A Bayesian Analysis of Return Dynamics with Stochastic Volatility and Lévy Jumps

Haitao Li,[1] Matin T. Wells,[2] and Cindy L. Yu[3]

January 2006

[1]Li is from the Stephen M. Ross School of Business, University of Michigan, Ann Arbor, MI 48109; E-mail: htli@umich.edu; phone: (734)-764-6409. [2]Wells is from the Department of Biological Statistics and Computational Biology and the Department of Social Statistics, Cornell University, Ithaca, NY 14853; E-mail: mtw1@cornell.edu; phone: (607)-255-4388; he gratefully acknowledges the support of NSF Grant DMS 02-04252. [3]Yu is from the Department of Statistics, Iowa State University, Ames, IA 50011; E-mail: cindyyu@iastate.edu; phone: (515)-294-3319. We would like to thank Torben Andersen and Luca Benzoni for their help in the early stage of this study. We thank Yacine Aït-Sahalia, Antje Berndt, Peter Carr, Bjorn Eraker, Bob Jarrow, Nour Meddahi, Ray Renken, Sidney Resnick, Ernst Schaumburg, Neil Shephard, George Tauchen, Liuren Wu, and seminar participants at Cornell University, the 2004 CIREQ/CIRANO financial econometrics conference, the 2004 International Chinese Statistical Association Applied Statistics Symposium, and the 2004 Institute of Mathematical Statistics Annual Meeting/6th Bernoulli World Congress for helpful comments. We are responsible for any remaining errors.

A Bayesian Analysis of Return Dynamics with Stochastic Volatility and Lévy Jumps

ABSTRACT

We develop Bayesian Markov chain Monte Carlo methods for inferences of continuous-time models with stochastic volatility and infinite-activity Lévy jumps using discretely sampled data. Simulation studies show that (i) our methods provide accurate joint identification of diffusion, stochastic volatility, and Lévy jumps, and (ii) affine jump-diffusion models fail to adequately approximate the behavior of infinite-activity jumps. In particular, the affine jump-diffusion models fail to capture the "infinitely many" small Lévy jumps which are too big for Brownian motion to model and too small for compound Poisson process to capture. Empirical studies show that infinite-activity Lévy jumps are essential for modeling the S&P 500 index returns.

The continuous-time finance literature in the past few decades has mainly relied on Brownian motion and compound Poisson process as basic model building blocks. Sophisticated models based solely on Brownian motion and compound Poisson process have been developed to capture important stylized behaviors of asset prices. One prominent example of such models is the popular affine jump-diffusion (hereafter AJD) models of Duffie, Pan, and Singleton (2000), in which affine diffusions capture continuous movements in asset prices and compound Poisson processes capture large, discontinuous jumps in asset prices.

Brownian motion and compound Poisson process, however, are only two special cases of a much broader class of stochastic processes, namely Lévy processes. Roughly speaking, a Lévy process is a continuous-time stochastic process with stationary and independent increments. Other than Brownian motion and compound Poisson process, there are many other members of the Lévy family that offer greater flexibility in modeling asset price dynamics. Specifically, the jump component of a general Lévy process is much more flexible than a compound Poisson process. For example, the so-called infinite-activity Lévy jumps have infinite jump arrival rates and can generate, in addition to large jumps, an infinite number of small jumps within any finite time interval.

Lévy processes have become increasingly popular in recent years, and various Lévy models have been developed in the asset pricing literature.[1] The important work of Carr and Wu (2004) has established Lévy processes as an attractive alternative to AJDs for modeling asset price dynamics. For example, they show that one can easily incorporate stochastic volatility by applying a stochastic time change to Lévy processes and obtain closed-form formulas (based on the characteristic function approach) for a wide variety of derivative securities in Lévy models. Therefore, Lévy processes are

[1] These include the inverse Gaussian model of Barndorff-Nielsen (1998); the generalized hyperbolic class of Eberlein, Keller, and Prause (1998); the variance-gamma model of Madan, Carr, and Chang (1998); the generalization of VG in Carr, Geman, Madan, and Yor (2002); and the finite moment log-stable model of Carr and Wu (2003) among others.

more flexible than AJDs for modeling purpose and are as tractable as AJDs for pricing purpose.

The recent developments of Lévy models, however, have raised some challenging theoretical and empirical issues in the literature.[2] The double-jump model of Eraker, Johannes, and Polson (2003) (hereafter EJP), one of the most sophisticated AJD models for stock returns, includes stochastic volatility, leverage effect, and compound Poisson jumps in both returns and volatility, and can capture many important stylized behaviors of major U.S. stock market indices. Therefore, it is not clear whether infinite-activity Lévy jumps, despite their *theoretical* appeals, have any significant *practical* advantages over such flexible AJD models. Our paper demonstrates the advantages of Lévy jump models by studying the following three closely related questions through numerical simulation and empirical analysis.

The first question we address is whether econometrically it is possible to jointly identify diffusion, stochastic volatility, and infinite-activity jumps using discretely sampled stock prices. This question is key to econometric analysis of Lévy processes and AJD models, because without the joint identification property it is not even possible to empirically compare the two classes of models. Unfortunately, estimation of stochastic volatility models with Lévy jumps is rather complicated. The probability densities of most Lévy processes are not known in closed form and for certain processes not all moments exist. It is also computationally challenging to deal with the high-dimensional latent volatility variables. We develop Bayesian Markov chain Monte Carlo (hereafter MCMC) methods that overcome these difficulties. Simulation evidence suggests that our MCMC methods provide accurate joint identification of diffusion, stochastic volatility, and Lévy jumps using discretely sampled data. Therefore, our methodology makes it practically feasible to examine the incremental contributions of Lévy jumps over existing AJD models.

