A Gaussian Approximation Scheme for Computation of Option Prices in Stochastic Volatility Models [1]

Ai-ru (Meg) Cheng
Department of Economics
University of California
Santa Cruz CA 95064

A. Ronald Gallant
Fuqua School of Business
Duke University
Durham NC 27708-0120

Chuanshu Ji
Department of Statistics
and Operations Research
University of North Carolina
Chapel Hill NC 27599-3260

Beom S. Lee
Department of Mathematics
University of Alabama
Tuscaloosa AL, 35487-0350.

May 2007

[1]Research supported by National Science Foundation Grants SES0000176 and SES0438174.

Abstract

We consider European options on a price process that follows the log-linear stochastic volatility model. There are two stochastic integrals in the option pricing formula that are costly to compute. We derive a central limit theorem that provides approximations to these integrals. For thirty day at the money options with parameter settings appropriate to foreign exchange data, our formulas improve computation speed by a factor of 1000 over brute force Monte Carlo at comparable accuracy. This improvement in speed gets computational efficiency to the point where Markov chain Monte Carlo (MCMC) statistical methods are feasible. We verify that MCMC statistical methods perform well on simulated joint price and options data that follow the log-linear stocastic volatility model. We provide estimates of model parameters from daily data on the Swiss Franc to Euro and Japanese Yen to Euro over the period 1999 to 2002.

Keywords: central limit theorem, option pricing, stochastic volatility, foreign exchange, Markov chain Monte Carlo.

1 Introduction

Stochastic volatility models are used extensively in economics and finance in a variety of contexts. Some early applications are Clark (1973), Taylor (1982), Tauchen and Pitt (1983), and Gallant, Hsieh, and Tauchen (1991). A review is provided by Ghysels, Harvey, and Renault (1996). When the application permits flexibility, the most common specification is the log-linear form as, for example, in Danielson (1994), Geweke (1994), Kim, Shephard, and Chib (1998), Gallant, Hsu, and Tauchen (1999), and Jacquier, Polson and Rossi (2004). Of special interest are Yu (2002), which examines the out-of-sample forecasting performance of the log-linear stochastic volatility model, Chernov, Gallant, Ghysels, and Tauchen (2003), which discusses alternative continuous-time stochastic volatility specifications and compares several to the log-linear specification using daily data on the Dow-Jones Industrial Average, and Scott (1987), which discusses the log-linear specification in connection with option pricing.

An important and computationally intesive application of the log-linear stochastic volatility model is the simultaneous extraction of volatility from a multivariate time series consisting of a returns series and option prices on the return series. Simultaneous extraction, where the specification and parameters of the volatility process of the returns series are forced to coincide with those of the volatility process used to price the options, both maximizes the statistical efficiency of volatility extraction and permits recovery of risk premia. For the arguments supporting this claim and discussion of the various statistical estimation algorithms available, see Chernov and Ghysels (2000), Pan (2002), Jones (2003), Polson and Stroud (2003), Eraker (2004), and Aït-Sahalia and Kimmel (2007).

Our work concerns estimating stochastic volatility models using financial data from the two aforementioned sources: underlying asset returns and option prices. Adding option data creates a significant computational challenge for the following reason. In a stochastic volatility model, a volatility time series consists of latent variables h_t, $t = 0, 1, \ldots, T$, where h_t is, e.g., the logarithm of volatility. Estimation of the series $\{h_t\}$ and the multidimensional parameter θ involved in a stochastic volatility model typically requires use of a Monte Carlo based algorithm such as efficient methods of moments (EMM) or Markov chain Monte Carlo

(MCMC). Those algorithms yield updated values $\theta^{(m)}$ and $h_t^{(m)}$, $t = 0, 1, \ldots, T$ in the mth iteration, $m = 1, \ldots, M$. In particular, as an important part of the mth iteration in fitting a stochastic volatility model, an option price $f_t^{(m)}$ is calculated by using a general Black-Scholes pricing formula based on the stochastic volatility model with the current parameter value $\theta^{(m-1)}$ and volatility value $h_t^{(m-1)}$ and is compared to the observed option price C_t. The comparison will lead to the adjustment of $h_t^{(m-1)}$ to $h_t^{(m)}$ for $t = 0, 1, \ldots, T$ and also the updating of $\theta^{(m-1)}$ to $\theta^{(m)}$. The generalzed Black-Scholes pricing formula (Garcia, Ghysels, Renault, 2007) for $f_t^{(m)}$ requires high-dimensional numerical integration to compute a conditional expectation over the space of sample paths of future volatility. This has to be done for every t and every m. The required computational time quickly adds up with large values of T (several hundred days) and M (typically more than $100,000$ iterations).

So far there have been two ways to handle the computation in practice. The first way is to use "brute force" simulation, i.e. to generate a large number of future volatility sample paths and average over these paths. Brute force can be accelerated somewhat using techniques such as antithetic variates. Even though we may neglect to mention it at times, all our "brute force" computations do use antithetic variates.

The second way is to abandon the log-linear specification and use the affine stochastic volatility model proposed in Heston (1993). Heston's closed-form option pricing formula avoids most of the numerical integration involved in non-affine models such as the log-linear specification considered here. Pan (2002) and Eraker (2004) add jumps to Heston's model.

Some approaches do not match option prices but rather rely on the the Chicago Board Options Exchange Market Volatility Index (VIX) as a direct measure of the latent volatility process. Jones (2003) follows this approach and stays within the CEV class. Aït-Sahalia and Kimmel (2007) analytically approximate the joint likelihood of prices and volatility. The approximations are derived form forward and backward Fokker-Plank-Kolmogorov partial differential equations. These approaches rely on the assumption that the VIX, which is an average of eight near-term near-the-money Black-Scholes implied volatilities from options on the S&P 100 index, is an adequate proxy to the latent volatility process defined by a stochastic volatility model. They explore the adequacy of this approximation within the CEV family and find that it depends mildly on the model considered. Of course, using option prices, as we do here, avoids the assumption entirely.

In this paper, we develop an integration scheme that is practicable for use in pricing options on a returns process represented by a log-linear stochastic volatility model. We propose an integration scheme based on a Gaussian approximation that is proved here. Its accuracy is verified by comparison to "brute force" Monte Carlo integration schemes. In principle the ideas are more generally applicable, but we do not make these extensions here.

In order to illustrate the ideas, we study the dynamics of two foreign exchange returns processes – Japanese Yen per Euro (JPY/EUR) and Swiss Franc per Euro (CHF/EUR) – and options on them. For this application, Melino and Turnbull (1990) argue that option prices based on a stochastic volatility model are more accurate than Black-Scholes prices and Jacquier, Polson and Rossi (2004), who fit only the returns process using statistical methods similar to those proposed here, provide estimates and characterizations against which we can cross check our results. One of our findings is that, indeed, incorporating the option price information provides more precise parameter estimates than using the information embedded in the return series alone, confirming the findings of Chernov and Ghysels (2000), Pastorello, Renault, and Touzi (2000), Polson and Stroud (2003), and Eraker (2004).

The rest of this paper is organized as follows: Section 2 describes the log-linear stochastic volatility model under the physical measure (P-measure), derives the transformation to the risk-neutral measure (Q-measure), and displays formulas for the computation of European options in terms of the risk neutral measure. Section 3 derives our numerical integration scheme. Section 4 implements MCMC estimation methods that use the scheme for data on returns alone and joint data on returns and options. Section 5 describes and reports simulation results. Section 6 presents empirical results for two foreign exchange applications. Section 7 concludes.

2 Option Prices for Stochastic Volatility Models

This section describes the model under the objective measure, derives the risk-neutral measure, and sets forth formulas for option price in terms of the risk-neutral measure.

2.1 The Stochastic Volatility Model

Let $S = \{S_t\}$ denote a continuous-time process that describes the evolution of the price of an asset. We presume that its logarithm follows the following stochastic volatility model:

$$
\begin{aligned}
y_t &= \log(S_t), &(1)\\
dy_t &= \mu\, dt + \exp\left(\frac{1}{2}\, h_t\right) \left[\sqrt{1-\rho^2}\, dW_{1t} + \rho\, dW_{2t}\right], &(2)\\
dh_t &= (\alpha + \beta h_t)\, dt + \sigma\, dW_{2t}, &(3)
\end{aligned}
$$

where $W = \{W_t\} = \{(W_{1t}, W_{2t})\}$ is a standard two-dimensional Wiener process defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $\{\mathcal{F}_t : 0 \leq t \leq T\}$ denote the filtration generated by W. The parameter ρ is the correlation of the asset return process and the volatility factor process. This model is analogous to the discrete-time logarithmic first order autoregressive stochastic volatility model of Jacquier, Polson and Rossi (2004).

2.2 Equivalent Martingale Measure

To price an option on S we follow established convention (Pan, 2002; Jones, 2003; Polson and Stroud, 2003; Eraker, 2004; etc.) and specify risk premia such that the stochastic differential equation describing the returns process has the same functional form under both the physical and risk neutral measures. They are

$$
\begin{aligned}
\tilde{\nu}_t &= \nu_1 + \nu_2 h_t, &(4)\\
\lambda_t &= \frac{1}{\sqrt{1-\rho^2}} \left\{ \exp\left(-\frac{1}{2}\, h_t\right) \left[(\mu - r) + \frac{1}{2} \exp\left(h_t\right)\right] - \rho \tilde{\nu}_t \right\}, &(5)
\end{aligned}
$$

where r is the short rate, presumed constant; $\tilde{\nu}_t$ is the risk premium process associated with $\{W_{2t}\}$ and λ_t is the risk premium process associated with $\{W_{1t}\}$. The time-t price of a contingent claim $g(S_{t'})$ (with $t' > t$) is the conditional expectation $\mathrm{E}^Q\left[e^{-r(t'-t)} g(S_{t'}) \mid \mathcal{F}_t\right]$ under the risk neutral measure Q, which can be expressed either as the Radon-Nikodým derivative with respect to the physical measure P, which is

$$
\xi_t = \left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \exp\Big(-\int_0^t \lambda_u\, dW_{1u} - \int_0^t \tilde{\nu}_u\, dW_{2u} - \frac{1}{2}\int_0^t \lambda_u^2\, du - \frac{1}{2}\int_0^t \tilde{\nu}_u^2\, du\Big), \qquad (6)
$$

or as a system of stochastic differential equations in terms of the Wiener processes

$$
W_{1t}^Q = W_{1t} + \int_0^t \lambda_u du, \qquad (7)
$$

$$W_{2t}^{Q} = W_{2t} + \int_0^t \tilde{\nu}_u du, \tag{8}$$

which is

$$dS_t = rS_t dt + \exp\left(\frac{1}{2}h_t\right) S_t \left[\sqrt{1-\rho^2} dW_{1t}^{Q} + \rho dW_{2t}^{Q}\right], \tag{9}$$

$$dh_t = \left[\alpha - \nu_1\sigma + (\beta - \nu_2\sigma)h_t\right] dt + \sigma dW_{2t}^{Q}. \tag{10}$$

2.3 European options

The price of a European call option under a stochastic volatility model can be computed as the conditional expectation under the risk-neutral measure Q of the call option price calculated according to the Black-Scholes (1973) formula. Let $\theta = \{\mu, \alpha, \beta, \sigma, \rho\}$ and $\nu = \{\nu_1, \nu_2\}$ denote the parameters of the stochastic volatility model described in (1)–(3) and (9)–(10). Then, as shown by Romano and Touzi (1997), for maturity (expiration) date T^* and strike price K, a European call option price can be computed as

$$C(S_t, h_t, \theta, \nu) = \mathrm{E}^{Q}\left[C^{BS}\left(S_t \exp Z_{t,T^*}, \overline{V}_{t,T^*}, r, K, T^* - t\right) | \mathcal{F}_t\right], \tag{11}$$

$$Z_{t,T^*} = \rho \int_t^{T^*} \exp\left(\frac{1}{2}h_u\right) dW_{2u}^{Q} - \frac{1}{2}\rho^2 \int_t^{T^*} \exp(h_u)\, du, \tag{12}$$

$$\overline{V}_{t,T^*} = \frac{1-\rho^2}{T^* - t} \int_t^{T^*} \exp(h_u)\, du, \tag{13}$$

where $C^{BS}(\cdot)$ is the Black-Scholes call option pricing formula; specifically,

$$C^{BS}(s, v, r, k, d) \tag{14}$$

$$= s\Phi\left(\frac{\log(s/k) + (r + v^2/2)\, d}{v\sqrt{d}}\right) - ke^{-rd}\Phi\left(\frac{\log(s/k) + (r + v^2/2)\, d}{v\sqrt{d}} - v\sqrt{d}\right),$$

where Φ is the cumulative standard normal distribution function. Notice that with constant volatility and $\rho = 0$, (11) is the original Black-Scholes call option price. A put option can be priced using the parity relation between a put (P_t) and a call (C_t), which is $P_t = C_t - S_t + K \exp[-r(T^* - t)]$.

