Journal of Financial Econometrics, 2007, Vol. 5, No. 1, 1–30

Estimating Latent Variables and Jump Diffusion Models Using High-Frequency Data

GEORGE J. JIANG
University of Arizona

ROEL C. A. OOMEN
University of Warwick

ABSTRACT

This article proposes a new approach to exploit the information in high-frequency data for the statistical inference of continuous-time affine jump diffusion (AJD) models with latent variables. For this purpose, we construct unbiased estimators of the latent variables and their power functions on the basis of the observed state variables over extended horizons. With the estimates of the latent variables, we propose a generalized method of moments (GMM) procedure for the estimation of AJD models with the distinguishing feature that moments of both observed and latent state variables can be used without resorting to path simulation or discretization of the continuous-time process. Using high frequency return observations of the S&P 500 index, we implement our estimation approach to various continuous-time asset return models with stochastic volatility and random jumps.

KEYWORDS: affine jump diffusion, generalized method of moments, high-frequency data, latent state variables, unbiased minimum-variance estimator

Jiang is from the Eller College of Management, University of Arizona, Tucson AZ 85721-0108, USA. Oomen is from the Department of Finance, Warwick Business School, The University of Warwick, Coventry CV4 7AL, United Kingdom. E-mail: roel.oomen@wbs.ac.uk. Oomen is also a research affiliate of the Department of Quantitative Economics at the University of Amsterdam, The Netherlands, and acknowledges financial support from the Netherlands Organization for Scientific Research (NWO). The authors would like to thank two referees, Søren Johansen and Nour Meddahi, for very helpful and constructive comments. They would also like to thank Elena Andreou, Hayne Leland, Nick Polson, Bas Werker, Hao Zhou, and the seminar participants at the 2003 North American Econometric Society Meetings in Washington DC, the Department of Economics at UC Riverside, the CentER at Tilburg University, and the participants of the New Frontiers in Financial Volatility Modelling Conference in Florence for helpful comments. Address correspondence to Roel C. A. Oomen, Department of Finance, Warwick Business School, The University of Warwick, Coventry CV4 7AL, United Kingdom, or e-mail: roel.oomen@wbs.ac.uk.

doi:10.1093/jjfinec/nbl007
Advance Access publication November 18, 2006

1 INTRODUCTION

Continuous-time jump diffusion models with latent state variables play an important role in finance as they are commonly used for the modeling of asset returns and provide a convenient framework for the pricing of derivatives and risk management. It is well known, however, that the statistical inference of such models poses a great challenge for the following two reasons. First, the transition density of the continuous-time jump diffusion process is, with the exception of a few special cases, generally unknown or unavailable in closed form so that the conventional likelihood-based inference is often inapplicable. Second, any inference procedure has to deal with the unobserved latent variables, which typically involves either using proxies for the latent variables or integrating out the latent variables in the likelihood function. Since the high-dimensional integral in the likelihood function mostly cannot be reduced to one (or substantially lower) dimensional integrals, the numerical integration often has to resort to path simulation of the discretized continuous-time process and is therefore computationally intensive. In addition to the challenge of model inference, the presence of latent variables can also complicate model applications. For instance, the implementation of risk management or portfolio allocation strategies often requires the conditional mean, conditional variance, and/or jump intensity of the asset price process as input parameters. Hence, accurate estimates of both model parameters *and* the latent state variables are crucial.

This article proposes a framework for the estimation of continuous-time jump diffusion models using unbiased estimates of the latent state variables. Specifically, we develop model-based estimators of latent state variables for the general affine jump diffusion (AJD) models of Duffie, Pan, and Singleton (2000). The key features of the latent variable estimator are that it is derived from the exact model structure without discretization of the continuous-time process, is unbiased with minimum-variance, and can exploit the information in high-frequency data. With the proposed latent variable estimators, we then develop a new generalized method of moments (GMM) estimation procedure for continuous-time jump diffusion models. To illustrate our approach, we apply the GMM method to various stochastic volatility models using S&P 500 high-frequency data.

The latent variable estimator we propose in this article is derived from the dynamic relation between the latent state variable and the observed state variable. For continuous-time AJD models, such relations can be readily uncovered from the conditional characteristic function. Intuitively, the resulting estimator can be viewed as a weighted sum of functionals of the observed state variables, with weights explicitly determined by the model structure. For example, for a stochastic volatility model, the conditional variance estimator is the weighted sum of re-scaled squared (high-frequency) returns. Owing to the analytical tractability of the AJD model, the latent state variable estimator is generally available in closed form and its finite sample properties can be established. In particular, we show that the estimator is conditionally unbiased and has minimum variance when optimal weights are used. A further advantage of our approach is that unbiased minimum-variance estimators of *power functions* of the latent state variable can also be constructed in a straightforward fashion. As demonstrated in this article, these estimators form the basis of the new estimation method for the continuous-time jump diffusion models.

The GMM procedure we propose in this article involves matching the sample moments of both observed and unobserved state variables to their corresponding population moments. With the unbiased estimators of the latent variables and its power functions, the high-order moments of both observed and unobserved state variables can be included in the estimation. Further, owing to the tractability of the AJD class, population moments are available in closed form so that implementation of the GMM procedure is straightforward. The approach is similar to that of Bollerslev and Zhou (2002) who match the model moments of integrated volatility to the corresponding sample moments calculated from the model-free realized variance or Pan (2002) who matches the model moments of spot volatility to the corresponding sample moments calculated from option-implied volatility. Our approach shares with these studies the advantage that with proxies of the latent variables, there is no path simulation involved in the estimation. The further advantage here is that with the unbiased estimators of the power functions of the latent variable, our approach can involve higher-order moment conditions without the potential error-in-variable problem. Moreover, the use of a model-based variance estimator, instead of a nonparametric proxy of integrated variance, allows for more efficient use of information through smoothing.

The remainder of this article is organized as follows. In the next section, we first develop the unbiased minimum-variance estimators for the latent state variables and their power functions under the AJD class of models. Using the stochastic volatility model, we illustrate the construction of the estimators and the computation of optimal weighting vector. Next, on the basis of the latent variable estimators, we outline the GMM estimation procedure for the AJD models. An empirical application is undertaken in Section 3 where the latent variables and parameters of various stochastic volatility with random jump models are estimated using the high-frequency data of the S&P 500 index. With the estimated asset return volatilities under a two-factor stochastic volatility model, we further examine their relation with other market variables. Our findings suggest that the two distinctly behaved volatility components reflect different types of information in the market. Section 4 concludes.

2 ESTIMATION OF AFFINE JUMP DIFFUSION MODELS WITH LATENT VARIABLES

The model considered in our study is the widely used AJD model with latent variables [see Duffie, Pan, and Singleton (2000)]. Let $X_t \in \mathcal{R}^n$, $t \geq 0$, denote an n-dimensional vector of state variables. Using the same notation as in Duffie, Pan, and Singleton (2000), we fix a probability space $(\Omega, \mathcal{F}, P)$ and an information filtration $(\mathcal{F}_t) = \{\mathcal{F}_t : t \geq 0\}$ and suppose that X is a Markov process relative to $\mathcal{F}_t$

in some state space $\mathcal{D} \in \mathbb{R}^n$ following the stochastic differential equation (SDE):

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dZ_t, \tag{1}$$

where W_t is an $(\mathcal{F}_t)$ −standard Brownian motion in $\mathbb{R}^n$, and $\mu(\cdot) : \mathcal{D} \to \mathbb{R}^n$ and $\sigma(\cdot) : \mathcal{D} \to \mathbb{R}^{n\times n}$ are respectively the drift function and diffusion function. Z is a pure jump process whose jumps have a fixed probability distribution $\mathcal{J}$ on $\mathbb{R}^n$ and arrive with intensity $\{\lambda(X_t) : t \geq 0\}$, for some $\lambda(\cdot) : \mathcal{D} \to [0, \infty)$. Following Duffie, Pan, and Singleton (2000) we assume that, conditional on the path of X_t, the jump times of Z_t are the jump times of a Poisson process with time varying intensity $\{\lambda(X_s) : 0 \leq s < t\}$, and the jump size Z_t at a jump time t is independent of $\{X_s : 0 \leq s < t\}$. The initial value of the stochastic process X_0 is assumed to be drawn from the marginal distribution of the process. For X_t to be a well-defined Markov process, regularity conditions on the filtration $(\mathcal{F}_t) = \{\mathcal{F}_t : t \geq 0\}$ and restrictions on the state space as well as on the coefficient functions, namely $(\mathcal{D}, \mu(\cdot), \sigma(\cdot), \lambda(\cdot), \mathcal{J})$, are required. For technical details, [see, for example, Ethier and Kurtz (1986), Duffie and Kan (1996), Duffie, Pan, and Singleton (2000), and Dai and Singleton (2000)].

For convenience and tractability, an affine structure (on $\mathcal{D}$) is often imposed on the coefficient functions $\mu(\cdot)$, $\sigma(\cdot)\sigma(\cdot)'$, and $\lambda(\cdot)$:

$$\begin{aligned} \mu(X_t) &= K_0 + K_1 X_t, \\ [\sigma(X_t)\sigma(X_t)']_{ij} &= [H_0]_{ij} + [H_1]_{ij} X_t, \\ \lambda(X_t) &= l_0 + l_1 X_t, \end{aligned} \tag{2}$$

where $K = (K_0, K_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}, H = (H_0, H_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}, l = (l_0, l_1) \in \mathbb{R}\times \mathbb{R}^n$. Let $g(c) = \int_{\mathbb{R}^n} \exp\{c \cdot z\} d\mathcal{J}(z)$ be the jump transform whenever the integral is well defined, where $c \in \mathbb{C}^n$ the set of n-tuples of complex numbers, $g(\cdot)$ determines the jump size distribution. It is clear that the set of parameters (K, H, l, g) completely specifies the AJD process and determines its statistical properties, given the initial condition X_0.

As mentioned in the introduction, statistical inference of continuous-time models has posed a great challenge to statisticians and econometricians as it requires the knowledge of dynamic properties or the transition density of the process. Under regularity conditions, the transition density function of the jump diffusion process as defined in Equation (1) satisfies both the Kolmogorov forward and backward (or Fokker−Planck) equations of the Markov process. However, the transition density functions of the multivariate diffusion and jump diffusion process in general do not have a closed analytical form. The presence of latent or unobserved variables, as considered in this article, presents an even greater challenge to statistical inference. Since the latent variables are not directly observed, any inference procedure must either integrate out the latent state variables from the model or rely on some proxies of the latent variables.

Despite these difficulties, remarkable progress has been made in recent years regarding the estimation of nonlinear latent variable models in general and the SV models in particular. A number of estimation procedures have been developed, including the QMLE [e.g. Ruiz (1994), Harvey and Shephard (1996)], MLE [e.g. Lo (1988), Bates (2006), Aït-Sahalia and Kimmel (2006)], SMM [Duffie and Singleton (1993)], indirect inference [e.g. Gouriéroux, Monfort, and Renault (1993), EMM [Gallant and Tauchen (1996)], Bayesian MCMC [e.g. Jacquier, Polson, and Rossi (1994), Eraker (2001), Chib, Nardari, and Shephard (2002)], ECF [e.g. Singleton (2001), Jiang and Knight (2002), Carrasco, Chernov, Florens, and Ghysels (2002)], and GMM [e.g. Melino and Turnbull (1990), Andersen and Sørensen (1996), Bollerslev and Zhou (2002), Pan (2002)].

This article contributes to this burgeoning literature by developing an alternative method for the estimation of continuous-time latent AJD models. In what follows, we first outline the unbiased latent variable estimator and then provide details of the GMM procedure.