The second question we study is whether AJD models can adequately approximate the behavior

[2]For example, Aït-Sahalia and Jacod (2004) note that although Lévy processes have been increasingly used in the literature, "relatively little is known about the(ir) inference problem, which is a difficult one."

of infinite-activity Lévy jumps through numerical simulation. We estimate AJD models using our MCMC methods based on data simulated from infinite-activity Lévy jump models. We find that in all simulations there are too many small movements in asset prices that cannot be captured by such fitted AJD models. Our simulation results show that the "infinitely many" small Lévy jumps fill a gap in AJD models by capturing price movements that are too large for Brownian motion to model but too small for compound Poisson process to capture. Our results also suggest that the perception that infinite-activity jumps should not be too much different from compound Poisson processes with very high arrival rates does not seem to be appropriate. Instead, we find that the above unique feature of infinite-activity Lévy jumps cannot be captured by even the most flexible AJD models.

The third question we investigate is whether Lévy jump models can outperform AJD models in capturing the time series dynamics of the S&P 500 index returns. We estimate both AJD and Lévy jump models using our MCMC methods based on daily returns of the S&P 500 index for the past twenty years. We find similar patterns of misspecifications of AJD models using actual data as that in our simulation studies. Even the most sophisticated AJD models cannot completely capture the many small movements in index returns. On the other hand, stochastic volatility models with Lévy jumps can capture the index returns very well.

Through our analyses of the three important questions, our paper has made methodological, numerical, and empirical contributions to the fast-growing literature on econometric analysis of Lévy processes. Although MCMC methods have been widely used in the existing literature, our paper is one of the first that deals with infinite-activity and infinite-variation Lévy jumps, and hence will be helpful for future developments of the literature.[3] Our simulation results complement the theoretical

[3]Earlier studies, such as Jacquier, Polson, and Rossi (1994); Kim, Shephard, and Chib (1998); and Chib, Nardari, and Shephard (2003), apply MCMC methods to estimate discrete-time stochastic volatility models. Other studies that apply MCMC methods to continuous-time models for stock price or interest rate include Jones (1998, 2003a, b); Elerian, Chib, and Shephard (2001); Eraker (2001, 2004); and EJP (2003), among others.

works of Aït-Sahalia (2004) and Aït-Sahalia and Jacod (2004). Aït-Sahalia (2004) shows through theoretical analysis that maximum likelihood method can perfectly identify diffusion variance in the presence of infinite-activity Cauchy jumps. Aït-Sahalia and Jacod (2004) further show that jump components with different power coefficients also can be separately identified as long as the power coefficients are sufficiently different. Our simulation results extend their findings by showing that the joint and accurate identifications of diffusion, stochastic volatility, and infinite-activity jumps are feasible when the underlying price dynamics approximate the time-series behavior of index returns. Our empirical results based on S&P 500 index returns also complement existing evidence that favors Lévy jump models over AJD models based on option data.[4] Overall, our results strongly suggest that the infinite-activity Lévy jumps are not only theoretically more appealing than compound Poisson jumps but also are empirically important for capturing financial data. Therefore, Lévy jumps can enrich existing AJD models by capturing certain important features of the data that are difficult for AJD models to capture.

This paper is organized as follows. In Section 1, we introduce pure Lévy jump processes and several stochastic volatility models with Lévy jumps for stock return dynamics. In Section 2, we develop MCMC methods for estimating model parameters and latent volatility and jump variables of the Lévy jump models. Section 3 presents simulation evidence on the performance of our MCMC methods in estimating Lévy jump models. Section 4 reports empirical results for the S&P 500 index returns. Section 5 concludes the paper with some final remarks. The Appendix provides detailed discussions of the implementations of the MCMC methods.

1. Return Dynamics with Lévy Jumps

1.1 Lévy Processes

Suppose the uncertainty of the economy is described by a probability space $(\Omega, \mathcal{F}, P)$ and a

[4]Existing empirical studies of Levy processes have mainly focused on option data. See, for example, Madan, Carr, and Chang (1998); Carr, Geman, Madan, and Yor (2002); Huang and Wu (2003); Carr and Wu (2003) among others.

filtration $\{\mathcal{F}_t\}$. Let X_t be a scalar Lévy process with respect to the filtration $\{\mathcal{F}_t\}$. Then X_t is adapted to $\mathcal{F}_t$, the sample paths of X_t are right-continuous with left limits, and $X_s - X_t$ is independent of $\mathcal{F}_t$ and distributed as X_{s-t} for $0 \leq t < s$. Lévy processes include Brownian motion and compound Poisson process as special cases. But they are much more general than the two because they allow discontinuous sample paths, non-normal increments, and more flexible jump structures that have (possibly) infinite arrival rates. Lévy processes are also semimartingales and are therefore suitable for modeling asset prices. The independence property of Lévy increments further guarantee that the underlying process is Markov and thus is consistent with the Efficient Market Hypothesis.

Although the probability densities of Lévy processes are generally not known in closed form, their characteristic functions $\phi_{X_t}(u)$ can be explicitly specified as follows,

$$\phi_{X_t}(u) = E\left[e^{iuX_t}\right] = e^{-t\psi_x(u)},\ t \geq 0,$$

where $\psi_x(u)$ is called the characteristic exponent and satisfies the following Lévy-Khintchine formula (see Bertoin, 1996, p. 12)

$$\psi_x(u) \equiv -i\mu u + \frac{\sigma^2 u^2}{2} + \int_{\mathbf{R}_0} \left(1 - e^{iux} + iux1_{|x|<1}\right) \pi(dx),$$

$u \in \mathbf{R}, \mu \in \mathbf{R}$, $\sigma \in \mathbf{R}^+$, and π is a measure on $\mathbf{R}_0 = \mathbf{R} \backslash \{0\}$ ($\mathbf{R}$ less zero) with

$$\int_{\mathbf{R}_0} \min\left(1, x^2\right) \pi(dx) < \infty.$$

The Lévy-Khintchine formula suggests that a Lévy process consists of three independent components: a linear deterministic drift part, a Brownian part, and a pure jump part. The triplet $\left(\mu, \sigma^2, \pi(\cdot)\right)$, often called the characteristics of the Lévy process, completely describe the probabilistic behavior of the process. The Lévy measure $\pi(dx)$ dictates the jump behavior of the process. It has the interpretation that $\pi(E)$, for any subset $E \subset \mathbf{R}$, is the rate at which the process takes jumps of size $x \in E$. In other words, $\pi(E)$ measures the number of jumps whose jump sizes falling in E per unit of time.

