Available online at www.sciencedirect.com
SCIENCE @ DIRECT®

Finance Research Letters 2 (2005) 201–209

Finance Research Letters

www.elsevier.com/locate/frl

Bayesian range-based estimation of stochastic volatility models

Michael W. Brandt [a,b,*], Christopher S. Jones [c]

[a] *Fuqua School of Business, Duke University, Durham, NC 27708, USA*
[b] *National Bureau of Economic Research, Cambridge, MA 02138, USA*
[c] *Marshall School of Business, University of Southern California, Los Angeles, CA 90089, USA*

Received 16 July 2005; accepted 15 September 2005

Available online 29 September 2005

Abstract

Alizadeh, Brandt, and Diebold [2002. Journal of Finance 57, 1047–1091] propose estimating stochastic volatility models by quasi-maximum likelihood using data on the daily range of the log asset price process. We suggest a related Bayesian procedure that delivers exact likelihood based inferences. Our approach also incorporates data on the daily return and accommodates a nonzero drift. We illustrate through a Monte Carlo experiment that quasi-maximum likelihood using range data alone is remarkably close to exact likelihood based inferences using both range and return data.

JEL classification: G10; G12; C11; C15

Keywords: Stochastic volatility models; Price range; Bayesian estimation

1. Introduction

Alizadeh, Brandt and Diebold (ABD, 2002) propose estimating stochastic volatility models by quasi-maximum likelihood using data on the daily range of the log asset price process. Their procedure is simple yet highly efficient, compared to standard daily return-based inferences, for two reasons. First, the daily range is a much less noisy measure of daily volatility than absolute or squared daily returns (see Parkinson, 1980, and Andersen and Bollerslev, 1998). Second, ABD

* Corresponding author. Fax: +1 919 882 9157.
E-mail address: mbrandt@duke.edu (M.W. Brandt).

1544-6123/$ – see front matter
doi:10.1016/j.frl.2005.09.001

demonstrate that the distribution of the log range is close to being Gaussian, which results in fairly accurate Kalman filter based quasi-maximum likelihood estimates.

We address three open issues in ABD's range-based estimation approach:

(1) the use of an approximate as opposed to exact likelihood function,
(2) the use of only daily range data when daily return data are also readily available, and
(3) the assumption of a zero drift of the log asset price process to derive the distribution of the log range.

We accomplish this by developing a closely related Bayesian procedure based on the exact joint distribution of the daily log range and log return, relying heavily on the simulation-based techniques introduced by Jacquier, Polson and Rossi (JPR, 1994).

We implement our procedure in the context of a Monte Carlo experiment to examine the extent to which exact likelihood inferences based on both range and return data dominates ABD's approximate likelihood approach based on range data alone. We find that the gains in efficiency are minimal. Our results confirm ABD's conjecture that the distribution of the log range is sufficiently close to being Gaussian to deliver virtually exact maximum likelihood estimates. They also suggest that, at least in the absence of a drift, absolute or squared daily returns contain little information not already revealed through the daily range.

2. Model

Consider the following stylized stochastic volatility model:

$$\frac{dS_\tau}{S_\tau} = \mu\, d\tau + h_\tau\, dW_{S\tau},$$
$$d\ln h_\tau = \kappa(\ln\bar{h} - \ln h_\tau)\, d\tau + \beta\, dW_{h\tau}. \tag{1}$$

In this model, the log volatility $\ln h_\tau$ evolves as a mean-reverting Ornstein–Uhlenbeck process with mean $\ln\bar{h}$ and mean reversion parameter $\kappa > 0$. The two Brownian motions are assumed independent.