2.1 Latent Variable Estimation for the Affine Jump Diffusion Models

For simplicity of notation, we partition the vector X_t into two sub-vectors, i.e. $X_t = \left(S_t', V_t'\right)'$, where $S_t \in \mathcal{R}^m$, $n > m > 0$, denotes the vector of *observed* state variables and $V_t \in \mathcal{R}^{n-m}$ denotes the vector of *unobserved or latent* variables. In financial models, S_t often represents observed asset prices, interest rates, exchange rates, trading volume etc., while V_t can be unobserved instantaneous volatility, instantaneous mean, jump intensity, information flow, etc. We also assume that observations of the state variables S_t are equally spaced with sampling interval δ over the time period $[0, T]$. That is, the sequence $\{S_{t+k\delta}\}_{k=0}^{N}$ for $t = 0, 1, \ldots, T-1$ is observed. In financial applications, the unit time interval usually corresponds to one trading day, thus with $\delta < 1$ we explicitly work with high-frequency or intraday data. It is noted that the approach outlined below also works with unequal but nonrandom sampling intervals, with the only difference being more cumbersome notations.

The key result derived in Duffie, Pan, and Singleton (2000), which we will exploit here, is that the conditional characteristic function (CCF) of the AJD process as defined in Equations (1) and (2) is of the following functional form:

$$\psi_t(u, X_{t+\delta}) = E\left[\exp\{iuX_{t+\delta}\} \mid \mathcal{F}_t\right] = \exp\left\{C(\delta, u) + D(\delta, u)' X_t\right\}, \tag{3}$$

where $D(\cdot)$ and $C(\cdot)$ are the solutions of complex-valued Ricatti equations:

$$\begin{aligned}\frac{\partial D(\delta, u)}{\partial \delta} &= K_1' D(\delta, u) + \frac{1}{2} D(\delta, u)' H_1 D(\delta, u) + l_1 \left(g\left(D(\delta, u)\right) - 1\right), \\ \frac{\partial C(\delta, u)}{\partial \delta} &= K_0' D(\delta, u) + \frac{1}{2} D(\delta, u)' H_0 D(\delta, u) + l_0 \left(g\left(D(\delta, u)\right) - 1\right),\end{aligned} \tag{4}$$

with boundary conditions $D(0, u) = iu$ and $C(0, u) = 0$. With certain specification of the coefficient function (K, H, l, g), explicit solutions of $D(\cdot)$ and $C(\cdot)$ can be found. In other cases, as noted in Duffie, Pan, and Singleton (2000), the solution would have to be found numerically by solving the relevant ODEs.

With the distinction made between observed and unobserved state variables, we can specialize Equation (3) as follows:

$$\begin{aligned}\psi_t\left(u_1,u_2,S_{t+\delta},V_{t+\delta}\right) &= E\left[\exp\left\{iu_1'S_{t+\delta}+iu_2'V_{t+\delta}\right\}|\mathcal{F}_t\right],\\ &= \exp\left\{C\left(\delta,u_1,u_2\right)+D1\left(\delta,u_1,u_2\right)'S_t+D2\left(\delta,u_1,u_2\right)'V_t\right\}, \quad (5)\end{aligned}$$

where $C\left(\delta,u_1,u_2\right)=C\left(\delta,u\right),\left(D1\left(\delta,u_1,u_2\right)',D2\left(\delta,u_1,u_2\right)'\right)=D\left(\delta,u\right)'$ and $u=\left(u_1',u_2'\right)'$. Setting $u_2=0$ in Equation (5) yields:

$$\begin{aligned}\psi_t\left(u_1,S_{t+\delta}\right) &= E\left[\exp\left\{iu_1'S_{t+\delta}\right\}|\mathcal{F}_t\right],\\ &= \exp\left\{C\left(\delta;u_1,0\right)+D1\left(\delta;u_1,0\right)'S_t+D2\left(\delta;u_1,0\right)'V_t\right\}. \quad (6)\end{aligned}$$

From this it follows (see Appendix B for a proof) that the l^{th} cumulant of $\Delta S_{t+k\delta}\equiv S_{t+k\delta}-S_{t+(k-1)\delta}$, for $k\geq 1$, conditional on $\mathcal{F}_t$ is given by:

$$K^{(l)}[\Delta S_{t+k\delta}|\mathcal{F}_t]=c^{(l)}\left(k\right)+d_1^{(l)}\left(k\right)'S_t+d_2^{(l)}\left(k\right)'V_t,\quad \text{for } l=1,2,\ldots \quad (7)$$

where

$$\begin{aligned}c^{(l)}\left(k\right) &= \frac{\partial^l}{i^l\partial(u_1')^l}\left\{C\left(\delta,u_1,0\right)+C\left((k-1)\delta,-iD1\left(\delta,u_1,0\right)-u_1,-iD2\left(\delta,u_1,0\right)\right)\right\}|_{u_1=0},\\ d_1^{(l)}\left(k\right) &= \frac{\partial^l}{i^l\partial(u_1')^l}\left\{D1\left((k-1)\delta,-iD1\left(\delta,u_1,0\right)-u_1,-iD2\left(\delta,u_1,0\right)\right)\right\}|_{u_1=0},\\ d_2^{(l)}\left(k\right) &= \frac{\partial^l}{i^l\partial(u_1')^l}\left\{D2\left((k-1)\delta,-iD1\left(\delta,u_1,0\right)-u_1,-iD2\left(\delta,u_1,0\right)\right)\right\}|_{u_1=0}.\end{aligned}$$

In particular, for $l=\{1,2\}$ we have

$$d_2^{(1)}\left(k\right)'V_t=E\left[\Delta S_{t+k\delta}|\mathcal{F}_t\right]-c^{(1)}\left(k\right)-d_1^{(1)}\left(k\right)'S_t, \quad (8)$$

$$d_2^{(2)}\left(k\right)'V_t=Var\left[\Delta S_{t+k\delta}|\mathcal{F}_t\right]-c^{(2)}\left(k\right)-d_1^{(2)}\left(k\right)'S_t. \quad (9)$$

The above expressions make explicit the link between the latent state variables and the conditional cumulants of the observed state variables. Note that when both $d_1^{(2)}=d_2^{(2)}=0$, higher-order cumulants are needed for the identification of latent variables. These results are similar to the conditional moments derived in Meddahi (2002) for the eigenfunction SV models proposed in Meddahi (2001). Note that because in financial models the observed state variable S_t is often a first-difference stationary process (e.g. log asset price) and the latent variable V_t often a stationary process (e.g. stochastic volatility), we express the above results in terms of ΔS_t and V_t.

The relation between the observed and unobserved variables through the conditional cumulants point to the possibility of constructing an unbiased estimator of the latent variables from observations of the observed state variables. In particular, we can construct an unbiased minimum-variance estimator of the latent variable V_t on the basis of the observations of the observed state variable S_t over the forward period $[t,t+\tau]$. For notational convenience, we present only the case where the dimensions of S_t and V_t are both equal to 1. Following Equations (8) and (9), and with discrete observations of S_t over the interval $[t, t+\tau]$, i.e. $\{S_{t+k\delta}\}_{k=0}^{N\tau}$, we define the following $N\tau \times 1$ vector

$$\vartheta_t = (\vartheta_t(1), \vartheta_t(2), \ldots, \vartheta_t(N\tau))', \tag{10}$$

where

$$\vartheta_t(k) = \begin{cases} d_2^{(1)}(k)^{-1}(\Delta S_{t+k\delta} - c^{(1)}(k) - d_1^{(1)}(k) S_t) & \text{if } d_2^{(1)} \neq 0 \\ d_2^{(2)}(k)^{-1}((\Delta S_{t+k\delta})^2 - (c^{(1)}(k) + d_1^{(1)}(k) S_t)^2 & \text{if } d_2^{(1)} = 0, d_2^{(2)} \neq 0 \\ \quad - c^{(2)}(k) - d_1^{(2)}(k) S_t) & \end{cases}$$

for $k = 1, 2, \ldots, N\tau$. Note that by construction we have $E[\vartheta_t(k)|\mathcal{F}_t] = V_t$. Thus, given a set of observed state variables from the AJD model in Equations (1) and (2) over time interval $[t, t+\tau]$, i.e. $\{S_{t+k\delta}\}_{k=0}^{N\tau}$, an unbiased estimator of the latent state variable[1] V_t can now be computed as:

$$\widehat{V}_t = W'\vartheta_t, \tag{11}$$

where W is an $N\tau \times 1$ weighting vector with weights summing to unity and ϑ_t is as given in Equations (10). For any choice of weighting vector W, the latent variable estimator specified by Equation (11) is unbiased, i.e. $E[\widehat{V}_t|\mathcal{F}_t] = V_t$, and has a variance equal to $Var[\widehat{V}_t|\mathcal{F}_t] = W'\Sigma_t W$, where $\Sigma_t = E\left[\vartheta_t\vartheta_t'|\mathcal{F}_t\right] - E[\vartheta_t|\mathcal{F}_t]\,E[\vartheta_t|\mathcal{F}_t]'$. The optimal weighting vector that leads to the minimum-variance estimator can be derived as:

$$W^* = \frac{\Sigma_t^{-1}\iota}{\iota'\Sigma_t^{-1}\iota}, \tag{12}$$

where ι is an $N\tau \times 1$ vector of ones. Note that when the weighting vector is chosen optimally, the variance of the estimator reduces to $Var[\widehat{V}_t|\mathcal{F}_t] = \left(\iota'\Sigma_t^{-1}\iota\right)^{-1}$ and we refer to the estimator as the unbiased minimum-variance (UMV) estimator throughout the remainder of this article. For the general case, diagonal elements of Σ_t can be obtained directly from the $\mathcal{F}_t$-conditional characteristic function of $\Delta S_{t+k\delta}, k > 0$. Expressions for the off-diagonal elements can be derived using the joint $\mathcal{F}_t$-conditional characteristic function of $\Delta S_{t+k\delta}$ and $\Delta S_{t+j\delta}$ for $j \neq k > 0$. In certain cases, the covariance matrix Σ_t may itself depend on the latent state variable that we aim to estimate. If that occurs, the optimal weights can be computed using a simple iterative procedure, which parallels the construction of the optimal GMM weighting matrix.

Compared to existing methods, the latent variable estimator outlined above has a number of distinguishing features. Different from the Kalman filter, the estimator is derived within the general AJD framework of Duffie, Pan, and

[1] It is also clear from here that the approach developed in this article can be extended to affine discrete-time processes, such as the compound autoregressive processes introduced by Darolles, Gouriéroux, and Jasiak (2001).

Singleton (2000) and therefore applies to a wider range of models. A key feature of the model-based approach is that it allows for the efficient use of neighboring information through smoothing over extended horizons by exploiting the model structure. In this aspect, the estimator is similar to the EWMA and Nelson ARCH filter for the spot variance and the nonparametric smoothing method developed in Andreou and Ghysels (2002) for the integrated variance. The important difference here is that the weighting vector is derived from the exact model structure and set optimally so as to ensure unbiasedness and minimum variance of the estimator. Perhaps most closely related to our article is the work by Foster and Nelson (1996) who develop variance and covariance smoothing estimators for a general continuous semi-martingale process. While their specification of the variance process is more general than ours, the weights (and smoothing windows) they derive are optimal only asymptotically, whereas ours are optimal even in a finite sample. Further, our approach remains valid in the presence of random jumps. This is particularly important since a number of recent articles find that jumps constitute a critical ingredient in asset return modeling [see for example Andersen, Benzoni, and Lund (2002), Bakshi, Cao, and Chen (1997), Bates (2000), Das (2002), Eraker, Johannes, and Polson (2003), Johannes (2004), and Pan (2002)]. Finally, in contrast to various simulation-based methods, our estimator is derived from the continuous-time model without discretization, and its computation is straightforward requiring no path simulations.