Depending on its Lévy measure $\pi(\cdot)$, a pure jump Lévy process can exhibit rich jump behaviors. For finite-activity jump processes, which have a finite number of jumps within any finite time interval, π needs to be integrable, that is,

$$\int_{\mathbf{R}_0} \pi(dx) = \lambda < \infty. \tag{1}$$

The classical example of a finite-activity jump process is the compound Poisson jump process of Merton (1976). In the Merton Jump (MJ) model, the integral in (1) defines the Poisson arrival intensity λ. Conditional on one jump occurring, the MJ model assumes that the jump magnitude is normally distributed with mean μ_y and variance σ_y^2. The Lévy measure of the MJ model is given by

$$\pi_{MJ}(dx) = \lambda \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x-\mu_y)^2}{2\sigma_y^2}\right) dx.$$

Obviously, one can choose any distribution, $F(x)$, for the jump size under the compound Poisson framework and obtain the Lévy measure $\pi(dx) = \lambda dF(x)$.

Unlike finite-activity jump processes, an infinite-activity jump process allows an (possibly) infinite number of jumps within any finite time interval. The integral of the Lévy measure in (1) is no longer finite. Within the infinite-activity category, the sample path of the jump process can exhibit either finite or infinite variation, meaning that the aggregate absolute distance traveled by the process is finite or infinite, respectively, over any finite time interval.

In our empirical analysis, we choose the relatively parsimonious variance-gamma (hereafter VG) model of Madan, Carr, and Chang (1988) as a representative of the infinite-activity but finite variation jump model. The VG process is obtained by subordinating an arithmetic Brownian motion with drift γ and variance σ by an independent gamma process with unit mean rate and variance rate ν, G_t^ν. That is,

$$X_{VG}(t|\sigma,\gamma,\nu) = \gamma G_t^\nu + \sigma W(G_t^\nu)$$

where $W(t)$ is a standard Brownian motion and is independent of G_t^ν. The Lévy measure of the VG

process is given by

$$\pi_{VG}(dx) = \frac{A_{\pm}^2 \exp\left(-\frac{A_{\pm}}{B_{\pm}}|x|\right)}{B_{\pm}|x|}(dx),$$

where $A_{\pm} = \frac{1}{\nu}\left(\sqrt{\frac{\gamma^2\nu^2}{4} + \frac{\sigma^2\nu}{2}} \pm \frac{\gamma\nu}{2}\right)$ and $B_{\pm} = A_{\pm}^2\nu$. The parameters with plus subscripts apply to positive jumps and those with minus subscripts apply to negative jumps. If $\gamma = 0$, then the jump structure is symmetric around zero and the subscripts are dropped. Note that as the jump size approaches zero, the arrival rate approaches infinity. Thus, an infinite-activity model incorporates (possibly) infinitely many small jumps. The Lévy measure of an infinite-activity jump process is singular at a zero jump size.

Another example of infinite-activity jump model is the Lévy α-stable process. In this process, jump sizes follow an α-stable distribution denoted as $S_\alpha(\beta, \delta, \gamma)$, with a tail index $\alpha \in (0, 2]$, a skew parameter $\beta \in [-1, 1]$, a scale parameter $\delta \geq 0$, and a location parameter $\gamma \in \mathbf{R}$. The parameter α determines the shape of the distribution, while β determines the skewness of the distribution. Stable densities are supported on either $\mathbf{R}$ or $\mathbf{R}^+$. The latter situation occurs only when $\alpha < 1$ and $\beta = \pm 1$. The characteristic function of an α-stable distribution S is given by

$$E\left[e^{iuS}\right] = \begin{cases} \exp\left(-\delta^\alpha |u|^\alpha \left[1 - i\beta\left(\tan\frac{\pi\alpha}{2}\right)(\text{sign } u)\right] + i\gamma u\right) & \alpha \neq 1 \\ \exp\left(-\delta|u|\left[1 + i\beta\frac{2}{\pi}(\text{sign } u)\ln|u|\right] + i\gamma u\right) & \alpha = 1. \end{cases}$$

For a standardized α-stable distribution, denoted as $S_\alpha(\beta, 1, 0)$, $\delta = 1$ and $\gamma = 0$.

All α-stable processes are built upon a fundamental process called α-stable motion. A process X_t is an α-stable motion if (i) $X_0 = 0$ a.s., (ii) X_t has independent increments, and (iii) the increment $X_t - X_s$ $(t > s)$ follows an α-stable distribution $S_\alpha\left(\beta, (t-s)^{\frac{1}{\alpha}}, 0\right)$. The role that α-stable motion plays for α-stable processes is similar to that of Brownian motion for diffusion processes. Among α-stable processes, we choose the finite moment log-stable (hereafter LS) process of Carr and Wu (2003) in our analysis. We obtain this process by multiplying an α-stable motion by a constant σ. Following Carr and Wu (2003), we set $\beta = -1$ to achieve finite moments for index levels and negative

skewness in the return density, a feature that cannot be captured by either a Brownian motion or a symmetric Lévy α-stable motion. We also restrict $\alpha \in (1,2)$ so that the process has the support of the whole real line. The α-stable process defined in this way is a Lévy process with infinite activity and infinite variation and has a Lévy measure

$$\pi_{LS}(dx) = c_{\pm} |x|^{-\alpha-1} dx,$$

where $c_{-} = \frac{-\sigma^{\alpha} \sec \frac{\pi\alpha}{2}}{\Gamma(-\alpha)}$. The parameters $c_{\pm}$ control both the scale and the asymmetry of the process. In the LS model, c_{+} is set to zero so that only negative jumps are allowed in the Lévy measure. However, it is important to point out that in addition to the pure jump part characterized by the Lévy measure $\pi_{LS}(dx)$, the LS process also has a deterministic drift part that compensates the negative jumps so that the whole process is a martingale. For infinite-variation jumps, the compensation is so much that the admissible domain of LS actually covers the whole real line, although there are only negative jumps. As a result, the LS process has an α-stable distribution with infinite p-th moment for $p > \alpha$.

1.2 Return Dynamics with Stochastic Volatility and Lévy Jumps

To examine the incremental contributions of Lévy jumps, we consider stochastic volatility models with either VG or LS jumps in returns. For comparison, we also consider some of the popular AJD models for return dynamics in the existing literature.

