ELSEVIER

Journal of Empirical Finance 10 (2003) 505–531

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

Univariate and multivariate stochastic volatility models: estimation and diagnostics

Roman Liesenfeld[a,*], Jean-François Richard[b]

[a] *Department of Economics, Eberhard-Karls-Universität, Mohlstrasse 36, 72074 Tübingen, Germany*
[b] *Department of Economics, University of Pittsburgh, 4S01 Wesley W. Posvar Hall, Pittsburgh, PA 15260, USA*

Abstract

A Maximum Likelihood (ML) approach based upon an Efficient Importance Sampling (EIS) procedure is used to estimate several extensions of the standard Stochastic Volatility (SV) model for daily financial return series. EIS provides a highly generic procedure for a very accurate Monte Carlo (MC) evaluation of the marginal likelihood which depends upon high-dimensional interdependent integrals. Extensions of the standard SV model being analyzed only require minor modifications in the ML-EIS procedure. Furthermore, EIS can also be applied for filtering which provides the basis for several diagnostic tests. Our empirical analysis indicates that extensions such as a semi-nonparametric specification of the error term distribution in the return equation dominate the standard SV model. Finally, we also apply the ML-EIS approach to a multivariate factor model with stochastic volatility.

JEL classification: C15; C22; C52
Keywords: Efficient importance sampling; Latent variables; Maximum likelihood

1. Introduction

The Stochastic Volatility (SV) model introduced by Taylor (1982, 1986) provides a successful alternative to the class of AutoRegressive Conditionally Heteroscedastic (ARCH) models in accounting for the time-varying and persistent volatility as well as for the leptokurtosis in financial return series. It arises from the mixture-of-distributions hypothesis which assumes that the volatility process is driven by the unobservable flow of

* Corresponding author. Tel.: +49-7071-2977439; fax: +49-7071-295546.
E-mail address: roman.liesenfeld@uni-tuebingen.de (R. Liesenfeld).

0927-5398/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0927-5398(02)00072-5

price-relevant information. The SV model is also closely related to continuous-time diffusion processes which are widely used in the option-pricing literature (see, for example, Hull and White, 1987).

While the SV model has theoretical appeal, its efficient estimation is not straightforward. The fundamental problem is that the latent volatility process enters the model nonlinearly, which leads to a likelihood function depending upon high-dimensional integrals. A variety of estimation procedures has been proposed to overcome this difficulty, including, for example, the Generalized Method of Moments (GMM) used by Melino and Turnbull (1990), the Quasi Maximum Likelihood (QML) approach followed by Harvey et al. (1994) and Ruiz (1994), the Efficient Method of Moments (EMM) applied by Gallant et al. (1997), and Markov-Chain Monte Carlo (MCMC) procedures used by Jacquier et al. (1994) and Kim et al. (1998). See, for example, Ghysels et al. (1996) for a survey of these estimation procedures and Andersen et al. (1999) for performance comparisons.

In the present paper, we use instead a Maximum Likelihood (ML) approach based upon the Efficient Importance Sampling (EIS) procedure of Richard and Zhang (1996, 1997, 1998) to estimate and test the SV model. EIS is a Monte Carlo (MC) technique for the evaluation of high-dimensional integrals, which is ideally suited for the computation of the likelihood for the SV model. It relies upon a sequence of simple low-dimensional least-squares regressions to obtain a very accurate global approximation of the integrand. This approximation leads to a MC sampler, which produces highly accurate MC estimates of the likelihood. This EIS procedure significantly generalizes the Accelerated Gaussian Importance Sampling (AGIS) technique proposed by Danielsson and Richard (1993) and applied, for example, by Danielsson (1994a), Liesenfeld (1998, 2001) and Liesenfeld and Jung (2000). While the AGIS technique is specifically designed for models with a latent Gaussian process, the EIS procedure in its most general form can be applied to models with arbitrary classes of distributions for the latent variables. Importance sampling techniques are now widely used across a broad range of statistical applications. See, for example, Liesenfeld and Richard (2001) for a recent survey and for a discussion of importance sampler selection. One particular approach, which is relevant for our paper, is that developed by Durbin and Koopman (1997, 2000) and applied to the SV model by Sandmann and Koopman (1998). Their approach is based upon a decomposition of the likelihood function into a component which can be evaluated by Kalman filter techniques and a remainder which is evaluated by importance sampling. In this respect, it constitutes an indirect implementation of importance sampling based upon local approximations of the model. In a recent work, Durham and Gallant (2001) analyze importance sampling techniques in the context of ML estimation of continuous-time diffusion processes. In particular, they apply importance sampling based upon local approximations to integrate out the unobserved states of the process at intermediate points between each pair of observation. In contrast, EIS relies upon global approximations of the integrand(s), designed for the specific purpose of minimizing the MC sampling variance of the approximation. A more explicit comparison between EIS and the approach of Durbin and Koopman (1997, 2000) is presented below.

In addition to its high numerical accuracy, the EIS approach offers a number of attractive features. Foremost, since it is used to evaluate the likelihood function itself, it allows one to rely upon the full range of likelihood-based inference techniques (estimation,

testing, Bayesian inference). Furthermore, in contrast, for example, to the QML and GMM approach, ML-EIS allows for asymptotically efficient estimation of fully parametric models.[1] Equally importantly, the EIS algorithm is highly generic and its basic structure does not depend upon a specific model. Hence, changes in the model being analyzed can easily be accommodated and only require minor adjustments in the algorithm. Finally, EIS can be used to calculate filtered estimates of arbitrary functions of the latent volatility process, which can serve as the basis for several diagnostic tests.

The objective of our paper is twofold. Firstly, we propose to illustrate the flexibility of ML-EIS by applying it to several univariate and multivariate extensions of the standard SV model under which returns follow a Gaussian distribution conditional on the volatility, which is itself modeled as a lognormal first-order autoregressive process. As shown by Gallant et al. (1997), the standard SV model fails to capture several salient features of financial return series suggesting a need for alternative specifications of the volatility process or the conditional return distribution. Secondly, we develop and apply diagnostic tests for the SV model based upon residuals obtained by a simple EIS filtering procedure. In particular, these residuals can be used to detect outliers and to check the distributional assumptions as well as the dynamic structure of the model.

The remainder of the paper is organized as follows. Section 2 briefly reviews the standard version of the SV model and provides a description of the EIS procedure and its implementation. Univariate extensions of the standard SV model are analyzed in Section 3 and a multivariate factor SV model in Section 4. Section 5 summarizes our results and concludes.

2. The ML-EIS approach

2.1. The basic SV model

The basic SV model used, for example, by Taylor (1986, 1994), Jacquier et al. (1994) and Kim et al. (1998) is given by

$$r_t = \exp(\lambda_t/2)\epsilon_t \tag{1}$$

$$\lambda_t = \gamma + \delta\lambda_{t-1} + v\eta_t, \tag{2}$$

where r_t is the return on day t:$1 \rightarrow T$. The processes $\{\epsilon_t\}$ and $\{\eta_t\}$ are mutually independent iid Gaussian random variables with zero means and unit variances and (γ,δ,v) are the parameters to be estimated. The log volatility λ_t, which is assumed to follow an AR(1) process with a persistence parameter δ, is unobservable. For $|\delta| < 1$, the returns r_t are strictly stationary. The parameter v measures the standard deviation of volatility shocks and is assumed to be greater than zero.

[1] Note that, as shown by Danielsson (1994b) in the context of the standard SV model, the finite sample properties of the ML-estimator based upon importance sampling are virtually identical to those of Bayesian posterior means evaluated by MCMC as in Jacquier et al. (1994), the latter being regarded in the literature as a benchmark for efficiency.

The evaluation of the likelihood associated with the observable r_t's requires that the latent λ_t's be integrated out from the joint density of the observed and latent variables. Since λ_t is serially dependent and enters the model nonlinearly, this integration problem cannot be solved by standard integration procedures. To overcome this difficulty, we apply the EIS procedure developed by Richard and Zhang (1996, 1997, 1998). Based upon this MC integration technique, likelihood functions of (nonlinear) dynamic latent variable models can be evaluated with high numerical accuracy allowing for accurate ML estimation of the parameters.

In the following subsections, we provide a description of the general principle of EIS and of its implementation for the basic SV model. Once the baseline algorithm is implemented, it can be adapted easily to various extensions of the basic SV model.

2.2. EIS

Let r_t, $t{:}1 \rightarrow T$ denote an n-dimensional vector of observable random variables and λ_t a q-dimensional vector of latent variables. Let $f(R,\Lambda;\theta)$ represent the joint density of $R=\{r_t\}_{t=1}^{T}$ and $\Lambda=\{\lambda_t\}_{t=1}^{T}$, indexed by the unknown parameter vector θ. The likelihood associated with the observable variables R is given by the $(T\times q)$-dimensional integral

$$L(\theta;R)=\int f(R,\Lambda;\theta)\mathrm{d}\Lambda. \tag{3}$$

Let $R_t=\{r_\tau\}_{\tau=1}^{t}$ and $\Lambda_t=\{\lambda_\tau\}_{\tau=1}^{t}$. The integrand in Eq. (3) can be factorized into the sequence of conditional density functions $f(\cdot)$ for (r_t,λ_t) given (R_{t-1},Λ_{t-1}). Based upon this factorization, the integral in Eq. (3) can be rewritten as

$$L(\theta;R)=\int \prod_{t=1}^{T} f(r_t,\lambda_t \mid \Lambda_{t-1},R_{t-1},\theta)\mathrm{d}\Lambda. \tag{4}$$

For ease of notation it is assumed that the initial conditions $(\lambda_0,\lambda_{-1},\ldots)$ and $(r_0,r_{-1},\ldots)$ are known constants but EIS can easily accommodate alternative (stochastic) assumptions on those. Assumptions relative to the dynamic structure of dynamic latent variable models are typically formulated in terms of the following additional factorization:

$$f(r_t,\lambda_t \mid \Lambda_{t-1},R_{t-1},\theta)=g(r_t \mid \lambda_t,R_{t-1},\theta)p(\lambda_t \mid \Lambda_{t-1},R_{t-1},\theta), \tag{5}$$

where $g(\cdot)$ denotes the conditional density of r_t given (λ_t,R_{t-1}) and $p(\cdot)$ the conditional density of λ_t given (Λ_{t-1},R_{t-1}). Note that this factorization implies that r_t is independent of Λ_{t-1}, given (λ_t,R_{t-1}), an assumption which can be relaxed if necessary. For the basic SV model, $g(\cdot)$ is a conditional Gaussian density defined by Eq. (1) and $p(\cdot)$ is the density associated with the Gaussian AR process given in Eq. (2).

A natural MC estimate of $L(\theta;R)$ based upon this factorization is given by

$$\hat{L}_N(\theta;R)=\frac{1}{N}\sum_{i=1}^{N}\left[\prod_{t=1}^{T} g(r_t \mid \tilde{\lambda}_t^{(i)}(\theta),R_{t-1},\theta)\right], \tag{6}$$

where $\{\tilde{\lambda}_t^{(i)}(\theta)\}_{t=1}^{T}$ denotes a trajectory drawn from the sequence of p densities. Specifically, $\tilde{\lambda}_t^{(i)}(\theta)$ is drawn from the conditional density $p(\lambda_t|\tilde{\Lambda}_{t-1}^{(i)}(\theta),R_{t-1},\theta)$. Note that this MC estimate is based upon a sequence of sampling densities p which is directly obtained

from the statistical specification of the model and which ignores the fact that the observation of R conveys critical information on the underlying latent process. Therefore, the trajectories drawn from the p process, which we shall refer to as natural samplers, have virtually no relation to the actual sequence of λ_t's that generated the observed sample of r_t's. It follows that the natural MC estimator (Eq. (6)) is highly inefficient. In fact, the MC sampling variance of Eq. (6) increases dramatically with the sample size T. For all practical purposes, a prohibitively large MC sample size N would be required in order to obtain reasonably accurate estimates of $L(\theta;R)$. See, for example, Danielsson and Richard (1993).