3 A Gaussian Approximation for Pricing Options

We approximate the option price $C(S_t, h_t, \theta, \nu)$ in two stages. Stage 1 is a Euler discretization approximation, $U_n\ \Delta \approx \int_t^{T^*} e^{h_u}\, du$ and $V_n\ \sqrt{\Delta} \approx \int_t^{T^*} e^{h_u/2}\, dW_{2u}^{Q}$, to the two integrals appearing in (12)–(13), where

$$U_n = \sum_{j=0}^{n-1} e^{h_{t+j\Delta}}, \tag{15}$$

$$V_n = \sum_{j=0}^{n-1} e^{\frac{1}{2}h_{t+j\Delta}} \epsilon_{t+(j+1)\Delta}, \tag{16}$$

$\epsilon_{t+(j+1)\Delta}$ are iid $N(0,1)$, and $n = (T^* - t)/\Delta$. Stage 2 is an approximation to the distribution of (U_n, V_n) by a Gaussian random vector (U, V) obtained via a central limit argument. The error in Stage 1 is of order $O(\Delta)$ for small Δ and fixed $T^* - t$. Stage 2 accepts this Δ and then requires large $n = (T^* - t)/\Delta$ for good accuracy. The error in Stage 2 is of order $O(n^{-\frac{1}{2}} \log n)$ (Dam, 1998). To get the option price, the quantity $C^{BS}\left(S_t \exp Z_{t,T^*}, \overline{V}_{t,T^*}, r, K, T^* - t\right)$, which appears in (11), is integrated with respect to the Gaussian distribution with appropriately scaled (U, V) replacing $\int_t^{T^*} e^{h_u}\, du$ and $\int_t^{T^*} e^{h_u/2}\, dW_{2u}^Q$ in (12)–(13).[1]

Stage 2 is not a fill-in asymptotic result because, while decreasing Δ does increase n, it also increases the correlation between adjacent points so that reducing Δ does not necessarily improve the accuracy of the approximation to (12)–(13). In applications to data, the available option expiries are limited so that Δ is the only choice variable. It then becomes an issue of whether or not accuracy in approximating $C(S_t, h_t, \theta, \nu)$ is good enough for the chosen Δ, which can be checked in a given application as we do in ours.

The obvious (and usual) approach to computing (11) is to discretize (9)–(10), generate N sample paths according to the recursions implied by the discretization, evaluate (11)–(13) over each sample path, and average (11) over the N repetitions. This is how we check the accuracy of our proposed Gaussian scheme approach in Section 5. Duffie and Glynn (1995) relates to this computation. They studied the trade-off between small Δ and large N for discretization schemes such as Euler, Milshtein, or Talay. Decreasing Δ reduces bias while increasing N reduces variance. Thus one can approximate to arbitrary accuracy and one can choose a pair $(\Delta,\, N)$ that minimizes computation time for given accuracy. For most purposes, even the minimizing choice is too computationally intensive to be practicable. From this perspective, what we propose is to replace the N Monte Carlo repetitions with a Gaussian approximation with Δ chosen to provide acceptable accuracy. Our proposal

[1] The explicit formula is (29) expressed in terms of the discretization (17), Black-Scholes formula (14), and equations (18)–(22) of Theorem 1.

follows.

The discretization of the volatility process under the risk-neutral measure Q, i.e. (10), by means of an Euler scheme (Kloeden and Platen, 1992, p.341) yields the recursion

$$h_{t+\Delta} = a + b\,h_t + c\,\epsilon_{t+\Delta}, \tag{17}$$

where $a = (\alpha - \nu_1\sigma)\Delta$, $b = 1 + (\beta - \nu_2\sigma)\Delta$, $c = \sqrt{\Delta}\sigma$, and $\epsilon_{t+\Delta}$ is an independent $N(0,1)$ random variable. Notice that our notation departs from convention because we have incorporated Δ into the expressions for a, b, and c in order to simplify later formulas. For fixed t, let $\mathrm{E}_t(\,\cdot\,)$, $\mathrm{Var}_t(\,\cdot\,)$ and $\mathrm{Cov}_t(\,\cdot\,)$ denote the conditional expectation, variance and covariance operators, respectively, given h_t.

Theorem 1 [2] *Assume $|b| < 1$, fix $\Delta > 0, t > 0$, and an arbitrary initial state h_t. As $n \to \infty$, the limiting distribution of $n^{-1/2}\,(U_n - \mathrm{E}_t U_n,\ V_n - \mathrm{E}_t V_n)$ conditioning on h_t is a bivariate normal distribution with mean $(0,0)$ and covariance matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ provided $a_{11}a_{22} > a_{12}^2$, with*

$$\begin{aligned} a_{11} &= \lim_{n\to\infty} n^{-1}\mathrm{Var}_t(U_n), \\ a_{12} &= a_{21} = \lim_{n\to\infty} n^{-1}\mathrm{Cov}_t(U_n,\ V_n), \\ a_{22} &= \lim_{n\to\infty} n^{-1}\mathrm{Var}_t(V_n); \end{aligned}$$

where

$$\mathrm{E}_t U_n = \sum_{i=0}^{n-1} \exp\left[\frac{a(1-b^i)}{1-b} + b^i h_t + \frac{c^2(1-b^{2i})}{2(1-b^2)}\right]; \tag{18}$$

$$\mathrm{E}_t V_n = 0; \tag{19}$$

$$\mathrm{Var}_t(U_n) = \sum_{i=0}^{n-1} \mathrm{Var}_t(e^{h_{t+i\Delta}}) + 2\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} \mathrm{Cov}_t(e^{h_{t+i\Delta}}, e^{h_{t+j\Delta}}), \tag{20}$$

$$\begin{aligned} \mathrm{Cov}_t(U_n,\ V_n) = \sum_{i=1}^{n-1}\sum_{j=0}^{i-1} c\ b^{i-j-1}\ \exp\Bigg[&\left(b^i + b^j/2\right) h_t + \frac{a\,(3/2 - b^i - b^j/2)}{1-b} \\ &+ \frac{c^2\,(5/4 - b^{2i} - b^{2j}/4 + b^{i-j} - b^{i+j})}{2(1-b^2)}\Bigg] \end{aligned} \tag{21}$$

$$\mathrm{Var}_t(V_n) = \mathrm{E}_t U_n; \tag{22}$$

with

$$\begin{aligned} &\mathrm{Var}_t(e^{h_{t+i\Delta}}) \\ &\quad = \exp\left(2b^i h_t + \frac{2a(1-b^i)}{1-b} + \frac{c^2(1-b^{2i})}{1-b^2}\right)\left[\exp\left(\frac{c^2(1-b^{2i})}{1-b^2}\right) - 1\right], \end{aligned}$$

[2]Details of the proof are provided in Appendix A.

$$
\begin{aligned}
&\mathrm{Cov}_t(e^{h_{t+i\Delta}}, e^{h_{t+j\Delta}}) \\
&\quad = \exp\left((b^i + b^j)h_t + \frac{a(2 - b^i - b^j)}{1 - b} + \frac{c^2(2 - b^{2i} - b^{2j})}{2(1 - b^2)}\right) \\
&\qquad \times \left[\exp\left(\frac{c^2(b^{j-i} - b^{j+i})}{1 - b^2}\right) - 1\right].
\end{aligned}
$$

4 Estimation

The estimation strategy follows Eraker (2004). The idea is to discretize a diffusion model such as (1)–(3) according to an Euler scheme (Kloeden and Platen, 1992, p.341) which provides a likelihood. With a likelihood available, several estimation strategies become feasible. Our choice is to use a Bayesian inference strategy implemented with MCMC methods because it is computationally stable and efficient for the class of problems considered here. MCMC methods generate a discrete-time Markov chain whose stationary distribution is the joint posterior distribution of model parameters and latent variables, hence samples from the posterior distribution of any marginal are easy to obtain by selecting elements from the joint series.

Discretization to get a likelihood can introduce bias. This bias may be eliminated by using any one of the fill-in methods proposed by Elarian, Chib, and Shephard (2001), Eraker (2001), or Durham and Gallant (2002). But as verified by Eraker (2004), for the type of data and class of models considered here, accuracy is adequate for a Δ of one day so there is no need to use fill-in methods if data sampled at a daily frequency are available.

The method proposed by Eraker (2004) is a “Metropolis within Gibbs” approach. In the manner of Gibbs, one block of parameters is sampled conditional on the rest and one cycles through the blocks. But, because the normalizing constants for the conditional distributions involved are not known, a Metropolis sampler is used to generate the draws. The conditional distributions themselves are obtained by using the fact that when the data and a block of parameters are held fixed, the conditional density for the remaining parameters is proportional to the likelihood times the prior. The proposal densities required for the block moves must be obtained by analyzing the problem. This, regrettably, is an art.

4.1 Estimation without options

Consider a time series of the fundamental asset returns Y, a set of latent state variables h, and the model parameters θ under the objective measure P. Because we will take Δ to be the sampling frequency of the data, Y has elements $y_{t+i\Delta}$ for $i = 0, \ldots, n_y - 1$ with distribution determined by (2)–(3). Similarly, the elements $h_{t+i\Delta}$ of h are determined by (3). The parameters are $\theta = \{\mu, \alpha, \beta, \rho, \sigma\}$.

The posterior distribution for parameters and latent variables is the conditional density $P(h, \theta \,|\, Y)$. Holding Y fixed, this density is proportional to the likelihood times the prior

$$P(h, \theta \,|\, Y) \propto P(Y, h \,|\, \theta)\, P(\theta) = P(Y \,|\, h, \theta)\, P(h \,|\, \theta)\, P(\theta) \tag{23}$$

with

$$P(Y|h, \theta)\ P(h|\theta) \approx \prod_{i=0}^{n-1} p(y_{t+(i+1)\Delta}|y_{t+i\Delta}, h_{t+i\Delta}, \theta)\ p(h_{t+(i+1)\Delta}|h_{t+i\Delta}, \theta), \tag{24}$$

where $p(y_{t+(i+1)\Delta}|y_{t+i\Delta}, h_{t+i\Delta}, \theta)$ is the normal density of $y_{t+(i+1)\Delta}$ with the conditional mean $\mu\Delta + y_{t+i\Delta}$ and conditional variance $e^{h_{t+i\Delta}}\Delta$, and $p(h_{t+(i+1)\Delta}|h_{t+i\Delta}, \theta)$ is the normal density of $h_{t+(i+1)\Delta}$ with the conditional mean $\alpha\Delta + (1 + \beta\Delta)h_{t+i\Delta}$ and conditional variance $\sigma^2\Delta$, following the discretization of (2)–(3).

The Gibbs blocking is as follows. The elements within the block of h are moved one at a time. The coordinate variable $h_{t+i\Delta}$ at each i is generated from the conditional density $p(h_{t+i\Delta} \,|\, h_{t+(i-1)\Delta}, h_{t+(i+1)\Delta}; y_{t+(i-1)\Delta}, y_{t+i\Delta}, y_{t+(i+1)\Delta}; \theta)$ due to the Markovian structure in the dynamics of (2)–(3). The drift parameters μ and $\{\alpha, \beta\}$ are moved as two separate blocks and the diffusion parameters $\{\sigma, \rho\}$ as another. The priors are uniform on their supports with the exception for μ which has a conjugate normal prior, allowing μ to be updated directly without Metropolis-Hastings acceptance/rejection intermediation. All priors are proper with boundary parameters set in the experiments to be effectively uninformative. The details are in Appendix B.

The single-move Gibbs sampler is not as efficient as some other Bayesian MCMC inference methods for estimating stochastic volatility models using only returns data Y. In particular, the multi-move Gibbs sampler with Kalman filters introduced in Kim, Shephard, and Chib (1998), Pitt and Shephard (1999), Chib, Nardari, and Shephard (2002) is more efficient computationally. We use the single-move Gibbs sampler in this paper for estimating stochastic volatility models based on both returns and option data. To the best of our knowledge, no multi-move MCMC methods are available for that task. Given that we must use it for that purpose, we also use it for returns alone to reduce our coding and debugging burden.

4.2 Estimation with options

We now consider the additional information provided by options on Y. The option prices C we consider are on homogeneous instruments sampled at the same time as $y_{t+i\Delta}$ so that C has elements $c_{t+j\Delta}$ for $j = 0, 1, \ldots, n_c - 1$, where $c_{t+j\Delta}$ is the logarithm of the observed option price.

As is common in derivative pricing applications (Polson and Stroud, 2003; Eraker, 2004), the distribution of C is determined by both an option pricing formula, which is (11) in our case, and an assumption of pricing errors. Pricing errors are not only required to permit application of Bayesian and similar likelihood based methods but are also plausible reflections of bid-ask bounce and similar micro structure considerations (Renault, 1997; Jacquier and Jarrow, 2000). We assume that errors are additive in logs,

$$c_{t+j\Delta} = f_{t+j\Delta} + \delta\widetilde{\epsilon}_j, \quad j = 0, 1, ..., n_c - 1, \tag{25}$$

where, from (11),

$$f_{t+j\Delta} = \log\left[C\left(S_{t+j\Delta}, h_{t+j\Delta}, \theta, \nu\right)\right], \tag{26}$$

and the errors $\widetilde{\epsilon}_j$, $j = 0, 1, ..., n_c - 1$ are iid $N(0, 1)$ random variables, independent of the processes S and h. Note that (25) treats observed option data as a noisy version of the option prices predicted based on a corresponding pricing formula. In Jones (2003), a different but somewhat equivalent approach is taken by considering the implied volatility data (VIX) to be a noisy version of the latent volatility variables.

There are several unknowns now, including $\theta = \{\mu, \alpha, \beta, \rho, \sigma\}$ and h that appear in the returns dynamics (2)–(3), the risk premia parameters $\nu = \{\nu_1, \nu_2\}$ that appear in the risk-neutral dynamics (9)–(10), and the pricing error standard deviation δ from (25).