Let Y_t represent the logarithm of stock price, i.e., $Y_t = \log(S_t)$. Then all models considered in our paper can be summarized by the following system of stochastic differential equations:

$$\begin{pmatrix} dY_t \\ dv_t \end{pmatrix} = \begin{pmatrix} \mu \\ \kappa(\theta - v_t) \end{pmatrix} dt + \sqrt{v_{t-}} \begin{pmatrix} 1 & 0 \\ \rho\sigma_v & \sqrt{(1-\rho^2)}\sigma_v \end{pmatrix} dW_t + \begin{pmatrix} dJ_t^y \\ dJ_t^v \end{pmatrix}, \tag{2}$$

where W_t is a standard Brownian motion in $\mathbf{R}^2$, J_t^y and J_t^v represent jumps in returns and volatility, respectively, μ measures the mean return, v_t the instantaneous volatility of return, θ the long-run mean of stochastic volatility, κ the speed of mean reversion of volatility, σ_v the so-called volatility of

volatility variable, and ρ the correlation between returns and instantaneous volatility.

The above model nests some of the most important AJD models for equity returns. For example, without jumps in returns and volatility, the above model reduces to the stochastic volatility model of Heston (1993). With Merton jumps only in returns, we have the stochastic volatility and Merton jump model of Bates (1996) and Andersen, Benzoni, and Lund (2002). With Merton jumps in both returns and volatility, we have the double-jump models of EJP (2003). In our analysis, we allow VG and LS jumps in returns in the above model.

To simplify our empirical analysis, we consider the first-order Euler discretized version of the above continuous-time model at daily frequency. Simulation studies in EJP (2003) show that the discretization bias of daily data is not significant. In total, we consider the following four discretized models in our empirical analysis.

- *Stochastic Volatility Model with Merton Jumps in Returns* (**SVMJ**)

The discretized version of the SVMJ model, which allows Merton jumps in returns, is

$$\begin{cases} Y_{t+1} = Y_t + \mu\Delta + \sqrt{v_t\Delta}\epsilon^y_{t+1} + J^y_{t+1}, \\ v_{t+1} = v_t + \kappa(\theta - v_t)\Delta + \sigma_v\sqrt{v_t\Delta}\epsilon^v_{t+1}, \end{cases} \tag{3}$$

where both ϵ^y_{t+1} and ϵ^v_{t+1} follow $N(0,1)$ with $corr(\epsilon^y_{t+1}, \epsilon^v_{t+1}) = \rho$; $J^y_{t+1} = \xi^y_{t+1}N^y_{t+1}, P(N^y_{t+1} = 1) = \lambda_y\Delta$, $\xi^y_{t+1} \sim N(\mu_y, \sigma^2_y)$ and is independent of ϵ^y_{t+1} and ϵ^v_{t+1}. For this process, we have observations $(Y_t)^T_{t=0}$; latent volatility variables $(v_t)^T_{t=0}$, jump times $(N^y_t)^T_{t=1}$, and jump sizes $(\xi^y_t)^T_{t=1}$; and model parameters $\mathbf{\Theta} = \{\mu, \kappa, \theta, \sigma_v, \rho, \lambda_y, \mu_y, \sigma_y\}$.

- *Stochastic Volatility Model with Correlated Merton Jumps in Returns and Volatility* (**SVCMJ**)

The discretized version of the SVCMJ model (the preferred model of EJP (2003)) allows correlated

compound Poisson jumps in returns and volatility, is

$$\begin{cases} Y_{t+1} = Y_t + \mu\Delta + \sqrt{v_t\Delta}\epsilon^y_{t+1} + J^y_{t+1}, \\ v_{t+1} = v_t + \kappa(\theta - v_t)\Delta + \sigma_v\sqrt{v_t\Delta}\epsilon^v_{t+1} + J^v_{t+1}, \end{cases} \tag{4}$$

where both ϵ^y_{t+1} and ϵ^v_{t+1} follow $N(0,1)$ with $corr(\epsilon^y_{t+1}, \epsilon^v_{t+1}) = \rho$; $J^y_{t+1} = \xi^y_{t+1}N_{t+1}$, $J^v_{t+1} = \xi^v_{t+1}N_{t+1}$, $P(N_{t+1} = 1) = \lambda\Delta$, $\xi^v_{t+1} \sim exp(\mu_v)$, $\xi^y_{t+1}|\xi^v_{t+1} \sim N(\mu_y + \rho_J\xi^v_{t+1}, \sigma_y^2)$. $(\xi^y_{t+1}, \xi^v_{t+1})$ is independent of ϵ^y_{t+1} and ϵ^v_{t+1}. For this process, we have observations $(Y_t)_{t=0}^T$; latent volatility variables $(v_t)_{t=0}^T$, jump times $(N_t)_{t=1}^T$, and jump sizes $(\xi^y_t)_{t=1}^T$ and $(\xi^v_t)_{t=1}^T$; and model parameters, $\mathbf{\Theta} = \{\mu, \kappa, \theta, \sigma_v, \rho, \lambda, \mu_v, \mu_y, \sigma_y, \rho_J\}$. Note that SVMJ is nested within SVCMJ.