A further feature of the present framework is that unbiased estimators of *power functions* of the latent state variables can also be constructed conveniently from the dynamic relationship between observed variables and latent variables. In general, the moments of a random variable are related to the cumulants through the following relationship [Kendall and Stuart (1977), p70]:

$$E[(\Delta S_{t+k\delta})^L|\mathcal{F}_t] = \sum_{m=1}^{L}\sum_{\pi}\left(\frac{K^{(p_1)}\,(\Delta S_{t+k\delta}|\mathcal{F}_t)}{p_1!}\right)^{\pi_1}\cdots\left(\frac{K^{(p_m)}\,(\Delta S_{t+k\delta}|\mathcal{F}_t)}{p_m!}\right)^{\pi_m}\frac{L!}{\pi_1!\cdots\pi_m!},$$

for $k > 0$, where the second summation extends over all nonnegative integer values of the π's such that

$$p_1\pi_1 + p_2\pi_2 + \cdots + p_m\pi_m = L.$$

By substituting the conditional cumulants derived in Equation (7) into the above relationship, we obtain:

$$\begin{pmatrix} E[\Delta S_{t+k\delta}|\mathcal{F}_t] \\ E[(\Delta S_{t+k\delta})^2|\mathcal{F}_t] \\ \vdots \\ E[(\Delta S_{t+k\delta})^L|\mathcal{F}_t] \end{pmatrix} = \begin{pmatrix} c_1(k) & b_{11}(k) & 0 & \cdots & \cdots & 0 \\ c_2(k) & b_{21}(k) & b_{22}(k) & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_L(k) & b_{L1}(k) & b_{L2}(k) & \cdots & \cdots & b_{LL}(k) \end{pmatrix} \begin{pmatrix} 1 \\ V_t \\ V_t^2 \\ \vdots \\ V_t^L \end{pmatrix}. \tag{13}$$

The expressions of each element in the above matrix can be solved via iteration. For instance, we have:

$$c_1(k) = c^{(1)}(k) + d_1^{(1)}(k) S_t,$$
$$b_{11}(k) = d_2^{(1)}(k),$$
$$c_2(k) = (c^{(1)}(k) + d_1^{(1)}(k) S_t)^2 + c^{(2)}(k) + d_1^{(2)}(k) S_t,$$
$$b_{21}(k) = 2(c^{(1)}(k) + d_1^{(1)}(k) S_t) d_2^{(1)}(k) + d_2^{(2)}(k),$$
$$b_{22}(k) = (d_2^{(1)}(k))^2.$$

From Equation (13), it is straightforward to see that unbiased estimators of $V_t^2, \ldots, V_t^L$ can be constructed recursively on the basis of the moments of lower orders. For instance, when $d_2^{(1)}(k) \neq 0$, from Equation (10) we have:

$$V_t^2 = \left(d_2^{(1)}(k)\right)^{-2} (E[(\Delta S_{t+k\delta})^2 \,|\mathcal{F}_t] + a(k) V_t + c(k)),$$

for $k = 1, 2, \ldots, N\tau$, $a(k) = -d_2^{(2)}(k) - 2d_2^{(1)}(k)(c^{(1)}(k) + d_1^{(1)}(k)S_t)$ and $c(k) = -(c^{(1)}(k) + d_1^{(1)}(k)S_t)^2$. With the unbiased estimator of V_t and the observed state variables, the unbiased minimum-variance estimator of V_t^2 can be constructed. It is noted that the statistical inference of the model often involves high-order moments and therefore nonlinear functions of latent variables. As we will see later on, the unbiased estimators of power functions of latent variables facilitate nicely the implementation of moments-based estimation methods.

A few remarks are in order at this stage. First, it is emphasized that the UMV estimator has minimum variance for *given* sampling frequency N and sampling horizon τ. In general, increasing N or τ can further improve the accuracy of the estimator. In practice, the choice of sampling frequency and horizon is often based on the structure of the process and subject to the availability of data. For instance, for a transient and volatile variance process, increasing the sampling frequency will be more helpful for the estimation of the spot variance than increasing the sampling horizon. On the other hand, for a highly persistent variance process, increasing the sampling horizon may be helpful to uncover its long run dynamics. Second, we note that the efficiency gain associated with any model-based estimation approach, including the one outlined here, comes at the cost of potential model misspecification or model risk. Thus, when implementing the estimator it is important to ensure that the model provides a reasonable description of the data and is sufficiently flexible to capture its salient features.

2.2 Illustration: Latent Variable Estimation for the SVJ Model

In continuous-time asset return models the observed state variable S_t typically denotes the logarithmic asset price, while the latent or unobserved variable V_t often represents the conditional volatility of returns. Below, we use the stochastic volatility with random jump (SVJ) model to illustrate the latent state estimation approach. Let the logarithmic asset price, $S_t = \ln P_t$, follow the stochastic differential equation:

$$dS_t = \left(\mu - \mu_J \lambda\right) dt + \sqrt{V_t} dW_t^s + \ln J_t dq_t\left(\lambda\right),$$
$$dV_t = \beta\left(\alpha - V_t\right) dt + \sigma\sqrt{V_t} dW_t^v, \tag{14}$$

where $dW_t^s dW_t^v = \rho dt$, $\ln J_t \sim \text{iid } \mathcal{N}\left(\mu_J, \sigma_J\right)$, and $dq_t(\lambda) \sim \text{iid Poisson}\left(\lambda dt\right)$. In this model the instantaneous asset return is driven by stochastic volatility and a random jump component. The stochastic volatility component follows a square-root process that can be correlated with the asset return and has a reflecting barrier at 0 which is attainable when $2\alpha\beta < \sigma^2$. The parameter β measures the mean reversion of the volatility process, and the correlation between dW^s and dW^v measures the level of asymmetry of the conditional volatility, or the so-called "leverage effect". The jump term is driven by a Poisson process with constant jump intensity and the jump size follows an independent log-normal distribution. It is noted that our approach can be easily extended to the case where the jump intensity is specified as a linear function of the state variable V_t as considered in Andersen, Benzoni, and Lund (2002) and Pan (2002). The SVJ model in Equation (14) has an associated closed-form expression for European option prices [see for example Heston (1993), Bakshi, Cao, and Chen (1997)] and is widely used in the finance literature, [e.g. Bates (1996, 2000), Bakshi, Cao, and Chen (1997), Bakshi, Cao, and Chen (2000), Bakshi and Madan (2000), and Scott (1997)].

For the SVJ specification in Equation (14), the joint CCF of $(S_{t+\delta}, V_{t+\delta})$ can be obtained analytically by solving the complex-valued Ricatti equations in Equation (4). The solution is given in Appendix A for the general model specification in Equation (19). On the basis of this, closed form expressions for the conditional cumulants of any order can be derived for both the observed and the latent variable. In particular, the second-order conditional cumulant of the observed variable $\Delta S_{t+k\delta}$ is derived as:

$$Var\left[\Delta S_{t+k\delta} | \mathcal{F}_t\right] = a\left(k\right) + c\left(k\right) V_t,$$

where $a(k) = \alpha(\delta + \beta^{-1} e^{-k\beta\delta}(1 - e^{\beta\delta})) + (\sigma_J^2 + \mu_J^2)\lambda\delta$ and $c\left(k\right) = \beta^{-1} e^{-k\beta\delta}(e^{\beta\delta} - 1)$. Referring to the general notations in Equation (7), for this specific example we have $d_2^{(1)} = 0$ and $d_2^{(2)} = c(k) \neq 0$. On the basis of Equations (10) and (11), for $d_2^{(1)}\left(k\right) = 0$ and $d_2^{(2)}\left(k\right) \neq 0$, the unbiased estimator of the latent spot variance is then given as:

$$\widehat{V}_t = W'\vartheta_t \quad \text{with} \quad \vartheta_t\left(k\right) = \frac{\left(\Delta S_{t+k\delta}\right)^2 - \left(\left(\mu - \mu_J\lambda\right)\delta\right)^2 - a\left(k\right)}{c\left(k\right)}, \tag{15}$$

for $k = 1, 2, ..., N\tau$, where the $N\tau \times 1$ weighting vector W can be chosen optimally, i.e. $W \propto \Sigma_t^{-1}\iota$, so as to minimize the variance of the estimator. On the basis of the CCF, a closed-form expression for Σ_t can be derived as a function of the model parameters and the latent state variable V_t. Since the optimal weights depend on the latent state variable that we aim to estimate, the UMV estimator is infeasible. In practice, this issue can be resolved through the use of an iterative procedure that parallels the construction of the GMM optimal weighting matrix. Jiang and Oomen (2006) report simulation results for the SV and SVJ model, which indicate that for a reasonable choice of initial weighting vector (e.g. flat) the efficiency loss due to the use of a feasible optimal weighting vector is negligible. They also provide extensive simulations on the performance of the spot variance estimator and find that it compares favorably to various alternative methods, including EWMA and the Nelson ARCH filter in the presence of jumps. As can be seen from Equation (15), the estimator of the time-t spot variance is equal to a constant plus the weighted sum of squared (high-frequency) returns over the interval $[t, t+\tau]$. When W is chosen optimally, the weighting vector is downward sloping so that future return observations receive progressively less weight. In particular, the closer β is to 0 (i.e. the more persistent the volatility process is), the slower is the decay in these weights.

From Equation (13) above, unbiased estimators of power functions of the latent spot variance can also be constructed. For instance, an unbiased estimator of V_t^2 can be derived using an unbiased estimate of V_t in conjunction with the following expression:

$$V_t^2 = c(k)^{-2}\left(\frac{1}{3}E[(\Delta S_{t+k\delta})^4\,|\mathcal{F}_t] - a(k)^2 - 2a(k)\,c(k)\,V_t - b(k)\right),$$

for $k = 1, 2, \ldots, N\tau$ and $b(k) = K^{(4)}[\Delta S_{t+k\delta}|\mathcal{F}_t]$ which is available in closed form from the CCF (for simplicity of notation we have assumed that $\mu = \mu_J\lambda$).

2.3 GMM Estimation Using the Latent State Variable Estimator

Using the latent state variable estimator developed above, we now outline the details of a novel GMM estimation approach. A key feature of the proposed procedure here is that in addition to matching moments of observed state variables, the unbiased estimator of latent state variables and its power functions also allows us to match moments of the latent variables. With the empirical application in mind and in line with Bollerslev and Zhou (2002), we focus on the case where the moment conditions are evaluated using low-frequency data. Namely, while the latent variable estimation is based on the high-frequency or intraday data (with sampling interval $\delta < 1$), the moment conditions in the GMM estimation are evaluated at a daily frequency (with sampling interval $\delta = 1$). Also, consistent with the discussion above, we consider the case where the observed state variables S_t are first difference-stationary and the latent variables V_t are stationary. Hence, the moment conditions are derived from the following CCF, which is based on Equation (5):

$$\psi_t(u_1, u_2, \Delta S_{t+1}, V_{t+1}) = \psi_t(u_1, u_2, S_{t+1}, V_{t+1})\, e^{-iu_1' S_t}. \tag{16}$$

The lemma below justifies the use of moment conditions involving power functions of both observed and unobserved variables.

Lemma 1.1. *Let $X_t = (S_t', V_t')$ follow the AJD process as in Equations (1) and (2) for a given set of parameters $\theta_0 \in \Theta$. On the basis of the joint CCF in Equation (16) the*

$\mathcal{F}_t$-conditional moments of ΔS_{t+1} and V_{t+1} can be derived as:

$$E[(\Delta S_{t+1})^{L_s} V_{t+1}^{L_v} | \mathcal{F}_t] = \left. \frac{\partial^{L_s+L_v} \psi_t\left(u_1, u_2, \Delta S_{t+1}, V_{t+1}\right)}{i^{L_s+L_v} \partial u_1^{L_s} \partial u_2^{L_v}} \right|_{u=0}$$
$$= \sum_{0 \le l_s + l_v \le L_s + L_v} c_{L_s, L_v}\left(l_s, l_v\right) \left(\Delta S_t\right)^{l_s} V_t^{l_v}$$

where $l_s, l_v, L_s, L_v \in \{0, 1, \ldots\}$, $t = 0, \ldots, T-1$ and $c_{L_s,L_v}(\cdot)$ is a function of the model parameters as characterized in the proof. Let $\widehat{V_t^{l_v}}$ denote the unbiased estimator of $V_t^{l_v}$ developed in Section 2.1 and define:

$$\epsilon_t\left(L_s, L_v\right) = \left(\Delta S_{t+1}\right)^{L_s} \widehat{V_{t+1}^{L_v}} - \sum_{0 \le l_s + l_v \le L_s + L_v} c_{L_s, L_v}\left(l_s, l_v\right) \left(\Delta S_t\right)^{l_s} \widehat{V_t^{l_v}}. \tag{17}$$

It then holds that:

$$E\left[\epsilon_t\left(L_s, L_v\right) | \mathcal{F}_t\right] = 0.$$

As special cases, we have

$$E\left[\widehat{V_t^{l_v}} | \mathcal{F}_t\right] = V_t^{l_v} \quad \text{or equivalently} \quad E\left[\widehat{V_t^{l_v}} - V_t^{l_v} | \mathcal{F}_t\right] = 0$$
$$\forall \quad 0 \le t \le T-1, 1 \le l_v \le L_v.$$

Proof See Appendix B.