We assume the three blocks $\{Y, h, \theta\}$, ν, and δ are mutually independent. Under this assumption, the joint posterior distribution of interest becomes

$$P(h, \theta, \nu, \delta | C, Y) \propto P(C | Y, h, \theta, \nu, \delta) P(Y | h, \theta) P(h | \theta) P(\theta) P(\nu) P(\delta). \tag{27}$$

We have expressions for $P(Y|h,\theta)$ and $P(h|\theta)$ from Subsection 4.1. From (25), the likelihood of the option prices is

$$P(C|Y, h, \theta, \nu, \delta) \propto \prod_{j=0}^{n_c - 1} \delta^{-1} \exp\left(-\frac{\left(c_{t+j\Delta} - f_{t+j\Delta}\right)^2}{2\,\delta^2}\right) \tag{28}$$

The Gibbs blocking is as follows. As in Subsection 4.1, the elements $h_{t+i\Delta}$ of h are moved one at a time, the drift parameters μ and $\{\alpha, \beta\}$ are moved as two separate blocks, and the diffusion parameters $\{\sigma, \rho\}$ as another. Priors for these are as in Subsection 4.1. The risk premium parameters ν and the pricing error parameter δ are moved as two separate blocks. The prior on ν is uniform. The prior for δ is the conjugate, so that δ can be directly sampled from a product of inverse gammas without Metropolis-Hastings acceptance/rejection intermediation. All priors are proper with parameters set to be effectively uninformative. The details are in Appendix B.

5 Simulation study

As a theoretical result, Theorem 1 given in Section 3 requires large $T^* - t$. In practice, large values of $T^* - t$ are restricted to long-term options and moments become harder to compute accurately for large values of $T^* - t$ due to the powers involved in the moment expressions. For that concern, we perform a series of simulation studies in this section to check the applicability of the Gaussian approximation scheme for different terms to maturity.

Table 1 shows the accuracy of option pricing computations that use brute force Monte Carlo and the accuracy of methods based on the Gaussian scheme (GS) for various maturities and strike prices. To construct the table, a bivariate time series of security prices S and volatilities h of length 600, which is the (rounded) length of our data, were generated according to equations (9)–(10). The parameter settings are given in the table legend and are roughly the same as found in our data except for a risk free rate of zero. At this point, S, h, and all parameter values are known. Then 600 option prices that are regarded as the true values were computed for each of the strike prices and maturities shown in the table using brute force Monte Carlo with 1,000,000 repetitions.

Brute force Monte Carlo uses a simulation of equation (10) from the current date to expiry to approximate the integrals in equations (12) and (13). With these, C^{BS} in equation (11) can be computed. $\mathrm{E}^Q(\cdot)$ is computed by averaging C^{BS} over Monte Carlo repetitions of these computations. The notation BF(N) in the table indicates that the number of Monte Carlo repetitions is N. We use an antithetic variates method to reduce variance in our Monte Carlo computations. The antithetic variates method uses a pair of draws from the elemental random nummber generator of opposite sign. This pair counts as two draws in reporting N because drawing the random numbers themselves is a negligable cost. Computation of C^{BS} for a given path is the costly part of the computation.

To implement the Gaussian scheme, substitute U_n and V_n into equation (11) using equations (12), (13), (15), and (16) to obtain the function

$$C(u,v) = C^{BS}\left[S_t \exp\left(\rho v - \frac{1}{2}\rho^2 u\right), \left(\frac{1-\rho^2}{T^*-t}\right)u, r, K, T^*-t\right], \tag{29}$$

where $C^{BS}(s,v,r,k,d)$ is the Black-Scholes formula given by (14). The integration to be performed is

$$\iint C(u,v)\, n(u,v \,|\, m, A)\, du\, dv, \tag{30}$$

where $m = (\mathrm{E}_t U_n,\ \mathrm{E}_t V_n)$ and A are the mean vector and the covariance matrix of the bivariate normal density $n(u,v \,|\, m, A)$ with $\mathrm{E}_t U_n$, $\mathrm{E}_t V_n$, and A given by equations (18)–(22) of Theorem 1.

GS Monte Carlo uses N draws of (U_n, V_n) from the bivariate Gaussian distribution with mean vector and covariance matrix given by equations (18)–(22) to compute the integral (30) by averaging $C(U_n, V_n)$ over the N draws. If a draw of U_n is negative, then we discard that C^{BS} and repeat. The notation GS(N) in the table indicates that the number of Monte Carlo repetitions is N. An antithetic variates technique is applied so that one single draw from a bivariate standard Gaussian distribution generates a pair of samples from the distribution of (U_n, V_n).

GS quadrature uses a Gauss-Hermite rule to compute the integral (30) directly. It is implemented as follows. Factoring A as $A = RR'$, (30) can be rewritten in a form suitable for Gauss-Hermite quadrature as

$$\iint \pi^{-1} C\left(m + \sqrt{2}\ tR\right) \exp(-t_1^2 - t_2^2)\, dt_1\, dt_2 \tag{31}$$

where $t = (t_1, t_2)$. A Gauss-Hermite rule has the form

$$\int \phi(s) \exp(-s^2)\, ds \approx \sum_{i=1}^{L} W_i \phi(s_i). \tag{32}$$

The L abscissae s_i and L weight factors W_i can be obtained from tabulations such as Table 25.10 of Abramowitz and Segun (1964) or can be computed as needed using an algorithm due to Golub (1973); see also Golub and Welsh (1969). Thus, we would use

$$\iint C(u,v)\, n(u,v \,|\, m, A)\, du\, dv \approx \sum_{i=1}^{L} \sum_{j=1}^{L} \pi^{-1} C\left(m + \sqrt{2}\ sR\right) W_i W_j \tag{33}$$

where $s = (s_i, s_j)$ except for the fact that the first element of $m+\sqrt{2}\ sR$ must be positive. We deal with this as follows. Let $\mathcal{S}$ denote the set of indexes (i,j) of the double sum appearing in equation (33) for which the first element of $m + \sqrt{2}\ sR$ is positive. Then the rule we actually use is

$$\iint C(u,v)\, n(u,v \,|\, m, A)\, du\, dv \approx \frac{\sum_{(i,j)\in\mathcal{S}} C\left(m + \sqrt{2}\ sR\right) W_i W_j}{\sum_{(i,j)\in\mathcal{S}} W_i W_j} \tag{34}$$

(The π in (33) arises from $\sum\sum W_i W_j = \pi$. Therefore it does not appear in (34)). In the work reported here, we use a five point rule; i.e. $L = 5$. These are reported in lines labeled GS(Q) in Table 1.

When quadrature is used, computation of the double sums in equations (20) and (21) become the main computational burden. If one plots the elements $S(i,j)$ of the sum as a function defined on the integers (i,j) one observes that it is quite smooth. This suggests that one could approximate by a polynomial and sum that instead. Consider equation (21). Our strategy is to select four equally spaced integers $i_1 = 1 < i_2 < i_3 < i_4 = n-1$, get the marginal totals $S(i_k) = \sum_{j=0}^{i_k - 1} S(i_k, j)$, fit the cubic

$$S(i_k) = a_0 + a_1(i_k) + a_2(i_k)^2 + a_2(i_k)^3, \tag{35}$$

and compute the sum by using the expression

$$\sum_{i=1}^{n-1} S(i) = (i_4)\, a_0 + \left(\frac{i_4^2}{2} + \frac{i_4}{2}\right) a_1 + \left(\frac{i_4^3}{3} + \frac{i_4^2}{2} + \frac{i_4}{6}\right) a_2 + \left(\frac{i_4^4}{4} + \frac{i_4^3}{2} + \frac{i_4^2}{4}\right) a_3. \tag{36}$$

Similarly for (20) using $S(i_k - 1)$ instead of $S(i_k)$ to adjust for the fact that the index starts at one rather than zero. (The same idea can be used to compute the marginal totals $S(i_k) = \sum_{j=0}^{i_k-1} S(i_k, j)$ and we do so.) Results for quadrature using these approximations to the double sums to compute the elements A of equation (30) are reported as GS(QI) in Table 1.

As a practical matter, only GS(1000), GS(Q), and GS(QI) are fast enough for use in the estimations reported in Section 6. As seen from Table 1, their overall accuracy is roughly the same. The entries in Table 1 are the root mean square error in the approximation to log option prices over the 600 option prices available for each entry in the table. These can be roughly interpreted as the average percent errors after multiplication by 100.

The time in seconds averaged over ten repetitions to compute an at the money option price using each of the methods used in Table 1 is shown as Table 2. As just stated and as indicated in Table 2, only GS(1000), GS(Q), and GS(QI) are fast enough to be practicable. Of these, GS(QI) is by far the fastest and the one that we shall use in Section 6.

The effectiveness of the estimation strategy proposed in Section 4 is examined in Table 3. Table 3 reports posterior means and standard deviations for estimation carried out as described in Subsections 4.1 and 4.2 for simulated data on at-the-money call options and the underlying. The simulated options prices are computed by "brute force" and the estimation involving options is conducted using the Gaussian scheme as described in the table legend. One concludes from the entries in the table that the GS(QI) scheme is adequate for application in an estimation context. Also, the finding of Jacquier and Jarrow (2000), Eraker, Johannes, and Polson (2003) and Eraker (2004) that estimation of model parameters from joint data on options and the underlying returns series rather than returns alone substantially reduces the standard deviations of estimates is verified.

6 Empirical Study

In this section, we apply the methods developed to currency price and options data.

6.1 Data

The exchange rate data are daily closing rates as determined by trades from January 1999 through December 2002 on the Japanese Yen to Euro (JPY/EUR) and Swiss Franc to Euro (CHF/EUR) obtained from the Pacific Exchange Rate Service web site (`fx.sauder.ubc.ca`). The returns computed from these series are geometric, i.e., $r_t = 100 \log(p_t/p_{t-1})$, where p_t is the closing quote. The sample size is 1003 for each. Summary statistics for these two series are shown in Table 4.

The options data are the average of the bid and ask closing quotes of over-the-counter (OTC) options written on JPY/EUR and CHF/EUR expressed as implied volatilities on a one month European contract with the strike price set to the one month forward currency rate, referred to as at-the-money-forward (ATMF) contracts hereafter. The sample period is September 2000 through December 2002 and the sampling frequency is daily, which yields a sample size of 592 for JPY/EUR and 594 for CHF/EUR. These data were kindly provided by Chien-te Hsu, CSFB. Because the OTC currency options are traded and quoted as implied volatilities, one needs to use a modified Black-Scholes formula to convert the data to the option premium (price). The convention for converting volatilities to prices in the OTC foreign exchange market is the Garman-Kohlhagen (1983) option pricing formula. Mathematically, the formula is identical to Merton's (1973) formula for options on dividend-paying stocks. One replaces the stock's dividend yield in Merton's formula by the risk free rate on the Euro and the risk free rate in Merton's formula by the risk free rate on the Japanese Yen for JPY/EUR or the Swiss Franc for CHF/EUR. Here, the 3-month LIBOR rates for the Euro, Japanese Yen, and Swiss Franc obtained from Datastream at the daily frequency were used as the risk-free rates in the formula. The exchange rate, exchange rate returns, ATMF implied volatilities, and ATMF option premiums are plotted over the sample period 2000–2003 for the Japanese Yen to Euro in Figure 1 and for the Swiss Franc to Euro in Figure 2. OTC options are used here rather than exchange-traded options because the OTC market is usually more liquid. The bid-ask spread on quoted implied volatilities for these contracts is about 4% (Chong, Ding, and Tan, 2003).

6.2 Empirical results based on returns data

Table 5 reports the posterior means and standard deviations of the log SV model over the sample period of January 1999 through December 2002 computed as described in Subsection 4.2.

The long-term mean of the volatility of the JPY/EUR series is $-\frac{\alpha}{\beta} = -0.6078$ which is higher than the value -3.3572 for the CHF/EUR series. The volatility persistence parameter β estimates of -0.1341 (autocorrelation of $1 + \beta = 0.8659$) for JPY/EUR and -0.1034 (autocorrelation 0.8966) for CHF/EUR are smaller and standard deviations are larger than those reported in Jaquier, Polson, and Rossi (2004). This seems to be an effect of the amount of information available. Later, in Section 6.3, we see that when the sample period is shortened estimates and standard deviations diverge further but that when options data is added deviations are much smaller and the estimates much closer to those reported by and Jaquier, Polson, and Rossi. The "leverage effect" parameter ρ is significantly negative in both series, indicating that the return and the volatility processes are negatively correlated during this sample period.

Figure 3 plots the returns on both exchange rate series and their annualized estimated volatility paths (often called smoothed volatility). Comparing the visual impression of volatility from the returns with the estimated volatility, the qualitative impression is that the estimated volatility is reasonable. Notice in particular that the effects of the event of September 11, 2001, on the Swiss Franc seems to have been captured by the volatility estimates.

6.3 Empirical results based on joint data

Table 6 reports the posterior means and standard deviations of the log SV model over the sample period of January 2000 through December 2002 both for returns alone, computed as described in Subsection 4.1, and for returns with options, computed as described in Subsection 4.2. For the JPY/EUR series there are 592 observations and for CHF/EUR series there are 594 observations.