- *Stochastic Volatility Model with Variance-Gamma Jumps in Returns* (**SVVG**)

In the SVVG model, we allow jumps in returns to follow a VG process. The discretized version of the SVVG model is

$$\begin{cases} Y_{t+1} = Y_t + \mu\Delta + \sqrt{v_t\Delta}\epsilon^y_{t+1} + J^y_{t+1}, \\ v_{t+1} = v_t + \kappa(\theta - v_t)\Delta + \sigma_v\sqrt{v_t\Delta}\epsilon^v_{t+1}, \end{cases} \tag{5}$$

where both ϵ^y_{t+1} and ϵ^v_{t+1} follow $N(0,1)$ with $corr(\epsilon^y_{t+1}, \epsilon^v_{t+1}) = \rho$. The jump process follows a VG process whose discretized version, J^y_{t+1}, is

$$J^y_{t+1} = \gamma G_{t+1} + \sigma\sqrt{G_{t+1}}\epsilon^J_{t+1},$$

where $\epsilon^J_{t+1} \sim N(0,1)$ and $G_{t+1} \sim \Gamma(\frac{\Delta}{\nu}, \nu)$. ϵ^J_{t+1} and G_{t+1} are independent of each other and are independent of ϵ^y_{t+1} and ϵ^v_{t+1}. The parameterization of the Gamma distribution, $\Gamma(\alpha, \beta)$, used in this paper has density form $\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{x/\beta}$. For this process, we have observations $(Y_t)_{t=0}^T$; latent volatility variables $(v_t)_{t=0}^T$, jump times/sizes $(J^y_t)_{t=1}^T$, and time-change variables $(G_t)_{t=1}^T$; and model parameters $\mathbf{\Theta} = \{\mu, \kappa, \theta, \sigma_v, \rho, \gamma, \sigma, \nu\}$.

- *Stochastic Volatility Model with Log-Stable Jumps in Returns* (**SVLS**)

In the SVLS model, we allow jumps in returns to follow a finite moment log-stable process. The discretized version of the SVLS model is

$$\begin{cases} Y_{t+1} = Y_t + \mu\Delta + \sqrt{v_t\Delta}\epsilon^y_{t+1} + J^y_{t+1}, \\ v_{t+1} = v_t + \kappa(\theta - v_t)\Delta + \sigma_v\sqrt{v_t\Delta}\epsilon^v_{t+1}, \end{cases} \tag{6}$$

where both ϵ^y_{t+1} and ϵ^v_{t+1} follow $N(0,1)$ with $corr(\epsilon^y_{t+1}, \epsilon^v_{t+1}) = \rho$. The jump size J^y_{t+1}, independent of ϵ^y_{t+1} and ϵ^v_{t+1}, follows a stable distribution with shape parameter α, skewness parameter -1, zero drift, and scale parameter $\sigma\Delta^{\frac{1}{\alpha}}$. That is, $J^y_{t+1} \sim S_\alpha(-1, \sigma\Delta^{\frac{1}{\alpha}}, 0)$. For this process, we have observations $(Y_t)_{t=0}^T$; latent volatility variables $(v_t)_{t=0}^T$, and jump times/sizes $(J^y_t)_{t=1}^T$; and model parameters $\mathbf{\Theta} = \{\mu, \kappa, \theta, \sigma_v, \rho, \alpha, \sigma\}$.

2. Estimating Lévy Jump Models via Markov Chain Monte Carlo

We face several challenges in estimating the above models. Unlike Cauchy process, the likelihood functions of VG and LS are not known in closed-form. For VG, the difficulty is that we need to integrate out the latent time-change variable $(G_t)_{t=1}^T$ to obtain the likelihood based only on observables. For LS, the difficulty is that the density of an α-stable distribution is unknown. Moreover, for LS, certain moments of asset returns do not even exist, which renders moment-based methods inapplicable. It is also computationally challenging to integrate out other high-dimensional latent variables, such as stochastic volatility, jump sizes, and jump times when implementing either likelihood or moment-based approaches. To overcome these difficulties, we develop a computational Bayesian MCMC approach for estimating stochastic volatility models with Lévy jumps.

MCMC conducts inferences by simulating efficiently from (potentially complicated) posterior distributions of model parameters and latent variables given the observed prices. MCMC samples from the typically high-dimensional and complex posterior distribution by generating a Markov Chain over parameters and latent variables whose equilibrium distribution is the desired posterior distribution. The Monte Carlo method uses these samples for numerical integration for parameter and state estimation. For a detailed discussion of MCMC, see Johannes and Polson (2003).

Since MCMC analysis of SVMJ and SVCMJ has been considered in previous studies, such as EJP (2003), we focus our discussions of MCMC methods on SVVG and SVLS. We mainly discuss how to derive the joint posterior distributions of model parameters and latent variables for the two models and briefly explain how to obtain posterior samples for individual parameters and latent variables by simulating from the complicated joint posterior distributions. More detailed discussions of our MCMC methods are provided in the appendix.

In SVVG, conditioning on v_t and J_{t+1}, $Y_{t+1}-Y_t$ and $v_{t+1}-v_t$ follow a bivariate normal distribution

$$\begin{aligned}\begin{pmatrix} Y_{t+1}-Y_t \\ v_{t+1}-v_t \end{pmatrix} | v_t, J_{t+1} &\sim N\left(\begin{pmatrix} \mu\Delta + J_{t+1} \\ \kappa(\theta - v_t)\Delta \end{pmatrix}, v_t\Delta \begin{pmatrix} 1 & \rho\sigma_v \\ \rho\sigma_v & \sigma_v^2 \end{pmatrix}\right), \\ J_{t+1}|G_{t+1}, \mathbf{\Theta} &\sim N(\gamma G_{t+1}, \sigma^2 G_{t+1}) \text{ and } G_{t+1}|\mathbf{\Theta} \sim \Gamma(\frac{\Delta}{\nu}, \nu).\end{aligned}$$

To simplify notation, we denote the index returns as $\mathbf{Y} = \{Y_t\}_{t=0}^T$, the volatility variables as $\mathbf{V} = \{v_t\}_{t=0}^T$, the jump times/sizes as $\mathbf{J} = \{J_t\}_{t=1}^T$, and the time-change variables as $\mathbf{G} = \{G_t\}_{t=1}^T$. The joint posterior distribution of parameters and latent variables, $p(\mathbf{\Theta}, \mathbf{V}, \mathbf{J}, \mathbf{G}|\mathbf{Y})$, can be decomposed into products of individual conditionals