Lemma 1.1 above underlines the distinguishing feature of the GMM estimator, i.e. the use of exact moments and unbiased estimators of the latent variables as well as their power functions. It is clear from Lemma 1.1 that for a given choice of L_s and L_v, the function $\epsilon_t\left(L_s, L_v\right)$ constitutes a valid moment restriction. We denote the vector of moment conditions by $f_t\left(\theta\right)$, with $q = \dim(f_t\left(\theta\right))$. Instrumental variables $z_t \in \mathcal{F}_t$ that are uncorrelated with $f_t\left(\theta\right)$ can be used to generate additional moment restrictions. Given the conditional information set in our model, the natural candidates of the instrumental variables are the historical returns and τ-day period lagged latent variable estimates.[2] Identification of the parameter vector θ requires the specification of a vector of moment restrictions $h_t(\theta) = f_t\left(\theta\right) \otimes \left(1, z_t\right)$ such that it has expectation zero when evaluated at the true population parameter θ_0, i.e. $E\left[h_t(\theta_0) | \mathcal{F}_t\right] = 0$, and has a dimension that exceeds the number of parameters to be estimated, i.e. $\dim(h) = q^* = q(1 + \dim(z_t)) \ge \dim(\theta)$. Note that Lemma 1.1

[2] The study of optimal instruments in the literature has focused on linear models. For instance, in the absence of conditional heteroskedasticity, the explicit expressions for the optimal instruments are derived in [e.g. Hansen (1985) and Hansen and Singleton (1991), Hansen and Singleton 1996]. When both serial correlation and conditional heteroskedasticity are present, Hansen (1985) and Hansen, Heaton, and Ogaki (1988) present a characterization of the efficiency bound for GMM estimators that corresponds to a given system of conditional moment restrictions.

implies that $\{h_t(\theta_0)\}$ is a martingale difference sequence and therefore serially uncorrelated.

Practical implementation of the GMM estimator involves replacing the population moments by their unbiased sample counterpart and then minimizing a quadratic form of the moment restrictions over the admissible parameter space, i.e.

$$\widehat{\theta} = \arg\min_{\theta\in\Theta}\{J(\theta) = \overline{h}(\theta)' W \overline{h}(\theta)\}, \tag{18}$$

where $\overline{h}(\theta) = \frac{1}{T}\sum_{t=1}^{T-1} h_t(\theta)$, Θ denotes the admissible parameter space and W is a positive semidefinite weighting matrix. Hansen (1982) shows that under regularity conditions, the GMM estimator in Equation (18) is consistent and asymptotically normal. Further, the asymptotic variance of $\widehat{\theta}$ can be minimized through an optimal choice of weighing matrix, i.e. $W = S(\theta)^{-1}$ where $S(\theta) = E\left[h(\theta)\, h(\theta)'\right]$. In practice, this efficient GMM estimator is typically implemented using a two-step procedure: in the first step a consistent estimate $\widehat{\theta}_{(1)}$ is obtained using an arbitrary weighting matrix, which is then used in the second step to obtain the efficient estimator using $\widehat{W} = \widehat{S}(\widehat{\theta}_{(1)})^{-1}$. A consistent estimator of the asymptotic covariance matrix of the GMM estimator is given by $\frac{1}{T}(h_\theta(\theta)'\widehat{W}h_\theta(\theta))^{-1}$ where $h_\theta(\theta) = \partial h(\theta)/\partial\theta$ is evaluated at the estimated parameters. Further, because the GMM criterion function $J(\theta)$ measures the distance between $h(\theta)$ and zero, it serves as a goodness-of-fit test for the model, i.e. a high value of this test suggests that the model is misspecified. As shown by Hansen (1982), under the null hypothesis of correct model specification we have $J(\theta) \overset{a}{\sim} \mathcal{X}^2_{q^*-p}$.

As mentioned earlier, a distinguishing feature of the proposed GMM approach is that the moments of both observed *and* latent state variables can be included. As a result, the set of moment conditions is greatly expanded leading to more robust and efficient parameter estimates. Since the population moments are derived directly from the CCF in Equation (3), they are exact in the sense that they correspond to the continuous-time model without any discretization or approximation error. Further, compared to the simulation-based methods, our approach is computationally less demanding since it does not rely on path simulation or numerical integration over latent state variables. It is noted that the use of unconditional and conditional moments of returns and volatility to estimate SV models has also been suggested, but not formally pursued, by Das and Sundaram (1999), Meddahi (2002), Meddahi and Renault (2004), among others. Meddahi (2002), in particular, derives a number of conditional and unconditional moments for the eigenfunction SV models that can be used for model estimation in a similar fashion as the method proposed in this article.

3 EMPIRICAL APPLICATION

The GMM estimation approach proposed in this article can be applied to AJD models for asset return dynamics, interest rate dynamics, etc. In finance literature, continuous-time models with latent state variables, such as stochastic mean and stochastic volatility, have been proposed for both asset return and term structure dynamics. Our application focuses on the estimation of equity index return models with alternative specifications. The general model specification incorporates a random jump component and two stochastic volatility components (hereafter the SV2J model). The SV2J model specifies the dynamics of the logarithmic asset price $S_t = \ln P_t$ as follows:

$$\begin{aligned} dS_t &= (\mu - \mu_J \lambda)dt + \sqrt{V_t}dW_t^{sv} + \sqrt{U_t}dW_t^{su} + \ln J_t dq_t(\lambda), \\ dV_t &= \beta\,(\alpha - V_t)\,dt + \sigma\sqrt{V_t}dW_t^{v}, \\ dU_t &= \kappa\,(\theta - U_t)\,dt + \gamma\sqrt{U_t}dW_t^{u}, \end{aligned} \tag{19}$$

where $dW_t^{sv}dW_t^{v} = \rho dt$, $dW_t^{su}dW_t^{u} = \eta dt$, $\ln J_t \sim \text{iid}\,\mathcal{N}\left(\mu_J, \sigma_J^2\right)$, and $dq_t(\lambda) \sim \text{iid Poisson}\,(\lambda dt)$. In the SV2J model, the asset return is driven by a random jump and two stochastic volatility components. Each stochastic volatility component follows a square-root process and can be correlated with the asset return. The jump term is driven by a Poisson process with constant jump intensity, and the jump size follows an iid log-normal distribution. The model nests many other models as special cases. In particular, when $\kappa = \gamma = \theta = 0$, we have the SVJ model as discussed in Section 2.2. When $\lambda = 0$, we have the SV(2) model, which has been considered in Alizadeh, Brandt, and Diebold (2002), Bates (2000), and Chacko and Viceira (2003).

With a single stochastic volatility component, the UMV estimate of the instantaneous variance, and its power functions, can be obtained following the procedure outlined for the SVJ model in Section 2.2. When there are multiple stochastic volatility components to identify, we need to rely on distinct moment conditions of the observed asset returns. Obviously, for the SV2J model various combinations of moments can be employed for the identification of V_t and U_t. In practice, the natural choices are those among the lowest order cumulants. On the basis of the SV2J model, we first simulate the cumulants of various orders with given parameter values and then compare them with the analytical cumulants derived from the model. The difference between the numerical cumulants calculated from the simulated paths and their theoretical counterparts indicates the robustness of the particular cumulant. We find that among the first four orders of conditional cumulants, the second and fourth-order cumulants are much more stable than the first- and third-order cumulants. Thus, we rely on the second-order conditional cumulants of asset returns for the identification and estimation of the instantaneous variance V_t and U_t and the fourth-order conditional cumulants of asset returns for the identification and estimation of V_t^2 and U_t^2, respectively. On the basis of the joint CCF of $(S_{t+\delta}, V_{t+\delta}, U_{t+\delta})$ for the SV2J model given in Appendix A, the conditional second-order cumulant of $\Delta S_{t+k\delta}$ can be derived as:

$$K^{(2)}[\Delta S_{t+k\delta}|\mathcal{F}_t] = a(\delta,k) + c_v(\delta,k)V_t + c_u(\delta,k)U_t,$$

where $c_v(\delta,k) = \beta^{-1}e^{-k\beta\delta}(e^{\beta\delta}-1)$, $c_u(\delta,k) = \kappa^{-1}e^{-k\kappa\delta}(e^{\kappa\delta}-1)$, and $a(\delta,k) = \alpha(\delta - \beta^{-1}e^{-k\beta\delta}\,(e^{\beta\delta}-1)) + \theta(\delta - \kappa^{-1}e^{-k\kappa\delta}(e^{\kappa\delta}-1)) + (\sigma_J^2 + \mu_J^2)\lambda\delta$. Keeping in mind that the two volatility series are most likely to capture the random components over different frequency of asset returns [see for example Alizadeh, Brandt, and Diebold (2002), Chacko and Viceira (2003)], we use return observations at two different sampling frequencies to construct the estimators, namely $\delta = 5$ minutes and $\delta^* =$ half-day. In other words, the estimation of V_t and U_t is based on

$$c_v\,(\delta,k)\,V_t + c_u\,(\delta,k)\,U_t = K^{(2)}[S_{t+k\delta+\delta} - S_{t+k\delta}|\mathcal{F}_t] - a\,(\delta,k)\,, \quad k = 0,1,...N\tau - 1,$$

and

$$c_v\left(\delta^*,k\right)V_t + c_u\left(\delta^*,k\right)U_t = K^{(2)}[S_{t+k\delta+\delta^*} - S_{t+k\delta}|\mathcal{F}_t] - a\left(\delta^*,k\right), \quad k = 0,1,...N\tau - 1.$$

Simulations show that these two frequencies are effective in identifying the two stochastic volatility components as long as the forward period used in estimation is sufficiently long (up to 5 days or a week). The other condition is that there is sufficient differential in the persistence level (or the mean reversion) of the two distinct volatility processes. Relevant conditional variance and covariance expressions of asset returns for the construction of UMV estimators are derived from the CCF in Appendix A. In addition, $K^{(4)}[\Delta S_{t+k\delta}|\mathcal{F}_t]$ is derived from the same CCF and used to construct the estimators of V_t^2 and U_t^2. Similar to the identification of V_t and U_t, we use return observations with two different frequencies to construct the estimators, namely $\delta = 5$ minutes and $\delta^* =$ half day. For the product term of V_t and U_t (i.e. $V_t \cdot U_t$) in the V_t^2 and U_t^2 estimation, we simply use the product of their respective estimates (i.e. $\widehat{V}_t \cdot \widehat{U}_t$).