Since we observe stock prices at discrete times $t \in \{0, 1, 2, \ldots, T\}$, we discretize the model by assuming that the return volatility is constant at h_t over the interval $t-1 < \tau \leqslant t$. Within each interval, the stock price then follows a standard geometric Brownian motion:

$$\frac{dS_\tau}{S_\tau} = \mu\, d\tau + h_t\, dW_{S\tau} \quad \text{for } t-1 < \tau \leqslant t, \tag{2}$$

and the conditional distribution of log volatility from one discrete interval to the next is approximately:

$$\ln h_t \mid \ln h_{t-1} \sim \mathrm{N}\big[\ln h_{t-1} + \kappa(\ln\bar{h} - \ln h_{t-1}), \beta^2\big]. \tag{3}$$

This discretization of the continuous time stochastic volatility model is not original. It is shared by ABD (2002), Duffie and Singleton (1993), Gallant et al. (1997), Harvey et al. (1994), Jacquier et al. (1994), Kim et al. (1998), and Taylor (1994), among others. Like us, these authors assume that for discrete time periods the volatility is constant and that from one period to the next the log volatility is conditionally normal and mean reverting.

For $t-1 \leqslant \tau \leqslant t$, let $x_\tau = \ln S_\tau - \ln S_{t-1}$, so that x_t is the log return for the period ending at time t. Define the *range* D_t as

$$D_t = \max_{t-1 \leqslant \tau \leqslant t} x_\tau - \min_{t-1 \leqslant \tau \leqslant t} x_\tau. \tag{4}$$

Under the assumption that x_τ is a martingale, ABD derive the density function of D_t and illustrate that it is approximately lognormal.

Our approach differs from that of ABD in that we focus on the *joint* distribution of x_τ and D_τ and in that we allow for a nonzero drift in x_τ. The drift is assumed to be constant and is denoted by a.

We decompose the joint distribution of the range and return as

$$p(D_t, x_t \mid a, h_t) = p(x_t \mid a, h_t)\, p(D_t \mid x_t, a, h_t). \tag{5}$$

The first term on the right is Gaussian and is therefore trivial to evaluate. The second term, the density of the range conditional on the contemporaneous return, is nonstandard and is derived in Appendix A. The appendix further shows that $p(D_t \mid x_t, a, h_t)$ is in fact independent of a.

3. Econometric method

We rewrite the discretized volatility model in more familiar notation as[1]

$$\ln h_t = \alpha + \delta \ln h_{t-1} + \sigma \nu_t. \tag{6}$$

We will assume that we have diffuse prior information about the parameters of the stochastic volatility process, meaning

$$p(\alpha, \delta, \sigma) \propto \frac{1}{\sigma} 1_{\sigma \geqslant \bar{\sigma}}. \tag{7}$$

The lower bound on σ is set to a small positive number and is required to insure proper posteriors (see Johannes and Polson, 2005). In practice, as long as the algorithm is initialized with a reasonable value of σ, the bound is never hit.

We analyze this model following Tanner and Wong's (1987) principle of data augmentation. In this framework, we calculate the posterior distribution of the model parameters numerically using Markov chain Monte Carlo methods, in which a Markov chain is constructed to generate the posterior as its invariant distribution. By generating a sufficiently long chain, the posterior distribution may be characterized with arbitrary precision.

The chain is constructed by alternatively drawing latent data (in this case, the T return volatilities) given the model parameters and drawing model parameters given latent data.

Given latent volatilities, drawing the model parameters in this case is simple, since the setup conforms to the standard Gaussian linear regression framework. Parameter draws are based on standard OLS estimates, $\hat{\alpha}$, $\hat{\delta}$, and $\hat{\sigma}$. Specifically, σ is drawn from the inverted gamma distribution $IG\left(\frac{2}{\hat{\sigma}^2(T-3)}, \frac{T-1}{2}\right)$. Given the draw of σ, the distribution of $[\alpha, \delta]'$ is bivariate normal with mean $[\hat{\alpha}, \hat{\delta}]'$ and covariance matrix $\sigma(H'H)^{-1}$, where H is a $(T-1) \times 2$ matrix whose tth row is equal to $[1, \ln h_t]$.

[1] Other studies, such as JPR, equivalently choose to model the log variance process.

The draw of the latent volatility path is more complicated because it is from a nonstandard distribution. In addition, the high dimensionality of the latent data makes the application of standard random number generation methods difficult. We therefore adapt the approach of JPR and decompose the multivariate draw into a sequence of univariate draws.

