approximation to the system of stochastic differential equations in (2) and (3). Therefore, the basic underlying assumptions about the behavior of asset prices are the same as in standard volatility models; however, by employing (7) instead of (4) we are able to coherently incorporate information about the observed price extremes in the inference of the price process. After a simple transformation $r_t = b_t - a_t$ and $w_t = a_t$ and marginalization over w_t, we recover the likelihood function described by Brandt and Jones (2005),

$$\begin{aligned} p(r_t,y_t \mid y_{t-1}) &= \frac{1}{\sqrt{2\pi}\sigma_t}\exp\left\{-\frac{(y_t-y_{t-1}-\mu)^2}{2\sigma_t^2}\right\} \\ &\quad \times \sum_{n=-\infty}^{\infty}\left[\frac{4n^2}{\sqrt{2\pi}\sigma_t^2}\left\{\frac{(2nr_t-|y_t-y_{t-1}|)^2}{\sigma_t^2}-1\right\}\right. \\ &\quad \times\exp\left\{\frac{(2nr_t-|y_t-y_{t-1}|)^2}{2\sigma_t^2}\right\} \\ &\quad +\frac{2n(n-1)}{\sqrt{2\pi}\sigma_t^2}(2nr_t-|y_t-y_{t-1}|) \\ &\quad \times\exp\left\{\frac{(2nr_t-|y_t-y_{t-1}|)^2}{2\sigma_t^2}\right\} \\ &\quad +\frac{2n(n-1)}{\sqrt{2\pi}\sigma_t^2}(2(n-1)r_t+|y_t-y_{t-1}|) \\ &\quad \left.\times\exp\left\{\frac{(2(n-1)r_t+|y_t-y_{t-1}|)^2}{2\sigma_t^2}\right\}\right], \end{aligned} \tag{8}$$

for $r_t > |y_t - y_{t-1}|$, while a further marginalization over the closing price y_t yields the (exact) likelihood underlying the range-based model of Alizadeh *et al.* (2002), which is independent of the opening price y_{t-1},

$$p(r_t \mid y_{t-1}) = p(r_t) = 8\sum_{n=1}^{\infty}(-1)^{n-1}\frac{n^2 r_t}{\sqrt{2\pi}\sigma_t}\exp\left\{-\frac{n^2 r_t^2}{2\sigma_t^2}\right\}. \tag{9}$$

Using either (8) or (9) as likelihoods entails a loss of information with respect to the full joint likelihood in (7). In the case of (9), the effect is clear as the range is an ancillary statistic for the drift of the diffusion, and therefore provides no information about it. This leads Alizadeh *et al.* (2002) to assume that the drift is zero, which has little impact on the estimation of the volatility of the process, but might have important consequences for other applications of the model. In the case of (8), although there is information about the drift contained in the opening and closing prices, the model ignores the additional information about the drift contained in the actual levels of the extremes.

In order to emphasize the importance of the information provided by the minimum and maximum returns, we present in figure 1 plots of the likelihood function for the first observation of the S&P500 dataset discussed in section 6.2 (the week ending on April 21, 1997). When only the closing price is available, the likelihood provides information about the drift of the process, but not about the volatility (note that, in this case, the likelihood is unbounded in a neighborhood of $\sigma_t = 0$, and therefore the maximum likelihood estimator does not exist). Therefore, information about the volatility in this type of model is obtained solely through the evolution of prices, and is therefore strongly influenced by the underlying smoothing process. In other words, the volatility parameters are only weakly identifiable, with the identifiability begin provided by the autoregressive prior in (5). In practice, this means that formally comparing alternative models for the volatility is extremely difficult. However, when the maximum and minimum are included in the analysis, the likelihood for a single time period does provide information about the volatility in that period, which can greatly enhance our ability to infer and test volatility models.

Although dealing with an infinite sum can seem troublesome, our experience suggests that a small number of terms (usually less than 20) suffice to provide an accurate approximation. In addition, we note that since the general term of the sum is strictly decreasing for

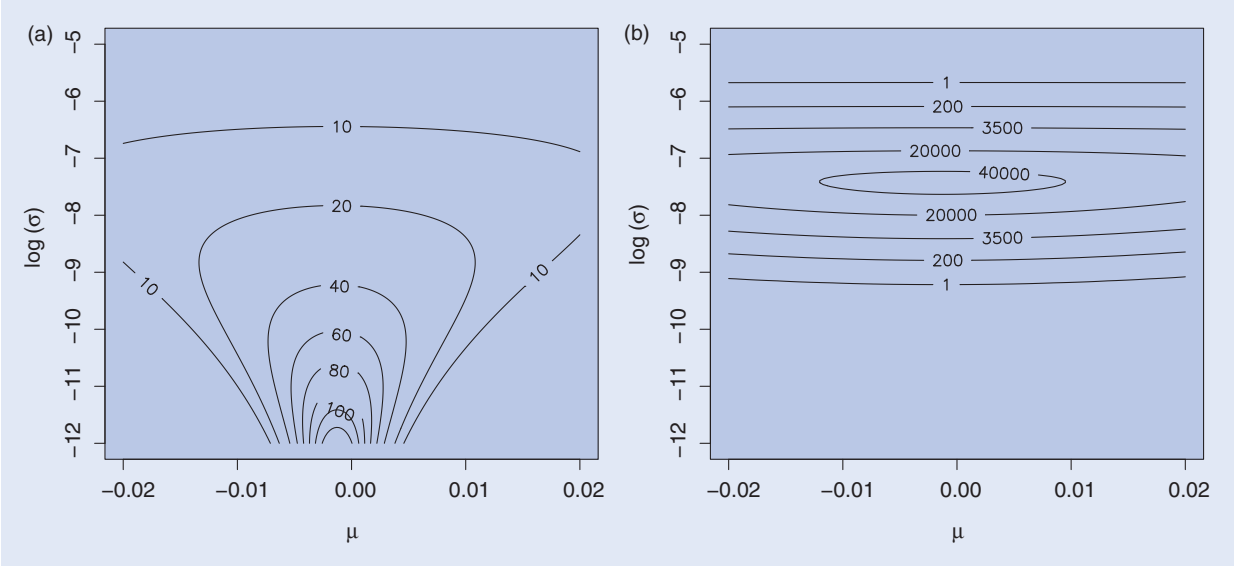

Figure 1. Likelihood functions for a single observation. The left panel corresponds to the Gaussian likelihood obtained solely from the closing price (equation (4)), while the right panel corresponds to the joint likelihood (equation (7)).

all the likelihoods discussed above, it is easy to implement an adaptive scheme that stops adding terms once the change in the value is smaller than a given tolerance.

Note that equation (6) assumes that the log closing price at period $t-1$, y_{t-1}, is the same as the log opening price at period t. However, in some cases these two prices can differ; for example, wars or unexpected news can happen during holidays whenever markets are closed, provoking jumps that are not governed by the diffusion processes. As with other stochastic volatility models, we can easily include this type of 'weekend effect' by realizing that equations (6) and (7) use y_{t-1} simply as the initial value of the GBM; therefore, they remain valid if we substitute y_{t-1} by the opening price at time t.