Regarding the choice of moment conditions in the GMM estimation, we follow the guidance provided by the Monte Carlo study of Andersen and Sørensen (1996) on the estimation of a discrete-time SV model. Firstly, in determining the number of moments used in the estimation, we keep in mind the following fundamental trade-off: inclusion of additional moments improves estimation performance for a given degree of precision in the estimation of the weighting matrix, but in finite samples this must be balanced against the deterioration in the estimate of the weighting matrix as the number of moments increases. Secondly, very-high-order moments should be avoided owing to their erratic finite sample behavior caused by the presence of fat tails in the asset return distribution. Asymptotic normality of the GMM estimator requires finite variance of the moment conditions and good estimates of these quantities in finite samples. Thus our moment selection focuses on the lower-order moments, which is consistent with Andersen and Sørensen (1996) and Jacquier, Polson, and Rossi (1994). Using the notation above, the vector of moment conditions are specified as:

$$f_t(\theta) = \left\{ \begin{array}{ccc} (\Delta S_{t+1})^{L_s^1} \widehat{V_{t+1}^{L_v^1}} & - & E[(\Delta S_{t+1})^{L_s^1} V_{t+1}^{L_v^1} | \mathcal{F}_t] \\ (\Delta S_{t+1})^{L_s^2} \widehat{V_{t+1}^{L_v^2}} & - & E[(\Delta S_{t+1})^{L_s^2} V_{t+1}^{L_v^2} | \mathcal{F}_t] \\ \vdots & - & \vdots \\ (\Delta S_{t+1})^{L_s^q} \widehat{V_{t+1}^{L_v^q}} & - & E[(\Delta S_{t+1})^{L_s^q} V_{t+1}^{L_v^q} | \mathcal{F}_t] \end{array} \right\}$$

where L_s^i, L_v^i are nonnegative integers for $i = 1, 2, \ldots, q$ and $t = 0, 1, \ldots T - 1$.

It is noted that our approach admits the moment conditions involving the latent variables in the estimation as the unbiased estimators of the power functions of latent variables are available. This is different from Bollerslev and Zhou (2002), in which a nuisance parameter needs to be included in order to adjust for the potential bias. The exact conditional moment conditions used in the estimation are determined with further consideration of the particular model specification:

- For the SV model, we set $\{(L_s^1, L_v^1); \ldots; (L_s^3, L_v^3)\} = \{(1, 1); (0, 1); (0, 2)\}$ with the lagged stochastic volatility as instrument variable, resulting in six moment conditions for four parameters. The cross moment of asset return and variance is important for the identification of the correlation between asset return and stochastic variance, and the first and second conditional moments of instantaneous variance is important for the identification of the mean-reversion parameter and the conditional variance of the stochastic volatility process. For the SV2 model, we use the moment conditions of the SV model for *each* SV component, resulting in 12 moment conditions for 8 parameters.
- For the SVJ model, the extra moment conditions in addition to those of the SV model are the third and fourth conditional moments of asset returns, resulting in 10 moment conditions for 7 parameters. Since the random jump component is intended to capture the skewness and excess kurtosis of asset return distribution, these high moments are important for the identification of jump parameters. For the SV2J model, we use the same moment conditions as for the SV2 model, plus the third and fourth conditional moments of asset returns as used in the SVJ model estimation, resulting in 16 moment conditions for 11 parameters.

3.1 The Data

The data set we use in this article consists of the 5-minute S&P 500 index returns over the period January 2, 1987 to June 30, 1995. To mitigate concerns about the impact of market microstructure noise [e.g. Bandi and Russell (2006)], we also perform our analysis using 10-minute data but find that the results are similar. Daily returns are extracted on the basis of index levels at the beginning of each trading day. Table 1 reports summary statistics of both intraday 5-minute and daily returns. As the summary statistics suggest, the distribution of high-frequency

Table 1 Summary statistics S&P 500 Index returns

	Returns		Absolute returns	
	5-min	Daily	5-min	Daily
Mean	4.79×10^{-4}	3.77×10^{-2}	4.29×10^{-2}	6.30×10^{-1}
Standard deviation	0.105	1.007	0.096	0.787
Skewness	−1.52	−3.68	18.55	8.57
Kurtosis	429.8	74.66	583.1	167.4
Minimum	−5.96	−19.5	0.00	0.00
Maximum	4.34	8.00	5.96	19.5
Autocorrelation				
lag 1	−0.193	0.053	0.601	0.218
lag 2	0.204	−0.057	0.451	0.231
lag 3	−0.083	−0.048	0.379	0.278
lag 4	0.058	−0.038	0.341	0.180
lag 5	−0.080	0.060	0.345	0.237

This table reports the summary statistics of the S&P 500 index returns and absolute returns over the period from January 1987 to June 1995.

returns is extremely fat-tailed as measured by the kurtosis (429.8). The daily returns exhibit less kurtosis but more negative skewness (−3.68 for daily return versus −1.52 for 5-minute return). The kurtosis of daily returns, however, is still very high at 74.66, suggesting the distribution is also fat-tailed. Finally, the highly significant positive autocorrelations of absolute returns suggest that both the intraday and daily return volatilities are highly persistent over time.

Another well-known feature of intraday return data is the intraday seasonality pattern in volatility. More specifically, the volatility of intraday returns demonstrates a U-shaped pattern during the trading hours, with increased market activity around the market opening and closing and diminished market activity around lunch time. Since the intraday seasonality may affect the optimal weights, we remove it by following Engle and Russell (1998) to fit a cubic spline with nodes set at half-hour intervals.

Before turning to the estimation results, we note that since the estimator of stochastic volatility relies on the parameter values and volatility estimates, in each iteration of the GMM optimization procedure the parameter values and the volatility estimates are updated accordingly. The updating procedure is very similar to the implied state generalized method of moments (IS-GMM) proposed in Pan (2002) and is also closely related to latent backfitting approach proposed by Pastorello, Patilea, and Renault (2003). To simplify the estimation procedure, we estimate the constant return parameter μ first using a robust estimation procedure. The estimate of the constant return parameter is 0.095 with a standard error of 0.043. Further, for the estimation of the spot variance, we need to set the time horizon, τ, over which the high-frequency return observations are used. With the

Table 2 GMM estimation results of alternative return models for S&P 500 Index

	SV	SVJ	SV2	SV2J
$\sqrt{\alpha}$	0.166 (0.089)	0.150 (0.064)	0.104 (0.058)	0.097 (0.057)
β	7.550 (3.570)	5.641 (2.421)	1.003 (0.467)	0.897 (0.469)
σ	2.971 (0.561)	2.003 (0.328)	0.820 (0.121)	0.990 (0.117)
ρ	−0.430 (0.112)	−0.316 (0.097)	0.052 (0.036)	0.028 (0.034)
$\sqrt{\theta}$			0.129 (0.079)	0.112 (0.077)
κ			8.437 (4.003)	6.230 (2.889)
γ			3.680 (0.499)	2.711 (0.323)
η			−0.694 (0.191)	−0.478 (0.135)
λ		0.562 (0.351)		0.751 (0.367)
μ_J		−0.071 (0.101)		−0.067 (0.090)
σ_J		0.064 (0.009)		0.051 (0.007)
GMM test of overidentifying restrictions				
χ^2	13.67	13.21	14.11	13.24
d.o.f.	2	3	4	5
p-value	0.11%	0.42%	0.70%	2.12%

This table reports the GMM estimation results for various specifications of the SV2J model in Equation (19) using S&P 500 index returns over the period from January 1987 to June 1995. The moment conditions used for the GMM estimation are given in Section 3. Standard errors are reported in parenthesis next to the parameter estimates. The blank cells indicate that the parameter is restricted to be zero.

guidance of simulations, we set $\tau = 5$ trading days (or equivalently one week) for the models with only stochastic volatility and $\tau = 10$ trading days (or equivalently two weeks) for models with random jumps. Increasing the horizon does not change the results much because the weights given to observations outside this interval is fairly low.

3.2 Estimation Results and Diagnostic Tests for SVJ models

The estimation results for the SV, SV2, SVJ, and SV2J models are reported in Table 2, together with the Hansen-J test. The estimation results for the SV model have higher values for both the mean-reversion parameter and the conditional variance of the SV process than those reported in the literature, see Bates (2006), albeit that our sample covers a much shorter time period owing to limited availability of high-frequency return observations. It is noted that a smaller value of the mean-reversion parameter in the SV process implies a stronger intertemporal persistence of volatility. We also experiment by including higher-order (third and fourth) return moments in the estimation. It results in a more negative estimate of ρ when the third-order return moment is included, and a smaller estimate of the mean-reversion parameter β when the fourth-order return moment is included. This suggests that the estimation results of the SV model are sensitive to the choice of moment conditions, a clear indication of model misspecification. The boundary condition $\sigma^2 < 2\beta\alpha$ is also clearly violated.

The SV2 model identifies two distinct stochastic volatility components with markedly different behavior. Namely, the first component, V_t, is highly persistent (with a half-life of about 170 days) and relatively less volatile than the second component, U_t, which is much less persistent (with a half-life of about 20 days) but more volatile. Interestingly, there appears to be a significant negative correlation between returns and the second volatility component but no significant correlation between returns and the first component. Finally, in terms of model specification, the SV2 model improves over the SV model, albeit that the p-value of the J-test remains below 1%.

The estimation results of both SVJ and SV2J models indicate that the jump component is highly significant, suggesting the importance of including random jump in the asset return model. However, the model specification test (J-test) shows only modest improvement, with the SV2J model having a p-value of just over 2% for the J-test. This may be partially due to the fact that higher-order moment conditions are used in the estimation of the SVJ and SV2J models. It is also noted that inclusion of jumps to the SV models leads to a less volatile and slightly more persistent stochastic variance process. The jump parameter estimates suggest infrequent, highly volatile, and, on average, negative jumps. This is hardly surprising since our sample includes the 1987 crash, which is likely to drive much of the significance of the jump component. The results are overall in line with Bates (2006) for the S&P 500 index and Chernov, Gallant, Ghysels, and Tauchen (2003) for the DJIA index.

As noted by Bollerslev and Zhou (2002), the standard Hansen J-test reported in Table 2 is not particularly informative about the sources of model misspecification. Following Tauchen (1985), we compute a series of diagnostic t-tests associated with each of the moment conditions used in the estimation. Using the notations in Section 3, we have:

$$t = \text{diag}\{[\widehat{W} - h_\theta(\widehat{\theta})(h_\theta(\widehat{\theta})'\widehat{W}^{-1}h_\theta(\widehat{\theta}))^{-1}h_\theta(\widehat{\theta})']\}^{-1/2}\sqrt{T}h(\widehat{\theta}), \tag{20}$$

where $\widehat{W}$ refers to the estimated weighting matrix. The diagnostic t-test in Equation (20) explicitly takes into account the variations resulting from the errors in parameter estimation. Intuitively, a large t-statistic indicates that the associated moment condition contributes significantly towards the model misspecification. Table 3 reports these diagnostic t-tests for all the moment conditions used in the estimation of the various stochastic volatility models. Overall, it is clear that the t-statistics are higher for higher-order moments, confirming the findings in earlier studies such as Andersen and Sørensen (1996) and Jacquier, Polson, and Rossi (1994) and justifying the rationale of focusing on the lower-order moments in the estimation. For the SV model, the apparent poor fit comes from the moment involving the squared variance, namely $E[\widehat{V}_t^2|\mathcal{F}_t]$ with a t-statistic of -4.773. This suggests that the SV model is not flexible enough to incorporate the dynamics of the volatility or the high volatility-of-volatility. The observation is consistent with that in Bollerslev and Zhou (2002) for the exchange rate models. This is also consistent with the fact that the boundary condition $\sigma^2 < 2\beta\alpha$ of the SV model based on parameter estimates is clearly violated. Considering the fact that only the very low orders of moment conditions are used in the SV model estimation, the poor fit of moment conditions suggests severe misspecification of the SV model. With the additional jump component in the SVJ model, the moment conditions associated with the stochastic volatility process are all improved as reflected in the lower t-statistics. However, the SV component remains severely misspecified for the S&P 500 index return variance. Further, the diagnostic t-statistics associated with the third and fourth return moments are highly significant. This suggests that even the SVJ model cannot generate sufficient negative skewness and excess kurtosis to fully account for the features of daily S&P 500 index returns. Re-estimating the model using returns after the 1987 market crash leads to substantially lower t-statistics for both the third and fourth return moment conditions. This is an indication that the poor fit of the third and fourth return moments is mainly due to the market crash. With constant jump intensity and normally distributed jumps, the SVJ model's apparent poor fit to extreme returns is probably not surprising.