$$\begin{aligned} p(\mathbf{\Theta}, \mathbf{V}, \mathbf{J}, \mathbf{G}|\mathbf{Y}) &\propto p(\mathbf{Y}, \mathbf{V}, \mathbf{J}, \mathbf{G}, \mathbf{\Theta}) = p(\mathbf{Y}, \mathbf{V}|\mathbf{J}) p(\mathbf{J}|\mathbf{G}, \mathbf{\Theta}) p(\mathbf{G}|\mathbf{\Theta}) p(\mathbf{\Theta}) \\ &\propto \prod_{t=0}^{T-1} \frac{1}{\sigma_v v_t \Delta \sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left((\epsilon_{t+1}^y)^2 - 2\rho\epsilon_{t+1}^y\epsilon_{t+1}^v + (\epsilon_{t+1}^v)^2\right)\right\} \\ &\times \prod_{t=0}^{T-1} \frac{1}{\sigma\sqrt{G_{t+1}}} \exp\left\{-\frac{(J_{t+1}-\gamma G_{t+1})^2}{2\sigma^2 G_{t+1}}\right\} \times \prod_{t=0}^{T-1} \frac{1}{\nu^{\frac{\Delta}{\nu}}\Gamma(\frac{\Delta}{\nu})} G_{t+1}^{\frac{\Delta}{\nu}-1} \exp\{-\frac{G_{t+1}}{\nu}\} \times p(\mathbf{\Theta}), \end{aligned}$$

where $\epsilon_{t+1}^y = (Y_{t+1} - Y_t - \mu\Delta - J_{t+1}) / \sqrt{v_t\Delta}$ and $\epsilon_{t+1}^v = (v_{t+1} - v_t - \kappa(\theta - v_t)\Delta) / (\sigma_v\sqrt{v_t\Delta})$.

In SVLS, conditioning on v_t and S_{t+1}, $Y_{t+1}-Y_t$ and $v_{t+1}-v_t$ follow a bivariate normal distribution

$$\begin{aligned}\begin{pmatrix} Y_{t+1}-Y_t \\ v_{t+1}-v_t \end{pmatrix} | v_t, S_{t+1} &\sim N\left(\begin{pmatrix} \mu\Delta + S_{t+1} \\ \kappa(\theta - v_t)\Delta \end{pmatrix}, v_t\Delta \begin{pmatrix} 1 & \rho\sigma_v \\ \rho\sigma_v & \sigma_v^2 \end{pmatrix}\right), \\ S_{t+1} &\sim S_\alpha(-1, \sigma\Delta^{\frac{1}{\alpha}}, 0).\end{aligned}$$

In SVLS, we model jumps using stable process which can exhibit skewness and heavier tails than normal distributions. Unfortunately, the probability density of S_{t+1}, $p\left(S_{t+1}|\mathbf{\Theta}\right)$, is unknown. This makes it difficult to explicitly write down the joint likelihood function of $(Y_{t+1}, v_{t+1}, S_{t+1})$, because $p\left(Y_{t+1}, v_{t+1}, S_{t+1}|\mathbf{\Theta}\right) = p\left(Y_{t+1}, v_{t+1}|S_{t+1}, \mathbf{\Theta}\right) p\left(S_{t+1}|\mathbf{\Theta}\right)$. Consequently, it is difficult to obtain the joint posterior distribution for SVLS.

Buckle (1995) provides a representation of stable variable which makes it possible to estimate parameters of stable distributions using MCMC. The basic observation of Buckle (1995) is that although the density of a stable variable is generally unknown, the joint density of the stable variable and a well-chosen auxiliary variable is explicitly known. This joint density in turn leads to known joint posterior density of the stable variable and the auxiliary variable, which can be used in our MCMC algorithm.

For the LS process we consider, we set $\alpha \in (1, 2]$, $\beta = -1$, $\gamma = 0$ and $\delta = \sigma\Delta^{\frac{1}{\alpha}}$. We denote the index returns as $\mathbf{Y} = \{Y_t\}_{t=0}^{T}$, the volatility variables as $\mathbf{V} = \{v_t\}_{t=0}^{T}$, the jump times/sizes as $\mathbf{S} = \{S_t\}_{t=1}^{T}$, and the auxiliary variables as $\mathbf{U} = \{U_t\}_{t=1}^{T}$. Based on Buckle's (1995) result, we obtain the joint posterior distribution of $\mathbf{V}$, $\mathbf{S}$, $\mathbf{U}$ and $\mathbf{\Theta}$ as

$$
\begin{aligned}
p\left(\mathbf{\Theta}, \mathbf{V}, \mathbf{S}, \mathbf{U}|\mathbf{Y}\right) &\propto p(\mathbf{Y}, \mathbf{V}, \mathbf{S}, \mathbf{U}, \mathbf{\Theta}) = p(\mathbf{Y}, \mathbf{V}|\mathbf{S}) p(\mathbf{S}, \mathbf{U}|\mathbf{\Theta}) p(\mathbf{\Theta}) \\
&\propto \prod_{t=0}^{T-1} \frac{1}{\sigma_v v_t \Delta \sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left(\left(\epsilon_{t+1}^{y}\right)^2 - 2\rho\epsilon_{t+1}^{y}\epsilon_{t+1}^{v} + \left(\epsilon_{t+1}^{v}\right)^2 \right) \right\} \\
&\times \left(\frac{\alpha}{|\alpha-1|\Delta^{\frac{1}{\alpha}}\sigma}\right)^T \times \exp\left\{ -\sum_{t=0}^{T-1} \left| \frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})} \right|^{\frac{\alpha}{\alpha-1}} \right\} \times \prod_{t=0}^{T-1} \left\{ \left| \frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})} \right|^{\frac{\alpha}{\alpha-1}} \frac{1}{\left| \frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}} \right|} \right\} \\
&\times \prod_{t=0}^{T-1} \left[\mathbf{1}_{S_{t+1}\in(-\infty,0)\cap U_{t+1}\in(-\frac{1}{2}, l_\alpha)} + \mathbf{1}_{S_{t+1}\in(0,\infty)\cap U_{t+1}\in(l_\alpha, \frac{1}{2})} \right] \times p(\mathbf{\Theta})
\end{aligned}
$$