Rather than drawing the h vector all at once, we instead draw each element h_t conditional on the rest of the vector (denoted h_{-t}). Since this draw is also conditional on the current draw of the parameter vector (denoted θ for shorthand), we can write our "target density" as

$$f_t(h_t) = p(h_t \mid h_{-t}, x, D, \theta), \tag{8}$$

where x and D denote the time series of returns and log ranges.

The Markovian nature of the process greatly reduces the relevant conditioning set, so that the density may be written as

$$f_t(h_t) = p(h_t \mid h_{t-1}, h_{t+1}, x_t, D_t, \theta). \tag{9}$$

As in JPR, we "cycle" through the h vector drawing a new value to replace the old for each t. At the end of each full cycle, a new draw from the parameter vector's conditional distribution replaces the previous draw.

Using Bayes rule, we can decompose the target distribution into the product of some simpler kernels:

$$\begin{aligned} f_t(h_t) &= p(h_t \mid h_{t-1}, h_{t+1}, x_t, D_t, \theta) \\ &\propto p(h_t \mid h_{t-1}, \theta) p(h_{t+1} \mid h_t, \theta) p(x_t \mid h_t, \theta) p(D_t \mid h_t, x_t, \theta) \\ &\propto \frac{1}{h_t} \exp\left\{-\frac{1}{2}\left(\frac{\ln h_t - m_1}{s_1}\right)^2\right\} \frac{1}{h_t} \exp\left\{-\frac{1}{2}\left(\frac{x_t}{h_t}\right)^2\right\} p(D_t \mid h_t, x_t), \end{aligned} \tag{10}$$

where

$$m_1 = \frac{\alpha(1-\delta) + \delta(\ln h_{t+1} + \ln h_{t-1})}{1+\delta^2} \quad \text{and} \quad s_1^2 = \frac{\sigma^2}{1+\delta^2} \tag{11}$$

result from completing the square of the first two lognormal kernels.

Because this is the density of a nonstandard distribution, we perform draws from it using the Metropolis–Hastings algorithm. We therefore form what JPR term a "cyclic Metropolis chain."

Since the Metropolis algorithm is an accept/reject procedure, we must select a proposal density to generate candidate values of h_t. Ideally, this proposal density should approximate the target density closely. To the extent that the two differ, the proposal density should "blanket" the target, meaning that it sufficiently often generates candidate values that are in the tails of the target distribution. If the blanketing is poor or if it is excessive (too many draws in the tail of the target), then the convergence of the chain may be slow.

As a proposal density we use an approximation of $p(h_t \mid h_{t-1}, h_{t+1}, D_t, \theta)$. Because we do not condition on x_t, the proposal density should on average have a higher variance than the target density. This could potentially result in a slow-mixing chain, but it is otherwise unproblematic. In any case, since x_t contains relatively little information about h_t when compared with D_t, this effect will prove to be minor and of no real concern.

We approximate the density $p(h_t \mid h_{t-1}, h_{t+1}, D_t, \theta)$ by first decomposing it, as we did with the target density, into the product of simpler kernels, or

$$p(h_t \mid h_{t-1}, h_{t+1}, D_t, \theta) \propto p(h_t \mid h_{t-1}, \theta) p(h_{t+1} \mid h_t, \theta) p(D_t \mid h_t, \theta), \tag{12}$$

where $p(h_t \mid h_{t-1}, \theta)$ and $p(h_{t+1} \mid h_t, \theta)$ are lognormal densities for h_t and h_{t+1}, respectively. The third kernel, $p(D_t \mid h_t, \theta)$, is of a nonstandard distribution. ABD show, however, that D_t is approximately lognormal with mean $\ln h_t + 0.43$ and standard deviation 0.29.[2]