3.2. *Variance evolution and prior specification*

The previous discussion focused on the characteristics of the likelihood function for the discretized process *conditional* on its volatility. The full specification of the model also requires that we define the evolution of the volatility in time. For illustrative purposes, this paper focuses on the simple autoregressive process for the log volatility described in section 2,

$$\begin{aligned} &\log(\sigma_t) \mid \log(\sigma_{t-1}) \sim \mathsf{N}(\alpha + \phi[\log(\sigma_{t-1}) - \alpha], \tau^2), \\ &\log(\sigma_0) \sim \mathsf{N}\left(\alpha, \frac{\tau^2}{1-\phi^2}\right). \end{aligned} \tag{10}$$

The model is completed by introducing priors for the unknown structural parameters in the models. Following standard practice, we assume that

$$\begin{aligned} &\mu \sim \mathsf{N}(d_\mu, D_\mu), \quad \alpha \sim \mathsf{N}(d_\alpha, D_\alpha), \\ &\phi \sim \mathsf{Beta}(q_\phi, r_\phi), \quad \tau^2 \sim \mathsf{IG}(u_\tau, v_\tau), \end{aligned} \tag{11}$$

where by $\mathsf{IG}(u_\tau, v_\tau)$ we mean an inverse gamma distribution with parameters u_τ and v_τ, which has the density

$$p(\tau \mid u_\tau, v_\tau) = \frac{v_\tau^{u_\tau}}{\Gamma(u_\tau)} \frac{1}{\tau^{u_\tau+1}} \exp\left\{-\frac{v_\tau}{\tau}\right\},$$

i.e. $1/\tau$ has a Gamma distribution with parameters u_τ and v_τ (Gelman *et al*. 2003). This choice of priors ensures that all parameters have the right support; in particular, the beta prior for ϕ ensures that $0 \leq \phi \leq 1$ (remember our discussion in section 2). Also, the choice of proper priors enables us to implement a particle filter for computation. The eight hyperparameters d_μ, D_μ, d_α, D_α, q_ϕ, r_ϕ, u_τ and v_τ have to be chosen according to the available prior information about the problem at hand. For example, it is natural to choose d_μ to be close to the market risk-free rate, while choosing d_α close to the logarithm of the long-term average volatility for the asset. In most practical applications, enough prior information is available so that eliciting these hyperparameters is relatively straightforward.

4. Inference using particle filters

The use of simulation algorithms to explore the posterior distribution of complex Bayesian hierarchical models has become popular in the last 20 years. In particular, Markov chain Monte Carlo (MCMC) algorithms, which generate a sequence of dependent samples from the posterior distribution of interest, have become ubiquitous. For inference in nonlinear state space models, sequential Monte Carlo algorithms, and in particular particle filters, have become a standard tool. Particle filters use a finite discrete approximation to represent the distribution of the state parameters at time t given the observations up to t, which in our case reduces to $p(\sigma_t \mid \{y_l, b_l, a_l\}_{l=1}^t))$, and sequentially updates it to obtain $p(\sigma_{t+1} \mid \{y_l, b_l, a_l\}_{l=1}^{t+1})$. As with other

Monte Carlo approaches to inference, the resulting samples can be used to obtain point and interval estimates, as well as to test hypotheses of interest. However, unlike MCMC algorithms, there is no need to check for convergence of the algorithm, study its mixing properties, or devise proposal distributions. Doucet *et al.* (2001) and Cappe *et al.* (2007) provide excellent reviews of sequential Monte Carlo methods.

Most sequential Monte Carlo algorithms are unable to handle structural parameters that do not evolve in time, and assume them fixed and known in advance. However, in our stochastic volatility model the structural parameters μ, α, ϕ and τ^2 are unknown and need to be estimated from the data. Designing particle filter algorithms that also allow for sequential structural learning in the case of CHLO models is extremely challenging because almost all algorithms that exploit conditional linearity and normality (such as those of Storvik 2002, Polson *et al.* 2008 and Carvalho *et al.* 2009) are not applicable because of the nonlinear structure in (7). Therefore, in the sequel we concentrate on a version of the auxiliary particle filter (Pitt and Shephard 1999) developed by Liu and West (2001). Their algorithm introduces an artificial Gaussian perturbation in the structural parameters and applies an auxiliary particle filter to the modified problem. In order to correct for the information loss generated by the artificial perturbation, the authors introduce a shrunk kernel approximation constructed in such a way as to preserve the mean and covariance of the distribution of the structural parameters. This kernel density approximation usually works well in practice, producing accurate reconstructions of the posterior distribution of both structural and state parameters while avoiding loss of information that plagues the self-organizing state space models of Kitagawa (1998).

Since the artificial evolution is assumed to follow a Gaussian distribution, it is typically necessary to transform the structural parameters so that the support of their distribution is the whole real line. Therefore, we describe our algorithm in terms of the transformed structural parameter $\eta = (\eta_1, \eta_2, \eta_3, \eta_4) = (\mu, \alpha, \text{logit}(\phi), \log(\tau^2)) \in \mathbb{R}^4$. In the sequel, we denote by $\sigma_t^{(j)}$ for $j = 1, \ldots, N$ the particles associated with σ_t, while we slightly abuse notation and write $\eta_t^{(j)}$ for $j = 1, \ldots, N$ to denote the particles used to represent the distribution of η conditional on the information available up to time t. After choosing a discount factor $0.5 < \epsilon < 1$ (controlling both the size of the perturbation and the level of shrinkage in the density estimator) and generating a sample of N particles from the prior distributions in (11), the algorithm proceeds by repeating the following steps for $t = 1, \ldots, T$.

(1) For each particle $j = 1, \ldots, N$, identify prior point estimates $(z_{t+1}^{(j)}, m_{t+1}^{j})$ for the joint vector of state and structural parameters $(\sigma_{t+1}^{(j)}, \eta_{t+1}^{(j)})$ such that

$$z_{t+1}^{(j)} = \exp\{\eta_{2,t}^{(j)} + \text{expit}(\eta_{3,t}^{(j)})[\log(\sigma_t^{(j)}) - \eta_{2,t}^{(j)}]\},$$

$$m_{t+1}^{(j)} = a\eta_t^{(j)} + (1-a)\bar{\eta}_t,$$

where $a = (3\epsilon - 1)/(2\epsilon)$, $\eta_t^{(j)} = (\eta_{1,t}^{(j)}, \eta_{2,t}^{(j)}, \eta_{3,t}^{(j)}, \eta_{4,t}^{(j)}) = (\mu_t^{(j)}, \alpha_t^{(j)}, (\phi)_t^{(j)}, \log(\tau^2)_t^{(j)})$ and $\bar{\eta}_t = \sum_{j=1}^{N} \eta_t^{(j)}$.

(2) Sample auxiliary indicators $\xi_t^{(1)}, \ldots, \xi_t^{(N)}$, each one also taking values in the set $\{1, \ldots, N\}$, so that

$$\Pr(\xi_t^{(j)} = k) = p(y_{t+1}, b_{t+1}, a_{t+1} \mid y_t, \mu = m_{1,t+1}^{(k)}, \sigma = z_{t+1}^{(k)}),$$

where $p(y_{t+1}, b_{t+1}, a_{t+1} \mid y_t, \mu, \sigma_t^2)$ is given in (7).

(3) Generate a set of new structural parameters $\eta_{t+1}^{(j)}$ by sampling

$$\eta_{t+1}^{(j)} \sim \mathsf{N}(m_{t+1}^{\xi_t^{(j)}}, (1-a^2)V_t),$$

where $V_t = \sum_{j=1}^{N} (\eta_t^{(j)} - \bar{\eta}_t)'(\eta_t^{(j)} - \bar{\eta}_t)/N$.

(4) Sample a value of the current state vector

$$\log(\sigma_{t+1}^{(j)}) \sim \mathsf{N}(\eta_{2,t+1}^{(j)} + \text{expit}(\eta_{3,t+1}^{(j)})[\log(\sigma_t^{(\xi_t^{(j)})}) - \eta_{2,t+1}^{(j)}], \exp(\eta_{4,t+1}^{(j)})).$$

(5) Resample the particles according to the probabilities

$$\omega_{t+1}^{(j)} \propto \frac{p(y_{t+1}, b_{t+1}, a_{t+1} \mid y_t, \mu = \eta_{1,t+1}^{(j)}, \sigma = \sigma_{t+1}^{(j)})}{p(y_{t+1}, b_{t+1}, a_{t+1} \mid y_t, \mu = m_{1,t+1}^{(j)}, \sigma = z_{t+1}^{(j)})}.$$

Although particle filters are not iterative algorithms and therefore mixing and convergence are not issues, sequential Monte Carlo algorithms might suffer from particle impoverishment. Particle impoverishment happens when the particle approximation at time t differs significantly from the approximation at time $t+1$, leading to a small number of particles receiving most of the posterior weight; in the worst case, a single particle receives all the posterior weight. A similar issue arises in importance sampling algorithms, where efficiency decreases as the distribution of the importance weights becomes less uniform. In the sequential Monte Carlo literature it is common to use the effective sample size (ESS) to monitor particle impoverishment,

$$ESS_t = \frac{N}{1 + (\mathsf{V}(\omega_t)/[\mathsf{E}(\omega_t)]^2)}.$$

Values of the ESS close to N point to well-behaved samplers with little particle impoverishment, while small values of N usually indicate uneven weights and particle representations that might be missing relevant regions of the parameter space.