In order to resolve this efficiency problem, EIS searches for a sequence of samplers that exploits the sample information on the λ_t's conveyed by the r_t's. Let, therefore, $\{m(\lambda_t|\Lambda_{t-1},a_t)\}_{t=1}^{T}$ denote a sequence of auxiliary samplers, typically a straightforward parametric extension of the natural samplers $\{p(\lambda_t|\Lambda_{t-1},R_{t-1},\theta)\}_{t=1}^{T}$, indexed by the auxiliary parameters $A=\{a_t\}_{t=1}^{T}$. For any given values of the auxiliary parameters, the integral (4) is rewritten as

$$L(\theta;R)=\int\prod_{t=1}^{T}\left[\frac{f(r_t,\lambda_t\mid\Lambda_{t-1},R_{t-1},\theta)}{m(\lambda_t\mid\Lambda_{t-1},a_t)}\right]\prod_{t=1}^{T}m(\lambda_t\mid\Lambda_{t-1},a_t)\mathrm{d}\Lambda, \tag{7}$$

and the corresponding importance sampling MC estimate of the likelihood is given by

$$\tilde{L}_N(\theta;R,A)=\frac{1}{N}\sum_{i=1}^{N}\left\{\prod_{t=1}^{T}\left[\frac{f(r_t,\tilde{\lambda}_t^{(i)}(a_t)\mid\tilde{\Lambda}_{t-1}^{(i)}(a_{t-1}),R_{t-1},\theta)}{m(\tilde{\lambda}_t^{(i)}(a_t)\mid\tilde{\Lambda}_{t-1}^{(i)}(a_{t-1}),a_t)}\right]\right\}, \tag{8}$$

where $\{\tilde{\lambda}_t^{(i)}(a_t)\}_{t=1}^{T}$ denotes a trajectory drawn from the sequence of auxiliary (importance) samplers m.

Note that, if we could find a sequence of m densities such that $\prod_t m(\lambda_t|\Lambda_{t-1},a_t)$ is proportional to $\prod_t f(r_t,\lambda_t|\Lambda_{t-1},R_{t-1},\theta)$, then the MC estimate $\tilde{L}_N$ would be equal to the actual likelihood function for all N's and its MC sampling variance would be zero. More generally, EIS aims at selecting values of the auxiliary parameters $\{a_t\}_{t=1}^{T}$ which provide a good match between the product in the numerator and that in the denominator in Eq. (8) in order to minimize the MC sampling variance of $\tilde{L}_N$. This minimization problem, which has to be solved for each relevant value of θ, is large dimensional and has to be broken down into manageable separate low-dimensional subproblems. The factorized form of Eqs. (7) and (8) whereby observations are essentially added one at the time suggests constructing a sequence of subproblems for each period t. Note, however, that, for any given values of R_t and θ, the integral of $f(r_t,\lambda_t|\Lambda_{t-1},R_{t-1},\theta)$ with respect to λ_t does depend upon Λ_{t-1}, while that of $m(\lambda_t|\Lambda_{t-1},a_t)$ equals one by definition. Therefore, it is not possible to secure a good match between the f's and m's period by period, independently from one another.

Instead, for any given (R,θ), EIS requires constructing a (positive) functional approximation $k(\Lambda_t;a_t)$ for the density $f(r_t,\lambda_t|\Lambda_{t-1},R_{t-1},\theta)$ with the requirement that it be analytically integrable with respect to λ_t. In Bayesian terminology, $k(\Lambda_t;a_t)$ serves as a density kernel for $m(\lambda_t|\Lambda_{t-1},a_t)$, which is then given by

$$m(\lambda_t\mid\Lambda_{t-1},a_t)=\frac{k(\Lambda_t;a_t)}{\chi(\Lambda_{t-1},a_t)},\ \text{where}\ \chi(\Lambda_{t-1},a_t)=\int k(\Lambda_t;a_t)\mathrm{d}\lambda_t. \tag{9}$$

Note that a good match between $f(r_t,\lambda_t|\Lambda_{t-1},R_{t-1},\theta)$ alone and $k(\Lambda_t;a_t)$ would leave $\chi(\Lambda_{t-1};a_t)$ unaccounted for. Since, however, $\chi(\Lambda_{t-1};a_t)$ does *not* depend on λ_t, it can be transferred back into the period $t-1$ minimization subproblem. Taken all together, EIS requires solving a simple back-recursive sequence of low-dimensional least-squares problems of the form:[2]

$$\hat{a}_t(\theta) = \arg\min_{a_t} \sum_{i=1}^{N} \left\{ \ln\left[f\left(r_t, \tilde{\lambda}_t^{(i)}(\theta) \mid \tilde{\Lambda}_{t-1}^{(i)}(\theta), R_{t-1}, \theta\right) \cdot \chi\left(\tilde{\Lambda}_t^{(i)}(\theta); \hat{a}_{t+1}(\theta)\right)\right] - c_t - \ln\, k\left(\tilde{\Lambda}_t^{(i)}(\theta); a_t\right) \right\}^2, \tag{10}$$

for $t:T\rightarrow 1$, with $\chi(\Lambda_T;a_{T+1})\equiv 1$. As in Eq. (6), $\{\tilde{\lambda}_t^{(i)}(\theta)\}_{t=1}^{T}$ denotes a trajectory drawn from the p densities, and the c_t's are unknown constants to be estimated jointly with the a_t's. Observe that, if the density kernel $k(\Lambda_t;a_t)$ is chosen within the exponential family of distributions, the EIS least-squares problems become linear in a_t under the canonical representation of exponential kernels. Finally, the EIS estimate of the likelihood function for a given value of θ is obtained by substituting $\{\hat{a}_t(\theta)\}_{t=1}^{T}$ for $\{a_t\}_{t=1}^{T}$ in Eq. (8). Note that the overall dimension of the auxiliary parameter $\{a_t\}_{t=1}^{T}$ is proportional to T.

The least-squares problems (Eq. (10)) are based on trajectories drawn from the natural samplers. Even if these are highly inefficient for the likelihood estimate itself, they typically produce a sequence $\{\hat{a}_t(\theta)\}_{t=1}^{T}$, which leads to vastly improved importance samplers. Nevertheless, in order to obtain maximally efficient importance samplers, a small number of iterations of the EIS algorithm is required, where the natural samplers p are replaced by the previous stage importance samplers. For such iterations to converge to fixed values of the auxiliary parameters $\hat{a}_t$, which are expected to produce optimal importance samplers, it is necessary to apply the technique known as that of Common Random Numbers (CRNs). The use of CRNs means that the N trajectories $\{\tilde{\Lambda}^{(i)}(\cdot)\}_{i=1}^{N}$ for any given sampling density are obtained by transformations of a fixed set of $T{\cdot}N$ canonical random numbers $\{\tilde{U}^{(i)}\}_{i=1}^{N}$, typically uniforms or standardized normals (see, for example, Devroye, 1986).

ML-EIS estimates of θ are obtained by maximizing $\tilde{L}_N(\theta;R,A)$ with respect to θ, using an iterative numerical optimizer. The convergence of such an optimizer, especially one based on numerical gradients, requires itself the use of CRNs, in order to ensure that $\tilde{L}_N(\theta;R,A)$ is a smooth function in θ.

It turns out that EIS provides a very flexible and numerically extremely efficient procedure for the likelihood evaluation of high-dimensional nonlinear latent variable models such as SV models. Our experience is that the integrand in Eq. (4) for such models is a well-behaved function of Λ given R which can be very accurately approximated by sequential EIS importance samplers as indicated by the fact that the R^2's associated with the EIS least-squares problems (Eq. (10)) are typically greater than 0.999.

As mentioned above, an alternative importance sampling approach which can be used to evaluate the likelihood of SV models is that developed by Durbin and Koopman (1997,

[2] An alternative version of EIS also includes the multiplicative weights $g(r_t|\tilde{\lambda}_t^{(i)}(\theta),R_{t-1},\theta)$. However, Eq. (10) is to be preferred as long as these weights exhibit high MC variance, typically the case with draws from the natural samplers.

2000) for the analysis of nonlinear and non-Gaussian state-space models. This approach is based upon an approximation of the original model for the observable variables R and the latent variables Λ by a linear Gaussian state-space model. Let $p(\Lambda;\theta)$ and $f(R|\Lambda,\theta)$ denote the marginal density of Λ and the conditional density of R given Λ under the original model, respectively. Furthermore, let $f_g(R|\Lambda,\theta)$ and $p_g(\Lambda|R,\theta)$ denote the conditional densities of R given Λ and Λ given R under the approximating linear Gaussian model, respectively. Finally, let $L_g(\theta;R)$ denote the likelihood for the approximating model which can be computed by standard Kalman filter techniques. Then, under the assumption that the marginal distribution for Λ is the same in the original as well as in the approximating Gaussian model, the actual likelihood

$$L(\theta;R) = \int f(R \mid \Lambda,\theta)p(\Lambda;\theta)\mathrm{d}\Lambda. \tag{11}$$

can be written as[3]

$$L(\theta;R) = L_g(\theta;R)\int \frac{f(R \mid \Lambda,\theta)}{f_g(R \mid \Lambda,\theta)}p_g(\Lambda \mid R,\theta)\mathrm{d}\Lambda. \tag{12}$$

Note that the Eq. (12) only requires numerical integration to evaluate the departure of the actual likelihood from the likelihood of the approximating model, rather than the actual likelihood itself. The Gaussian density $p_g(\Lambda|R,\theta)$ plays the role of an importance sampler, whereby numerical efficiency requires a selection of the approximating model such that the match between $f_g(R|\Lambda,\theta)$ and $f(R|\Lambda,\theta)$ is as close as possible. In order to achieve such a good match, Durbin and Koopman (1997, 2000) use a local approximation of $f(R|\Lambda,\theta)$ by $f_g(R|\Lambda,\theta)$ in the neighborhood of $E(\Lambda|R)$ under the approximating model.

Compared to EIS, where an importance sampler is obtained directly, the Durbin–Koopman approach is based upon an *indirect* selection of an importance sampler via the construction of an operational approximation to the original model. A further difference between the two methods is that EIS provides a *global* approximation of the integrand in Eq. (11), whereas the Durbin–Koopman method is based upon a corresponding *local* approximation. This is important in so far as global approximations typically lead to more efficient samplers than local ones. Finally, the Durbin–Koopman procedure is specifically designed for models which can be reasonably approximated by a linear Gaussian state-space model leading to a Gaussian sampler, whereas application of EIS, which can accommodate essentially arbitrary classes of samplers, is not restricted to such a type of model.[4] Additional advantages of the EIS algorithm are (1) that it can be iterated producing thereby a critical test of the robustness of the optimal sampler against the choice of an initial sampler; (2) that it produces direct measures of the numerical accuracy of the EIS approximations as, for example, the R^2's of the EIS least squares; and (3) that, as illustrated below, variations of a baseline model require only minor modifications of the initial computer program.