Letting $p^*(D_t \mid h_t, \theta)$ denote this approximate lognormal distribution for D_t, our proposal density is

$$\begin{aligned} f_p(h_t) &\propto p(h_t \mid h_{t-1}) p(h_{t+1} \mid h_t) p^*(D_t \mid h_t) \\ &\propto \frac{1}{h_t} \exp\left\{-\frac{1}{2}\left(\frac{\ln h_t - m_1}{s_1}\right)^2\right\} \frac{1}{D_t} \exp\left\{-\frac{1}{2}\left(\frac{\ln D_t - 0.43 - \ln h_t}{0.29}\right)^2\right\} \\ &\propto \frac{1}{h_t} \exp\left\{-\frac{1}{2}\left(\frac{\ln h_t - m_1}{s_1}\right)^2\right\} \exp\left\{-\frac{1}{2}\left(\frac{\ln h_t - m_2}{s_2}\right)^2\right\} \\ &\propto \frac{1}{h_t} \exp\left\{-\frac{1}{2}\left(\frac{\ln h_t - m}{s}\right)^2\right\}, \end{aligned} \tag{13}$$

where m_1 and s_1 are the same as before, $m_2 = \ln D_t + 0.43$, $s_2 = 0.29$,

$$m = \frac{m_1 s_1^{-2} + m_2 s_2^{-2}}{s_1^{-2} + s_2^{-2}} \quad \text{and} \quad s = \frac{1}{\sqrt{s_1^{-2} + s_2^{-2}}}. \tag{14}$$

In summary, the proposal density for h_t is lognormal with mean parameter m and variance parameter s^2. Furthermore, the density does not depend on the previous draw of h_t.

The Metropolis–Hastings algorithm dictates that the probability of replacing the old draw h_t with the candidate draw h_t^* is equal to

$$\min\left[\frac{f_t(h_t^*) f_p(h_t)}{f_t(h_t) f_p(h_t^*)}, 1\right]. \tag{15}$$

If the draw is not accepted, the new value of h_t is set equal to the previous draw. Draws performed in this way, when combined with draws from the conditional distribution of $\theta = [\alpha, \delta, \sigma]$, produce a Markov chain whose invariant distribution is the joint distribution of θ and h.

Exceptions to this procedure occur for the draws of the first and last values of the volatility process, h_1 and h_T. For h_T, the draw is the same, except that the conditioning variable h_{T+1} no longer exists. For h_1, the conditioning variable h_0 does not exist. We will assume, however, that the process has been running for a sufficient period prior to the beginning of our sample so that the value h_1 is distributed according to the stationary distribution of the process, if such a distribution exists. Appendix B details these variations of the procedure.

4. Monte Carlo results

Here we present simulation evidence to compare our approach with those of both ABD and JPR. Parameter values are from JPR, where we consider three sets of parameters corresponding to their "intermediate" case for the coefficient of variation of h_t. Under these parameters, we simulate volatility paths h_t of length 500 according to (6). Within each period, we simulate

[2] In fact, the skewness and kurtosis of the true density are slightly lower than those of the lognormal. The approximate density should therefore be expected to perform well in "blanketing" the target density, which should help avoid the slow-mixing problem identified by Mengersen and Tweedie (1996).

Brownian motion with mean zero and volatility h_t by discretizing the interval into 1000 subperiods. The range and return are computed for each period from these simulated values.

We generate 5000 Monte Carlo samples for both the method outlined in Section 3 and a Bayesian version of the method presented in ABD. The two approaches are identical except that the ABD method replaces the true likelihood for both D_t and x_t with the approximate lognormal likelihood for D_t only. Thus, the ABD approach differs both because it uses an approximation and because it does not use returns data.

For each sample, we run the Gibbs sampler for 11,000 iterations and discard the first 1000. The remaining draws are averaged to compute posterior means, and these estimates are used to compute the means and root mean squared errors reported in Table 1. Results for the return-based estimates are taken directly from JPR, where we have halved their reported numbers for α and σ since we are working with a log volatility rather than log variance model.

The results illustrate the benefits of using the range, particularly for estimating the volatility of volatility parameter σ. For this parameter, the RMSEs of the two range-based estimators is approximately half as large as those for the JPR approach in all three cases. For the other two parameters, the improvement is limited to the low and medium persistence cases. For the high persistence case, the range-based estimates of α and δ are actually less precise (both more biased and variable) than the return based estimates.