5. Missing data

In some instances (for example, non-exchange traded assets), data on the observed extremes might be unavailable or unreliable for some trading periods. When MCMC algorithms are used for inference, their structure can be easily exploited to deal with this type of situation by adding an additional sampling step in which the missing or unreliable data is imputed conditionally on the current value of the parameters (Gelman *et al.* 2003). However, this type of iterative procedure is not available in particle filter algorithms such as the one described in section 4. This means that implementation under missing

data requires that we compute the corresponding marginal distributions from the joint density derived from (6), which are to be used in steps (2) and (5) of the particle filter instead of (6).

As discussed in section 3, the marginal density of the log closing price is simply a normal distribution with mean μ and variance σ_t^2. Therefore, if both the maximum and minimum are missing, the likelihood for the period simply becomes the likelihood for a standard stochastic volatility model in (4). When the minimum is not available, the marginal density for the maximum and the closing price is simply given by

$$q(y_t, b_t \mid y_{t-1}) = \frac{2(2b_t - (y_t + y_{t-1}))}{\sqrt{2\pi\sigma_t^6 t^3}} \times \exp\left\{-\frac{(2b_t - (y_t + y_{t-1}))^2}{2\sigma_t^2 t} + \frac{\mu(y_t - y_{t-1})}{\sigma_t^2} - \frac{\mu^2 t}{2\sigma_t^2}\right\},$$

while if the maximum is missing, the marginal density for the minimum and the closing is given by

$$q(y_t, a_t \mid y_{t-1}) = \frac{2(y_t + y_{t-1} - 2a_t)}{\sqrt{2\pi\sigma_t^6 t^3}} \times \exp\left\{-\frac{(y_t + y_{t-1} - 2a_t)^2}{2\sigma_t^2 t} + \frac{\mu(y_t - y_{t-1})}{\sigma_t^2} - \frac{\mu^2 t}{2\sigma_t^2}\right\}.$$

The density of the maximum and minimum can also be computed by integrating equation (7) from a to b with respect to the variable y_t. Proof of these results is given by Dana *et al.* (2007), or can be obtained directly by computing the corresponding limits in expression (7). For a thorough review on missing data in general, see Dempster *et al.* (1997), Hu *et al.* (1998) and Scheffer (2002)

6. Illustrations

6.1. Simulation study

In this section we use a simulation study to compare the performance of four stochastic volatility models: STSV, which uses the Gaussian likelihood in (1) and therefore employs solely the information contained in the opening and closing prices; RASV, which uses the likelihood (9) and represents a Bayesian version of the range-only model described by Alizadeh *et al.* (2002); RCSV, which uses the likelihood in (8) and therefore corresponds to the model based on the range and the opening/closing prices described by Brandt and Jones (2005); and EXSV, our proposed model using the full likelihood (7) and employing all the information contained in the CHLO prices.

Our simulation study uses 1000 random samples, each comprising 156 periods of returns generated under the stochastic volatility model described in equations (4) and (5). The setup of the simulation attempts to reproduce values that are representative of the weekly volatility process associated with the S&P500 index. First, we generate the sequence of volatilities using the Ornstein–Uhlenbeck model in (5), with parameters $\alpha = -3.75$, $\phi = 0.95$ and $\tau = 0.11$. Then, for each period, we generate a sample path for the geometric Brownian motion in (4) over a grid with 1000 nodes using the corresponding value for the volatility, assuming that $\mu = 0.000961$ (for weekly returns, this corresponds approximately to an average 5% annual return). The maximum, minimum, opening and closing prices over the period were computed from this discretized sample path. We assume that the value of the assets at the beginning of each simulation S_0 is \$100 and compute both the root mean squared deviation (RMSD) of the estimated volatility with respect to their true value, and their median absolute deviation (MAD).

With very minor modifications, the particle filter algorithm described in section 4 was used to fit all four models, and was implemented in MATLAB using 30,000 particles. Unless noted otherwise, the results we report below correspond to prior hyperparameters $d_\mu = 0$, $D_\mu = 0.0001$, $q_\phi = 1$ and $r_\phi = 1$, $d_\alpha = -3.75$, $D_\alpha = 0.025$ and $u_\tau = 6$, $v_\tau = 0.06$, so that the prior distributions for μ, α and τ are centered around the true parameter values for all models and the prior for ϕ is uniform. In addition, a subsample of 20 equally spaced observations from the volatility path is retained and used to compute a realized-volatility estimate by computing the sum of squares of squared infra-period returns (labeled REAZ in table 1). This scenario reflects a typical scenario in highly liquid markets where some high-frequency data is available but a large number of transactions are carried out between high-frequency observations.

Table 1 compares the performance of the different estimation approaches in this simulation scenario. The first number in each cell corresponds to the median ratio of the pertinent performance measure across all 1000 simulations, while the numbers in parentheses correspond to the 5% and 95% quantiles, respectively. Ratios close to one indicate that the two models in question have similar performances, while ratios much larger than one indicate that the first model has a much larger RMSD (or MAD) than the second, and therefore performs worse.

In agreement with the results reported by Alizadeh *et al.* (2002), the first row of table 1 shows that using the range alone as a volatility proxy produces consistently more accurate volatility estimates than using opening and closing prices alone. Similarly, the second row suggests that, although including the opening prices tends to

Table 1. Results from our simulation example. We show the median along with the 5% and 95% quantiles (in parentheses) of the ratio between deviation measures for three pairs of models, computed over a total of 1000 simulated data sets.

Model	RMSD	MAD
STSV/RASV	1.45 (1.20, 1.68)	1.40 (1.02, 1.82)
RASV/RCSV	1.03 (0.89, 1.12)	1.01 (0.84, 1.16)
RCSV/EXSV	1.06 (0.93, 1.23)	1.11 (0.97, 1.24)
RASV/EXSV	1.09 (0.99, 1.22)	1.10 (0.94, 1.28)
EXSV/REAZ	1.19 (1.11, 1.29)	1.17 (1.10, 1.26)

improve the volatility estimates over those obtained from the range-based model most of the time, it can sometimes decrease accuracy (especially when MAD is used to measure the accuracy of the reconstruction). A similar situation, although much less severe, happens with our model based on the full likelihood; our model tends to improve over RASV and RCSV most of the time (with an average efficiency gain in the range of 5–10%), but can underperform in some data sets. Finally, the last line of the table suggests that, as expected, estimators based on realized volatility improve over estimators based solely on open/close prices. However, if only a relatively small number of infra-period prices are available, CHLO estimators seem to outperform realized volatility estimators. Formal t-tests on the mean value of the ratios confirm that all the differences discussed above are statistically significant at the 95% confidence level.

Although the impact of the full information on the estimation of the volatility is moderate, including the full information contained in CHLO data can greatly improve the estimation of the drift of the model. For example, the ratio of the absolute error in the estimate of the drift μ based on the 156 time points between RASV and EXSV has a median of 1.11, with the 5% and 95% quantiles being 0.91 and 1.47 over the 1000 simulations. This indicates an average gain in efficiency of about 11%. We also note that EXSV tends to be more robust to prior misspecification. Sensitivity analysis performed using different hyperparameters (results not shown) showed that estimates tend to be less affected by prior choice under our EXSV model.

Further insight into the behavior of these four models is provided by figure 2, which shows the true and reconstructed volatility paths for one simulation in our study. Note that paths generated by STSV tend to underestimate the volatility of volatility. Also, the paths from RASV, RCSV and EXSV tend to be quite similar to each other, especially RCSV and EXSV. These results suggest that much of the information about the volatility contained in the high and low prices is provided by the range, but also that the actual levels of the minimum and maximum can provide additional helpful information in most cases.