[3] Whereas Durbin and Koopman (1997) consider the case where the latent variables in Λ follow a linear Gaussian model under the original specification, Durbin and Koopman (2000) also allow for corresponding non-Gaussian models.

[4] An explicit and detailed discussion of the properties of the Durbin–Koopman method can be found in the discussion on Durbin and Koopman (2000) following the article.

2.3. The implementation of EIS for the basic SV model

The basic SV model given in Eqs. (1) and (2) assumes that the returns r_t given the log volatility λ_t as well as the log volatility given its past value λ_{t-1} are distributed as Gaussian random variables. Their densities are given by:

$$g(r_t \mid \lambda_t, \theta) \propto \exp\left\{-\frac{1}{2}\left[r_t^2 \exp(-\lambda_t) + \lambda_t\right]\right\} \tag{13}$$

$$p(\lambda_t \mid \lambda_{t-1}, \theta) \propto \exp\left\{-\frac{1}{2v^2}(\lambda_t - \gamma - \delta\lambda_{t-1})^2\right\}, \tag{14}$$

respectively, where multiplicative factors which do not depend upon λ_t are omitted.

The implementation of EIS starts with the selection of a class of density kernels k for the auxiliary importance samplers m capable of providing good functional approximations to the products $f \cdot \chi$. A natural choice for m is to use parametric extensions of the natural samplers p. For the basic SV model, this implies that $k(\Lambda_t;a_t)$ be a Gaussian density kernel for λ_t given λ_{t-1}. Furthermore, the fact that density kernels of Gaussian distributions are closed under multiplication (see DeGroot, 1970; Richard and Zhang, 1998) immediately suggests the following parametrization for k:

$$k(\Lambda_t; a_t) = p(\lambda_t \mid \lambda_{t-1}, \theta)\zeta(\lambda_t, a_t), \tag{15}$$

where the auxiliary function $\zeta(\lambda_t,a_t)$ is itself a Gaussian density kernel. Under this parametrization, the natural sampler p cancels out in the least-squares problem (Eq. (10)) to the effect that $\ln \zeta(\lambda_t,a_t)$ serves to approximate $\ln g(r_t|\lambda_t,R_{t-1},\theta) + \ln \chi(\Lambda_t,a_{t+1})$. In particular, the appropriate auxiliary function for the basic SV model is given by $\zeta(\lambda_t,a_t) = \exp(a_{1,t}\lambda_t + a_{2,t}\lambda_t^2)$, with $a_t=(a_{1,t},a_{2,t})$, and the density kernels of the importance samplers have the form

$$k(\Lambda_t; a_t) \propto \exp\left\{-\frac{1}{2}\left[\left(\frac{\gamma + \delta\lambda_{t-1}}{v}\right)^2 - 2\left(\frac{\gamma + \delta\lambda_{t-1}}{v^2} + a_{1,t}\right)\lambda_t + \left(\frac{1}{v^2} - 2a_{2,t}\right)\lambda_t^2\right]\right\}. \tag{16}$$

Accordingly, the conditional mean and variance of λ_t on m are given by

$$\mu_t = \sigma_t^2\left(\frac{\gamma + \delta\lambda_{t-1}}{v^2} + a_{1,t}\right), \quad \sigma_t^2 = \frac{v^2}{1 - 2v^2 a_{2,t}}, \tag{17}$$

respectively. Integrating $k(\Lambda_t;a_t)$ with respect to λ_t and omitting irrelevant multiplicative factors leads to the following expression for the integrating constant

$$\chi(\lambda_{t-1}, a_t) \propto \exp\left\{\frac{\mu_t^2}{2\sigma_t^2} - \frac{(\gamma + \delta\lambda_{t-1})^2}{2v^2}\right\}, \tag{18}$$

which is itself a Gaussian density kernel for λ_{t-1}. Based upon these functional forms, the computation of an EIS estimate of the likelihood for the basic SV model requires the following steps:

Step (0): Use the natural samplers p to draw N trajectories of the latent variable $\{\tilde{\lambda}_t^{(i)}(\theta)\}_{t=1}^{T}$.

Step (t); (t:$T \rightarrow 1$): Use these random draws to solve the back-recursive sequence of least-squares problems defined in Eq. (10). The step t least-squares problem is characterized by the following linear auxiliary regression:

$$-\frac{1}{2}\left[r_t^2 \exp(-\tilde{\lambda}_t^{(i)}(\theta)) + \tilde{\lambda}_t^{(i)}(\theta)\right] + \ln\chi(\tilde{\lambda}_t^{(i)}(\theta), \hat{a}_{t+1}(\theta))$$
$$= \text{constant} + a_{1,t}\tilde{\lambda}_t^{(i)}(\theta) + a_{2,t}[\tilde{\lambda}_t^{(i)}(\theta)]^2 + \eta_t^{(i)}, \quad i: 1 \rightarrow N, \tag{19}$$

where $\eta_t^{(i)}$ denotes the regression error term. The initial condition for the integrating constant (Eq. (18)) is given by $\chi(\lambda_T,\cdot) \equiv 1$.

Step ($T+1$): The EIS samplers $\{m(\lambda_t|\Lambda_{t-1},\hat{a}_t(\theta))\}_{t=1}^{T}$, which are characterized by the conditional mean and variance given in Eq. (17), are used to draw N trajectories $\{\tilde{\lambda}_t^{(i)}(\hat{a}_t(\theta))\}_{t=1}^{T}$ from which the EIS estimate of the likelihood is calculated according to Eq. (8).

As mentioned above, a small number of iterations of the EIS algorithm (from 3 to 5 iterations) are required to obtain maximally efficient importance samplers. Instead of starting these iterations with the natural samplers p, one can also begin with importance samplers which are based upon a second-order Taylor series expansion (TSE) of $\{\ln g(r_t|\lambda_t,\theta)\}$. In contrast to the natural samplers, these TSE importance samplers exploit sample information on the λ_t's conveyed by the r_t's and may, therefore, reduce the number of EIS iterations by one or two. For the basic SV model, the TSE of ln g around zero is given by

$$-\frac{1}{2}[r_t^2 \exp(-\lambda_t) + \lambda_t] \cong \text{constant} + \frac{1}{2}(r_t^2 - 1)\lambda_t - \frac{1}{4}r_t^2\lambda_t^2, \tag{20}$$

which implies the following coefficient values in the auxiliary function ζ: $a_{t,1}=(r_t^2-1)/2$ and $a_{t,2}=-r_t^2/4$. These values are inserted in Eq. (17) to obtain the moments of the corresponding Gaussian TSE importance samplers.

Once this EIS algorithm has been programed for the basic SV model, changes in the SV model only require minor modifications of the program. For example, a substitution of the conditional Gaussian distribution for r_t given λ_t by another suitable distribution changes only a few lines in the code which are associated with the definition of the density $g(r_t|\lambda_t, \theta)$.

Finally, it should be mentioned that, within the context of the QML approach of Harvey et al. (1994), the EIS samplers lead to a MC sampling variance of the likelihood estimate of zero. To estimate the basic SV model, they use the following transformation of the return Eq. (1): $\ln r_t^2 = -1.27 + \lambda_t + \epsilon_t^*$, where ϵ_t^* is iid with mean zero and variance $\pi^2/2$, and approximate ϵ_t^* by a normal random variable. If EIS is applied to evaluate the corresponding quasi likelihood $L(\theta;R^*)$, where $R^*=\{r_t^*\}_{t=1}^{T}$ and $r_t^* = \ln r_t^2 - 1.27$, the conditional density given in Eq. (13) has to be replaced by $g(r_t^*|\lambda_t) \propto \exp\{-(r_t^*-\lambda_t)^2/\pi^2\}$. It follows that the functions g, χ and ζ are all Gaussian density kernels allowing for a

perfect fit in the corresponding auxiliary regressions and, therefore, resulting in zero MC variance for the EIS quasi likelihood estimates.

2.4. Filtering and diagnostics

Once the parameters of the SV model have been estimated, EIS can be used to produce appropriate sequences of filtered estimates of the latent variable λ_t or of functions thereof which can serve as the basis for diagnostic checking. Alternative approaches of computing filtered estimates of the latent variables in SV models are the reprojection techniques of Gallant and Tauchen (1998) based upon a projection of the data onto a semi-nonparametric transition density and the particle filter algorithm used by Kim et al. (1998).[5]

Let $h(\lambda_t)$ denote a function such as, for example, $\exp(\lambda_t)$, which represents the conditional return variance. Then the sequence of conditional expectations of $h(\lambda_t)$ given R_{t-1}, the past observations of returns, provides a sequence of filtered estimates of $h(\lambda_t)$. For the SV model these expectations take the form of a ratio of integrals:

$$E[h(\lambda_t) \mid R_{t-1}] = \frac{\int h(\lambda_t) p(\lambda_t \mid \Lambda_{t-1}, R_{t-1}, \theta) f(R_{t-1}, \Lambda_{t-1}; \theta) \mathrm{d}\Lambda_t}{\int f(R_{t-1}, \Lambda_{t-1}; \theta) \mathrm{d}\Lambda_{t-1}}. \tag{21}$$

In the sequel we assume that the parameter vector θ is known. In practice, θ is set equal to its ML-EIS estimate. The integral in the denominator corresponds to the likelihood function $L(\theta; R_{t-1})$ associated with the first $t-1$ observations and can be evaluated by MC integration using the sequence of samplers $\{m(\lambda_\tau|\Lambda_{\tau-1}, \hat{a}_\tau^{t-1})\}_{\tau=1}^{t-1}$, where $\{\hat{a}_\tau^{t-1}\}_{\tau=1}^{t-1}$ denotes the values of the auxiliary parameters characterizing the EIS samplers for $L(\theta; R_{t-1})$. An importance MC estimate of the integral in the numerator is given by

$$\frac{1}{N}\sum_{i=1}^{N}\left\{h[\tilde{\lambda}_t^{(i)}(\theta)]\prod_{\tau=1}^{t-1}\left[\frac{f(r_\tau, \tilde{\lambda}_\tau^{(i)}(\hat{a}_\tau^{t-1}) \mid \tilde{\Lambda}_{\tau-1}^{(i)}(\hat{a}_{\tau-1}^{t-1}), R_{\tau-1}, \theta)}{m(\tilde{\lambda}_t^{(i)}(\hat{a}_\tau^{t-1}) \mid \tilde{\Lambda}_{\tau-1}^{(i)}(\hat{a}_{\tau-1}^{t-1}), \hat{a}_\tau^{t-1})}\right]\right\}, \tag{22}$$

where $\{\tilde{\lambda}_\tau^{(i)}(\hat{a}_\tau^{t-1})\}_{\tau=1}^{t-1}$ denotes a trajectory drawn from the sequence of EIS samplers associated with $L(\theta; R_{t-1})$ and $\tilde{\lambda}_\tau^{(i)}(\theta)$ is a draw from $p(\lambda_t|\tilde{\Lambda}_{t-1}^{(i)}(\hat{a}_{t-1}^{t-1}), R_{t-1}, \theta)$. Actually, using the same draws to evaluate the numerator and denominator of the ratio (Eq. (21)) contributes reducing further its MC sampling variance (see, for example, Geweke, 1989). Note that the evaluation of the sequence of filtered estimates requires the determination of a new sequence of EIS samplers for each $E[h(\lambda_t)|R_{t-1}]$, $t{:}1 \to T$. In the following applications, we rerun the complete EIS algorithm for each t to obtain such updated sequences of EIS samplers. However, the individual EIS samplers for λ_τ, $\tau{:}1 \to t-1$ for $L(\theta; R_{t-1})$ and $L(\theta; R_t)$ only differ for observations close enough to t. In the future, we intend to exploit this important feature to develop efficient algorithms for computing sequentially updated EIS samplers, whereby observations are added one at the time.