Comparison between the SV2 and SV models suggests that decomposing the return variance into two components improves the model specification, which is consistent with the Hansen J-test reported in Table 2. While the affine SV model

Table 3 Moment condition tests

	SV	SVJ	SV2	SV2J
Panel A: Moment restrictions				
$E[\widehat{V}_{t+1}\vert\mathcal{F}_t] - \widehat{V}_{t+1}$	−0.032	−0.011	0.007	0.003
$E[\widehat{V}^2_{t+1}\vert\mathcal{F}_t] - \widehat{V}^2_{t+1}$	−4.773	−3.170	−0.410	−0.308
$E[\widehat{V}_{t+1}\Delta S_{t+1}\vert\mathcal{F}_t] - \widehat{V}_{t+1}\Delta S_{t+1}$	1.410	0.319	0.779	0.616
$E[\widehat{U}_{t+1}\vert\mathcal{F}_t] - \widehat{U}_{t+1}$			−0.045	−0.061
$E[\widehat{U}^2_{t+1}\vert\mathcal{F}_t] - \widehat{U}^2_{t+1}$			−4.848	−4.020
$E[\widehat{U}_{t+1}\Delta S_{t+1}\vert\mathcal{F}_t] - \widehat{U}_{t+1}\Delta S_{t+1}$			1.532	0.302
$E[(\Delta S_{t+1})^3\vert\mathcal{F}_t] - (\Delta S_{t+1})^3$		2.950		2.972
$E[(\Delta S_{t+1})^4\vert\mathcal{F}_t] - (\Delta S_{t+1})^4$		−6.641		−5.879
Panel B: Moment restrictions with instruments				
$E[\widehat{V}_{t+1}\widehat{V}_{t-\tau}\vert\mathcal{F}_t] - \widehat{V}_{t+1}\widehat{V}_{t-\tau}$	−2.339	−2.122	−0.876	−0.879
$E[\widehat{V}^2_{t+1}\widehat{V}_{t-\tau}\vert\mathcal{F}_t] - \widehat{V}^2_{t+1}\widehat{V}_{t-\tau}$	−2.732	−2.501	−1.035	−1.170
$E[\widehat{V}_{t+1}\Delta S_{t+1}\widehat{V}_{t-\tau}\vert\mathcal{F}_t] - \widehat{V}_{t+1}\Delta S_{t+1}\widehat{V}_{t-\tau}$	1.660	1.213	0.951	0.764
$E[\widehat{U}_{t+1}\widehat{U}_{t-\tau}\vert\mathcal{F}_t] - \widehat{U}_{t+1}\widehat{U}_{t-\tau}$			−2.661	−1.970
$E[\widehat{U}^2_{t+1}\widehat{U}_{t-\tau}\vert\mathcal{F}_t] - \widehat{U}^2_{t+1}\widehat{U}_{t-\tau}$			−2.010	−1.221
$E[\widehat{U}_{t+1}\Delta S_{t+1}\widehat{U}_{t-\tau}\vert\mathcal{F}_t] - \widehat{U}_{t+1}\Delta S_{t+1}\widehat{U}_{t-\tau}$			1.733	0.790
$E[(\Delta S_{t+1})^3\widehat{Z}_{t-\tau}\vert\mathcal{F}_t] - (\Delta S_{t+1})^3\widehat{Z}_{t-\tau}$		3.220		3.355
$E[(\Delta S_{t+1})^4\widehat{Z}_{t-\tau}\vert\mathcal{F}_t] - (\Delta S_{t+1})^4\widehat{Z}_{t-\tau}$		−9.071		−8.156

This table reports the diagnostic $t-$test statistic (given by Equation 20) associated with each of the moment conditions used in the GMM estimation of various models. Note that the instrument of return moments is set as $\widehat{Z}_t = \widehat{V}_t$ in the SVJ model and $\widehat{Z}_t = \widehat{V}_t + \widehat{U}_t$ in the SV2J model.

is still severely misspecified for the more volatile and less persistent volatility component, it is not rejected for the less volatile and more persistent volatility component. The diagnostic tests of the SV2J model suggest that the SV components and random jump component are to certain extent complementary to each other. Relative to the SV2 model, the additional jump component in the SV2J model improves the fit of moment conditions associated with the volatility process, in particular those of the more volatile and less persistent volatility component. The *t*-statistics of these moment condition are in general reduced. Compared to the SVJ model, the additional volatility component in the SV2J model seems to improve the fit of the fourth return moment, even though not so for the third return moment. This is likely because the mixture of distribution of the additional SV component can further incorporate a high level of kurtosis in returns. Overall, the moment condition tests confirm the specification tests reported in Table 2 and further pinpoint the sources of model misspecification.

3.3 Further Examination of the Stochastic Volatility Components

As noted earlier, a useful by-product of the model inference proposed in this article is the estimates of stochastic volatility. That is, the estimates of daily spot variance associated with all four models are obtained as a result of the estimation procedure. In Figure 1, we plot the estimated volatility series for the SV and SV2 models in Panels A and B, respectively. The volatility estimates of the SVJ and SV2J models behave similarly and are therefore not plotted. Consistent with the model parameter estimates, the first volatility component of the SV2 model appears to be quite stable, while the second component is highly volatile and behaves similar to that of the univariate component of the SV model in Panel A. To further understand the characteristic of each volatility component in the following we investigate their relation to various market variables.

In finance literature, asset price volatility is believed to be directly related to the rate of flow of market information [e.g. Clark (1973), Tauchen and Pitts (1983), Ross (1989), Andersen (1996)]. Various studies have focused on the relationship between asset price volatility and market activity. For instance, the relation between trading volume and price volatility has been investigated in many empirical studies [e.g. Clark (1973), Tauchen and Pitts (1983), Lamoureux and Lastrapes (1990), Gallant, Rossi, and Tauchen (1992)] or using theoretical models [e.g. Karpoff (1986), Harris and Raviv (1993), He and Wang (1995)]. In particular, the model in Harris and Raviv (1993) predicts that the trading in speculative markets is mainly driven by the differential opinions among traders. At the same time, return volatility has also been linked to returns through the ''leverage effect'' [e.g. Black (1976), Christie (1982), Tauchen, Zhang, and Liu (1996)] and the risk premium or ''feedback effect'' [e.g. Merton (1980), Engle, Lilien, and Robins (1987), French, Schwert, and Stambaug (1987), Ghysels, Santa-Clara, and Valkanov (2005), Brandt and Kang (2004)]. Interestingly, while the ''leverage effect'' suggests a negative correlation between asset return and conditional volatility, the risk premium or ''feedback effect'' predicts a positive relation

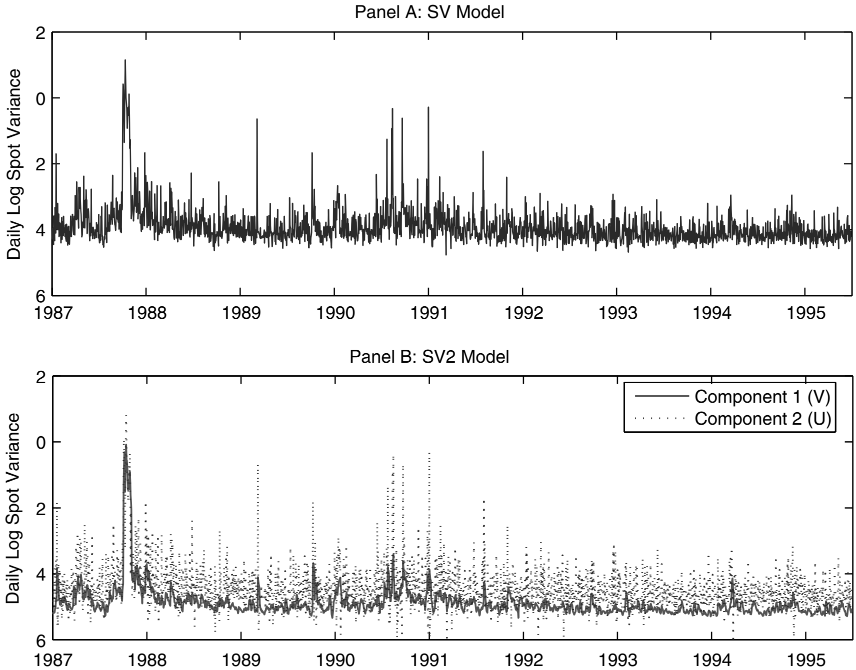

Figure 1 Estimated volatility components for SV and SV2 model. The figure plots the logarithmic estimated spot volatility for the S&P 500 index over the period from January 1987 to June 1995 under the SV model in Panel A, and the two volatility components under the SV2 model in Panel B. The SV2 model is specified in Equation (14) with $\lambda = 0$, while the SV model is obtained by further imposing the restrictions $\kappa = \gamma = \theta = 0$.

between asset return and conditional volatility. It is noted, however, that the extensive empirical literature on volume−volatility and return−volatility relations has, to the best of our knowledge, concentrated exclusively on a single component of volatility measure. Below, we use the two volatility components, identified from the SV2 model, to examine the volume−volatility and return−volatility relationship. The motivation of our analysis is similar to that of Chan and Fong (2000) who investigate the volatility−volume relationship by decomposing trade volume into two distinct components, namely number of transactions and trade size. The aim of our analysis is to provide further understanding or interpretation of the decomposed volatility series in terms of their respective market information content.

Using the time series estimates of the two volatility components, we perform linear regressions for both the trading volume and return of the S&P 500 index. The results are reported below, with Newey−West t-statistics in parenthesis under the coefficient estimates:

$$\text{Volume}_t = \underset{(3.50)}{-0.003} + \underset{(2.53)}{0.201} V_t + \underset{(1.30)}{0.059} U_t \qquad R^2 = 18.1\% \tag{21}$$

$$\text{Return}_t = \underset{(2.02)}{-0.058} + \underset{(1.64)}{7.354} V_t - \underset{(1.77)}{3.176} U_t \qquad R^2 = 3.86\% \tag{22}$$

where V_t and U_t are the two daily spot volatility components estimated at the beginning of the day under the SV2 model, Return_t and Volume_t are the daily return and the time-detrended daily trading volume of the S&P 500 index, respectively. In the volume regression, while the coefficients of both volatility components are positive, only the first coefficient is significant at conventional critical levels. That is, trading volume is mainly driven by the first volatility component V_t. In contrast, the coefficients of the two volatility components in the return regression have opposite signs, with both coefficients marginally significant at about 10% critical level. The first volatility component, V_t, has a relation with return as predicted by the risk premium or feedback effect, while the second volatility component, U_t, has a relation with return as predicted by the leverage effect.

The volume regression results provide evidence to support the conjecture of the model in Harris and Raviv (1993), i.e. the trading activity is mainly driven by a subset of the market information. Such information is believed to be related to the heterogeneous beliefs among investors. That is, heterogeneous beliefs among investors tend to drive up the trading activities or increase the trading volume. The regression results in Equation (21) identifying the more persistent and less volatile stochastic volatility component V_t as the measure of such information set. It is also believed that homogeneous beliefs among investors tend to have a directional impact on asset prices as investors adjust the asset valuation and expected market returns according to the common belief. That is, such information tends to have significant impact on asset return but may not induce higher trading activities. The regression results in Equations (21) and (21) identify the less persistent and more volatile stochastic volatility component U_t as the measure of such information set. The results in Equation (21) further suggest that such volatile or less anticipated informational shock has an asymmetric relation with the market return. That is, a negative informational shock tends to have a larger impact on market return than a positive one, or the so-called leverage effect. On the other hand, the positive coefficient of the stochastic volatility component V_t suggests that a highly persistent or more anticipated information flow comes with a positive risk premium required by investors to compensate the expected risk factor. The results seem to reconcile the seemingly contradicting return–volatility relations as predicted by the leverage effect and the risk premium or feedback effect.