6.2. *Estimating the volatility in the S&P500 index*

In this section we consider the series of the weekly S&P500 prices covering the 10-year period between April 21, 1997 and April 9, 2007, for a total of 520 observations.

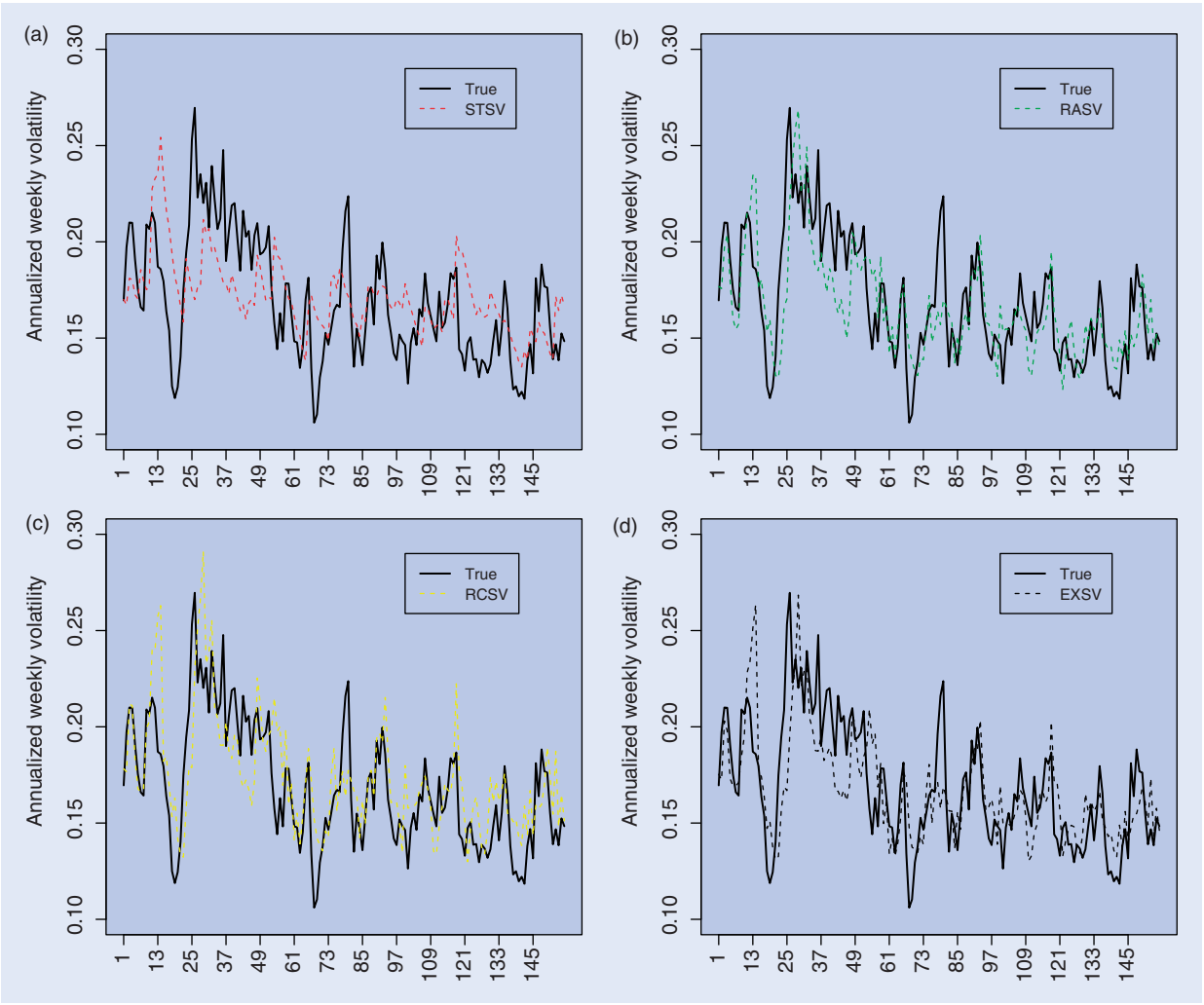

Figure 2. True and reconstructed posterior mean paths of the volatility for one simulation in our study for STSV (top left panel), RASV (top right), RCSV (bottom left) and EXSV (bottom right).

Figure 3 shows the evolution of the log returns at closing, as well as the observed ranges in log returns (computed as the maximum observed log return minus the minimum observed log return over the week). Note that both plots provide complementary but distinct information about the volatility in prices. The series of closing returns does not exhibit any long-term trend, but different levels of volatility can be clearly seen from both plots, with the period 1997–2002 presenting a higher average volatility than the period 2003–2007.

The data were analysed using three of the models considered in the previous section: STSV, RASV and EXSV. Prior hyperparameters were chosen so that $d_\mu = 0$, $D_\mu = 0.0001$ (therefore, we expect the average weekly returns to be between -0.03 and 0.03 with high probability), $q_\phi = 1$ and $r_\phi = 1$ (expressing the lack of information about the correlation parameter), $d_\alpha = -3.75$, $D_\alpha = 0.025$ and $u_\tau = 6$, $v_\tau = 0.06$ (so that the median of the annualized long-term volatility is *a priori* around 20%). These distributions for α and τ, although informative, imply relatively wide ranges for the parameters *a priori*, which essentially cover all likely values we reasonable expect for the model parameters. Our sensitivity analysis varied d_α between -4.5 and -3.0, and v_τ between 0.03 and 0.12, with little impact in the parameter estimates.

The models were fitted using the particle filter algorithm described in section 4. We used a total of 100,000 particles for each model, and monitored particle impoverishment by computing the effective sample sizes at each point in time. Although not a serious issue in any of the three models, particle impoverishment is more pronounced for EXSV. This is probably due to the additional information provided by the extremes, which makes the likelihood tighter and the weights associated with them less uniform.

Point and interval estimates for the structural parameters under EXSV are shown in figure 4. These are filtered estimates, which means that they incorporate information available only until the time they were computed. We note that there is substantial learning about the structural parameters. After $T = 520$ observations, the posterior mean for the median of the stationary distribution of volatility, $\nu = \exp\{\alpha\}$, is 0.1546, with a symmetric 90% credible interval (0.1510, 0.1670), while the autoregressive coefficient for the volatility has a posterior mean of 0.8935 with credible interval (0.8832, 0.9134). The results from STSV (not shown) are similar, but tend to produce a larger value of the autocorrelation coefficient (posterior mean 0.9690, 90% credible interval (0.9245, 0.9873)). One surprising feature of these estimates is the pronounced drop in the autocorrelation coefficient ϕ in the week of January 24, 2000, which is accompanied by an increase in the volatility of the volatility, τ. This can be explained by the large negative return observed for this week (around -11% in the week). This observation suggests that a model that includes volatility jumps would be more appropriate for this data, however the development of such a model is beyond the scope of this paper and will be discussed elsewhere.