Intuitively, and as ABD illustrate, the improvement for estimating the volatility of volatility parameter σ comes from the fact that the higher signal to noise ratio of the range allows the range-based estimators to better differentiate noise in the volatility proxy from fluctuations in volatility. The more volatility fluctuates, or the lower the persistence of volatility, the greater is the potential benefit from range-based volatility estimation. The improved precision for α

Table 1
Monte Carlo simulation results

Low persistence			Medium persistence			High persistence		
α	δ	σ	α	δ	σ	α	δ	σ
			True parameter values					
−0.368	0.900	0.182	−0.184	0.950	0.130	−0.074	0.980	0.083
			Return-based estimates (from JPR)					
−0.435	0.880	0.175	−0.280	0.920	0.140	−0.110	0.970	0.115
(0.170)	(0.046)	(0.034)	(0.170)	(0.046)	(0.033)	(0.070)	(0.020)	(0.040)
			Range-based estimates using ABD method					
−0.429	0.884	0.187	−0.238	0.936	0.135	−0.122	0.967	0.088
(0.122)	(0.033)	(0.018)	(0.097)	(0.026)	(0.015)	(0.079)	(0.021)	(0.013)
			Range and return-based estimates using exact likelihood					
−0.437	0.882	0.190	−0.242	0.935	0.137	−0.124	0.967	0.089
(0.125)	(0.033)	(0.017)	(0.099)	(0.026)	(0.015)	(0.080)	(0.021)	(0.012)

Notes. The table reports means and root mean squared errors (in parentheses) for three separate Monte Carlo exercises. Return-based estimates are values taken from JPR. Since their model is written in terms of log variance, we divide their α and σ parameters by two for comparison with our results, which are in terms of log volatility. The final two sets of results are each based on our own calculations, where each statistic is computed from 5000 Monte Carlo simulations. In every case, volatility time series are generated using the parameters reported in the top row, and intraday price paths are then simulated conditional on those volatilities using the Euler discretization with 1000 time steps per day. Range-based estimates use the approach of ABD, which relies on a Gaussian approximation of the density of the log range. Range and return-based estimates use the Bayesian method developed in this paper, which differs from the ABD approach both because it incorporates returns data and because it uses the exact likelihood function.

and δ are a by-product of estimating σ more precisely (since δ and σ are linked through the unconditional volatility of volatility). It appears from our results that for the high persistence case this indirect effect on the estimates of α and δ is swamped by other differences between the estimators. For example, JPR use loose but proper conjugate priors while we use improper diffuse priors. Although the effects of this difference should be very minor, it could explain the results for α and δ in the high-persistence case.

Finally, the results suggest that little is to be gained from using the exact likelihood of the range and return data relative to the approximate likelihood of the range data alone employed by ABD. One caveat here is that we simulated log prices by assuming a zero drift in order to be consistent with JPR and ABD. In a situation where the drift is likely to be large, such as with alternative asset classes, our new method may prove more useful since it is able to account for nonzero drift without difficulty.

Acknowledgments

We thank Dmitry Livdan for excellent research support and Entropia for providing computational resources.

Appendix A. Conditional distributions of the range

Consider a driftless Brownian motion x_τ with volatility h over the interval $\tau \in [0, 1]$. We want

$$\text{Prob}[D \in dD,\ x_1 \in d\bar{x} \mid h], \tag{A.1}$$

where $D = M - m$ and

$$M = \max_{t-1 \leqslant \tau \leqslant t} x_\tau, \qquad m = \min_{t-1 \leqslant \tau \leqslant t} x_\tau. \tag{A.2}$$

Assume $x_0 = 0$, so that x_1 can be interpreted as a log return.