[5] In general, the use of particle filter algorithms to produce estimates of latent state variables in nonlinear and non-Gaussian state-space models are discussed, for example, in Kitagawa (1996) and Pitt and Shephard (1999a).

Filtering enables us to compute a sequence of standardized residuals of the form:

$$z_t = [r_t - E(r_t \mid R_{t-1})]\mathrm{Var}(r_t \mid R_{t-1})^{-1/2}. \tag{23}$$

For the basic SV model, the conditional mean and variance of r_t given R_{t-1} are given by zero and $E[\exp(\lambda_t)|R_{t-1}]$, respectively. If the model is correctly specified, z_t has mean zero and unit variance and is serially uncorrelated in the first- and second-order moments. Hence, the series of z_t can be used for diagnostic checking of the assumed dynamic structure in the mean and variance. Note that the moments in Eq. (23) are marginalized with respect to λ_t. Therefore, the distribution of z_t under the correct specification is unknown and can display serial dependence in higher order moments (see, for example, Gallant et al., 1991).

In order to check the distributional assumptions of the SV model, we use an approach similar to that followed by Kim et al. (1998), which requires computing the conditional probability that r_t be less than the actually observed return r_t^o, i.e.[6] $u_t = \Pr(r_t \le r_t^o|R_{t-1})$. This probability corresponds to the following choice for the function $h(\lambda_t)$ in Eq. (21):

$$h(\lambda_t) = F\Big([r_t^o - E(r_t \mid \lambda_t)]\mathrm{Var}(r_t \mid \lambda_t)^{-1/2}\Big), \tag{24}$$

where $F(\cdot)$ denotes the distribution function of the variable $[r_t - E(r_t|\lambda_t)]\mathrm{Var}(r_t|\lambda_t)^{-1/2}$. For the basic SV model, this variable is given by $r_t\exp(-\lambda_t/2)$ and follows a standard normal distribution. If the model is correctly specified, u_t is a serially independent random variable following a uniform distribution on [0,1] (see Rosenblatt, 1952; Kim et al., 1998). Using the inverse of a standard normal distribution function, the variable u_t can be mapped into a standard normal distribution:

$$z_t^* = F_N^{-1}(u_t). \tag{25}$$

Under the hypothesis of a correctly specified model, these normalized residuals are serially independent random variables following a standard normal distribution.

2.5. Application

The ML-EIS approach is applied to estimate the SV model for daily data of the following financial time series: The IBM stock price, the S&P500 stock price index and the exchange rate for the US Dollar/Deutsche Mark ($/DM). The daily prices s_t are transformed into continuously compounded rates centered around their sample mean: $r_t = 100[\ln(s_t/s_{t-1}) - (1/T)\sum_{t=1}^{T}\ln(s_t/s_{t-1})]$. The IBM series, which is also analyzed in Liesenfeld and Jung (2000), are closing prices from January 1, 1973 to December 31, 1991, leaving a sample of 4694 observations. The S&P500 series is taken from the dataset of Danielsson (1998) and consists of 4671 prices from January 1, 1975 to December 31, 1993. The $/DM data are noon spot prices from the Chicago Federal Reserve Bank with a sample period from January 1, 1980 to December 31, 1998 for a total of 4774 observations.

[6] In the context of density forecast evaluation, the use of the probability u_t is discussed in Diebold et al. (1998, 1999).

Table 1
ML-EIS estimation results for the basic SV model

	IBM	S&P500	$/DM
γ	0.017 (0.005) [0.0001]	− 0.009 (0.003) [0.0001]	− 0.038 (0.008) [0.0002]
δ	0.964 (0.009) [0.0003]	0.977 (0.005) [0.0002]	0.962 (0.008) [0.0002]
v	0.175 (0.021) [0.0010]	0.138 (0.015) [0.0006]	0.186 (0.018) [0.0006]
Log-likelihood	− 7959.6 [0.296]	− 5804.5 [0.256]	− 4687.2 [0.162]

Asymptotic (statistical) standard errors obtained from a numerical approximation to the Hessian are in parentheses and MC (numerical) standard errors are in brackets. The ML-EIS estimates are based on a MC sample size $N = 50$ and three EIS iterations.

ML-EIS estimation results for the basic SV model given in Eqs. (1) and (2), based upon a simulation sample size $N = 50$, are given in Table 1. Each EIS likelihood evaluation requires approximately 8 s on a Pentium II Celeron 466 MHz notebook for a code written in GAUSS. A full ML estimation requires approximately 13 BFGS iterations and takes of the order of 12 min. The ML-EIS parameter estimates are numerically very accurate, as indicated by the MC (numerical) standard deviations reported between brackets, which were computed from 20 ML-EIS estimations conducted under different sets of CRNs. The estimates of the persistence parameter δ are always highly significant and lie between 0.96 and 0.98. Such a high persistence of volatility shocks is in full accordance with results reported in the previously cited studies of SV models.

The filtered estimates of the volatility $E[\exp(\lambda_t)|R_{t-1}]$ obtained from the basic SV model for IBM are displayed in Fig. 1 together with the time series and the quantile–quantile plot of the normalized residuals z_t^*. Unsurprisingly, the movements of the filtered volatility estimates follow closely the spread in the return series. Nevertheless, the graphs for the normalized residuals indicate that there are a few large values of the observed returns such as, for example, that of the October 1987 crash which cannot be explained by the model. Table 2 presents the results of diagnostic checks on the residuals z_t^* and z_t. As seen from this table, the Kolmogorov–Smirnov z-statistic KS(z^*) indicates a rejection of normality for the IBM residuals z_t^*, while the S&P500 and $/DM residuals pass this test at the 1% significance level. However, the values of the kurtosis of the IBM and S&P500 residuals z_t^* considerably exceed 3, which is the benchmark value for a normal random variable. This suggests that the model has difficulties explaining the observed leptokurtosis in the returns. In this context, it is important to note that the kurtosis statistic is much more sensitive to deviations in the tail area of the distribution than the Kolmogorov–Smirnov statistic. The Ljung–Box statistic $Q_{30}(\cdot)$ for the squared residuals z_t^{*2} and z_t^2 including 30 lags indicates that the model successfully accounts for the serial correlation in the volatility of all three return series, while the Ljung–Box statistic for z_t^* and z_t points towards residual autocorrelation for S&P500 and $/DM. This suggests including an autoregressive component in the mean Eq. (1) for these two series.[7] Taken all together,

[7] For the S&P500, we reestimated the basic SV model as well as the other SV specifications discussed below including a first-order autoregressive mean component. The corresponding residuals z_t^* and z_t pass the Ljung–Box test at conventional significance levels. Since the parameter estimates and all other test statistics are virtually the same as those obtained without an autoregressive mean component, we will not consider further this extension in what follows.

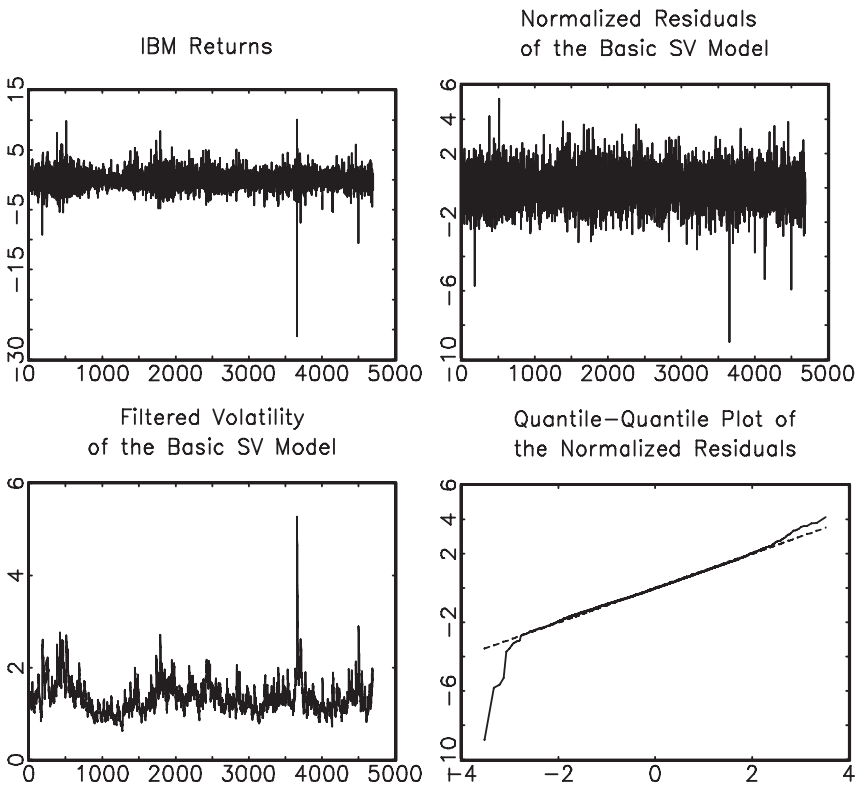

Fig. 1. Top row: time series of the IBM returns r_t (left) and of the normalized residuals z_t^* obtained from the basic SV model (right); bottom row: time series of $E[\exp(\lambda_t)|R_{t-1}]^{1/2}$ (left) and quantile−quantile plot of z_t^* (right); the dashed line plots the quantiles of the standard normal distribution against the quantiles of the standard normal and the solid line plots the sorted values of z_t^* against the quantiles of the standard normal.

the empirical evidence suggests that the return series exhibit properties, especially distributional characteristics, which are not completely explained by the basic SV model. Hence, extensions of the basic SV model appear to be required in order to approximate the return series. Several extensions are considered in the following section.

Table 2
Diagnostics for the basic SV model

	Skewness	Kurtosis	KS(z^*)	$Q_{30}(z^*)$	$Q_{30}(z^{*2})$	$Q_{30}(z)$	$Q_{30}(z^2)$
IBM	−0.08	5.12	1.710 (0.006)	25.9 (0.679)	15.8 (0.984)	26.6 (0.642)	10.8 (0.999)
S&P500	−0.34	6.43	1.549 (0.017)	82.6 (<0.001)	27.7 (0.588)	79.5 (<0.001)	19.4 (0.931)
$/DM	−0.05	3.38	1.422 (0.035)	56.8 (0.002)	25.2 (0.716)	59.4 (0.001)	22.2 (0.846)

Probability values are given in parentheses.