Figure 5 shows point estimates for the volatility of returns under all three models. Again, these are filtered estimates. Although the overall level of the volatility series seems to be similar in all three instances, and the estimates obtained from RASV and EXSV are similar, there are striking differences in the behavior of STSV with respect to RASV and EXSV. For example, note that the peaks during the high volatility period of 1997–2002 are much more pronounced under RASV and EXSV, while the average volatility between 2003 and 2007 seems to be lower under both RASV and EXSV than under STSV. Therefore, including information on extreme values seems to help to correct not only for underestimation, but also for overestimation of the volatility. To complement the information in figure 5, we present in figure 6 the posterior distribution for the volatility of returns on December 6, 2000 and April 2, 2007 under all three models. All distributions are right skewed, but the

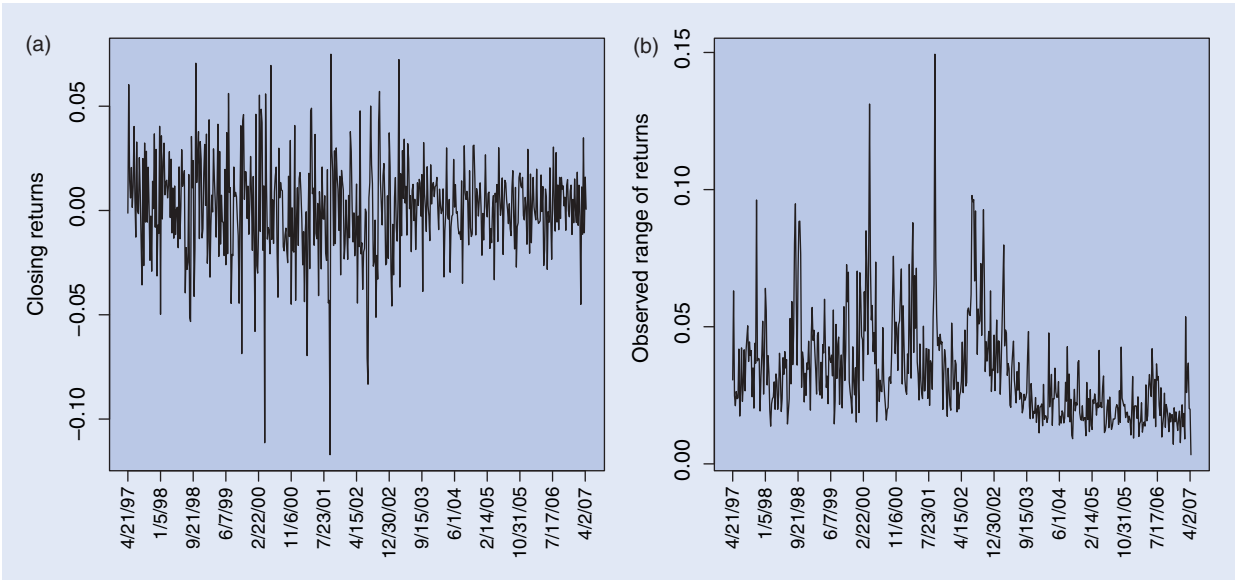

Figure 3. Observed weekly closing returns (left panel) and return ranges (right panel) in the S&P500 data.

distributions under RASV and EXSV are shifted to the left, which suggest that STSV overestimates the true volatility of the model in both cases. In addition, for December 6, 2000 the distributions for RASV and EXSV are almost identical; however, for April 2, 2007, they are noticeable different. More important, the posterior under both EXSV and RASV has a lower variability than that under STSV, reflecting the additional information contained in the extreme values. In particular, including information about the observed extremes in prices seems to increase the volatility in the volatility (and therefore, increases the kurtosis in the distribution of returns).

6.3. *Estimating the volatility of Google stock*

We now consider the longer series of the daily prices for Google (Nasdaq GS:GOOG) covering the five-year period between August 19, 2004 and December 24, 2009 for a total of 1348 observations. The raw log returns based on opening/closing prices are presented in the left panel of figure 7, while the right panel shows the daily ranges. Both plots reflect large changes in the volatility over the five-year period, with a particularly significant increase at the beginning of the third quarter of 2008.

This dataset was analysed through only two models, STSV and EXSV, and we used the same priors and sensitivity analysis we described in section 6.2. Again, inferences are based on 100,000 particles for each model, and we monitored particle impoverishment by computing the effective sample sizes at each point in time.

Figure 8 presents point and interval estimates for the structural parameters under EXSV, while figure 9 shows the posterior mean of the volatility path. The first set of figures suggest that there is substantial learning about the

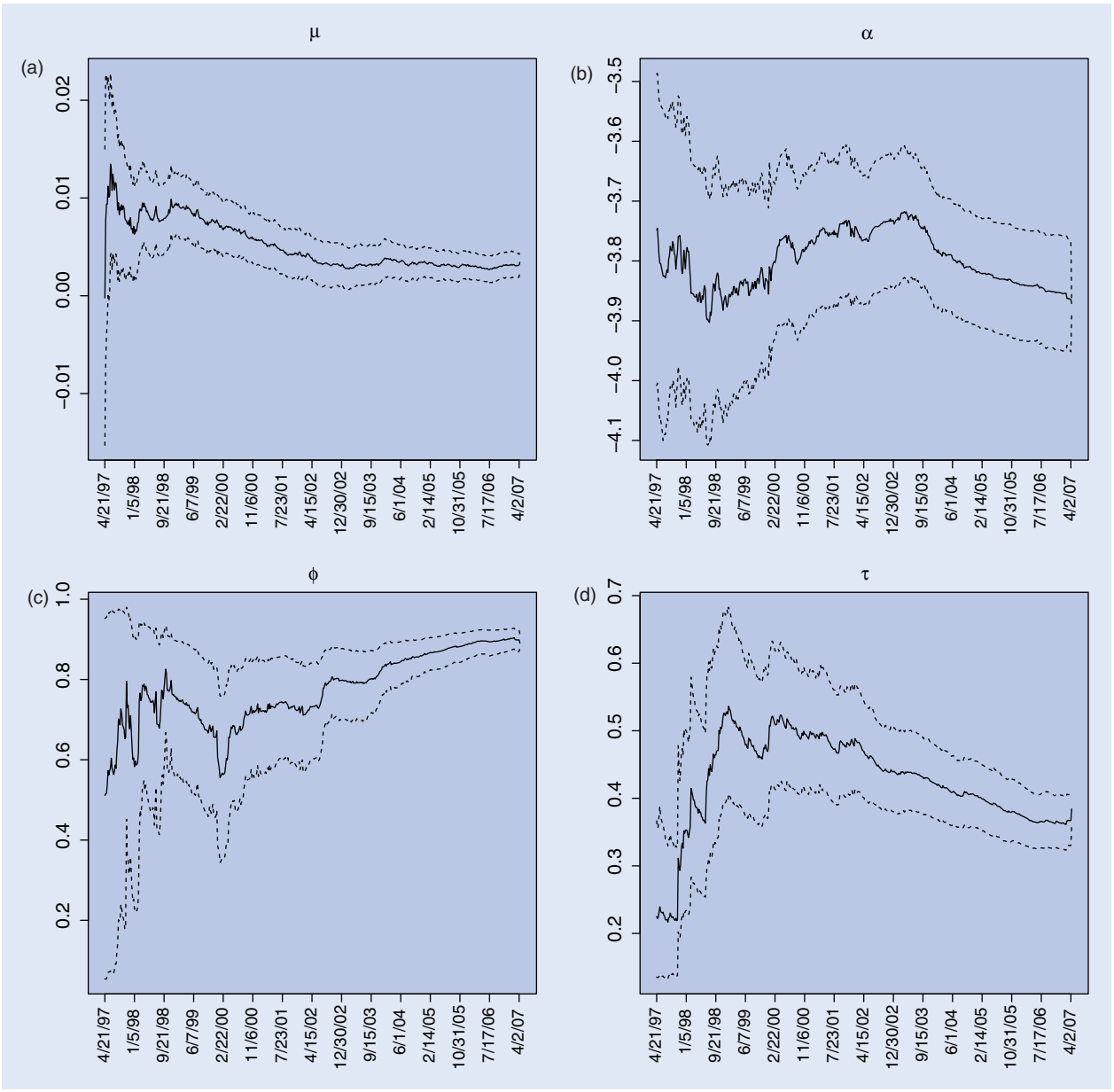

Figure 4. Filtered posterior means and 5% and 95% posterior quantiles for the structural parameters under EXSV for the S&P500 data.

structural parameters, particularly for α and ϕ. Also, as with the S&P500 data, the point estimate of the correlation coefficient under EXSV is smaller than under STSV. On the other hand, the smoother structure of the STSV estimate in figure 9 suggests that EXSV is able to capture the high-frequency structure in the volatility path.