As seen in Table 6 there are large discrepancies in the estimates computed with and without options data. Although the long-run means of the volatilities for each series with joint data (JPY/EUR: $-\frac{\alpha}{\beta} = -0.2548$; CHF/EUR: $-\frac{\alpha}{\beta} = -3.7990$) are reasonably close to their counterparts with returns data alone (JPY/EUR: $-\frac{\alpha}{\beta} = -1.0032$; CHF/EUR: $-\frac{\alpha}{\beta} = -3.2765$), taking into account the standard deviations; the estimates of the volatility mean reversion speed parameter β differ considerably, implying that the estimated spot volatility process is much more persistent when estimated from the joint data than when estimated from the returns data alone. As remarked in Section 6.2, this seems to be an effect of the amount of information available. When the series becomes longer or options data are added, estimates become more similar to those reported Jaquier, Polson, and Rossi (2004), who use a long time series.

Estimates of the volatility of the volatility σ are smaller with the joint data than using the returns data alone for both series. Notice that the "leverage effect" parameter ρ of the CHF/EUR series is negative, suggesting that the underlying return distribution is skewed toward right. This finding coincides with the unconditional moments reported in Table 4 and a plot of the risk-reversal premiums for these data (not shown).

Figure 4 plots the estimated spot volatility paths based on the joint data (dotted line) and returns data alone (solid line). We observe that the volatility paths estimated using joint data are visually less volatile and more persistent than their counterparts based on the returns data, which is consistent with the interpretation of the estimates in Table 6 immediately above.

The estimates for δ, which is the root mean square pricing error, are below the bid-ask bounce for these data.

The most striking observation that one draws from Table 6 is the large reduction in the standard deviations of model parameters when returns data are augmented by options data, which agrees with results reported by Chernov and Ghysels (2000), Ge (2000), Pastorello, et. al. (2000), Eraker (2004), and Polson and Stroud (2003).

It is well known that adequately representing daily returns dynamics with a continuous time model within the class of stochastic volatility models requires two factors: one to capture the persistence in volatility and one to thicken tails. Moreover, the process that thickens tails can be either an extra diffusion added to the volatility process or a jump process of one form or another. On this see Bates (1996), Eraker (2001), and Chernov, Gallant, Ghysels, and Tauchen (2003). From this perspective, our model is misspecified. It is interesting to see how this misspecification manifests itself.

To assess model specification, we plot the estimated residuals for the return and volatility processes based on the returns data alone and the joint data in Figures 5 and 6, respectively. In each panel is also shown the raw kurtosis for that series.

What one sees in comparing Figures 5 and 6 is that adding options causes the volatility process to become both more persistent and less able to add the variation that fattens the tails in returns. The fact that an extra factor is needed, be it a nearly white noise diffusion or a jump process, becomes fairly obvious.

It does not seem possible to decide whether the extra factor should be a jump process or a diffusion from data at the daily frequency because both models do about as well statistically (Chernov, Gallant, Ghysels, and Tauchen, 2003). Nonparametric methods using high frequency data seem to favor the presence of jumps (Huang and Tauchen, 2005).

7 Conclusion

A Gaussian approximation was established that can be used to compute European option prices on securities that follow the log-linear stochastic volatility models. It is efficient enough to apply MCMC methods to the estimation of model parameters from joint returns and options data.

Accuracy is adequate for one common application — foreign exchange options — as verified by simulation and two applications. Accuracy is relatively easy to check in a give application and one should.

Calculation of required moments in Gaussian approximations consumes nearly the entire cost of the algorithm. The simple interpolation strategy used in Section 5 (see (35)–(36)) reduces the computational time by at least an order of magnitude with slight sacrifice in accuracy. Possible future work includes more efficient methods for computing moments in Gaussian approximations. Accuracy could be further improved if conditional skewness and kurtosis can be calculated and included in Gaussian-like approximations (Edgeworth expansions). More tedious computation will be involved.

Another task we are engaged in concerns two-factor log-linear stochastic volatility models.

This direction appears more useful based on the comparison studies in Chernov, Gallant, Ghysels and Tauchen (2003). Two-factor log-linear stochastic volatility models perform much better wherein one slowly mean reverting factor provides volatility persistence and one rapidly mean reverting factor provides fat tails. The evidence gathered in Chernov, Gallant, Ghysels and Tauchen (2003) is only based on returns data (Dow Jones Industrial Average – DJIA). Addition of option data in similar studies will surely need efficient methods for computing option prices, such as Gaussian approximations introduced in this paper and appropriate modifications.

A Appendix

Proof of Theorem 1: For $i = 0, 1, 2, \ldots,$

$$h_{t+i\Delta} = a\,\frac{1-b^i}{1-b} + b^i h_t + c\sum_{m=1}^{i} b^{i-m}\epsilon_{t+m\Delta}. \tag{37}$$

Hence (18) follows from

$$\mathrm{E}_t U_n = \sum_{i=0}^{n-1} \mathrm{E}_t[\exp(h_{t+i\Delta})]$$

and

$$\mathrm{E}_t[\exp(h_{t+i\Delta})] = \exp\left[\frac{a(1-b^i)}{1-b} + b^i h_t + \frac{c^2(1-b^{2i})}{2(1-b^2)}\right].$$

To verify (20), note that for $0 \le i \le j \le n-1$,

$$\begin{aligned}
&\mathrm{E}_t(e^{h_{t+i\Delta}} e^{h_{t+j\Delta}}) \\
=\;& \mathrm{E}_t\left[e^{h_{t+i\Delta}}\ \mathrm{E}_{t+i\Delta}(e^{h_{t+j\Delta}})\right] \\
=\;& \exp\left[\frac{a(1-b^{j-i})}{1-b} + \frac{c^2(1-b^{2j-2i})}{2(1-b^2)}\right]\ \mathrm{E}_t\left\{\exp\left[(1+b^{j-i})h_{t+i\Delta}\right]\right\} \\
=\;& \exp\left\{\left[\frac{a(1-b^{j-i})}{1-b} + \frac{c^2(1-b^{2j-2i})}{2(1-b^2)}\right] + (1+b^{j-i})\left(\frac{a(1-b^i)}{1-b} + b^i h_t\right)\right. \\
&\quad \left. +(1+b^{j-i})^2\ \frac{c^2(1-b^{2i})}{2(1-b^2)}\right\} \\
=\;& \exp\left[(b^i+b^j)h_t + \frac{2a(1+b)(2-b^i-b^j) + c^2(2+2b^{j-i}-2b^{j+i}-b^{2i}-b^{2j})}{2(1-b^2)}\right];
\end{aligned}$$

hence,

$$\mathrm{Var}_t(e^{h_{t+i\Delta}})$$

$$
\begin{aligned}
&= \mathrm{E}_t(e^{2h_{t+i\Delta}}) - \Big[\mathrm{E}_t(e^{h_{t+i\Delta}})\Big]^2 \\
&= \exp\left[2b^i h_t + \frac{2a(1+b)(1-b^i) + 2c^2(1-b^{2i})}{1-b^2}\right] \\
&\quad - \exp\left[2b^i h_t + \frac{2a(1+b)(1-b^i) + c^2(1-b^{2i})}{1-b^2}\right] \\
&= \exp\left(2b^i h_t + \frac{2a(1-b^i)}{1-b} + \frac{c^2(1-b^{2i})}{1-b^2}\right)\ \left[\exp\left(\frac{c^2(1-b^{2i})}{1-b^2}\right) - 1\right],
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathrm{Cov}_t(e^{h_{t+i\Delta}}, e^{h_{t+j\Delta}}) \\
&= \mathrm{E}_t(e^{h_{t+i\Delta}} e^{h_{t+j\Delta}}) - \mathrm{E}_t(e^{h_{t+i\Delta}})\ \mathrm{E}_t(e^{h_{t+j\Delta}}) \\
&= \exp\left((b^i + b^j)h_t + \frac{a(2-b^i-b^j)}{1-b} + \frac{c^2(2-b^{2i}-b^{2j}+2b^{j-i}-2b^{j+i})}{2(1-b^2)}\right) \\
&\quad - \exp\left((b^i + b^j)h_t + \frac{a(2-b^i-b^j)}{1-b} + \frac{c^2(2-b^{2i}-b^{2j})}{2(1-b^2)}\right) \\
&= \exp\left((b^i + b^j)h_t + \frac{a(2-b^i-b^j)}{1-b} + \frac{c^2(2-b^{2i}-b^{2j})}{2(1-b^2)}\right) \\
&\quad \left[\exp\left(\frac{c^2(b^{j-i}-b^{j+i})}{1-b^2}\right) - 1\right].
\end{aligned}
$$

To verify (21), note that for $i \le j$,

$$
\mathrm{E}_t(e^{h_{t+i\Delta}} e^{h_{t+j\Delta}/2} \epsilon_{t+(j+1)\Delta}) = 0;
$$

For $i > j$, recall that

$$
h_{t+i\Delta} = a\ \frac{1-b^{i-j}}{1-b} + b^{i-j} h_{t+j\Delta} + c \sum_{m=j+1}^{i} b^{i-m} \epsilon_{t+m\Delta},
$$

which implies that

$$
\exp(h_{t+i\Delta} + h_{t+j\Delta}/2) = \exp\left[\frac{a(1-b^{i-j})}{1-b} + (b^{i-j} + 1/2)h_{t+j\Delta} + c \sum_{k=j+1}^{i} b^{i-k} \epsilon_{t+k\Delta}\right];
$$

moreover,

$$
\begin{aligned}
&\mathrm{E}_{t+j\Delta}\left[\epsilon_{t+(j+1)\Delta} \exp\left(c \sum_{k=j+1}^{i} b^{i-k} \epsilon_{t+k\Delta}\right)\right] \\
&= \mathrm{E}_{t+j\Delta}\left[\epsilon_{t+(j+1)\Delta} \exp(c b^{i-j-1} \epsilon_{t+(j+1)\Delta})\right]\ \mathrm{E}_{t+(j+1)\Delta}\left[\exp\left(c \sum_{k=j+2}^{i} b^{i-k} \epsilon_{t+k\Delta}\right)\right] \\
&= c\ b^{i-j-1}\ \exp\left[\frac{c^2(1-b^{2i-2j})}{2(1-b^2)}\right],
\end{aligned}
$$

in particular, if $i = j+1$, then

$$\mathrm{E}_{t+j\Delta}\left[\epsilon_{t+(j+1)\Delta}\exp\left(c\sum_{k=j+1}^{i} b^{i-k}\epsilon_{t+k\Delta}\right)\right] = c\, e^{c^2/2}.$$

Hence for $i > j$,

$$\begin{aligned}
&\mathrm{E}_t(e^{h_{t+i\Delta}} e^{h_{t+j\Delta}/2}\epsilon_{t+(j+1)\Delta}) \\
= \;& c\, b^{i-j-1}\exp\left[\frac{c^2(1-b^{2i-2j})}{2(1-b^2)} + \frac{a(1-b^{i-j})}{1-b}\right]\; \mathrm{E}_t\left\{\exp\left[(b^{i-j}+1/2)h_{t+j\Delta}\right]\right\} \\
= \;& c\, b^{i-j-1}\;\exp\left[\left(b^i + b^j/2\right)h_t + \frac{a\,(3/2 - b^i - b^j/2)}{1-b}\right. \\
&\left. + \frac{c^2\,(5/4 - b^{2i} - b^{2j}/4 + b^{i-j} - b^{i+j})}{2(1-b^2)}\right],
\end{aligned}$$

which implies (21).

Although V_n itself satisfies a martingale CLT, Theorem 1 requires the *joint* asymptotic normality for (U_n, V_n), not just for the two components separately. Using the Cramér-Wold device, it suffices to show that

($\star$) for any $u, v \in I\!\!R$, $n^{-1/2}\left[u(U_n - \mathrm{E}_t U_n) + vV_n\right]$ converges in distribution to a normal random variable with mean zero and variance $a_{11}\, u^2 + 2a_{12}\, uv + a_{22}\, v^2$.

We invoke some CLT for strong mixing sequences. See the nice survey papers Bradley (1986) and Peligrad (1986). Recall the definition of strong mixing. Let $\{X_i,\ i = 0, 1, 2, \ldots\}$ be a stochastic process, and $\mathcal{F}_m^n$ the σ-algebra generated by $X_i,\ m \le i \le n$. $\{X_i\}$ is said to be strong mixing with rate α_n if

$$\alpha_n = \sup_{t\ge 0}\ \sup_{A\in\mathcal{F}_0^t,\ B\in\mathcal{F}_{t+n}^\infty} |P(A\cap B) - P(A)P(B)| \longrightarrow 0$$

as $n \to \infty$. For fixed $t \ge 0$, $\Delta \ge 0$, and each $j = 0, 1, \ldots$, let $\xi_j = u(e^{h_{t+j\Delta}} - \mathrm{E}_t e^{h_{t+j\Delta}}) + ve^{h_{t+j\Delta}/2}\epsilon_{t+(j+1)\Delta}$, then $u(U_n - \mathrm{E}_t U_n) + vV_n = \sum_{j=0}^{n-1}\xi_j$. Note that $\{h_{t+j\Delta},\ j = 0, 1, \ldots\}$ is a strong mixing sequence with an exponential mixing rate, so is the sequence $\{\xi_j\}$ [since for each j, ξ_j is a function of $(h_{t+j\Delta}, h_{t+(j+1)\Delta})$]. Theorem 1 is a CLT conditioning on h_t, hence $\{\xi_j\}$ is not a stationary sequence. Based on the comments in Peligrad (1986) (p216, section 4) and the calculation we have done, ($\star$) will follow from that the sequence $\{\xi_n^2\}$ is uniformly integrable, and the sequence

$$B_n = \frac{\sup_{k\geq 0} \mathrm{Var}_t(\xi_{t+k\Delta} + \cdots + \xi_{t+(k+n-1)\Delta})}{\mathrm{Var}_t(\xi_t + \cdots + \xi_{t+(n-1)\Delta})}, \quad n = 1, 2, \ldots$$

is bounded. QED.