3. Univariate extensions of the basic SV model

3.1. Fat-tailed error distribution

The diagnostics for the basic SV model suggest that it cannot completely account for the leptokurtosis in the IBM and S&P500 returns. Hence, a natural extension of the SV model consists of assuming that the distribution of ϵ_t in Eq. (1) has fat tails. One natural candidate is the scaled Student's t-distribution as used, for example, by Gallant et al. (1997), Sandmann and Koopman (1998) and Liesenfeld and Jung (2000). Under this specification, the density of the conditional distribution of r_t given λ_t is given by

$$g(r_t \mid \lambda_t, \theta) \propto \exp(-\lambda_t/2)\left[1+\frac{r_t^2\exp(-\lambda_t)}{\omega-2}\right]^{-(\omega+1)/2}, \quad \omega > 2, \tag{26}$$

where ω denotes the degrees of freedom (for $\omega \to \infty$, the t-distribution approaches the normal distribution). As mentioned above, the implementation of EIS for the likelihood evaluation of such an extension only requires the replacement of density (13) by density (26) in the EIS procedure initially designed for the basic SV model, a trivial modification.

The ML-EIS estimation results for the SV-t model are presented in Table 3. Likelihood ratio statistics (LR) of the hypothesis H_0: $1/\omega = 0$ for all three series indicate a strong rejection of the model with Gaussian errors against the SV-t specification. Estimated degrees of freedom vary between 7.81 (S&P500) and 10 ($/DM), indicating large deviations from normality. Furthermore, estimates of the persistence parameter δ under the SV-t model are uniformly greater than those obtained under the basic SV model, whereas estimated variance parameters v are smaller. As discussed in Liesenfeld and Jung (2000), these differences can be explained by the findings that large return observations are interpreted as transitory shocks in the fat-tail mean process under the t-specification and, in sharp contrast, as persistent shocks in the volatility process under the Gaussian specification. Finally, we note that the statistical and MC standard errors are smaller under the SV-t model than under the basic model for all parameters common to both specifications, as expected from the fact that the SV-t model more easily accommodates outliers in the data.

Fig. 2 displays the time series and the quantile–quantile plot of the IBM residuals z_t^* under the SV-t model. Obviously, t-errors are most useful in the tail areas of the error

Table 3
ML-EIS estimation results for the SV model with t-errors

	IBM	S&P500	$/DM
γ	0.004 (0.002) [<0.0001]	− 0.002 (0.002) [<0.0001]	− 0.024 (0.006) [0.0001]
δ	− 0.993 (0.003) [<0.0001]	0.992 (0.003) [<0.0001]	0.975 (0.006) [0.0001]
v	− 0.066 (0.011) [0.0001]	0.072 (0.011) [0.0001]	0.140 (0.016) [0.0002]
$1/\omega$	0.116 (0.013) [0.0001]	0.128 (0.013) [0.0001]	0.100 (0.017) [0.0001]
Log-likelihood	− 7922.3 [0.038]	− 5757.0 [0.046]	− 4666.9 [0.082]
LR	74.5	95.1	40.6

Asymptotic (statistical) standard errors obtained from a numerical approximation to the Hessian are in parentheses and MC (numerical) standard errors are in brackets. The ML-EIS estimates are based on a MC sample size $N = 50$ and three EIS iterations. The likelihood ratio test follows a $\chi^2_{(1)}$-distribution with the 1% critical value being 6.63.

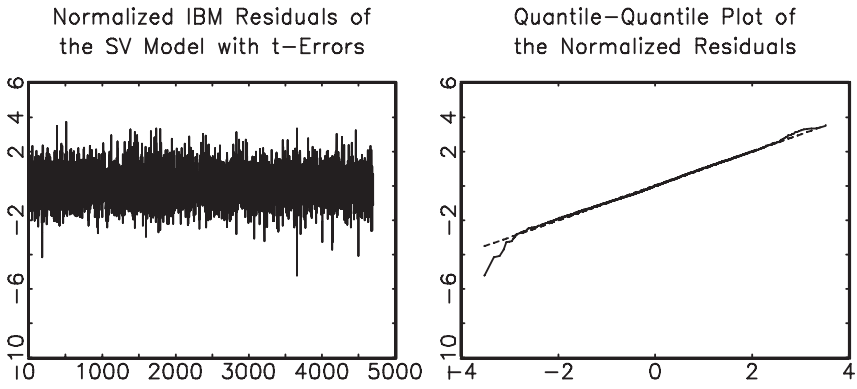

Fig. 2. Normalized IBM residuals z_t^* obtained from the SV model with t-errors (left) and quantile−quantile plot of z_t^* (right); the dashed line plots the quantiles of the standard normal distribution against the quantiles of the standard normal and the solid line plots the sorted values of z_r^* against the quantiles of the standard normal.

distribution, as confirmed by the values of the kurtosis for the z_t^*'s presented in Table 4. In fact, in all three cases the kurtosis of z_t^* is very close to that of a normal distribution. However, the Kolmogorov–Smirnov test still rejects normality for the IBM filtered residuals, while the S&P500 and \$/DM filtered residuals pass this test at conventional significance levels. Furthermore, the Ljung–Box statistic for the squared IBM residuals z_t^2 and of the squared S&P500 residuals z_t^{*2} and z_t^2 suggests a failure to account for the dynamics in the volatility.

3.2. Semi-nonparametric error distribution

A more flexible alternative to the normal distribution is offered by the class of semi-nonparametric (SNP) distributions of Gallant and Nychka (1987). Let, therefore, z_t denote a random variable with the following density function

$$f(z_t) = \frac{1}{\kappa}\left(\sum_{j=0}^{K}\alpha_j z_t^j\right)^2 \phi(z_t), \quad \text{with } \kappa = \int\left(\sum_{j=0}^{K}\alpha_j z_t^j\right)^2 \phi(z_t)\mathrm{d}z_t, \tag{27}$$

where $\phi(\cdot)$ denotes the standard normal density and $\sum_{j=0}^{K}\alpha_j z_t^j$ is a polynomial of degree K in z_t with coefficients α_j, $j{:}1 \rightarrow K$. The constant term α_0 in the polynomial is set equal to 1

Table 4
Diagnostics for the SV model with t-errors

	Skewness	Kurtosis	KS(z^*)	$Q_{30}(z^*)$	$Q_{30}(z^{*2})$	$Q_{30}(z)$	$Q_{30}(z^2)$
IBM	0.06	3.17	1.753 (0.004)	25.7 (0.688)	49.6 (0.014)	30.8 (0.427)	58.3 (0.001)
S&P500	−0.06	3.14	0.789 (0.561)	83.1 (<0.001)	59.4 (0.001)	76.4 (<0.001)	72.6 (<0.001)
\$/DM	−0.05	2.98	0.755 (0.619)	56.3 (0.003)	19.6 (0.926)	60.6 (0.001)	15.6 (0.986)

Probability values are given in parentheses.

Table 5
ML-EIS estimation results for the SV model with SNP-errors

	IBM	S&P500	$/DM
γ	0.006 (0.002) [0.0002]	− 0.003 (0.002) [< 0.0001]	− 0.025 (0.006) [0.0001]
δ	− 0.989 (0.003) [0.0001]	0.989 (0.003) [0.0001]	0.974 (0.006) [0.0001]
ν	− 0.086 (0.012) [0.0009]	0.092 (0.010) [0.0003]	0.143 (0.016) [0.0005]
α_1	0.131 [0.0064]	0.150 [0.0015]	0.059 [0.0008]
α_2	− 0.190 [0.0017]	− 0.161 [0.0010]	− 0.021 [0.0006]
α_3	− 0.007 [0.0006]	− 0.013 [0.0001]	− 0.005 [0.0001]
α_4	0.017 [0.0001]	0.014 [< 0.0001]	0.015 [0.0001]
Log-likelihood	− 7911.7 [0.214]	− 5765.7 [0.129]	−4662.9 [0.177]
LR	95.7	77.5	48.7

Asymptotic (statistical) standard errors obtained from a numerical approximation to the Hessian are in parentheses and MC (numerical) standard errors are in brackets. The ML-EIS estimates are based on a MC sample size $N = 50$ and three EIS iterations. The likelihood ratio test follows a $\chi^2_{(4)}$-distribution with the 1% critical value being 13.28.

in order to identify the α_j's. The integrating constant κ as well as the mean and variance of z_t, denoted by μ_z and σ_z^2, respectively, can be computed by applying the recursive formulas for the moments of the standard normal distribution (see, for example, Patel and Read, 1982). For $K = 0$, $f(z_t)$ is a Gaussian density. As illustrated in Fenton and Gallant (1996), the class of SNP densities can produce density functions with a very broad range of different shapes, including densities with fat-t like tails, multimodal and skewed densities.

The standardized error term in the mean Eq. (1) is now defined as $\epsilon_t = (z_t - \mu_z)/\sigma_z$, where z_t has a SNP density function. The resulting density of the conditional distribution of r_t given λ_t has the form

$$g(r_t \mid \lambda_t, \theta) \propto \exp(-\lambda_t/2)\left(\sum_{j=0}^{K} \alpha_j [\mu_z + \sigma_z r_t \exp(-\lambda_t/2)]^j\right)^2 \phi[\mu_z + \sigma_z r_t \exp(-\lambda_t/2)]. \tag{28}$$

In very high-dimensional problems such as those discussed here, one cannot afford letting the ML optimizer search in regions of the parameter space that would be either inadmissible (for example, implying negative variances) or totally implausible (generating numerical underflows or other numerical pathologies). Experience indicates that sufficient flexibility is preserved by restricting the α_j's to an appropriate neighborhood of zero. In the present paper, we use the transformation $\alpha_j = -c_j + 2c_j[\sin(\alpha_j^\star)]^2$, with $c_j > 0$ and $\alpha_j^\star \in (-\infty, \infty)$, so that $\alpha_j \in (-c_j, c_j)$. The c_j's are chosen large enough to be non-binding at convergence.[8]

We use Schwarz's (1978) information criterion (SIC) in order to select the degree of the polynomial K. For all three series, SIC selects a model with $K = 4$. The corresponding ML-EIS estimation results are presented in Table 5. As for the specifications discussed above, the parameter estimates of the SV-SNP model are numerically very accurate. Furthermore,

[8] Following Fenton and Gallant (1996), we also estimated the SNP model under a formulation, in which the density (28) is rewritten in terms of orthogonal Hermite polynomials in order to enhance numerical stability. It turned out that the results (not presented here) are the same as those under the formulation used in Eq. (28).

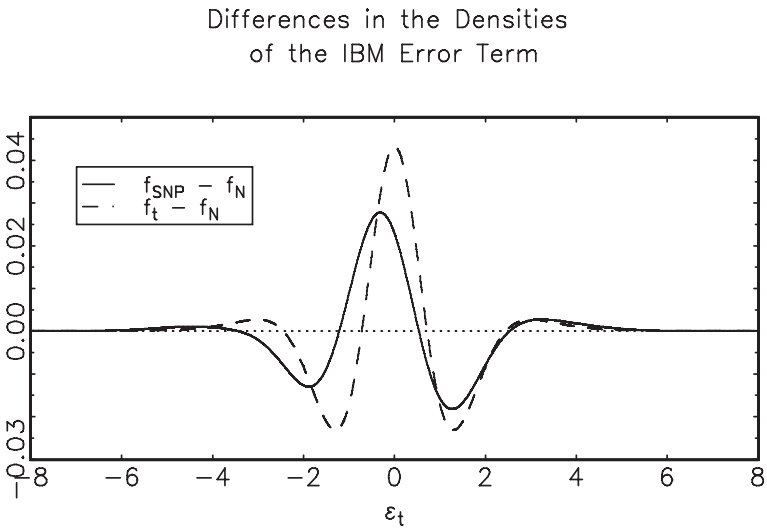

Fig. 3. Estimated density functions of the IBM errors ϵ_t under the SV-t model (dashed line) and under the SV-SNP model (solid line), both expressed in deviations from the normal density function of the basic SV model.

the values of the likelihood ratio test statistic for the null hypothesis of normality ($\alpha_j = 0$; $j{:}1 \rightarrow 4$) indicate that the SNP-SV specification represents a significant improvement over the Gaussian specification for all three series. Based upon a comparison of the log likelihood values, however, the SNP model is preferable to the t model only for IBM and \$/DM. Finally, we note that the estimates of the α_j's, which shape deviations from normality, are fairly stable across all three series and that the estimated values of the volatility parameters δ and ν lie between those previously obtained under the basic SV model and the SV-t model.