7. Discussion and future work

This paper adds to the growing body of literature suggesting that the information contained in the observed extremes of asset prices can greatly contribute to increase the accuracy and efficiency of volatility estimates. In addition, our results also suggest that the information contained in the level of the extreme returns (which is lost when using the observed ranges for inferences) can also contribute to more efficient estimation of the volatility, and almost as important in certain applications, to the estimation of the drift of the process. This paper also provides a unifying framework in which to understand several CHLO models available in the literature, allowing for a direct comparison of different assumptions.

The methodology described in this paper relies on the use of particle filters for computation, and in particular the augmented auxiliary particle filter described by Liu and West (2001). Both the Liu and West (2001) algorithm we employ and the use of particle filters to fit open/close stochastic volatility models have been around for a while. However, to the best of our knowledge, nobody has explored the use of particle filters to fit CHLO models before. Although the Liu and West (2001) method tends to degenerate quickly in certain applications (especially in high-dimensional state spaces), the evidence from our empirical evaluation suggests that it generates reasonable samples from the posterior distribution in the context of our univariate CHLO models. In addition, the algorithm is easier to implement and *much faster* than MCMC samplers for online problems.

CHLO models such as those discussed in this paper complement methods based on high-frequency realized volatilities (see, for example, Barndorff-Nielsen and Shephard (2002)). In most markets, information about high and low prices is readily available even though tic-by-tic, five-minutes and even hourly data is not available, or available only at a high cost. In those circumstances, methods based on CHLO data are extremely useful while methods that use realized volatilities are simply not applicable. In addition, note that CHLO models can be applied to high-frequency data at any frequency above tick data, since with any such constant frequency models, the data normally used, even in realized volatility models, is the close-to-close data between periods.

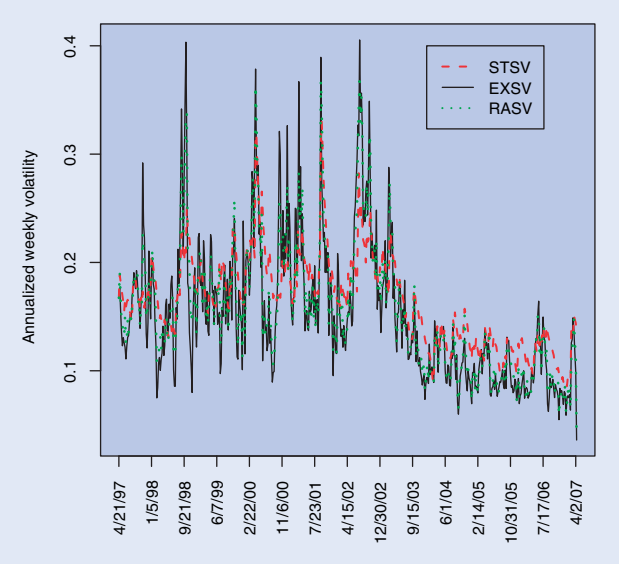

Figure 5. Filtered posterior means of the volatilities of S&P500 data using three different stochastic volatility models.

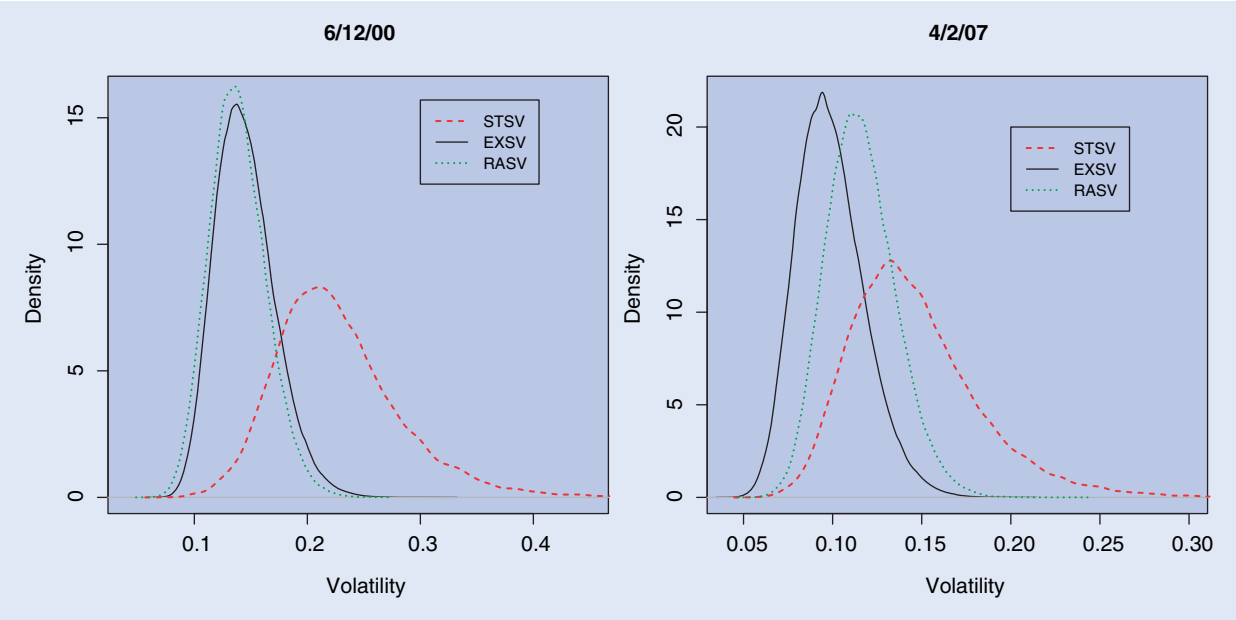

Figure 6. Posterior distribution for the volatility of returns for December 6, 2000 (left panel) and April 2, 2007 (right panel).

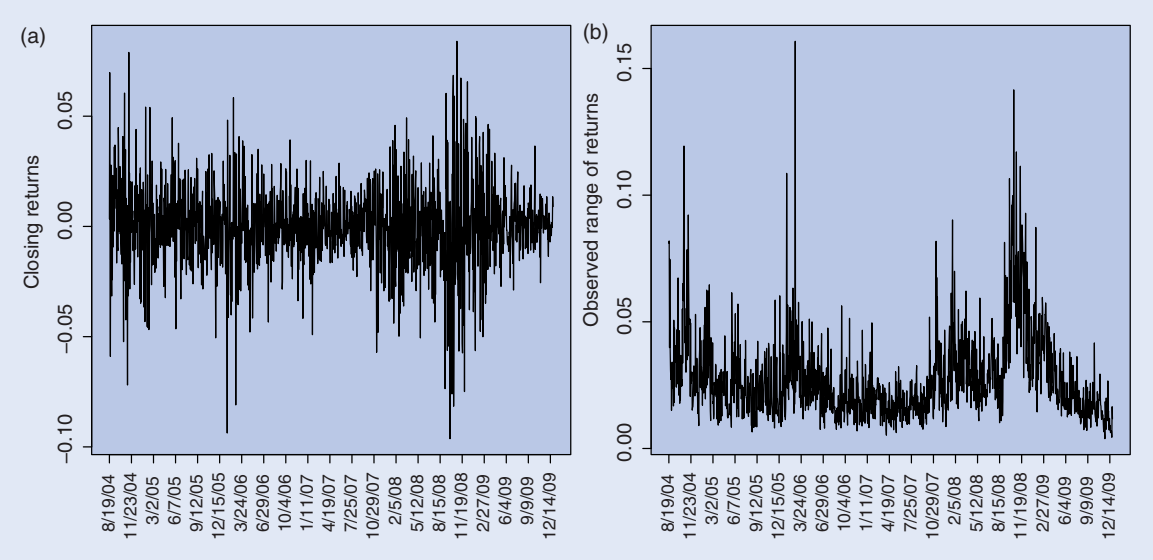

Figure 7. Observed weekly closing returns (left panel) and return ranges (right panel) for the Google data.