B Appendix

In this section, we provide details of the MCMC algorithms given in Chapter 4 that calibrate the SV models. The unknowns include the volatility sequence h and the model parameters $\Theta = \{\theta, \nu, \delta\}$ with $\theta = \{\mu, \alpha, \beta, \sigma, \rho\}$ and $\nu = \{\nu_1, \nu_2\}$. Our goal is to sample from the joint posterior distribution of the unknowns conditioning on the data $\{Y, C\}$. The general strategy of the Gibbs sampler, as a special class of MCMC algorithms applied to current context, is to cycle through several disjoint blocks of the unknowns, with each block simulated from the conditional distribution given all other blocks. Here is a list of those conditionals where the omitted variables are due to conditional independence:

- $P(h|Y, C, \Theta)$
- $P(\sigma, \rho|Y, C, h, \mu, \alpha, \beta, \nu, \delta)$
- $P(\alpha, \beta|Y, C, h, \mu, \sigma, \rho, \nu, \delta)$
- $P(\nu|Y, C, h, \alpha, \beta, \sigma, \rho, \delta)$ (μ omitted)
- $P(\mu|Y, h, \alpha, \beta, \sigma, \rho)$ (C, ν, δ omitted)
- $P(\delta|Y, C, h, \alpha, \beta, \sigma, \rho, \nu)$ (μ omitted)

Note here and in what follows, the capital notation $P(\cdot|\cdot)$ denotes conditional probabilities that involve the entire sequences Y and C, while the lower-case notation $p(\cdot|\cdot)$ denotes conditional probabilities that involve some coordinates of Y and C.

Due to the non-linear forms of the option pricing formulas (11)-(14), the first four conditional distributions listed above are not available in closed forms. Therefore, we apply Metropolis-Hastings algorithms (MH) to sample h and parameters $\{\alpha, \beta, \sigma, \rho, \nu\}$. The parameters μ and δ can be updated directly via Bayesian conjugacy without resorting to

MCMC. Using MH to update a variable amounts to simulate a candidate value from a proposal density then accept the proposed value with a certain probability expressed in terms of the related MH ratio.

Updating volatility sequence h

Let $g = 1, ..., G$ denote the successive MCMC iterations and Δ be the time increment between two adjacent data points. In the g-th iteration, the coordinates $h_{t+j\Delta}^{(g)}$, $j = 1, ..., n$ are updated sequentially, following a single-move MH scheme. We first consider the intermediate points $j = 2, ..., n-1$ then present a slightly different treatment to the two boundary points $j = 1$ and $j = n$.

At each point $j = 2, ..., n-1$, we generate $h_{t+j\Delta}^{*}$ from a proposal density q of normal distribution with mean $\alpha^{(g-1)}\Delta + (1+\beta^{(g-1)}\Delta)h_{t+(j-1)\Delta}^{(g)}$ and variance $s\ (\sigma^{(g-1)})^2\Delta$ following the discretization of (3), where $s > 0$ is a tuning parameter for controlling the acceptance/rejection rate in the MH chain. We set $s = 0.9$ in our experiment.

To determine the acceptance probability, note that $P\left(h_{t+j\Delta}|h_{t+k\Delta}, k \neq j; Y, C, \Theta\right) \propto P\left(C, Y, h, \Theta\right) \propto P\left(C|Y, h, \Theta\right) P\left(Y, h|\theta\right)$. If we write $P\left(C|Y, h, \Theta\right) P\left(Y, h|\theta\right)$ as a product of n factors, and collect terms that involve $h_{t+j\Delta}$, we will define the acceptance probability ω as

$$\omega = \min\{\omega_1\omega_2\omega_3\omega_4,\ 1\}, \tag{38}$$

where

$$\begin{aligned}
\omega_1 &= \frac{p(c_{t+j\Delta}|y_{t+j\Delta}, h_{t+j\Delta}^{*}, \Theta^{(g-1)})}{p(c_{t+j\Delta}|y_{t+j\Delta}, h_{t+j\Delta}^{(g-1)}, \Theta^{(g-1)})}, \\
\omega_2 &= \frac{p(y_{t+(j+1)\Delta}, h_{t+(j+1)\Delta}^{(g-1)}|y_{t+j\Delta}, h_{t+j\Delta}^{*}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}{p(y_{t+(j+1)\Delta}, h_{t+(j+1)\Delta}^{(g-1)}|y_{t+j\Delta}, h_{t+j\Delta}^{(g-1)}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}, \\
\omega_3 &= \frac{p(y_{t+j\Delta}, h_{t+j\Delta}^{*}|y_{t+(j-1)\Delta}, h_{t+(j-1)\Delta}^{(g)}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}{p(y_{t+j\Delta}, h_{t+j\Delta}^{(g-1)}|y_{t+(j-1)\Delta}, h_{t+(j-1)\Delta}^{(g)}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}, \\
\omega_4 &= \frac{q(h_{t+j\Delta}^{*}|h_{t+(j-1)\Delta}^{(g)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)})}{q(h_{t+j\Delta}^{(g-1)}|h_{t+(j-1)\Delta}^{(g)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)})}.
\end{aligned}$$

For $j = 1$, we adopt the stationary distribution of h as a proposal density of $h^*_{t+\Delta}$, i.e.

$$h^*_{t+\Delta} \sim N\left(-\frac{\alpha^{(g-1)}}{\beta^{(g-1)}},\ \frac{(\sigma^{(g-1)})^2}{\beta^{(g-1)}(2+\beta^{(g-1)}\Delta)}\right); \tag{39}$$

and the acceptance probability ω is given by

$$\omega = \min\{\omega_1\omega_2\omega_3,\ 1\}, \tag{40}$$

where

$$\begin{aligned}
\omega_1 &= \frac{p(c_{t+\Delta}|y_{t+\Delta}, h^*_{t+\Delta}, \Theta^{(g-1)})}{p(c_{t+\Delta}|y_{t+\Delta}, h^{(g-1)}_{t+\Delta}, \Theta^{(g-1)})}, \\
\omega_2 &= \frac{p(y_{t+2\Delta}, h^{(g-1)}_{t+2\Delta}|y_{t+\Delta}, h^*_{t+\Delta}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}{p(y_{t+2\Delta}, h^{(g-1)}_{t+2\Delta}|y_{t+\Delta}, h^{(g-1)}_{t+\Delta}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}, \\
\omega_3 &= \frac{q(h^{(g-1)}_{t+\Delta}|\alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)})}{q(h^*_{t+\Delta}|\alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)})}.
\end{aligned}$$

For $j = n$, we propose $h^*_{t+n\Delta}$ such that

$$h^*_{t+n\Delta} \sim N\left(\alpha^{(g-1)}\Delta + (1+\beta^{(g-1)}\Delta)h^{(g)}_{t+(n-1)\Delta},\ s(\sigma^{(g-1)})^2\Delta\right), \tag{41}$$

where $s = 0.9$ is the same tuning parameter; then accept $h^*_{t+n\Delta}$ with probability

$$\omega = \min\{\omega_1\omega_2\omega_3,\ 1\}, \tag{42}$$

where

$$\begin{aligned}
\omega_1 &= \frac{p(c_{t+n\Delta}|y_{t+n\Delta}, h^*_{t+n\Delta}, \Theta^{(g-1)})}{p(c_{t+n\Delta}|y_{t+n\Delta}, h^{(g-1)}_{t+n\Delta}, \Theta^{(g-1)})}, \\
\omega_2 &= \frac{p(y_{t+n\Delta}, h^*_{t+n\Delta}|y_{t+(n-1)\Delta}, h^{(g)}_{t+(n-1)\Delta}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}{p(y_{t+n\Delta}, h^{(g-1)}_{t+n\Delta}|y_{t+(n-1)\Delta}, h^{(g)}_{t+(n-1)\Delta}, \mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}, \\
\omega_3 &= \frac{q(h^{(g-1)}_{t+n\Delta}|h^{(g)}_{t+(n-1)\Delta}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)})}{q(h^*_{t+n\Delta}|h^{(g)}_{t+(n-1)\Delta}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)})}.
\end{aligned}$$

Updating parameters $\{\sigma, \rho\}$

At the g-th MH iteration ($g = 1, ..., G$), a pair $\{\sigma^*, \rho^*\}$ is generated by

$$\begin{pmatrix} \sigma^* \\ \rho^* \end{pmatrix} = \begin{pmatrix} \sigma^{(g-1)} \\ \rho^{(g-1)} \end{pmatrix} + \begin{pmatrix} s_1\epsilon_1 \\ s_2\epsilon_2 \end{pmatrix}, \tag{43}$$

where ϵ_1 and ϵ_2 are iid $N(0, 1)$ random variables, while s_1 and s_2 are two tuning factors. We set $s_1 = s_2 = 0.002$ in our experiment. This proposal density q, referred to as a random walk MH, is a symmetric function of $\{\sigma^*, \sigma^{(g-1)}\}$ and of $\{\rho^*, \rho^{(g-1)}\}$, hence $q(\cdot|\cdot)$ does not appear in the following MH ratio.

To determine the acceptance probability, note that $P(\sigma, \rho|Y, C, h, \mu, \alpha, \beta, \nu, \delta) \propto P(C|Y, h, \Theta)P(Y, h|\mu, \alpha, \beta, \sigma, \rho)P(\sigma, \rho)$. Assuming a unform prior $\pi(\sigma, \rho)$ on the rectangle $[0.0, 1.5] \times [-1.0, 0.5]$, we accept $\{\sigma^*, \rho^*\}$ with probability

$$\omega = \min\{\omega_1\omega_2\omega_3,\ 1\}, \tag{44}$$

where

$$\begin{aligned}
\omega_1 &= \frac{P(C|Y, h^{(g)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^*, \rho^*, \nu^{(g-1)}, \delta^{(g-1)})}{P(C|Y, h^{(g)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)}, \nu^{(g-1)}, \delta^{(g-1)})}, \\
\omega_2 &= \frac{P(Y, h^{(g)}|\mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^*, \rho^*)}{P(Y, h^{(g)}|\mu^{(g-1)}, \alpha^{(g-1)}, \beta^{(g-1)}, \sigma^{(g-1)}, \rho^{(g-1)})}, \\
\omega_3 &= \frac{\pi(\sigma^*, \rho^*)}{\pi(\sigma^{(g-1)}, \rho^{(g-1)})}.
\end{aligned}$$

Updating parameters $\{\alpha, \beta\}$

A similar random walk MH is used to propose $\{\alpha^*, \beta^*\}$ by

$$\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} \alpha^{(g-1)} \\ \beta^{(g-1)} \end{pmatrix} + \begin{pmatrix} s_3\epsilon_3 \\ s_4\epsilon_4 \end{pmatrix} \tag{45}$$

with iid $N(0, 1)$ errors $\{\epsilon_3, \epsilon_4\}$ and tuning parameters $s_3 = s_4 = 0.002$.

Assuming a uniform prior $\pi(\alpha, \beta)$ on the rectangle $[-1.5,\ 1.0] \times [-0.9,\ 0.0]$, we accept $\{\alpha^*, \beta^*\}$ with probability

$$\omega = \min\{\omega_1\omega_2\omega_3,\ 1\}, \tag{46}$$

where

$$\begin{aligned}
\omega_1 &= \frac{P(C|Y,h^{(g)},\alpha^*,\beta^*,\sigma^{(g)},\rho^{(g)},\nu^{(g-1)},\delta^{(g-1)})}{P\big(C|Y,h^{(g)},\alpha^{(g-1)},\beta^{(g-1)},\sigma^{(g)},\rho^{(g)},\nu^{(g-1)},\delta^{(g-1)}\big)},\\
\omega_2 &= \frac{P(Y,h^{(g)}|\mu^{(g-1)},\alpha^*,\beta^*,\sigma^{(g)},\rho^{(g)})}{P\big(Y,h^{(g)}|\mu^{(g-1)},\alpha^{(g-1)},\beta^{(g-1)},\sigma^{(g)},\rho^{(g)}\big)},\\
\omega_3 &= \frac{\pi(\alpha^*,\beta^*)}{\pi\big(\alpha^{(g-1)},\beta^{(g-1)}\big)}.
\end{aligned}$$

Updating parameters $\{\nu_1,\nu_2\}$

We use a random walk MH with normal errors again to propose $\nu^*=\{\nu_1^*,\nu_2^*\}$.