where $\epsilon_{t+1}^{y} = \left(Y_{t+1} - Y_t - \mu\Delta - S_{t+1}\right) / \sqrt{v_t\Delta}$, $\epsilon_{t+1}^{v} = \left(v_{t+1} - v_t - \kappa(\theta - v_t)\Delta\right) / \left(\sigma_v\sqrt{v_t\Delta}\right)$, $l_\alpha = \frac{\alpha-2}{2\alpha}$, and $t_\alpha(U_{t+1}) = \left(\frac{sin[\pi\alpha U_{t+1} + \frac{(2-\alpha)\pi}{2}]}{cos[\pi U_{t+1}]}\right)\left(\frac{cos[\pi U_{t+1}]}{cos[\pi(\alpha-1)U_{t+1} + \frac{(2-\alpha)\pi}{2}]}\right)^{(\alpha-1)/\alpha}$. We obtain joint posterior samples of $\mathbf{\Theta}, \mathbf{V}$, $\mathbf{S}$, and $\mathbf{U}$ by simulating from the above joint posterior density. We then marginal-

ize $\mathbf{U}$ out to obtain the samples for $\mathbf{\Theta}, \mathbf{V}$, and $\mathbf{S}$. That is, we simply throw away the observations of $\mathbf{U}$ and retain the observations of $\mathbf{\Theta}, \mathbf{V}$, and $\mathbf{S}$.

In general, it is difficult to simulate directly from the above high-dimensional posterior distributions. Instead, we derive the complete conditional distributions for each individual parameter and latent variable and obtain posterior samples by simulating from these individual complete conditionals iteratively following standard MCMC procedure. For example, for SVVG, we obtain the posterior distribution $p\left(\Theta_i|\mathbf{\Theta}_{-i}, \mathbf{J}, \mathbf{G}, \mathbf{V}, \mathbf{Y}\right)$ for $i = 1, ..., k$, where Θ_i is the i-th element of $\mathbf{\Theta}$ and $\mathbf{\Theta}_{-i} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_k)$, the posterior distribution for jump times $p\left(J_t|\mathbf{\Theta}, \mathbf{G}, \mathbf{V}, \mathbf{Y}\right)$, jump sizes $p\left(G_t|\mathbf{\Theta}, \mathbf{J}, \mathbf{V}, \mathbf{Y}\right)$, and latent volatility variables $p\left(v_t|v_{t+1}, v_{t-1}, \mathbf{\Theta}, \mathbf{J}, \mathbf{G}, \mathbf{Y}\right)$, for all t. In estimation, we draw posterior samples from the above complete conditional distributions and use the means of the posterior samples as parameter estimates and the standard deviations of the posterior samples as standard errors of the parameter estimates. The appendix provides the priors, the posterior distributions, and the updating procedures for model parameters and latent variables for all four models.

3. Bayesian Inferences of Lévy Jump Models: Simulation Evidence

In this section, through numerical simulations, we investigate two fundamental questions on econometric analysis of Lévy processes. The first question is whether it is possible to jointly identify diffusion, stochastic volatility, and infinite-activity Lévy jumps using discretely sampled asset prices. The second question is whether compound Poisson processes can adequately approximate the behavior of infinite-activity Lévy jumps.

Our simulation studies in this section are closely related to Aït-Sahalia (2004) and Aït-Sahalia and Jacod (2004). Aït-Sahalia (2004) shows through theoretical analysis that maximum likelihood can perfectly disentangle Brownian motion from infinite-activity Cauchy jumps using discretely sampled data. Aït-Sahalia (2004) shows that the intuition behind this result is that while there is an infinite

number of small jumps in a Cauchy process, this "infinity" remains relatively small compared to the number of Brownian movements, and while the jumps are infinitesimally small, they remain relatively bigger than the increments of a Brownian motion during the same time interval. Aït-Sahalia and Jacod (2004) show that it is possible to distinguish not only the continuous part of a Lévy process from its jump part, but also different types of jumps.

While Aït-Sahlia (2004) focuses on estimating diffusion variance in the presence of Cauchy jumps, we study the general problem of jointly identifying diffusion, stochastic volatility, and Lévy jumps via numerical simulations. We also conduct simulation studies in which we fit AJD models to data simulated from Lévy jump models. These simulation studies allow us to examine the advantages of Lévy jumps over compound Poisson processes in the presence of stochastic volatility.

3.1 Joint Identification of Stochastic Volatility and Lévy Jumps

We first provide simulation evidence that the MCMC methods developed in the previous section can accurately estimate stochastic volatility models with Lévy jumps. In our simulation studies, for each of the following four models, SVMJ, SVCMJ, SVVG, and SVLS, we generate 100 samples of twenty years of daily data. For comparison, we use similar procedures and parameters as that in EJP (2003) to generate data from SVMJ and SVCMJ. Although the simulation of VG is straightforward, it is quite difficult to simulate from LS because there are no standard random number generators for stable distributions. We apply the method of Chambers, Mallows, and Stuck (1976) to simulate stable random variables through a nonlinear transformation of two independent uniform random variables. This method works for arbitrary characteristic exponent α $(0 < \alpha < 2)$ and skewness parameter β $(-1 \leq \beta \leq 1)$. When applying this method to the LS process, we set the skewness parameter $\beta = -1$ and transform the simulated stable variables to our target stable variables, which have a drift of zero and dispersion of $\sigma\Delta^{\frac{1}{\alpha}}$.

For each of the four models, we obtain MCMC estimates of model parameters and latent variables

using each of the 100 sample paths. In our estimation, we choose the initial values of model parameters randomly within their respective permissible ranges. We set the initial values of all volatility variables to 1 and the initial values of all jump variables of MJ and CMJ to zero, i.e., no jumps. We also choose the initial values for VG and LS jumps and their associated latent variables randomly within their respective permissible ranges.[5] For each sample path, we conduct 50,000 iterations in our MCMC procedure and use the means of the posterior samples after burn-in period (30,000 iterations) as parameter estimates. Therefore, for each parameter of each model, we have 100 estimates from the 100 sample paths. The four panels of Table 1 report the true parameters, the average of the 100 estimates, and the RMSEs of the 100 estimates for the four models.

We obtain very similar results as that of EJP (2003) in our simulation studies of SVMJ and SVCMJ models. Panels A and B of Table 1 show that our MCMC methods can accurately estimate the parameters of both models. The mean estimates over the 100 samples of most parameters are very close to the true parameters with small RMSEs. Similar to EJP (2003), we find that it is generally more difficult to estimate the jump parameters than the diffusion parameters due to the relative small jump intensity. Nonetheless, the jump parameters are still accurately estimated.