Decompose the joint density:

$$\begin{aligned} &\text{Prob}\big[(M-m) \in dD,\ x_1 \in d\bar{x} \mid h\big] \\ &\quad = \underbrace{\text{Prob}[x_1 \in d\bar{x} \mid h]}_{P_1} \times \underbrace{\text{Prob}\big[(M-m) \in dD \mid x_1 = \bar{x}, h\big]}_{P_2}. \end{aligned} \tag{A.3}$$

The density P_1 is just Gaussian. The joint density of M and m is

$$\begin{aligned} &\text{Prob}[M \in dv,\ m \in -du \mid x_1 = \bar{x}, h] \\ &\quad = \sum_{k=-\infty}^{\infty} \frac{4k^2}{h^2}\left[\frac{(2k(u+v)-\bar{x})^2}{h^2} - 1\right] \frac{\phi\big(\frac{2k(u+v)-\bar{x}}{h}\big)}{\phi\big(\frac{\bar{x}}{h}\big)} \\ &\qquad - \cdots \sum_{k=-\infty}^{\infty} \frac{4k(k-1)}{h^2}\left[\frac{(2k(u+v)+\bar{x}-2v)^2}{h^2} - 1\right] \frac{\phi\big(\frac{2k(u+v)+\bar{x}-2v}{h}\big)}{\phi\big(\frac{\bar{x}}{h}\big)}. \end{aligned} \tag{A.4}$$

A change of variable yields the density P_2:

$$\begin{aligned} &\text{Prob}\big[(M-m) \in dD \mid |x_1| = |\bar{x}|, h\big] \\ &\quad = \sum_{k=-\infty}^{\infty} \frac{4k^2}{h^2}\left[\frac{(2kD-|\bar{x}|)^2}{h^2} - 1\right] \frac{\phi\big(\frac{2kD-|\bar{x}|}{h}\big)}{\phi\big(\frac{|\bar{x}|}{h}\big)} \big(D - |\bar{x}|\big) \end{aligned}$$

$$
\begin{aligned}
&+\cdots \sum_{k=-\infty}^{\infty} \frac{2k(k-1)}{h^2}\big(2kD-|\bar{x}|\big)\frac{\phi\big(\frac{2kD-|\bar{x}|}{h}\big)}{\phi\big(\frac{|\bar{x}|}{h}\big)} \\
&-\cdots \sum_{k=-\infty}^{\infty} \frac{2k(k-1)}{h^2}\big(2(k-1)D+|\bar{x}|\big)\frac{\phi\big(\frac{2(k-1)D+|\bar{x}|}{h}\big)}{\phi\big(\frac{|\bar{x}|}{h}\big)}
\end{aligned}
\tag{A.5}
$$

for $D \geqslant |\bar{x}|$, where $\phi(z) = \exp(-z^2/2)/\sqrt{2\pi}$.

Finally, we note that the zero drift assumption was unnecessary, as the conditional density of the range given the contemporaneous return does not depend on the drift of the process. Only the Gaussian component P_1 depends on the drift of x, and that is in the obvious way.

To see the drift independence of D, assume instead that x_τ had drift a. Then if $0 < \tau < 1$:

$$
\begin{bmatrix} x_\tau \\ x_1 \end{bmatrix} \sim \mathrm{N}\left(a\begin{bmatrix} \tau \\ 1 \end{bmatrix}, \sigma^2\begin{bmatrix} \tau & \tau \\ \tau & 1 \end{bmatrix}\right).
$$

The conditional distribution of x_τ given x_1 has mean $a\tau + \tau(x_1 - a) = \tau x_1$ and variance $\sigma^2(\tau - \tau^2)$. Thus, the marginal distribution of x_τ is independent of a.

For $0 < \tau_1 < \tau_2 < 1$, the covariance between x_{τ_1} and x_{τ_2} is also independent of a simply because the conditional covariance of a multivariate normal is never a function of its mean. Thus, the entire *path* of x_τ between 0 and 1 is independent of a after conditioning on x_1, which means that D is independent of a as well.

Appendix B. Drawing latent volatility for $t = 1$ and $t = T$

The proposal and target densities presented in Section 2 must be modified to draw the first and last values of the latent volatility process.