To summarize, our estimation results suggest that at least three distinct factors are identified from the asset return dynamics of the S&P 500 index, namely a highly persistent and less volatile SV volatility component, a less persistent but more volatile SV component, and a random jump component. Further analysis also provides evidence that these distinct factors are associated with different types of information in the market and have different impact on market activities. The decomposition of such information may help us to further understand the links between information flow and trading activity as well as asset return dynamics in the financial market.

4 CONCLUSION

This article proposes a new approach for the statistical inference of continuous-time jump diffusion models with latent variables. We first construct unbiased minimum-variance estimators for the latent variables as well as their power functions. The estimators are based on the exact dynamic relationship between the observed state variables and latent variables and therefore are consistent with model specification. When applied to models with stochastic volatility, the estimators can exploit the high-frequency return observations with improved finite sample performance. With the estimators of latent variables as well as their power functions, we propose a new GMM procedure for continuous-time AJD models with latent variables. The distinguishing feature of the method is that moments of both observed and unobserved latent variables can be used. The estimation involves neither path simulation nor discretization of the continuous time process, and does not suffer from the error-in-variable problem even when the moments of latent variables are included. We apply our GMM estimator to a number of commonly used continuous-time asset return models with stochastic volatility and random jumps using the high-frequency S&P 500 index return observations. Our results suggest that at least three distinct factors are identified from the asset return dynamics, namely a highly persistent and stable stochastic volatility component, a less persistent and more volatile stochastic volatility component, as well as a random jump component. Finally, an exploratory analysis into the characteristics of the estimated volatility series suggests that the distinctly behaved volatility components represent different types of market information and are related to different aspects of market activities.

APPENDIX : A Conditional Characteristic Functions of the SV2J Model

For the SV2J model in Equation (19), the joint CCF of $(S_{t+\delta}, V_{t+\delta}, U_{t+\delta})$ can be written as:

$$\psi_t\left(u_1, u_2, u_3, S_{t+\delta}, V_{t+\delta}, U_{t+\delta}\right) = \exp\{C\left(\delta, u_1, u_2, u_3\right) + D1\left(\delta, u_1, u_2, u_3\right) S_t + D2\left(\delta, u_1, u_2, u_3\right) V_t + D3\left(\delta, u_1, u_2, u_3\right) U_t\},$$

where $C\left(\cdot\right), D1\left(\cdot\right)$ and $D2\left(\cdot\right)$ are solved from the Ricatti equations as:

$$\begin{aligned} C\left(\delta, u_1, u_2, u_3\right) = {} & \left(iu_1\mu + i\beta\alpha u_2 + i\kappa\theta u_3\right)\delta \\ & + \frac{\alpha\beta}{\sigma^2}\left[\left(b1 - h1\right)\delta - 2\ln\left(\frac{1 - g1e^{-h1\delta}}{1 - g1}\right)\right] \\ & + \frac{\theta\kappa}{\gamma^2}\left[\left(b2 - h2\right)\delta - 2\ln\left(\frac{1 - g2e^{-h2\delta}}{1 - g2}\right)\right] \\ & + \lambda\delta(\exp\{iu_1\mu_J - \frac{1}{2}u_1^2\sigma_J^2\} - 1 - iu_1\mu_J) \end{aligned}$$

$$D1\,(\delta, u_1, u_2, u_3) = iu_1$$

$$D2\,(\delta, u_1, u_2, u_3) = iu_2 + \frac{b1 - h1}{\sigma^2}\frac{1 - e^{-h1\delta}}{1 - g1e^{-h1\delta}}$$

$$D3\,(\delta, u_1, u_2, u_3) = iu_3 + \frac{b2 - h2}{\gamma^2}\frac{1 - e^{-h2\delta}}{1 - g2e^{-h2\delta}}$$

with

$$h1\,(u_1, u_2, u_3) = \left[b1^2 + \sigma^2(u_1^2 + 2\rho\sigma u_1 u_2 + 2i\beta u_2)\right]^{1/2}$$

$$h2\,(u_1, u_2, u_3) = \left[b2^2 + \gamma^2(u_1^2 + 2\eta\gamma u_1 u_3 + 2i\kappa u_3)\right]^{1/2}$$

and $b1\,(u_1, u_2, u_3) = \beta - \rho\sigma i u_1 - \sigma^2 u_2 i,$ $b2\,(u_1, u_2, u_3) = \kappa - \eta\gamma i u_1 - \gamma^2 u_3 i,$ $g1\,(u_1, u_2, u_3) = (b1 - h1)\,/\,(b1 + h1), g2\,(u_1, u_2, u_3) = (b2 - h2)\,/\,(b2 + h2).$

APPENDIX : B

Proof of Equation (7). Using Equations (5) and (6), the law of iterated expectations, and $D = i(-iD)$, we have:

$$\begin{aligned}
\psi_t\,(u_1, \Delta S_{t+k\delta}) &= E\left[\exp\left\{iu_1(S_{t+k\delta} - S_{t+(k-1)\delta})\right\} \mid \mathcal{F}_t\right] \\
&= E\left[E\left[\exp\left\{iu_1 S_{t+k\delta} - iu_1 S_{t+(k-1)\delta}\right\} \mid \mathcal{F}_{t+(k-1)\delta}\right] \mid \mathcal{F}_t\right] \\
&= E\left[\exp\left\{C\,(\delta;\, u_1, 0) + (D1\,(\delta;\, u_1, 0)' - iu_1)S_{t+(k-1)\delta}\right.\right. \\
&\quad \left.\left. + D2\,(\delta;\, u_1, 0)'\, V_{t+(k-1)\delta}\right\} \mid \mathcal{F}_t\right] \\
&= \exp\{C\,(\delta;\, u_1, 0)\}\psi_t\left(-iD1\,(\delta;\, u_1, 0)' - u_1, -iD2\,(\delta;\, u_1, 0)',\right. \\
&\quad \left. S_{t+(k-1)\delta}, V_{t+(k-1)\delta}\right)
\end{aligned}$$

The conditional cumulant expressions now follow directly. ■

Proof of Lemma 1.1. The result that $E\left[\epsilon_t\,(L_s, L_v)\,|\mathcal{F}_t\right] = 0$ follows directly from the unbiasedness of the estimates of power functions of the latent state variables.

In particular:

$$E\left[\epsilon_t\left(L_s,L_v\right)|\mathcal{F}_t\right]=E[(\Delta S_{t+1})^{L_s}\widehat{V_{t+1}^{L_v}}-\sum_{0\le l_s+l_v\le L_s+L_v}c_{L_s,L_v}\left(l_s,l_v\right)\left(\Delta S_t\right)^{l_s}\widehat{V_t^{l_v}}|\mathcal{F}_t]$$
$$=E[[(\Delta S_{t+1})^{L_s}\widehat{V_{t+1}^{L_v}}|\mathcal{F}_{t+1}]|\mathcal{F}_t]$$
$$-E\left[\sum_{0\le l_s+l_v\le L_s+L_v}c_{L_s,L_v}\left(l_s,l_v\right)\left(\Delta S_t\right)^{l_s}\widehat{V_t^{l_v}}|\mathcal{F}_t\right]$$
$$=E[(\Delta S_{t+1})^{L_s}V_{t+1}^{L_v}|\mathcal{F}_t]-\sum_{0\le l_s+l_v\le L_s+L_v}c_{L_s,L_v}\left(l_s,l_v\right)\left(\Delta S_t\right)^{l_s}V_t^{l_v}=0$$

It is easy to verify from the exponential affine characteristic function in Equation (5) or Equation (16) that $E[(\Delta S_{t+1})^{L_s}V_{t+1}^{L_v}|\mathcal{F}_t]$ is a bivariate polynomial in terms of ΔS_{t+1} and V_{t+1} up to orders of L_s+L_v. The coefficients $c_{L_s,L_v}\left(l_s,l_v\right)$ can thus be obtained recursively as illustrated below. For $L_s=0$ and $L_v=1$, we have:

$$E[V_{t+1}|\mathcal{F}_t]=\left.\frac{\partial\psi_t}{i\partial u_2}\right|_{u=0}=\left.\left(\frac{\partial C}{i\partial u_2}+\frac{\partial D1}{i\partial u_2}\Delta S_t+\frac{\partial D2}{i\partial u_2}V_t\right)\psi_t\right|_{u=0}$$
$$=c_{0,1}\left(0,0\right)+c_{0,1}\left(1,0\right)\Delta S_t+c_{0,1}\left(0,1\right)V_t$$

so that $c_{0,1}\left(0,0\right)=\partial C/(i\partial u_2)|_{u=0}$, $c_{0,1}\left(1,0\right)=\partial D1/(i\partial u_2)|_{u=0}$, and $c_{0,1}\left(0,1\right)=\partial D2/(i\partial u_2)|_{u=0}$. Similarly, for $L_s=1$ and $L_v=1$, we have

$$E[\Delta S_{t+1}V_{t+1}|\mathcal{F}_t]=\left.\frac{\partial^2\psi_t}{i^2\partial u_1\partial u_2}\right|_{u=0}=\left.\frac{\partial}{i\partial u_1}\left[\left(\frac{\partial C}{i\partial u_2}+\frac{\partial D1}{i\partial u_2}\Delta S_t+\frac{\partial D2}{i\partial u_2}V_t\right)\psi_t\right]\right|_{u=0}$$
$$=\left.\left(\frac{\partial^2 C}{i^2\partial u_1\partial u_2}+\frac{\partial^2 D1}{i^2\partial u_1\partial u_2}\Delta S_t+\frac{\partial^2 D2}{i^2\partial u_1\partial u_2}V_t\right)\right|_{u=0}$$
$$+\left(\frac{\partial C}{i\partial u_2}+\frac{\partial D1}{i\partial u_2}\Delta S_t+\frac{\partial D2}{i\partial u_2}V_t\right)$$
$$\times\left.\left(\frac{\partial C}{i\partial u_1}+\frac{\partial(D1-iu_1)}{i\partial u_1}\Delta S_t+\frac{\partial D2}{i\partial u_1}V_t\right)\right|_{u=0}$$
$$=c_{1,1}\left(0,0\right)+c_{1,1}\left(1,0\right)\Delta S_t+c_{1,1}\left(0,1\right)V_t+c_{1,1}\left(1,1\right)\Delta S_tV_t$$
$$+c_{1,1}\left(2,0\right)\left(\Delta S_t\right)^2+c_{1,1}\left(0,2\right)V_t^2$$

where

$$c_{1,1}\left(0,0\right)=2\left.\frac{\partial^2 C}{i^2\partial u_1\partial u_2}\right|_{u=0}$$
$$c_{1,1}\left(1,0\right)=\left.\left(\frac{\partial^2 D1}{i^2\partial u_1\partial u_2}+\frac{\partial C}{i\partial u_2}\left(\frac{\partial D1}{i\partial u_1}-1\right)+\frac{\partial C}{i\partial u_1}\frac{\partial D1}{i\partial u_2}\right)\right|_{u=0}$$
$$c_{1,1}\left(1,0\right)=\left.\left(\frac{\partial^2 D1}{i^2\partial u_1\partial u_2}+\frac{\partial C}{i\partial u_2}\frac{\partial D1}{i\partial u_1}+\frac{\partial C}{i\partial u_1}\frac{\partial D1}{i\partial u_2}\right)\right|_{u=0}$$

$$c_{1,1}(0,1) = \left(\frac{\partial^2 D2}{i^2 \partial u_1 \partial u_2} + \frac{\partial C}{i\partial u_2}\frac{\partial D2}{i\partial u_1} + \frac{\partial C}{i\partial u_1}\frac{\partial D2}{i\partial u_2} \right)\Big|_{u=0}$$

$$c_{1,1}(1,1) = \left(\frac{\partial D1}{i\partial u_2}\frac{\partial D2}{i\partial u_1} + \frac{\partial D1}{i\partial u_1}\frac{\partial D2}{i\partial u_2} \right)\Big|_{u=0}$$

$$c_{1,1}(2,0) = \frac{\partial^2 D1}{i^2 \partial u_1 \partial u_2}\Big|_{u=0}$$

$$c_{1,1}(0,2) = \frac{\partial^2 D2}{i^2 \partial u_1 \partial u_2}\Big|_{u=0}$$

The coefficients $c_{L_s,L_v}(l_s, l_v)$ for higher-order moments can be derived using a similar recursive procedure. ■

Received March 17, 2006; revised August 18, 2006; accepted September 12, 2006

REFERENCES

Aït-Sahalia, Y., and R. Kimmel. (2006). ''Maximum Likelihood Estimation of Stochastic Volatility Models.'' Forthcoming in *Journal of Financial Economics*.