Note that $P(\nu|Y,C,h,\alpha,\beta,\sigma,\rho,\delta)\propto P(C|Y,h,\Theta)P(\nu)$. Conditioning on $\sigma^{(g)}$, we use a uniform prior $\pi(\nu)$ supported on the square $\left[\frac{-0.5}{\sigma^{(g)}},\ \frac{0.2}{\sigma^{(g)}}\right]^2$, and accept ν^* with probability

$$\omega=\min\{\omega_1\omega_2,\ 1\}, \tag{47}$$

where

$$\begin{aligned}
\omega_1 &= \frac{P(C|Y,h^{(g)},\alpha^{(g)},\beta^{(g)},\sigma^{(g)},\rho^{(g)},\nu^*,\delta^{(g-1)})}{P\big(C|Y,h^{(g)},\alpha^{(g)},\beta^{(g)},\sigma^{(g)},\rho^{(g)},\nu^{(g-1)},\delta^{(g-1)}\big)},\\
\omega_2 &= \frac{\pi(\nu^*)}{\pi\big(\nu^{(g-1)}\big)}.
\end{aligned}$$

Updating parameters μ and δ

To update μ, we specify a conjugate normal prior $\pi(\mu)$ with mean $\mu_0=0$ and variance $1/p_0=1,000,000$. Then the posterior satisfies

$$P(\mu^{(g)}|Y,h^{(g)})\sim N\left(\frac{D'y+\mu_0p_0}{D'D+p_0},\frac{1}{D'D+p_0}\right), \tag{48}$$

where

$$\begin{aligned}
D &= \sqrt{\Delta}\left[\exp\big(-\frac{1}{2}h^{(g)}_{t+\Delta}\big),...,\exp\big(-\frac{1}{2}h^{(g)}_{t+(n-1)\Delta}\big)\right]',\\
y &= \frac{1}{\sqrt{\Delta}}\left[(Y_{t+2\Delta}-Y_{t+\Delta})\exp\big(-\frac{1}{2}h^{(g)}_{t+\Delta}\big),...,(Y_{t+n\Delta}-Y_{t+(n-1)\Delta})\exp\big(-\frac{1}{2}h^{(g)}_{t+(n-1)\Delta}\big)\right]'.
\end{aligned}$$

Similarly, we specify a conjugate inverse gamma (IG) prior for δ^2, $\pi(\delta^2) \sim IG(n_0, n_0 s_0^2)$ with $n_0 = 2.0$ and $s_0 = 0.005$. Therefore, the posterior of δ^2 can be derived as

$$P(\delta_{(g)}^2 | Y, C, h^{(g)}, \alpha^{(g)} \beta^{(g)}, \sigma^{(g)}, \rho^{(g)}, \nu^{(g)}) \sim IG\left(n_0 + \frac{(T-1)}{2}, n_0 s_0^2 + \frac{\left[\sum_{j=1}^{n} (c_{t+j\Delta} - f_{t+j\Delta})^2\right]}{2}\right) \tag{49}$$

where $\{c_{t+j\Delta}, j = 1, ..., n\}$ are the logarithms of the observed option prices and $\{f_{t+j\Delta}, j = 1, ..., n\}$ are the predicted option prices according to (11)-(14) and (26).

References

Abramowitz, M., and Stegun, I.A. (1964). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover.

Aït-Sahalia, Y., and Kimmel, R. (2007). Maximum likelihood estimaton of stochastic volatility models. Journal of Financial Economics 83, 413–452.

Bates, D.S. (1996). Jumps and stochastic volatility: Exchange rate process implicit in Deutsch Mark options. Review of Financial Studies 9, 69–107.

Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–659.

Bradley, R.C. (1986). Basic properties of strong mixing conditions. Dependence in Probability and Statistics (edited by E. Eberlein and M.S. Taqqu), 165–192. Birkhäuser.

Chernov, M. and Ghysels, E. (2000). A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. Journal of Financial Economics 56, 407–458.

Chernov, M., Gallant, A.R., Ghysels, E., and Tauchen, G. (2003). Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Chib, S., Nardari, F. and Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. Journal of Econometrics 108, 281-316.

Chong, B.-S., Ding, D.K. and Tan, K.-H. (2003). Maturity effect on bid-ask spreads of OTC currency options. Review of Quantitative Finance and Accounting 21, 5–15.

Clark, P.K. (1973). A subordinated stochastic process model with finite variances for speculative prices. Econometrica 41, 135–156.

Dam, B.K. (1998). Berry-Esseen theorem for stationary strong mixing sequences. Vietnam Journal of Mathematics 26, 185–187.

Danielson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. Journal of Econometrics 63, 375–400.

Durham, G.B. and Gallant, A.R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. Journal of Business and Economic Statistics 20, 297–316.

Duffie D. and Glynn, P. (1995). Efficient Monte Carlo Estimation of Security Prices. The Annals of Applied Probability 4, 897–905.

Elerian, O., Chib, S. and Shephard, N. (2001). Likelihood inference for discretely observed non-linear diffusions. Econometrica 69, 959–994.

Eraker, B. (2001). MCMC analysis of diffusion models with applications to finance. Journal of Business and Economic Statistics 19, 177–191.

Eraker, E., Johannes, M. and Polson, N. (2003). The impact of jumps in returns and volatility. Journal of Finance 53, 1269–1300.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. Journal of Finance 54, 1367–1404.

Gallant, A.R., Hsieh, D.A., and Tauchen, G.E. (1991). On fitting a recalcitrant series: The pound/dollar exchange rate, 1974–1983, in Barnett, W.A., Powell, J., and Tauchen, G.E. (eds.), Nonparametric and semiparametric methods in econometrics and statistics. Proceedings of the Fifth International Symposium in Economic Theory and Econometrics, 199–240. Cambridge University Press.

Gallant, A.R., Hsu, C.-T. and Tauchen, G. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. The Review of Economics and Statistics 81, 617–631.

Garcia, R., Ghysels, E., and Renault, E. (2007). The econometrics of option pricing. in Aït-Sahalia, Y., and Hansen, L.P., eds. (2007), Handbook of Financial Econometrics. Elsevier/North-Holland, Amsterdam, forthcoming.

Garman, M.B. and Kohlhagen, S.W. (1983). Foreign currency option values. Journal of International Money and Finance 2, 231–237.

Ge, X. (2000). Bayesian Calibration of Stochastic Volatility Models. Ph.D. dissertation. University of North Carolina at Chapel Hill.

Geweke, J. (1994). Bayesian comparison of econometric models. Working paper. Federal Reserve Bank of Minneapolis Research Department.

Ghysels, E., Harvey, A. and Renault, E. (1996). Stochastic volatility. in Maddala, G.S., and Rao, C.R. (eds.) Handbook of Statistics, Vol.14, Elsevier.

Golub, G.H. (1973). Some modified matrix eigenvalue problems. SIAM Reviews 15, 318–334.

Golub, G.H. and Welsch, J.H. (1969). Calculation of Gaussian quadrature rules. Mathematics of Computation 23, 221–230.

Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with application to bond and currency options. Review of Financial Studies 6, 327–343.

Huang, X., and Tauchen, G. (2005) The relative contribution of jumps to total price variance. Journal of Financial Econometrics 3, 456–499.

Jacquier, E., and Jarrow, R. (2000). Bayesian analysis of contingent claim model error. Journal of Econometrics 94, 145–180.

Jacquier, E., Polson, N., and Rossi, P. (2004). A Bayesian analysis fat-tailed stochastic volatility models with correlated errors. Journal of Econometrics 122, 185–212.

Jones, C.S. (2003). The dynamics of stochastic volatility: Evidence from underlying and option markets. Journal of Econometrics 116, 181–224.

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

Kloeden, P.E. and Platen, E. (1992). Numerical Solution of Stochastic Differential Equations. Springer-Verlag.

Melino, A. and Turnbull, S. (1990). Pricing foreign currency options with stochastic volatility. Journal of Econometrics 45, 239–265.

Merton, R. (1973). The theory of rational option pricing. Bell Journal of Economics and Management 7, 141–183.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. Journal of Financial Economics 63, 3–50.

Pastorello, S., Renault, E. and Touzi, N. (2000). Statistical inference for random-variance option pricing. Journal of Business and Economic Statistics 18, 358–367.

Peligrad, M. (1986). Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables. in Eberlein, E., and Taqqu, M.S. (eds.) Dependence in Probability and Statistics 193–223. Birkhäuser.

Pitt, M. and Shephard, N. (1999). Filtering via simulation: auxiliary particle filter. Journal of American Statistical Association 94, 590–599.

Polson, N. and Stroud, J. (2003). Bayesian inference for derivative prices. Working paper, University of Chicago.

Renault, E. (1997). Econometric models of option pricing errors. in Kreps, D., and Wallis, K. Advances in Economics and Econometrics, 223-278. Cambridge Univ. Press.

Romano, M. and Touzi, N. (1997). Contingent claims and market completeness in a stochastic volatility model. Mathematical Finance 7, 399–412.

Scott, L. (1987). Option pricing when the variance changes randomly: theory, estimation and an application. Journal of Financial and Quantitative Analysis 22, 419–438.

Tauchen, G.E. and Pitt, M. (1983). The price variability-volume relationship on speculative markets. Econometrica 51, 485–505.

Taylor, S.J. (1982). Financial returns modeled by the product of two stochastic processes — A study of the daily sugar prices 1961–1975. in Anderson, O.D. (ed.), Time Series Analysis: Theory and Practice 1, 203–226. North-Holland.

Yu, J. (2002). Forecasting volatility in the New Zealand stock market. Applied Financial Economics 12, 193–202.

Table 1. Accuracy of the Gaussian Scheme and Brute Force Monte Carlo

	Maturity in Days		
Method	30	90	180
In the Money: $K/S = 0.9$			
GS(1000)	0.0026	0.0044	0.0047
GS(10000)	0.0018	0.0032	0.0036
GS(100000)	0.0017	0.0031	0.0035
GS(Q)	0.0063	0.0041	0.0035
GS(QI)	0.0063	0.0044	0.0036
BF(1000)	0.0011	0.0021	0.0031
BF(10000)	0.0004	0.0007	0.0010
BF(100000)	0.0002	0.0003	0.0004
At the Money: $K/S = 1.0$			
GS(1000)	0.0039	0.0066	0.0063
GS(10000)	0.0024	0.0053	0.0052
GS(100000)	0.0023	0.0053	0.0051
GS(Q)	0.0048	0.0067	0.0051
GS(QI)	0.0048	0.0072	0.0049
BF(1000)	0.0014	0.0023	0.0033
BF(10000)	0.0005	0.0008	0.0011
BF(100000)	0.0003	0.0004	0.0005
Out of the Money: $K/S = 1.1$			
GS(1000)	0.0090	0.0058	0.0074
GS(10000)	0.0085	0.0043	0.0063
GS(100000)	0.0084	0.0043	0.0061
GS(Q)	0.0043	0.0058	0.0062
GS(QI)	0.0044	0.0069	0.0054
BF(1000)	0.0021	0.0025	0.0035
BF(10000)	0.0007	0.0008	0.0012
BF(100000)	0.0003	0.0004	0.0005

Shown are the root mean squared errors of the approximation to a series of 600 simulated log call option prices using the Gaussian scheme with 1,000, 10,000, 100,000 Monte Carlo repetitions and quadrature both with (GS(QI)) and without (GS(Q)) approximation to the double sums in equations (20) and (21). Also shown are the root mean square approximations by brute force Monte Carlo for 1,000, 10,000, 100,000 Monte Carlo repetitions. The option prices regarded as correct were computed from a simulation of length 600 of the underlying prices and volatilities using equations (1)–(3) by brute force Monte Carlo with 1,000,000 repetitions using equations (11)–(13) at the maturities and strikes shown in the table. Antithetic variates were used for all Monte Carlo computations. The parameter values were set as follows: $\mu = 0.0, \alpha = 0.0, \beta = -0.06, \rho = -0.3, \sigma = 0.2, \nu_1 = -0.5, \nu_2 = 0.0, r = 0.0$. Because of the use of logarithms, table entries are approximate root mean percentage errors after multiplication by 100.

Table 2. Computing Speeds of Numerical Integration Schemes

	Maturity in Days		
Method	30	90	180
GS(1000)	0.0030	0.0152	0.0594
GS(10000)	0.0196	0.0305	0.0763
GS(100000)	0.1743	0.1893	0.2336
GS(Q)	0.0023	0.0150	0.0589
GS(QI)	0.0002	0.0003	0.0005
BF(1000)	0.0169	0.0460	0.0907
BF(10000)	0.1630	0.4665	0.9071
BF(100000)	1.6313	4.6025	9.0705

Shown are the time in seconds to compute at the money log option prices using the central limit approximation with 1,000, 10,000, 100,000 Monte Carlo repetitons and quadrature.

Table 3. Posterior Means and Standard Deviations from Simulated Data

parameter	true value	returns only	returns and options GS(QI)
μ	0.0	0.0451	0.0566
		(0.0376)	(0.0291)
α	0.0	-0.0605	-0.0333
		(0.0370)	(0.0164)
β	-0.06	-0.2528	-0.1038
		(0.1009)	(0.0051)
σ	0.2	0.5032	0.3245
		(0.1045)	(0.0184)
ρ	-0.3	-0.0123	-0.2079
		(0.1063)	(0.0430)
ν_1	-0.5		-0.4929
			(0.0633)
δ	0.0		0.0145
			(0.0020)

Using an Euler scheme with $\Delta = 1$ applied to (9)–(10) at parameter settings as shown and r and ν_2 fixed at zero and not involved in the estimation, ATM option prices, where the strike prices are set to the asset price $K = S$, are computed for $n = 600$ and an expiration date of thirty days using "brute force" with 1,000,000 repetitions to integrate (11)–(13). To these simulated data an MCMC chain of 500,000 iterations (after discarding 500,000 iterations to dissipate transients) was generated as described in Subsecions 4.1 and 4.2 using the Gaussian scheme with quadrature and interpolated conditional means and variances to integrate (11)–(13). The means and standard deviations (in parentheses) of the chain are shown. The ratio of the posterior standard deviation of the volatility $exp(h_t/2)$ without options to the posterior standard deviation of the volatility with options was computed at each time point and averaged over observations 4 through 594, which eliminates end effects. This average is 5.45, indicating that volatility is much more precisely estimated when options are included.