Fig. 3 displays the estimated density functions for the IBM errors ϵ_t under the SNP model (f_{SNP}) and the t model (f_t), both expressed in deviations from the normal density of the basic SV model (f_{N}). As expected, f_t is more peaked around zero and exhibits fatter tails relatively to f_{N}. The density f_{SNP} is also more peaked near zero than f_{N}, though not as much as f_t. Furthermore, f_{SNP} is slightly skewed to the right and exhibits asymmetric behavior in the tails. Whereas the right tail behaves more like f_t, the left tail is thinner than that of f_{N} and shows some evidence of a side lobe around -5.

The time series and the quantile–quantile plot of the normalized IBM residuals z_t^* displayed in Fig. 4 indicate that the SNP model approximates the distributional properties of the IBM returns fairly well.[9] Notice, however, that the outlier associated with the October 1987 crash remains hard to account for. Diagnostic checks for the residuals are presented in Table 6. The normalized residuals for all three series pass the Kolmogorov–

[9] According to Eq. (24) the calculation of the normalized residuals z_t^* requires the computation of $F(r_t^{\circ} \exp\{-\lambda_t/2\})$, where $F(\cdot)$ is a SNP distribution function. This function is calculated analytically by applying the recursive formula for the incomplete moments of the standard normal distribution (see, for example, Patel and Read, 1982).

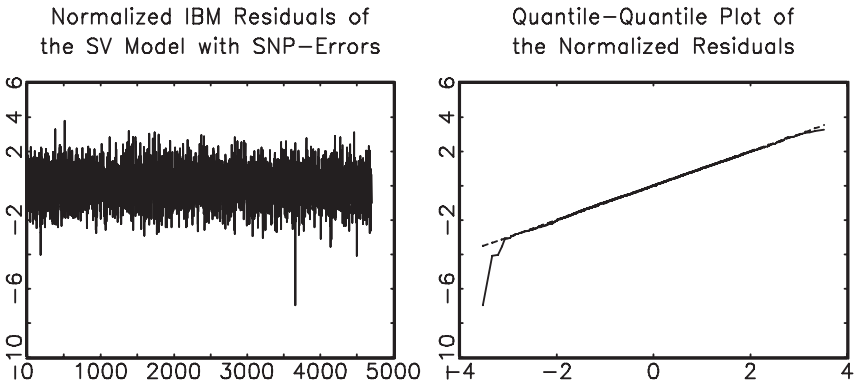

Fig. 4. Normalized IBM residuals z_t^* obtained from the SV model with SNP-errors (left) and quantile–quantile plot of z_t^* (right); the dashed line plots the quantiles of the standard normal distribution against the quantiles of the standard normal and the solid line plots the sorted values of z_t^* against the quantiles of the standard normal.

Smirnov test, whereas the kurtosis for the IBM and S&P500 residuals suggests that leptokurtosis is not fully accounted for. According to the Ljung–Box statistic for z_t^{*2} and z_t^2, the SV-SNP model fits the volatility dynamics of all three return series.

3.3. Two-component SV model

It is well-known that the Student's t-distribution can be interpreted as a mixture of normal distributions with a stochastic variance following an inverted gamma distribution (see, for example, Bollerslev, 1987). Hence, the SV model with t-errors can be interpreted as a SV model with two independent volatility processes: a dynamic process given by $\exp(\lambda_t/2)$ and a static process modeled as the square root of an inverted gamma random variable. Thus, another natural extension of the basic SV model given in Eqs. (1) and (2) consists of introducing a second dynamic volatility process as discussed, for example, in Shephard (1996).

Let r_t denote a univariate return process and λ_t a 2×1 vector containing the elements $\lambda_{1,t}$ and $\lambda_{2,t}$. A SV model with two independent dynamic volatility processes is then defined as

$$r_t = \exp(\iota'\lambda_t/2)\epsilon_t, \tag{29}$$

$$\lambda_t = \Gamma + \Delta\lambda_{t-1} + V\eta_t, \tag{30}$$

where $\iota=(1,1)'$ and $\eta_t=(\eta_{1,t},\eta_{2,t})'$ together with

$$\Gamma = (\gamma_1, 0)', \quad \Delta = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix}.$$

Table 6
Diagnostics for the SV model with SNP-errors

	Skewness	Kurtosis	KS(z^*)	$Q_{30}(z^*)$	$Q_{30}(z^{*2})$	$Q_{30}(z)$	$Q_{30}(z^2)$
IBM	−0.10	3.51	1.169 (0.130)	26.0 (0.674)	34.3 (0.269)	28.2 (0.560)	22.6 (0.830)
S&P500	−0.18	4.07	1.348 (0.053)	82.8 (<0.001)	43.9 (0.048)	77.9 (<0.001)	23.4 (0.797)
$/DM	0.007	3.02	0.624 (0.832)	55.2 (0.003)	19.1 (0.938)	60.3 (0.001)	15.0 (0.990)

Probability values are given in parentheses.

The innovations ϵ_t, $\eta_{1,t}$ and $\eta_{2,t}$ are assumed to be mutually independent iid Gaussian random variables with zero means and unit variances. In order to ensure stationarity of r_t, we impose the constraints $|\delta_1|<1$ and $|\delta_2|<1$. The absence of a constant term in $\lambda_{2,t}$ reflects an identifiability restriction. For complete identification, we impose the additional restriction $\delta_1>\delta_2$, ensuring that $\lambda_{1,t}$ and $\lambda_{2,t}$ are not interchangeable and that $\iota'\lambda_t$ does not collapse into an unidentified AR(1) specification for $\delta_1=\delta_2$. Finally, notice that if $\delta_2=0$ and $v_2\to 0$, then the second volatility component vanishes and the two-component model coincides with the basic SV model.

This two-component SV model is similar to the two-component GARCH model proposed by Engle and Lee (1999) where the long memory behavior of the return volatility is modeled as the sum of a long-run and a short-run GARCH component. By imposing the constraint $\delta_1>\delta_2$ in the two-component SV model, we assume that $\lambda_{1,t}$ represents the long-run component of volatility and $\lambda_{2,t}$ its short-run component.[10]

The two-component EIS procedure essentially requires replacing the univariate λ_t samplers used earlier by the corresponding bivariate samplers, once again a fairly straightforward extension of the initial EIS algorithm. Since in particular, the natural sampling densities p for the two-component model are bivariate Gaussian densities, the auxiliary functions used to construct the importance samplers are themselves kernels of bivariate Gaussian densities and are given by $\zeta(\lambda_t,a_t)=\exp(\lambda_t'B_t+\lambda_t'\mathbf{C}_t\lambda_t)$, where $B_t=(b_{1,t},b_{2,t})'$ and $\mathbf{C}_t$ is a symmetric 2×2 matrix with the elements $\{c_{j,l,t}\}$. The corresponding density kernels of the importance samplers, as given by $p(\lambda_t|\lambda_{t-1},\theta)\ \zeta(\lambda_t,a_t)$, have the form

$$k(\lambda_t,a_t)\propto\exp\Big\{-\frac{1}{2}[(\Gamma+\Delta\lambda_{t-1})'\mathbf{W}^{-1}(\Gamma+\Delta\lambda_{t-1})-2\lambda_t'[\mathbf{W}^{-1}(\Gamma+\Delta\lambda_{t-1})+B_t]+\lambda_t'(\mathbf{W}^{-1}-2\mathbf{C}_t)\lambda_t]\Big\}, \tag{31}$$

where $\mathbf{W}=\mathbf{V}\mathbf{V}'$. Accordingly, the mean and variance-covariance matrix of the importance samplers are given by $\mu_t=\sum_t[\mathbf{W}^{-1}(\Gamma+\Delta\lambda_{t-1})+B_t]$ and $\sum_t=(\mathbf{W}^{-1}-2\mathbf{C}_t)^{-1}$, respectively. Note that the specification of the model implies natural samplers p whereby $\lambda_{1,t}$ and $\lambda_{2,t}$ are uncorrelated. However, the conditional density function $g(r_t|\lambda_t,\theta)$ introduces interdependencies between $\lambda_{1,t}$ and $\lambda_{2,t}$. Hence, $k(\lambda_t,a_t)$ should be a density

[10] The statistical properties of the two-component SV model can be obtained by amalgamating the properties of a one-component SV model discussed, for example, in Shephard (1996) with those of aggregated AR processes analyzed, for example, in Granger (1980).

kernel for correlated random variables and the nondiagonal elements of the matrix $\mathbf{C}_t$ should not be zero. It follows that the EIS auxiliary regressions are characterized by

$$-\frac{1}{2}[r_t^2\exp(-\iota'\tilde{\lambda}_t^{(i)}(\theta))+\iota'\tilde{\lambda}_t^{(i)}(\theta)]+\ln\chi(\tilde{\lambda}_t^{(i)}(\theta),\hat{a}_{t+1}(\theta))$$
$$=\text{constant}+b_{1,t}\tilde{\lambda}_{1,t}^{(i)}(\theta)+b_{2,t}\tilde{\lambda}_{2,t}^{(i)}(\theta)+c_{1,1,t}[\tilde{\lambda}_{1,t}^{(i)}(\theta)]^2+c_{2,2,t}[\tilde{\lambda}_{2,t}^{(i)}(\theta)]^2$$
$$+2c_{1,2,t}[\tilde{\lambda}_{1,t}^{(i)}(\theta)\tilde{\lambda}_{2,t}^{(i)}(\theta)]+\eta_t^{(i)}. \quad (32)$$

The ML-EIS estimation results for the two-component model are presented in Table 7. As indicated by the MC standard errors, the parameter estimates are very accurate with as little as $N=50$ MC draws for integrals of dimensions greater than 8000! In fact, the statistical standard errors are at least 5 times as large as their MC counterparts. For all three series, the likelihood-ratio statistic of the joint hypothesis that $\delta_2=\nu_2=0$ indicates that the two-component model represents a significant improvement over the basic model with only one volatility process. The estimates of the shock impacts on the long-run and the short-run volatility components represented by ν_1 and ν_2, respectively, are always highly significant with ν_2 significantly greater than ν_1. The estimated values of the persistence parameter of the long-run component δ_1 range between 0.98 and 0.99, while those of the short-run component δ_2 lie between -0.14 and 0.39. However, none of the estimates of δ_2 is significantly different from zero. This suggests that the second volatility component $\lambda_{2,t}$ is serially uncorrelated (as it is implicitly assumed in the SV model with t-errors, though we are discussing here another distributional assumption). Therefore, it is not surprising that the estimates of δ_1 and ν_1 are very close to the volatility parameters under the SV-t model. A comparison of the log-likelihood values indicates that the SV-t model is preferable to the two-component SV model at least for IBM and S&P500.