Alizadeh, S., M. Brandt, and F. Diebold. (2002). ''Range-Based Estimation of Stochastic Volatility Models.'' *Journal of Finance 57*, 1047–1092.

Andersen, T. G. (1996). ''Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility.'' *Journal of Finance 51*(1), 169–204.

Andersen, T. G., and B. E. Sørensen. (1996). ''GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study.'' *Journal of Business and Economic Statistics*, *14*(3), 328–352.

Andersen, T. G., L. Benzoni, and J. Lund. (2002). ''An Empirical Investigation of Continuous-Time Equity Return Models.'' *Journal of Finance 57*(3), 1239–1284.

Andreou, E., and E. Ghysels. (2002). ''Rolling Sample Volatility Estimators: Some New Theoretical, Simulation and Empirical Results.'' *Journal of Business and Economic Statistics*, *20*, 363–376.

Bakshi, G., C. Cao, and Z. Chen. (1997). ''Empirical Performance of Alternative Option Pricing Model.'' *Journal of Finance 52*(5), 2003–2049.

Bakshi, G., and D. Madan. (2000). ''Spanning and Derivative-Security Valuation.'' *Journal of Financial Economics 55*(2), 205–238.

Bakshi, G., C. Cao, and Z. Chen. (2000). ''Pricing and Hedging Long-Term Options.'' *Journal of Econometrics 94*(1–2), 277–318.

Bandi, F. M., and J. R. Russell. (2006). ''Separating Microstructure Noise from Volatility.'' *Journal of Financial Economics 79*, 655–692.

Bates, D. S. (1996). ''Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options.'' *Review of Financial Studies 9*(1), 69–107.

Bates, D. S. (2000). ''Post-87 Crash Fears in the S&P 500 Futures Option Market.'' *Journal of Econometrics 94*(1/2), 181–238.

Bates, D. S. (2006). ''Maximum Likelihood Estimation of Latent Affine Processes.'' *Review of Financial Studies 19*(3), 909–965.

Black, F. (1976). ''Studies of Stock Price Volatility Changes.'' *Proceedings of the 1976 meetings of the American Statistical Association, Business and Economics Statistics Section*. Washington, DC: American Statistical Association, 177–181.

Bollerslev, T., and H. Zhou. (2002). ''Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility.'' *Journal of Econometrics 109*, 33–65.

Brandt, M. W., and Q. Kang. (2004). ''On the Relationship Between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach.'' *Journal of Financial Economics 72*, 217–257.

Carrasco, M., M. Chernov, J.-P. Florens, and E. Ghysels. (2002). ''Efficient Estimation of Jump Diffusions and General Dynamic Models with a Continuum of Moment Conditions.'' Forthcoming in *Journal of Econometrics*.

Chacko, G., and L. M. Viceira. (2003). ''Spectral {GMM} Estimation of Continuous-Time Processes.'' *Journal of Econometrics 116*, 259–292.

Chan, K., and W.-M. Fong. (2000). ''Trade Size, Order Imbalance, and the Volatility Volume Relation.'' *Journal of Financial Economics 57*, 247–273.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen. (2003). ''Alternative Models for Stock Price Dynamics.'' *Journal of Econometrics 116*, 225–257.

Chib, S., F. Nardari, and N. Shephard. (2002). ''Markov Chain Monte Carlo Methods for Stochastic Volatility Models.'' *Journal of Econometrics 108*, 281–316.

Christie, A. A. (1982). ''The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rates Effects.'' *Journal of Financial Economics 10*, 407–432.

Clark, P. K. (1973). ''A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices.'' *Econometrica 41*(1), 135–155.

Dai, Q., and K. J. Singleton. (2000). ''Specification Analysis of Affine Term Structure Models.'' *Journal of Finance 55*(5), 1943–1978.

Darolles, S., C. Gouriéroux, and J. Jasiak. (2001). ''Compound Autoregressive Models.'' CREST, D.P.

Das, S. R. (2002). ''The Surprise Element: Jumps in Interest Rates.'' *Journal of Econometrics 106*(1), 27–65.

Das, S. R., and R. K. Sundaram. (1999). ''Of Smiles and Smirks: A Term Structure Perspective.'' *Journal of Financial and Quantitative Analysis 34*(2), 211–239.

Duffie, D., and K. J. Singleton. (1993). ''Simulated Moments Estimation of Markov Models of Asset Prices.'' *Econometrica 61*, 929–952.

Duffie, D., and R. Kan. (1996). ''A Yield-Factor Model of Interest Rates.'' *Mathematical Finance 6*(4), 379–406.

Duffie, D., J. Pan, and K. J. Singleton. (2000). ''Transform Analysis and Asset Pricing for Affine Jump-Diffusions.'' *Econometrica 68*(6), 1343–1376.

Engle, R. F., and J. R. Russell. (1998). ''Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data.'' *Econometrica 66*(5), 1127–1162.

Engle, R., D. Lilien, and R. Robins. (1987). ''Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model.'' *Econometrica 55*(2), 391–407.

Eraker, B. (2001). ''MCMC Analysis of Diffusion Models with Applications to Finance.'' *Journal of Business and Economic Statistics 19*, 177–191.

Eraker, B., M. Johannes, and N. Polson. (2003). ''The Impact of Jumps in Equity Index Volatility and Returns.'' *Journal of Finance 58*, 1269–1300.

Ethier, S., and T. Kurtz. (1986). *Markov Processes, Characterization and Convergence*. New York: John Wiley and Sons.

Foster, D. P., and D. B. Nelson. (1996). ''Continuous Record Asymptotics for Rolling Sample Variance Estimators.'' *Econometrica 64*(1), 139–174.

French, K. R., G. W. Schwert, and R. Stambaugh. (1987). ''Expected Stock Returns and Volatility.'' *Journal of Financial Economics 19*, 3–29.

Gallant, A. R., and G. E. Tauchen. (1996). ''Which Moments to Match?'' *Econometric Theory 12*, 657–681.

Gallant, A. R., P. E. Rossi, and G. Tauchen. (1992). ''Stock Prices and Volume.'' *Review of Financial Studies 5*, 199–242.

Ghysels, E., P. Santa-Clara, and R. Valkanov. (2005). ''There is a Risk-Return Tradeoff After All.'' *Journal of Financial Economics 76*, 509–548.

Gouriéroux, C., A. Monfort, and E. Renault. (1993). ''Indirect Inference.'' *Journal of Applied Econometrics 8*, S85–S199.

Hansen, L. P. (1982). ''Large Sample Properties of Generalized Method of Moments Estimator.'' *Econometrica 50*(4), 1029–1054.

Hansen, L. P. (1985). ''A Method for Calculating Bounds on Asymptotic Covariance Matrices of Generalized Method of Moments Estimators.'' *Journal of Econometrics 30*, 203–238.

Hansen, L. P., and K. J. Singleton. (1991). ''Computing Semiparametric Efficiency Bounds for Linear Time Series Models.'' In W. A. Barnett, J. Powell, and G. E. Tauchen (eds.), *Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*. Cambridge, UK: Cambridge University Press, 388–411, Chapter 15.

Hansen, L., and K. Singleton. (1996). ''Efficient Estimation of Linear Asset Pricing Models with Moving Average Errors.'' *Journal of Business and Economic Statistics 114*, 53–68.

Hansen, L. P., J. C. Heaton, and M. Ogaki. (1988). ''Efficiency Bounds Implied by Multi-Period Conditional Moment Restrictions.'' *Journal of the American Statistical Association 83*, 863–871.

Harris, M., and A. Raviv. (1993). ''Differences of Opinion Make a Horse Race.'' *Review of Financial Studies 6*(3), 473–506.

Harvey, A. C., and N. G. Shephard. (1996). ''Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns.'' *Journal of Business and Economic Statistics 14*, 429–434.

He, H., and J. Wang. (1995). ''Differential Information and Dynamic Behavior of Stock Trading Volume.'' *Review of Financial Studies 8*(4), 919–972.

Heston, S. L. (1993). ''A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options.'' *Review of Financial Studies 6*, 327–344.

Jacquier, E., N. G. Polson, and P. Rossi. (1994). ''Bayesian Analysis of Stochastic Volatility Models.'' *Journal of Business and Economic Statistics 12*, 371–417, (with discussion).

Jiang, G. J., and J. L. Knight. (2002). ''Estimation of Continuous-Time Processes Via the Empirical Characteristic Function.'' *Journal of Business and Economic Statistics 20*(2), 198–212.

Jiang, G. J., and R. C. Oomen. (2006). ''Filtering and Smoothing Spot and Integrated Variance for the Affine Jump Diffusion Model.'' Manuscript University of Arizona, Department of Finance.

Johannes, M. (2004). ''The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models.'' *Journal of Finance 59*(1), 227–260.

Karpoff, J. (1986). ''A Theory of Trading Volume.'' *Journal of Finance 61*, 1069–1085.

Kendall, M., and A. Stuart. (1977). *The Advanced Theory of Statistics*, Vol. 1. London: Charles Griffin and Company Limited.

Lamoureux, C. G., and W. D. Lastrapes. (1990). ''Heteroskedasticity in Stock Return Data: Volume Versus {GARCH} Effects.'' *Journal of Finance*, *45*(1), 221–229.

Lo, A. W. (1988). ''Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data.'' *Econometric Theory*, *4*, 231–247.

Meddahi, N. (2001). ''An Eigenfunction Approach for Volatility Modeling.'' Working paper CRDE 29, University of Montreal.

Meddahi, N. (2002). ''Moments of Continuous Time Stochastic Volatility Models.'' Manuscript University of Montreal.

Meddahi, N., and E. Renault. (2004). ''Temporal Aggregation of Volatility Models.'' *Journal of Econometrics*, *119*(2), 355–379.

Melino, A., and S. M. Turnbull. (1990). ''Pricing Foreign Currency Options with Stochastic Volatility.'' *Journal of Econometrics*, *45*, 239–265.

Merton, R. C. (1980). ''On Estimating the Expected Return on the Market; An Exploratory Investigation.'' *Journal of Financial Economics 8*, 323–361.

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., V. Patilea, and E. Renault. (2003). ''Iterative and Recursive Estimation in Structural Nonadaptive Models.'' *Journal of Business and Economic Statistics 21*(4), 449–509, with discussion.

Ross, S. A. (1989). ''Information and Volatility: The No-Arbitrage Martingale Approach to Timing and Resolution Irrelevance.'' *Journal of Finance 44*, 1–17.

Ruiz, E. (1994). ''Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models.'' *Journal of Econometrics 63*, 289–306.

Scott, L. (1997). ''Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods.'' *Mathematical Finance 7*(4), 413–426.

Singleton, K. J. (2001). ''Estimation of Affine Asset Pricing Models using the Empirical Characteristic Function.'' *Journal of Econometrics 102*, 111–141.

Tauchen, G. E. (1985). ''Diagnostic Testing and Evaluation of Maximum Likelihood Models.'' *Journal of Econometrics 30*, 415–443.

Tauchen, G., and M. Pitts. (1983). ''The Price Variability-Volume Relationship on Speculative Markets.'' *Econometrica 51*, 485–505.

Tauchen, G., H. Zhang, and M. Liu. (1996). ''Volume, Volatility, and Leverage: A Dynamic Analysis.'' *Journal of Econometrics 74*, 177–208.