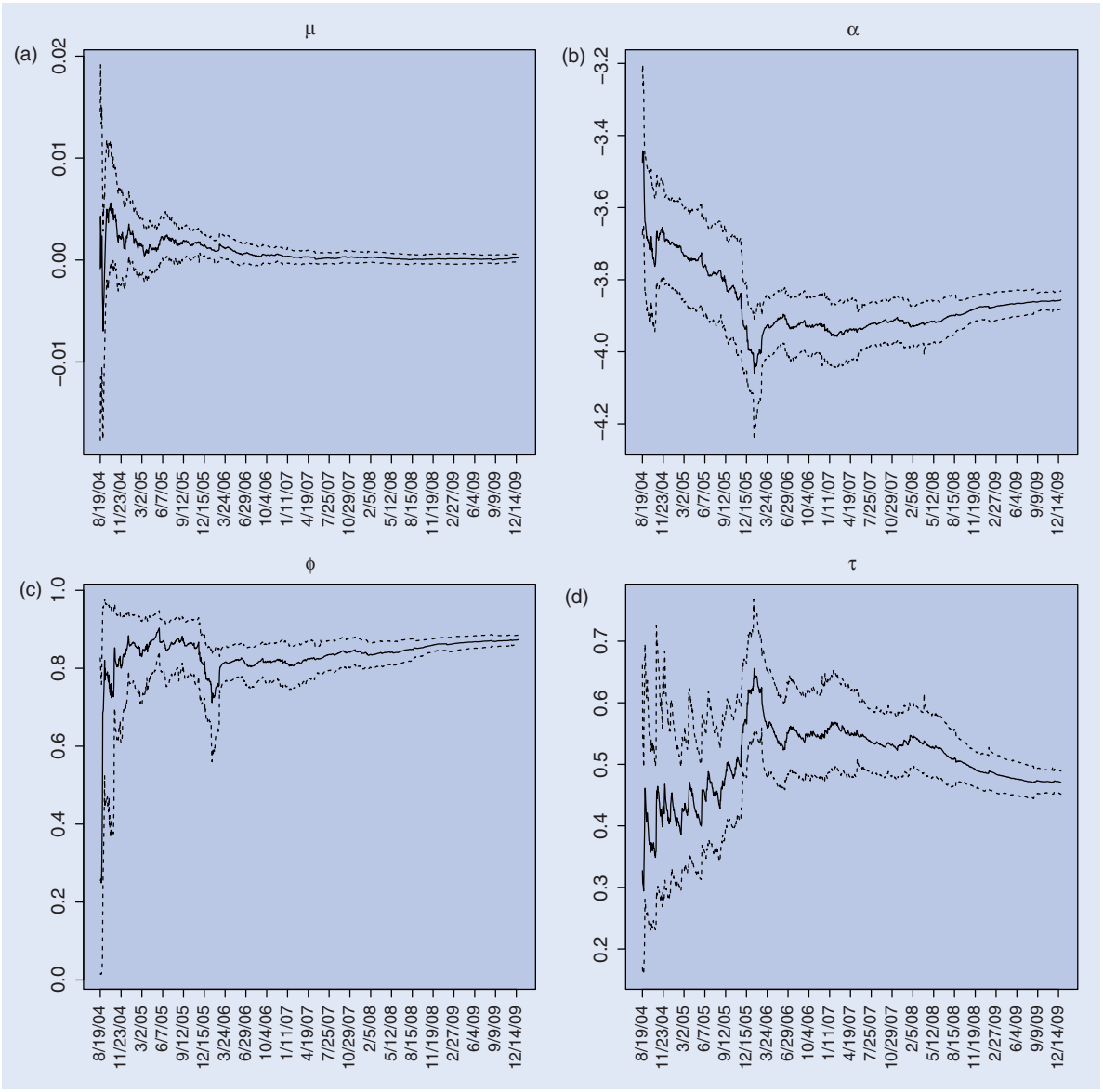

Figure 8. Filtered posterior means and 5% and 95% posterior quantiles for the structural parameters under EXSV for the Google data.

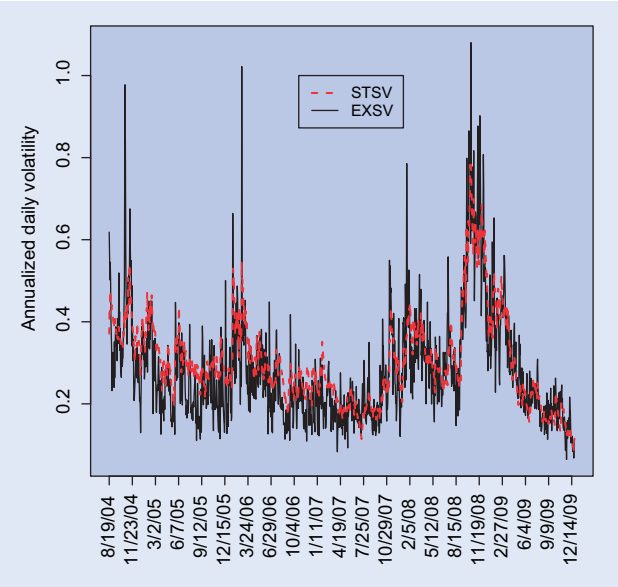

Figure 9. Filtered posterior means of the volatilities for Google stock using two different stochastic volatility models.

Some extensions of the simple stochastic volatility model discussed in this paper are immediate and have already been hinted at. For example, it is well known that leverage effects can be incorporated by introducing a correlation between the innovations in the price and volatility processes. In addition, non-equally spaced observations can easily be accommodated by slightly rewriting the likelihood equations in terms of non-unit-length intervals. The strong identifiability of model parameters provided by using the full information in the model allows us to reliably compare different evolution dynamics (e.g., Ornstein–Uhlenbeck, Markov-switching, jump and squared root processes). Indeed, model comparison in this setting is straightforward thanks to our reliance on particle filters for computation. Finally, we plan to investigate how the additional information provided by the range can help when reconstructing option prices. In this regard, note that, conditionally on the parameters of the underlying stochastic process, the pricing formulas of Heston (1993) can be used almost directly in our problem. Therefore, our model can be extended to generate a posterior distribution for the price of any option of interest, which can be compared with the prices observed in the market.

High-frequency information from covariates of interest can be easily included in the model. If only CHLO data on the covariate is available, the four values could be introduced in the model as independent regressors (either in a linear or kernel/spline regression) without changing the algorithm. If instead high-frequency information about the covariate is available, then introducing all high-frequency values can be difficult because the predictors are likely to be highly correlated. Some options in this case include adding only a small number of the most relevant principal components associated with the covariates, using a structured regression model such as MIDAS regression (Ghysels *et al*. 2002, 2007), or structured priors on the regression coefficients (Rodriguez and Puggioni 2010). Details on this type of model will be discussed elsewhere.

Last, but not least, another avenue that deserves future exploration is the use of extreme prices for the estimation of covariance across multiple asset prices. In that regard, we can start by writing a model for the multivariate log-asset prices $\mathbf{y}_t$ as a multivariate GBM, with time variance covariance matrix, and deriving the joint density for closing, high and low prices conditional on the opening prices, in a manner similar to what we did in this paper. This joint likelihood contains all relevant information in the extremes and avoids the calculation of cross ranges (Rogers and Zhou 2008), which scale badly to higher dimensions.

References

Abramov, V. and Klebaner, F., Estimation and prediction of a non-constant volatility. *Asia–Pacif. Financial Mkts*, 2007, **14**, 1–23.

Alizadeh, S., Brandt, M. and Diebold, F., Range-based estimation of stochastic volatility models. *J. Finance*, 2002, **57**, 1047–1092.

Andersen, T., Bollerslev, T., Diebold, F. and Labys, P., The distribution of exchange rate volatility. *J. Am. Statist. Assoc.*, 2001, **96**, 42–55.

Ball, C. and Tourus, W., The maximum likelihood estimation of security price volatility: theory, evidence, and application to option pricing. *J. Business*, 1984, **57**(1), 97–112.

Barndorff-Nielsen, O. and Shephard, N., Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *J. R. Statist. Soc. Ser. B*, 2002, **63**, 253–280.

Bollerslev, T., Generalized autoregressive conditional heteroscedasticity. *J. Econometr.*, 1986, **31**, 307–327.

Brandt, M. and Jones, C., Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Finance Res. Lett.*, 2005, **2**, 201–209.