For the final period volatility, h_T, the target density is somewhat simpler due to the lack of a subsequent period. For this value, the target density becomes

$$
\begin{aligned}
f_T(h_T) &= p(h_T \mid h_{T-1}, x_T, D_T) \\
&\propto p(h_T \mid h_{T-1})\, p(x_T \mid h_T)\, p(D_T \mid h_T, x_T) \\
&\propto \frac{1}{h_T}\exp\left\{-\frac{1}{2}\left(\frac{\ln h_T - m_1}{s_1}\right)^2\right\}\frac{1}{h_T}\exp\left\{-\frac{1}{2}\left(\frac{x_T}{h_T}\right)^2\right\}p(D_T \mid h_T, x_T),
\end{aligned}
\tag{B.1}
$$

where now

$$
m_1 = \alpha + \delta \ln h_{T-1} \quad \text{and} \quad s_1^2 = \sigma^2. \tag{B.2}
$$

As before, the proposal density used is a lognormal with mean and standard deviation parameters

$$
m = \frac{m_1 s_1^{-2} + m_2 s_2^{-2}}{s_1^{-2} + s_2^{-2}} \quad \text{and} \quad s = \frac{1}{\sqrt{s_1^{-2} + s_2^{-2}}}, \tag{B.3}
$$

where m_2 and s_2 are the same as before.

Similarly, the target density for initial volatility, h_1, reflects the fact that there is no preceding volatility. If the volatility process is stationary, however, then the draw of h_1 should be consistent with the stationary distribution, which is lognormal. We therefore write the target density as the product of the unconditional density and some more familiar terms:

$$\begin{aligned} f_1(h_1) &= p(h_1 \mid h_2, x_1, D_1) \\ &\propto p(h_1)p(h_2 \mid h_1)p(x_1 \mid h_1)p(D_1 \mid h_1, x_1) \\ &\propto \frac{1}{h_1}\exp\left\{-\frac{1}{2}\left(\frac{\ln h_1 - \alpha/(1-\delta)}{\sigma/\sqrt{1-\delta^2}}\right)^2\right\}\frac{1}{h_2}\exp\left\{-\frac{1}{2}\left(\frac{\ln h_2 - \alpha - \delta \ln h_1}{\sigma}\right)^2\right\} \\ &\quad \times \frac{1}{h_1}\exp\left\{-\frac{1}{2}\left(\frac{x_1}{h_1}\right)^2\right\}p(D_1 \mid h_1, x_1) \\ &\propto \frac{1}{h_1}\exp\left\{-\frac{1}{2}\left(\frac{\ln h_1 - m_1}{s_1}\right)^2\right\}\frac{1}{h_1}\exp\left\{-\frac{1}{2}\left(\frac{x_1}{h_1}\right)^2\right\}p(D_1 \mid h_1, x_1), \end{aligned} \tag{B.4}$$

where

$$m_1 = \frac{\alpha(1-\delta) + \delta(\ln h_{t+1} + \ln h_{t-1})}{1+\delta^2} \quad \text{and} \quad s_1^2 = \frac{\sigma^2}{1+\delta^2}. \tag{B.5}$$

Again, given the previous definitions of m_2 and s_2, the proposal density is lognormal with mean and standard deviation parameters

$$m = \frac{\alpha(1-\delta^2)}{1-\delta} + \delta(\ln h_2 - \alpha) \quad \text{and} \quad s = \sigma. \tag{B.6}$$

References

Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002. Range-based estimation of stochastic volatility models. Journal of Finance 57, 1047–1091.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905.

Duffie, D., Singleton, K., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929–952.

Gallant, A.R., Hsieh, D.A., Tauchen, G.E., 1997. Estimation of stochastic volatility models with diagnostics. Journal of Econometrics 81, 159–192.

Harvey, A., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Review of Economic Studies 61, 247–264.

Jacquier, E., Polson, N.G., Rossi, P.E., 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371–389.

Johannes, M., Polson, N.G., 2005. MCMC methods for financial econometrics. In: Aït-Sahalia, Y., Hansen, L. (Eds.), Handbook of Financial Econometrics.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. Review of Economic Studies 65, 361–393.

Mengersen, K.L., Tweedie, R.L., 1996. Rates of convergence of the Hastings and Metropolis algorithms. Annals of Statistics 24, 101–121.

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. Journal of Business 53, 61–65.

Tanner, M., Wong, W., 1987. The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82, 528–550.

Taylor, S.J., 1994. Modelling stochastic volatility. Mathematical Finance 4, 183–204.