Table 4. Exchange Rate Sample Statistics

	Series	
Statistic	JPY/EUR	CHF/EUR
Mean	-0.0067	-0.0112
Std. Dev.	0.8343	0.2351
Skewness	-0.0125	-0.8543
Kurtosis	4.5195	13.6625
Sample size	1003	1003
Autocorrelations:		
Lag 1	0.1109	0.2001
Lag 2	0.1360	0.1345
Lag 3	0.0694	0.1777
Lag 4	0.1081	0.1700
Lag 5	0.0116	0.2105

The data are daily exchange rates, Swiss Frank to Euro (CHF/EUR) and Japanese Yen to Euro (JPY/EUR), from January 1999 to December 2002 expressed as a geometric return.

Table 5. Posterior Means and Standard Deviations from Daily Returns Alone, 1999–2002.

parameter	JPY/EUR	CHF/EUR
μ	-0.0066	-0.0124
	(0.0089)	(0.0071)
α	-0.0815	-0.3468
	(0.0352)	(0.1055)
β	-0.1341	-0.1033
	(0.0496)	(0.0306)
σ	0.3567	0.4064
	(0.0778)	(0.0624)
ρ	-0.1672	-0.3136
	(0.0864)	(0.0822)

Values computed from MCMC chain of 40,000 (after discarding 20,000 to dissipate transients) run as described in Subsecion 4.1. The means and standard deviations (in parentheses) of the chain are shown.

Table 6. Posterior Means and Standard Deviations from Daily Returns and Options, 2000–2002.

	JPY/EUR		CHF/EUR	
parameter	returns alone	returns and options	returns alone	returns and options
μ	0.0003	0.0005	-0.0000	-0.0002
	(0.0196)	(0.0266)	(0.0063)	(0.0067)
α	-0.4323	-0.004	-1.1019	-0.0737
	(0.1374)	(0.009)	(0.2610)	(0.0219)
β	-0.4309	-0.0157	-0.3363	-0.0194
	(0.1262)	(0.0043)	(0.0786)	(0.0061)
σ	0.8142	0.1664	0.7458	0.2508
	(0.1052)	(0.0223)	(0.0930)	(0.0113)
ρ	-0.0705	0.0478	-0.0167	-0.4949
	(0.1076)	(0.0639)	(0.0749)	(0.0372)
ν_1		0.5135		0.5917
		(0.0865)		(0.1127)
ν_2		0.3867		0.2534
		(0.0418)		(0.0294)
δ		0.0160		0.0345
		(0.0029)		(0.0016)

Values computed from MCMC chain of 500,000 (after discarding 200,000 to dissipate transients) run as described in Subsection 4.1 for returns alone and in Subsection 4.2 for returns and options using the CLT with quadrature and interpolated contitional means and variances to integrate (11)–(13). The means and standard deviations (in parentheses) of the chain are shown.

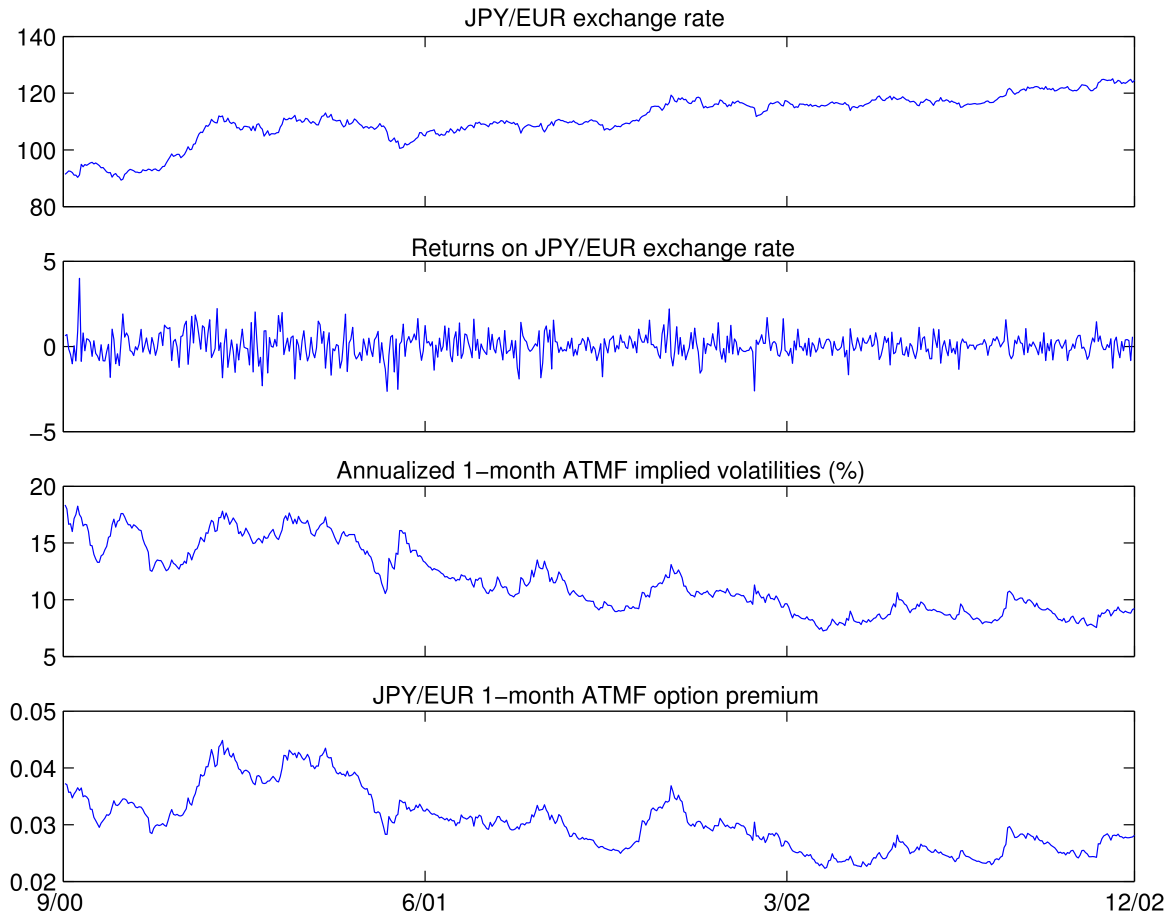

Figure 1. Underlying and Options Data, Japanese Yen to Euro. Shown in the first panel are daily JPY/EUR exchange rates at a daily frequency. The geometric return computed from the values in the first panel are shown in the second. The third panel shows the implied volatilities for an ATMF one month JPY/EUR option. The last panel shows option price computed from the implied volatilities and the three month LIBOR rates on the Japanese Yen and the Euro using the Garman-Kohlhagen option pricing formula.

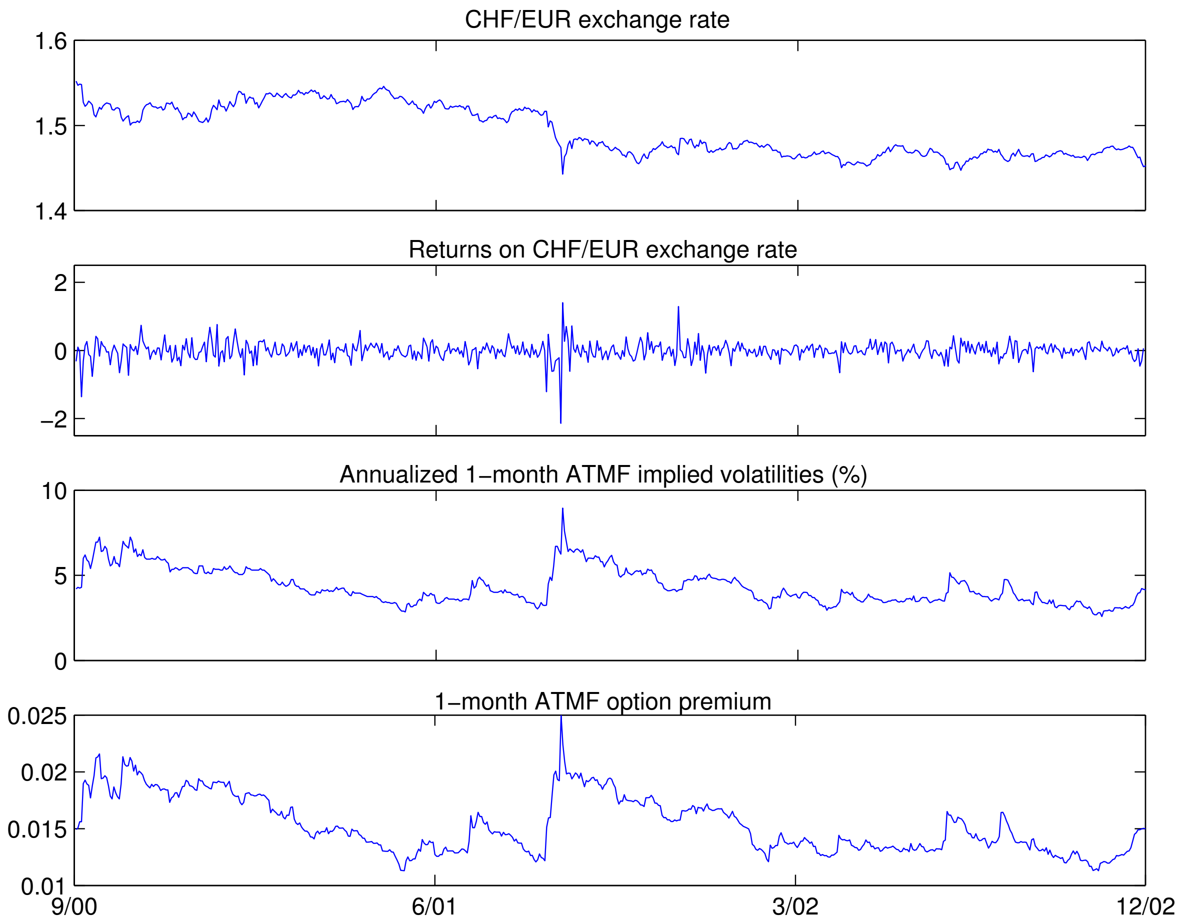

Figure 2. Underlying and Options Data, Swiss Frank to Euro. Legend as for Figure 1 but for Swiss Franc.

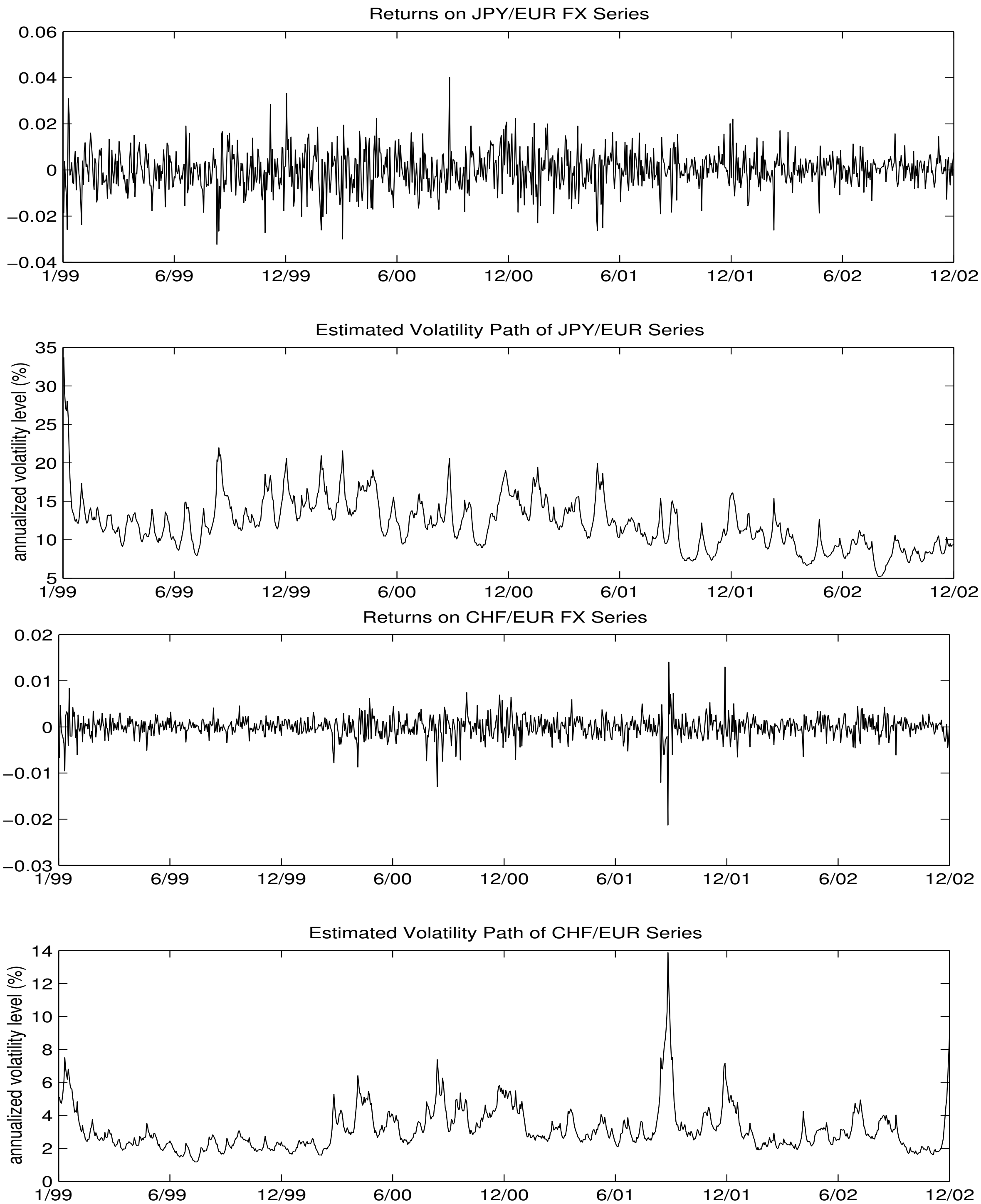

Figure 3. Volatility Estimated from Returns Alone. Geometric returns computed from the data and volatility estimates (often called smoothed volatility) computed from MCMC chain of 40,000 (after discarding 20,000 to dissipate transients) run as described in Subsecion 4.1.