Cappe, O., Godsill, S.J. and Moulines, E., An overview of existing methods and recent advances in sequential Monte Carlo. *Proc. IEEE*, 2007, **95**, 899–924.

Carvalho, C.M., Johannes, M., Lopes, H.F. and Polson, N.G., Particle learning and smoothing. Technical Report, University of Chicago – Booth School of Bussiness, 2009.

Courtault, J., Kabanov, Y., Bru, B., Crepel, P., Lebon, I. and Le Marchand, A., Louis Bachelier on the centenary of theorie de la speculation. *Math. Finance*, 2000, **10**, 341–353.

Dana, R., Jeanblanc, M. and Kennedy, A., *Financial Markets in Continuous Time*, 1st ed., 2007 (Springer: New York).

Dempster, A., Laird, N. and Rubin, D., Maximum-likelihood from incomplete data via the EM algorithm. *J. R. Statist. Soc. Ser. B*, 1997, **1**, 1–38.

Doucet, A., de Freitas, N. and Gordon, N., editors, *Sequential Monte Carlo Methods in Practice*, 2001 (Springer: Berlin).

Dym, H., McKean, H.P., Aldous, D. and Tong, Y.L., *Fourier Series and Integrals*, 1st ed., 1985 (Academic Press: San Francisco).

Engle, R.F., Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, 1982, **50**, 987–1007.

Fama, E.F., The behavior of stock market prices. *J. Business*, 1965, **38**, 34–105.

Freedman, D., *Brownian Motion and Diffusion* (Holden-Day Series in Probability and Statistics), 1st ed., 1971 (Holden-Day: New York).

Gallant, A., Hsu, C. and Tauchen, G., Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Rev. Econ. Statist.*, 1999, **81**, 617–631.
Garman, M. and Klass, M., On the estimation of security price volatilities from historical data. *J. Business*, 1980, **53**, 67–78.
Gatheral, J. and Oomen, R., Zero-intelligence realized variance estimation. *Finance Stochast.*, 2010, **14**, 249–283.
Gelman, A., Carlin, J., Stern, H. and Rubin, D., *Bayesian Data Analysis*, 1st ed., 2003 (Chapman & Hall/CRC: New York).
Ghysels, E., Santa-Clara, P. and Valkanov, R., The MIDAS touch: mixed data sampling regression models. Technical Report, University of North Carolina, Chapel Hill and University of California, Los Angeles, 2002.
Ghysels, E., Sinko, A. and Valkanov, R., MIDAS regressions: further results and new directions. *Econometr. Rev.*, 2007, **26**, 53–90.
Heston, S.L., A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Stud.*, 1993, **6**, 327–343.
Hofmann, M., Rate of convergence for parametric estimation in stochastic volatility models. *Stochast. Process. Applic.*, 2002, **97**, 147–170.
Hu, M., Salvucci, S. and Cohen, M., Evaluation of some popular imputation algorithms. In *Section on Survey Research Methods*, pp. 308–313, 1998 (American Statistical Association).
Hull, J. and White, A., The pricing of options on assets with stochastic volatilities. *J. Finance*, 1987, **42**, 281–300.
Jacquier, E., Polson, N.G. and Rossi, P.E., Bayesian analysis of stochastic volatility models. *J. Business Econ. Statist.*, 1994, **12**, 371–389.
Kim, S., Shephard, N. and Chib, S., Stochastic volatility: likelihood inference and comparison with ARCH models. *Rev. Econ. Stud.*, 1998, **65**, 361–393.
Kitagawa, G., A self-organizing state-space model. *J. Am. Statist. Assoc.*, 1998, **93**, 1203–1215.
Klebaner, F.C., *Introduction to Stochastic Calculus with Applications*, 2nd ed., 2005 (Imperial College Press: London).
Lildholdt, P.M., Estimation of GARCH models based on open, close, high, and low prices. Centre for Analytical Finance Working Papers Series 128, Department of Economics, University of Aarhus, and Center for Analytical Finance, 2002. Available online at: http://www.cls.dk/caf/wp/wp-128.pdf
Liu, J. and West, M., Combined parameter and state estimation in simulation-based filtering. In *Sequential Monte Carlo Methods in Practice*, edited by A. Doucet, N. de Freitas, and N. Gordon, 2001 (Springer: New York).
Magdon, M. and Atiya, A., A maximum likelihood approach to volatility estimation for a Brownian motion using high, low and close price data. *Quant. Finance*, 2003, **3**, 376–384.
Mandelbrot, B., The variation of certain speculative prices. *J. Business*, 1963, **36**, 394–419.
Officer, R.R., The variability of the market factor of the New York Stock Exchange. *J. Business*, 1973, **46**, 434–453.
Parkinson, M., The extreme value method for estimating the variance of the rate of return. *J. Business*, 1980, **53**, 61–65.
Perl, J., *DeMark Indicators*, 2008 (Bloomberg Press: New York).
Pitt, M.K. and Shephard, N., Filtering via simulation: auxiliary particle filters. *J. Am. Statist. Assoc.*, 1999, **94**, 590–599.
Polson, N.G., Stroud, J.R. and Müller, P., Practical filtering with sequential parameter learning. *J. R. Statist. Soc. Ser. B*, 2008, **70**, 413–428.
Rodriguez, A. and Puggioni, G., Mixed frequency models: Bayesian approaches to estimation and prediction. *Int. J. Forecast.*, 2010, **26**, 293–311.
Rogers, L., Volatility estimation with price quanta. *Math. Finance*, 1998, **8**, 61–65.
Rogers, L. and Satchell, S., Estimating variance from high, low and closing prices. *Ann. Appl. Probab.*, 1991, **1**, 504–512.
Rogers, L., Satchell, S. and Yoon, Y., Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Appl. Financial Econ.*, 1994, **4**, 241–247.
Rogers, L. and Zhou, F., Estimating correlation from high, low, opening and closing prices. *Ann. Appl. Probab.*, 2008, **18**, 813–823.
Rosenbaum, M., Etudes de quelques problèmes d' estimation statistique en finance. PhD Thesis, Université Paris-Est, 2007.
Russell, J. and Engle, R., Econometric analysis of discrete-valued irregularly-spaced financial transactions data using a new autoregressive conditional multinomial model. CRSP Working Paper 470, University of California at San Diego Working Paper No. 98-10, 1998. Available online at SSRN: http://ssrn.com/abstract = 106528
Scheffer, J., Dealing with missing data. *Res. Lett. Inf. Math. Sci.*, 2002, **3**, 153–160.
Shephard, N., editors, *Stochastic Volatility: Selected Readings*, 2005 (Oxford University Press: Oxford).
Storvik, G., Particle filters for state-space models with the presence of unknown static parameters. *IEEE Trans. Signal Process.*, 2002, **50**, 281–289.
Wilder, J., *New Concepts in Technical Trading Systems*, 1st ed., 1978 (Trend Research: McLeansville, NC).

Appendix A: Proof of theorem 3.1

Proof: Consider X_t a driftless Brownian motion with unit variance starting at zero. In chapter 3 of Klebaner (2005, p. 75) the joint cumulative distribution for the maximum $M_t = \max_{0\le s\le t}\{X_s\}$, minimum $m_t = \min_{0\le s\le t}\{X_s\}$ and final value of X_t is obtained, yielding

$$\begin{aligned}&\Pr(m_t \ge a_t, M_t \le b_t, X_t \le x_t)\\ &= \int_{-\infty}^{x_t} \sum_{n=-\infty}^{\infty} [\varphi(z - 2n(b_t - a_t))\\ &\quad - \varphi(z + (2n(b_t - a_t) - 2a_t)]\mathrm{d}z,\end{aligned}$$

where $\varphi(z)$ corresponds to the density of a standard Gaussian distribution. To include the drift and non-unit volatility, we can use Girsanov's theorem. Finally, information about the opening price can be incorporated by translating the Brownian motion, i.e. letting $a_t^* = a_t - y_{t-1}$, $b_t^* = b_t - y_{t-1}$ and $x_t = y_t - y_{t-1}$, which yields (7).