The time series and the quantile–quantile plot of the normalized IBM residuals z_t* under the two-component model are shown in Fig. 5. The introduction of a second volatility process improves the approximation of the IBM return distribution relative to that of the basic SV model (see Fig. 1). However, the two-component model still has

Table 7
ML-EIS estimation results for the two-component SV model

	IBM	S&P500	$/DM
γ_1	0.003 (0.001) [0.0001]	−0.004 (0.002) [0.0002]	−0.027 (0.006) [0.0004]
δ_1	0.993 (0.003) [0.0002]	0.991 (0.003) [0.0003]	0.975 (0.006) [0.0004]
ν_1	0.064 (0.012) [0.0009]	0.078 (0.011) [0.0013]	0.142 (0.016) [0.0015]
δ_2	0.363 (0.158) [0.0279]	−0.072 (0.105) [0.0155]	−0.139 (0.172) [0.0353]
ν_2	0.475 (0.048) [0.0080]	0.559 (0.039) [0.0054]	0.486 (0.050) [0.0086]
Log-likelihood	−7928.4 [0.564]	−5761.6 [0.923]	−4666.3 [0.501]
LR	62.3	85.7	41.7

Asymptotic (statistical) standard errors obtained from a numerical approximation to the Hessian are in parentheses and MC (numerical) standard errors are in brackets. The ML-EIS estimates are based on a MC sample size $N=50$ and three EIS iterations. The likelihood ratio test follows a $\chi^2_{(2)}$-distribution with the 1% critical value being 9.21.

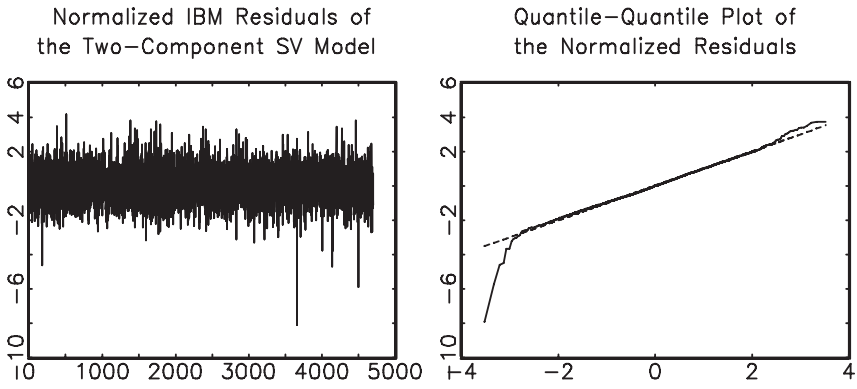

Fig. 5. Normalized IBM residuals z_t^* obtained from the two-component SV model (left) and quantile–quantile plot of z_t^* (right); the dashed line plots the quantiles of the standard normal distribution against the quantiles of the standard normal and the solid line plots the sorted values of z_t^* against the quantiles of the standard normal.

difficulties accounting for a few large returns. Table 8 presents the results of diagnostic checks of the two-component model. According to the Kolmogorov–Smirnov test, the model explains the distributional properties of the S&P500 and $/DM returns, but not those of the IBM data. The Ljung–Box test for the squared residuals shows that the model successfully accounts for the dynamics in the second-order moments for IBM and $/DM only.

Taken all together, the likelihood ratio test suggests that all the extensions of the standard SV model analyzed here represent significant improvements relative to the standard model. Furthermore, the diagnostics indicate that the SV model with a semi-nonparametric error distribution fits reasonably well the distributional properties and the volatility dynamics of all three return series, whereas the other SV specifications fail to do so for at least one of the return series. Finally, we note that the analysis can be extended in a number of further directions. For example, the EIS procedure can be modified to handle SV models with covariates in the volatility equation and/or non-zero correlation between the shocks in the mean and volatility equation in order to capture the so-called 'leverage-effect'. The SV model under these assumptions has already been analyzed using the AGIS procedure by Danielsson (1994a). Another possible extension, recently proposed by Chib et al. (2002), is to introduce a jump component into the mean equation to allow for large but transitory shocks.

Table 8
Diagnostics for the two-component SV model

	Skewness	Kurtosis	KS(z^*)	$Q_{30}(z^*)$	$Q_{30}(z^{*2})$	$Q_{30}(z)$	$Q_{30}(z^2)$
IBM	−0.04	4.22	1.727 (0.005)	26.6 (0.642)	29.2 (0.507)	29.2 (0.505)	20.83 (0.893)
S&P500	−0.19	4.49	0.901 (0.392)	78.2 (<0.001)	86.4 (<0.001)	74.4 (<0.001)	54.0 (0.005)
$/DM	−0.05	3.14	0.771 (0.593)	51.6 (0.008)	15.9 (0.983)	53.4 (0.005)	14.5 (0.992)

Probability values are given in parentheses.

4. A multivariate SV model

The SV models discussed above are univariate. However, there are theoretical as well as empirical reasons to study multivariate volatility models. Firstly, portfolio allocation and asset pricing can only be meaningfully discussed within a multivariate framework. Secondly, correlation across asset returns requires simultaneous multivariate estimation for full efficiency. Thirdly, multivariate structural volatility models can provide useful information about the factors driving the volatility process. Hence, we consider next a one-factor multivariate volatility model with a stochastic volatility process for single factor. Such a multivariate factor SV specification, originally proposed by Shephard (1996) and Jacquier et al. (1999), has several attractive features. Foremost, the dimension of the parameter space remains operational since it only increases linearly with the number of assets being modeled. Moreover, the model accounts not only for the volatility dynamics of individual asset returns, but also for the typically observed serial dependence in the correlation between asset returns, as measured by the cross products $r_{i,t}r_{i',t}$ calculated for asset i and i'. Alternative dynamic factor models are the latent factor ARCH model of Diebold and Nerlove (1989) and the factor ARCH representation used by Engle et al. (1990).

Let $r_t=(r_{1,t},\ldots,r_{n,t})'$ denote a vector of n asset returns. A multivariate one-factor SV model for r_t is

$$r_t = Dx_t + e_t \tag{33}$$

$$x_t = \exp(\lambda_t/2)\epsilon_t \tag{34}$$

$$\lambda_t = \gamma + \delta\lambda_{t-1} + v\eta_t, \tag{35}$$

where $D=(d_1,\ldots,d_n)'$ denotes a vector of factor loadings, x_t a latent factor following a univariate SV process, ϵ_t and η_t iid Gaussian random variables with zero means and unit variances and $e_t=(e_1,t,\ldots,e_{n,t})'$, a vector of serially independent idiosyncratic errors with $e_t \sim N(0,\Sigma_e)$ and $\Sigma_e=\mathrm{diag}(\sigma^2_{e,j})$. The error terms ϵ_t, η_t and $e_{j,t}$ are assumed to be mutually independent. In order to achieve identification, we impose the restriction $d_1=1$.

Under this model, the conditional joint distribution of the returns r_t given λ_t is a multivariate normal given by $r_t|\lambda_t \sim N[0,DD'\exp(\lambda_t)+\Sigma_e]$, where the time-series behavior of the conditional variance–covariance is driven by the volatility process of the common factor x_t. The unconditional variance–covariance matrix of r_t is given by $\mathrm{Var}(r_t)=DD'E[\exp(\lambda_t)]+\Sigma_e$. Hence, the overall variance–covariance is decomposed into a component which is due to the variation in the common factor and a component reflecting the variation in the idiosyncratic errors. Following an interpretation offered by Diebold and Nerlove (1989), the common factor reflects the flow of new information relevant to the pricing of all assets, upon which asset specific shocks represented by the idiosyncratic errors are superimposed.

Again, the implementation of EIS for the factor SV model only requires modifying the conditional density of r_t given λ_t in the baseline algorithm described in Section 2.2. For the factor SV model this density has the form

$$g(r_t \mid \lambda_t,\theta) \propto \det[DD'\exp(\lambda_t)+\Sigma_e]^{-1/2}\exp\left\{-\frac{1}{2}r_t'[DD'\exp(\lambda_t)+\Sigma_e]^{-1}r_t\right\}. \tag{36}$$

Table 9
ML-EIS estimation results for the factor SV model

	BP	DM	YEN	FF
d_j	1.000	1.375 (0.018) [0.00002]	0.855 (0.019) [0.00002]	1.303 (0.018) [0.00001]
$\sigma_{\varepsilon,j}$	0.442 (0.005) [<0.00001]	0.100 (0.008) [0.00004]	0.546 (0.006) [<0.00001]	0.197 (0.004) [0.00002]
	γ	δ	ν	Log-likelihood
	−0.064 (0.013) [0.00031]	0.962 (0.008) [0.00017]	0.192 (0.019) [0.00049]	11,035.7 [0.146]

Asymptotic (statistical) standard errors obtained from a numerical approximation to the Hessian are in parentheses and MC (numerical) standard errors are in brackets. The ML-EIS estimates are based on a MC sample size $N=50$ and three EIS iterations.

The data to which we fit this multivariate model consists of 4774 observations on four exchange rates: British Pound/US-Dollar (BP), Deutsche Mark/US-Dollar (DM), Yen/US-Dollar (YEN) and French Franc/US-Dollar (FF). These are daily noon spotprices from the Chicago Federal Reserve Bank from January 1, 1980 to December 31, 1998. The prices $s_{j,t}$, $j=$BP, DM, YEN, FF are transformed into continuously compounded rates centered around their sample mean: $r_{j,t} = 100[\ln(s_{j,t}/s_{j,t-1}) - (1/T)\sum_{t=1}^{T}\ln(s_{j,t}/s_{j,t-1})]$.

The ML-EIS estimation results are shown in Table 9. All parameter estimates are reasonable and significant at the 1% level. The estimates of the factor loadings d_j indicate that the European currencies load more heavily on the common factor than the YEN. Using the parameter estimates, we computed the corresponding estimate of the unconditional variance–covariance matrix $\widehat{\mathrm{Var}}(r_t)$, to be compared with the sample variance–covariance matrix of the returns S_{rr}. They are given by

$$\widehat{\mathrm{Var}}(r_t) = \begin{pmatrix} 0.44 & -0.33 & -0.21 & -0.32 \\ & 0.47 & 0.28 & 0.43 \\ & & 0.48 & 0.27 \\ & & & 0.45 \end{pmatrix},$$

$$S_{\mathrm{rr}} = \begin{pmatrix} 0.44 & -0.34 & -0.21 & -0.32 \\ & 0.48 & 0.29 & 0.44 \\ & & 0.48 & 0.27 \\ & & & 0.46 \end{pmatrix},$$

Table 10
Ljung-Box statistic with 30 lags for the cross products of the residuals ($z_{j,t}z_{k,t}$)

	BP	DM	YEN	FF
BP	369.9 (1271.7)	321.2 (1040.5)	322.4 (429.7)	82.5 (1011.5)
DM		583.4 (754.0)	102.4 (300.4)	59.9 (669.4)
YEN			1831.3 (959.5)	60.8 (262.5)
FF				112.5 (344.2)

The Ljung–Box statistic for the cross products of the returns ($r_{j,t}r_{k,t}$) including 30 lags are given in parentheses. The Ljung–Box statistic follows a $\chi^2_{(30)}$-distribution with the 1% critical value being 50.9.

respectively. Obviously, the model explains very well the unconditional variances and correlations of the data. Furthermore, we find that the common factor explains 55%, 97%, 37% and 90% of the overall variation in the returns of BP, DM, YEN and FF, respectively. The estimated values of the volatility parameters δ and ν are of the same order of magnitude as those typically obtained under univariate SV models. Finally, the sum of the individual log-likelihood values for four independent univariate SV models (as defined in Eqs. (1) and (2)) equals $-18{,}199$, which is substantially smaller than the log-likelihood value for the multivariate model given by $-11{,}036$. This huge difference is not surprising since the multivariate model accounts for the correlation between the returns, in sharp contrast with the univariate specifications.