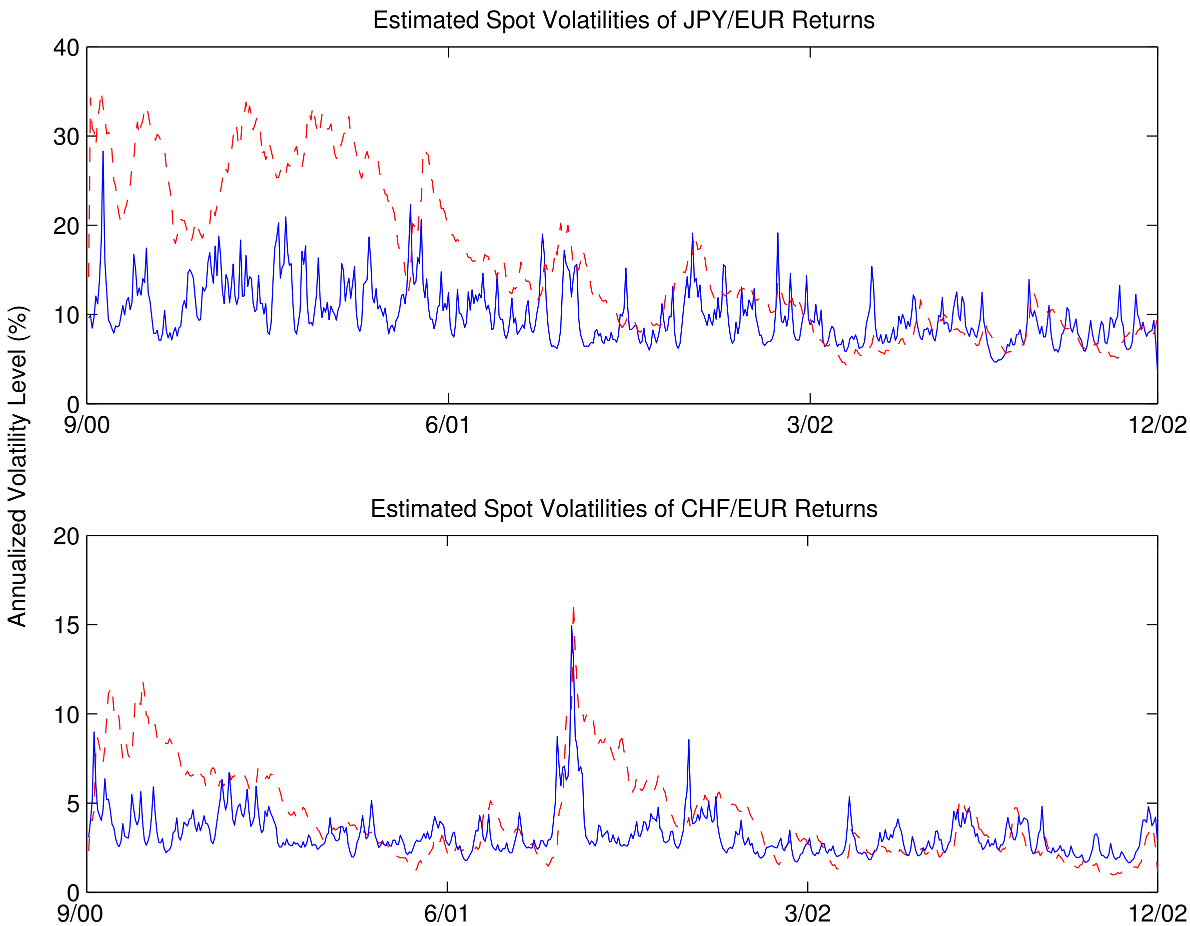

Figure 4. Volatility Estimated from Returns and Options. The solid line is estimated volatitility (often called smoothed volatitltiy) computed from returns alone using the MCMC methods described in Subsecion 4.1. The dotted line is the same computed from returns and options using the methods described in Subsecion 4.2 and the CLT with quadrature and interpolated conditional means and variances to integrate (11)–(13). In both instances values computed from an MCMC chain of 500,000 (after discarding 200,000 to dissipate transients).

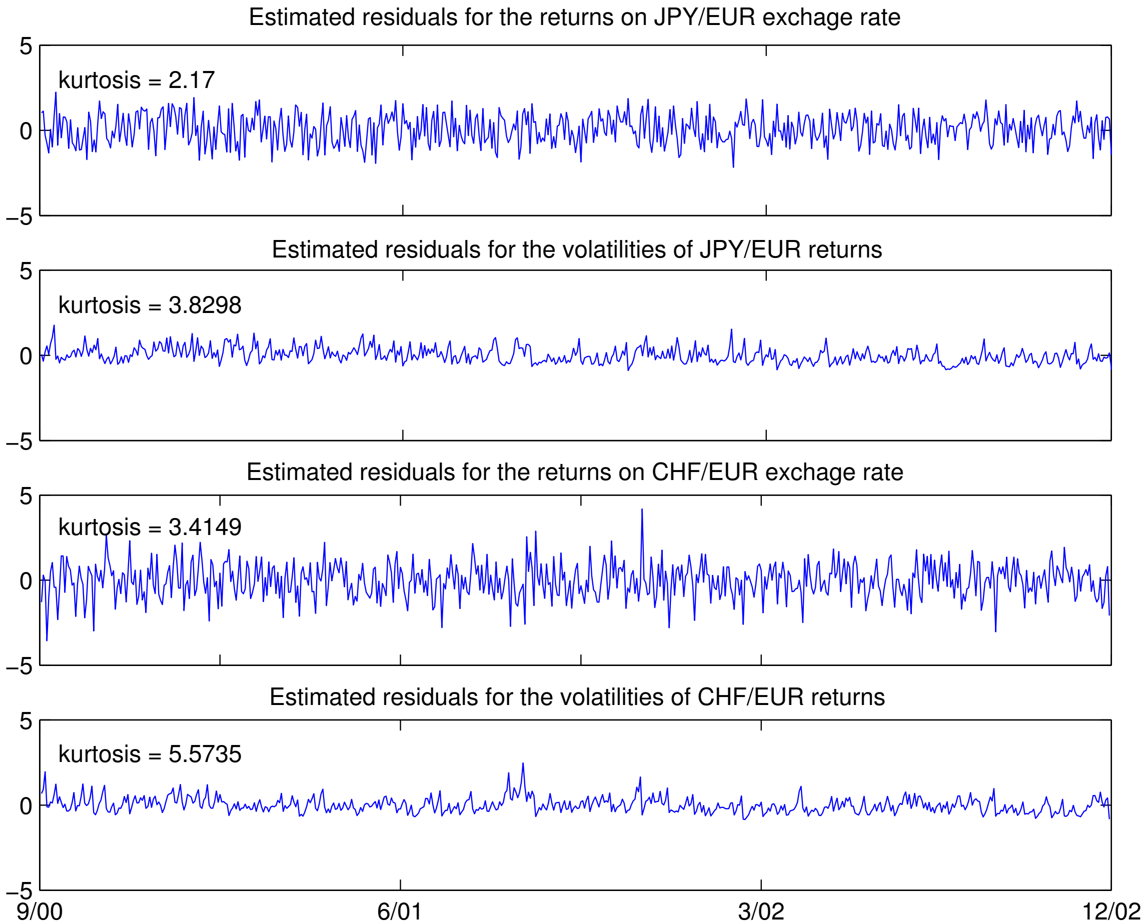

Figure 5. Estimated Residuals Based on Returns Data. Shown in the first panel is the estimated residuals from JPY/EUR return process (2) based on returns data. The second panel shows the residuals from JPY/EUR return volatility process (3) estimated based on returns data. The last two panels show the estimated residuals as in the first two but for CHF/EUR. Also shown in each panel is the kurtosis for the plotted residuals series. The returns residuals are normalized by the standard deviation computed from smooted volatility series at each time point. The volatility residuals are computed from the smoothed volatility process that is treated as if it were data that follow the autoregression (17).

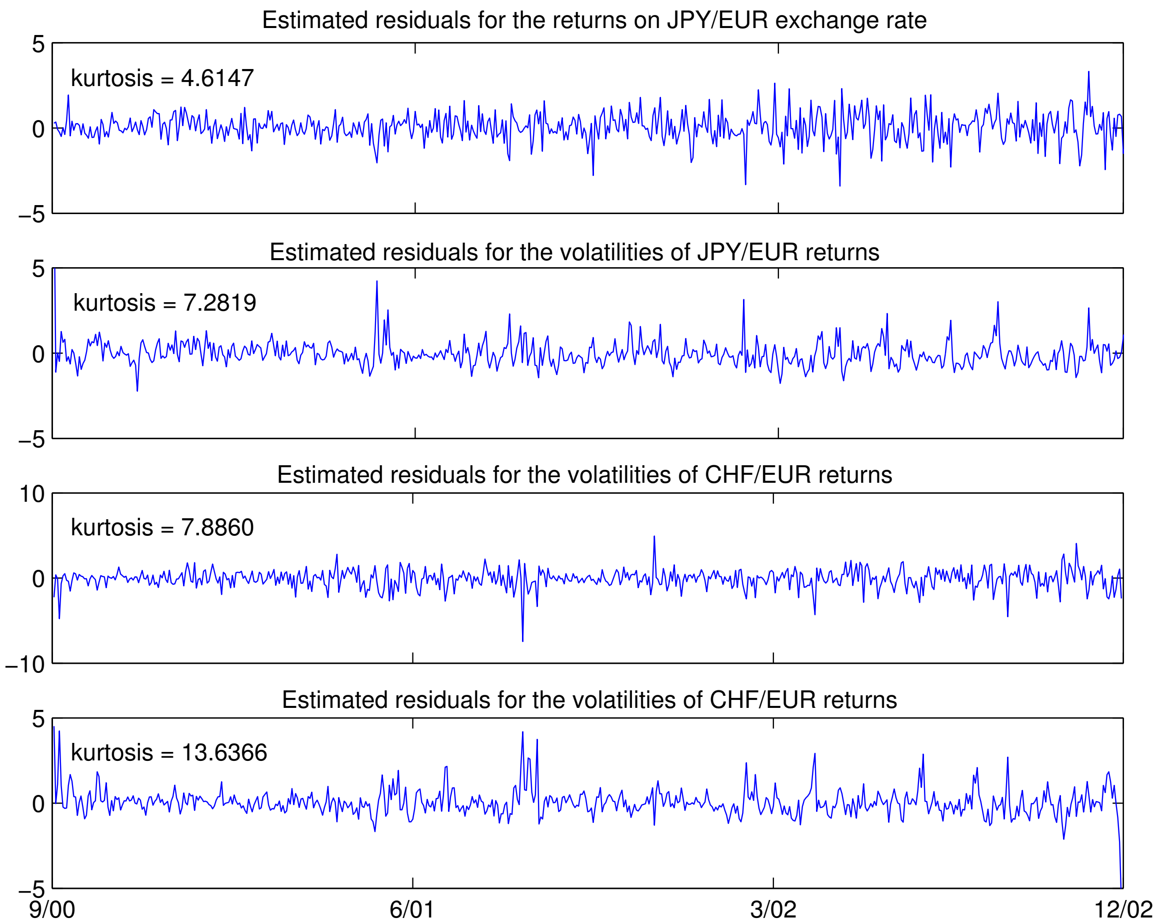

Figure 6. Estimated Residuals Based on Joint Data. Shown in the first panel is the estimated residuals from JPY/EUR return process (2) based on joint data. The second panel shows the residuals from JPY/EUR return volatility process (3) estimated based on joint data. The last two panels show the estimated residuals as in the first two but for CHF/EUR. Also shown in each panel is the kurtosis for the plotted residuals series. The returns residuals are normalized by the standard deviation computed from smooted volatility series at each time point. The volatility residuals are computed from the smoothed volatility process that is treated as if it were data that follow the autoregression (17).