In order to check the adequacy of the assumed dynamic structure in the second-order moments, we computed the vector of residuals $z_t = \mathrm{Var}(r_t|R_{t-1})^{-1/2} r_t$, where $\mathrm{Var}(r_t|R_{t-1})^{-1/2}$ denotes the inverse of the Cholesky factor of $\mathrm{Var}(r_t|R_{t-1})$. If the model is correctly specified, then the elements in the matrix $z_t z_t'$, which are given by the cross products $z_{j,t}z_{k,t}$, j,k = BP, DM, YEN, FF, are serially uncorrelated. Table 10 presents the values of the Ljung–Box statistic for these cross products including 30 lags together with the corresponding values for the returns $r_{j,t}r_{k,t}$. As seen from this table, the model removes most of the observed serial correlation in $r_{j,t}r_{k,t}$ in nearly all cases. Nevertheless, the hypothesis that $z_{j,t}z_{k,t}$ is serially uncorrelated is clearly rejected in all cases, indicating that the joint dynamics in the second-order moments of the returns cannot be fully explained by the simple one-factor SV model. Hence, it might be necessary to allow for a second latent factor and/or idiosyncratic errors following stochastic volatility processes, extensions we propose to pursue in the future and which have recently been analyzed using MCMC techniques by Pitt and Shephard (1999b) and Aguilar and West (2000).

5. Conclusion

In this paper we estimate and test various extensions of the standard SV model using a ML approach based upon the EIS procedure of Richard and Zhang (1996, 1997, 1998). EIS is a generic and numerically very efficient MC technique for the evaluation of high-dimensional (interdependent) integrals, which can be applied for accurate computation of the likelihood function for SV models. Hence, it allows one to rely upon the full range of likelihood based inference techniques including estimation

and testing. Moreover, EIS can be applied to produce filtered residuals which provide the basis for several diagnostic tests. We explored and compared the following models: A SV model with a Gaussian, a Student's t and a semi-nonparametric error distribution in the return equation, respectively, a two-component SV model with two independent dynamic volatility processes and a multivariate one-factor SV model. A major advantage offered by EIS for this type of exploratory analysis is that all these models can be analyzed under the very same baseline EIS algorithm with only very minor changes for each specific model.

Our empirical results based upon daily data on exchange rates, stock returns and stock index returns can be summarized as follows: Firstly, the ML-EIS parameter estimates are numerically very accurate even for likelihood functions depending on 8000-dimensional integrals. Secondly, the SV model with a semi-nonparametric error distribution fits the distributional and dynamic properties of the return data reasonably well, whereas the other SV specifications generally fail to do so. Thirdly, while a multivariate one-factor SV model performs quite well in explaining the joint dynamics in the volatility of a set of four exchange rates series, we might usefully consider adding additional dynamic factors in the future.

Acknowledgements

An earlier version of this paper was presented at the EC Squared Econometric Conference 2000 in Dublin. We are grateful to seminar participants at the University of Mannheim and at the Université catholique de Louvain (CORE) and two referees for helpful comments. The first author would like to acknowledge financial support from the Deutsche Forschungsgemeinschaft.

References

Aguilar, O., West, M., 2000. Bayesian dynamic factor models and portfolio allocation. Journal of Business and Economic Statistics 18, 338–357.

Andersen, T., Chung, H., Sørensen, B.E., 1999. Efficient method of moments estimation of a stochastic volatility model: a Monte Carlo study. Journal of Econometrics 91, 61–87.

Bollerslev, T., 1987. A conditionally heteroskedastic time series model for security prices and rates of return data. Review of Economics and Statistics 59, 542–547.

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

Danielsson, J., 1994a. Stochastic volatility in asset prices estimation with simulated maximum likelihood. Journal of Econometrics 64, 375–400.

Danielsson, J., 1994b. A comment on: Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 393–395.

Danielsson, J., 1998. Multivariate stochastic volatility models: estimation and a comparison with VGARCH models. Journal of Empirical Finance 5, 125–173.

Danielsson, J., Richard, J.F., 1993. Accelerated Gaussian importance sampler with application to dynamic latent variable models. Journal of Applied Econometrics 8, S153–S173.

DeGroot, M., 1970. Optimal Statistical Decisions. McGraw-Hill, New York.

Devroye, L., 1986. Non-Uniform Random Variate Generation. Springer Verlag, New York.

Diebold, F.X., Nerlove, M., 1989. The dynamics of exchange rate volatility: a multivariate latent-factor ARCH model. Journal of Applied Econometrics 4, 1–22.

Diebold, F.X., Gunther, T.A., Tay, A.S., 1998. Evaluating density forecasts with applications to financial risk management. International Economic Review 39, 863–883.

Diebold, F.X., Tay, A.S., Wallis, K.F., 1999. Evaluating density forecasts of inflation: the survey of professional forecasters. In: Engle, R.F., White, H. (Eds.), Cointegration Causality and Forecasting. Oxford Univ. Press, Oxford, pp. 76–90.

Durbin, J., Koopman, S.J., 1997. Monte Carlo maximum likelihood estimation for non-Gaussian state space models. Biometrika 84, 669–684.

Durbin, J., Koopman, S.J., 2000. Time series analysis of non-Gaussian observations based on state space models form both classical and Bayesian perspectives (with discussion). Journal of the Royal Statistical Society, Series B 62, 3–56.

Durham, G.B., Gallant, A.R., 2001. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. Unpublished manuscript, University of North Carolina, Chapel Hill.

Engle, R.F., Lee, G.G.J., 1999. A long-run and short-run component model of stock return volatility. In: Engle, H., White, H. (Eds.), Cointegration, Causality and Forecasting. Oxford Univ. Press, Oxford, pp. 475–497.

Engle, R.F., Ng, V., Rothschild, M., 1990. Asset pricing with a factor-ARCH covariance structure. Journal of Econometrics 45, 213–237.

Fenton, V.M., Gallant, A.R., 1996. Qualitative and asymptotic performance of SNP density estimators. Journal of Econometrics 74, 77–118.

Gallant, A.R., Nychka, D.W., 1987. Semi nonparametric maximum likelihood estimation. Econometrica 55, 363–390.

Gallant, A.R., Tauchen, G.E., 1998. Reprojecting partially observed systems with application to interest rate diffusions. Journal of the American Statistical Association 93, 10–24.

Gallant, A.R., Hsieh, D.A., Tauchen, G.E., 1991. On fitting a recalcitrant series: the pound/dollar exchange rate, 1974–1983. In: Barnett, W.A., Powell, J., Tauchen, G.E. (Eds.), Nonparametric and Semiparametric Methods in Economics and Statistics. Cambridge Univ. Press, Cambridge, pp. 199–240.

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

Geweke, J., 1989. Bayesian inference in econometric models using Monte Carlo integration. Econometrica 57, 1317–1339.

Ghysels, E., Harvey, A.C., Renault, E., 1996. Stochastic volatility. In: Maddala, G., Rao, C.R. (Eds.), Handbook of Statistics, vol. 14. Elsevier Sciences, Amsterdam.

Granger, C.W.J., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227–238.

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

Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatilities. Journal of Finance 42, 281–300.

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

Jacquier, E., Polson, N.G., Rossi, P.E., 1999. Stochastic volatility: univariate and multivariate extensions. Working paper 99s-26, CIRANO, Montreal.

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

Kitagawa, G., 1996. Monte Carlo filter and smoother for non-Gaussian non-linear state space models. Journal of Computational and Graphical Statistics 5, 1–25.

Liesenfeld, R., 1998. Dynamic bivariate mixture models: modeling the behavior of prices and trading volume. Journal of Business and Economic Statistics 16, 101–109.

Liesenfeld, R., 2001. A generalized bivariate mixture model for stock price volatility and trading volume. Journal of Econometrics 104, 141–178.

Liesenfeld, R., Jung, R.C., 2000. Stochastic volatility models: conditional normality versus heavy-tailed distributions. Journal of Applied Econometrics 15, 137–160.

Liesenfeld, R., Richard, J.F., 2001. Monte Carlo methods and Bayesian computation: importance sampling. In: Smelser, N.J., Baltes, P.B. (Eds.), International Encyclopedia of the Social and Behavioral Sciences. Elsevier, Oxford, pp. 9999–10004.

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

Patel, J.K., Read, C.B., 1982. Handbook of the Normal Distribution. Marcel Dekker, New York.

Pitt, M.K., Shephard, N., 1999a. Filtering via simulation: auxiliary particle filters. Journal of the American Statistical Association 94, 590–599.

Pitt, M.K., Shephard, N., 1999b. Time varying covariance: a factor stochastic volatility approach (with discussion). In: Bernardo, J., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics, vol. 6. Oxford Univ. Press, Oxford.

Richard, J.F., Zhang, W., 1996. Econometric modeling of UK house prices using accelerated importance sampling. Oxford Bulletin of Economics and Statistics 58, 601–613.

Richard, J.F., Zhang, W., 1997. Accelerated Monte Carlo integration: an application to dynamic latent variable models. In: Mariano, R., Weeks, M., Schuermann, T. (Eds.), Simulation-Based Inference in Econometrics: Methods and Application. Cambridge Univ. Press, Cambridge, pp. 47–70.

Richard, J.F., Zhang, W., 1998. Efficient high-dimensional Monte Carlo importance sampling. Unpublished manuscript, Department of Economics, University of Pittsburgh.

Rosenblatt, M., 1952. Remarks on a multivariate transformation. Annals of Mathematical Statistics 23, 470–472.

Ruiz, E., 1994. Quasi-maximum likelihood estimation of stochastic volatility models. Journal of Econometrics 63, 289–306.

Sandmann, G., Koopman, S.J., 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. Journal of Econometrics 87, 271–301.

Schwarz, G., 1978. Estimation the dimension of a model. Annals of Statistics 6, 461–464.

Shephard, N., 1996. Statistical aspects of ARCH and stochastic volatility. In: Cox, D.R., Hinkley, D.V., Barndorff-Nielsen, O.E. (Eds.), Time Series Models in Econometrics, Finance and Other Fields. Chapman & Hall, London, pp. 1–67.

Taylor, S.J., 1982. Financial returns modelled by the product of two stochastic processes—a study of daily sugar prices. In: Anderson, O.D. (Ed.), Time Series Analysis: Theory and Practice 1. North-Holland, Amsterdam, pp. 203–226.

Taylor, S.J., 1986. Modelling Financial Time Series. Wiley, Chichester.

Taylor, S.J., 1994. Modeling stochastic volatility: a review and comparative study. Mathematical Finance 4, 183–204.