More important, Panels C and D of Table 1 show that our MCMC methods also can accurately estimate the parameters of the two Lévy jump models. The mean estimates over the 100 samples of most parameters are again very close to the true parameters with small RMSEs. This is true for both the diffusion parameters and the jump parameters. Actually the RMSEs of Lévy jump parameters are much smaller than that of MJ parameters. Given that the Lévy jumps occur more frequently than compound Poisson jumps, we are able to identify the jump parameters more accurately. Simulation studies show that our MCMC algorithm can reach the true values of α and σ of

[5]The initial values chosen in this way are generally very different from their corresponding true values, suggesting that the excellent performance of our MCMC methods is not because we choose initial values that are too close to the true values.

SVLS very quickly even from initial values that are far from the true values. Once it reaches there, the algorithm only makes small updates on the two parameters. As a result, we see relatively small dispersions in the estimates of these two parameters. The RMSEs of some of the diffusion parameters in SVVG and SVLS are slightly higher, but not substantially, than that of SVMJ and SVCMJ. This suggests that the presence of infinite-activity jumps could make it more difficult to estimate diffusion parameters. However, our MCMC methods can still separate the two with sufficient accuracy from daily data. In results not reported here, we find that our MCMC algorithms can accurately estimate the latent volatility and jump variables, even though their initial values are chosen to be far from their corresponding true values.

The simulation results in this section show that our MCMC methods provide accurate joint identification of diffusion, stochastic volatility, and infinite-activity jumps using only underlying prices sampled at daily frequency. These simulation results strongly suggest that although the main conclusions of Aït-Sahalia (2004) are derived for Brownian motion with constant volatility and Cauchy jumps, they should hold for a much broader class of models: Models with stochastic volatility, an important feature of many financial time series, and models with other infinite-activity Lévy jumps.

3.2 Can Compound Poisson Processes Approximate Lévy Jumps?

One common perception in the literature is that infinite-activity jumps should not be too much different from compound Poisson processes with very high arrival rates. In this section, we examine the important question whether AJD models can adequately approximate the behavior of infinite-activity Lévy jumps via numerical simulations.

We conduct two sets of simulations. In the first set, we simulate data from models with Brownian motion (constant drift and volatility), and Cauchy, VG, and LS jumps, respectively, and we fit both the corresponding true model and the model with the same Brownian motion but Merton jump to

each simulated dataset. In the second set, we simulate data from SVVG and SVLS and we fit each dataset using the corresponding true model and SVMJ and SVCMJ models. For each model, we generate 100 sample paths of 20-year daily data. After obtaining the estimates of model parameters and latent variables for one sample path, the standardized residuals of a fitted model are calculated as:

$$\frac{Y_{t+1} - Y_t - \mu\Delta - J^y_{t+1}}{\sqrt{v_t\Delta}} = \epsilon^y_{t+1} \sim N(0,1).$$

Of course, for the first set of models, we replace v_t by the estimated constant volatility of the Brownian motion when calculating model residuals. Therefore, we obtain 100 sets of residuals from the 100 samples for each model. If a model is correctly specified, its residuals should follow $N(0,1)$ and we can use this fact to conduct specification analysis. If we fit AJD models to data simulated from Lévy models, the residuals of the fitted models might appear too frequently around zero because AJD models cannot capture the many small Lévy jumps.

Panels A and B of Figure 1 report the 95% confidence bands of kernel density estimators of model residuals obtained from fitting Cauchy and MJ models to the 100 sample paths simulated from Cauchy process, respectively. The 95% confidence bands are constructed based on the following procedure. For each of the 100 sample paths, a set of residuals in returns ϵ^y_t is computed based on the estimated model parameters and latent variables. Kernel density estimator is computed for each of the 100 sets of residuals. Then the interval $(-4, 4)$ is equally divided into 800 grid points. At each fixed grid point, we obtain a 95% confidence interval based on the 100 kernel density estimators at that point. Repeating this step for all grid points, we obtain a 95% confidence band of kernel density estimators of the residuals. We can see clearly from Panel A that the residuals of the Cauchy process are very close to $N(0,1)$. On the other hand, the residuals of the MJ model deviate from $N(0,1)$ significantly. The residuals exhibit a much higher peak and thinner wings than $N(0,1)$, suggesting that there are too many small jumps remaining in the residuals that cannot be captured by the MJ

model.

We repeat the above exercise for VG and LS and report the corresponding results in Figure 1 as well. Panels C and D of Figure 1 report the 95% confidence bands of the kernel estimators of the residuals of VG and MJ models estimated using data simulated from VG, respectively. It is clear that the residuals of VG resemble $N(0,1)$ closely, while the residuals of MJ model exhibit a much higher peak and thinner wings than $N(0,1)$. We obtain slightly different results for LS in Panels E and F on Figure 1: The residuals of MJ have not only a higher peak and thinner wings than $N(0,1)$, but also are left skewed. This is consistent with the fact that the LS model under our parameterization generates negatively skewed data.

In addition to graphical illustrations, we also formally test whether model residuals follow $N(0,1)$ using the Kolmogorov-Smirnov test. For each of the 100 sets of residuals constructed above, the Kolmogorov-Smirnov test compares the empirical cumulative distribution function (CDF) estimated from data with the CDF of $N(0,1)$ and rejects the null hypothesis if the maximum distance between the two CDFs is too big. Therefore, in total we have 100 test statistics and p-values for the 100 sets of residuals. We report the average p-values and the percentage of rejections of the null hypothesis out of the 100 residuals in Panel A of Table 2. The Kolmogorov-Smirnov test rejects the null hypothesis of $N(0,1)$ for all 100 sets of residuals of the three estimated MJ models, and the average p-values are all close to zero. In contrast, the Kolmogorov-Smirnov test fails to reject the null hypothesis of $N(0,1)$ for residuals obtained by fitting corresponding true models to simulated data.

As most financial time series exhibit time varying volatility, we conduct simulation studies for models with stochastic volatility and Lévy jumps. We generate data from SVVG and fit the simulated data using the true model and two misspecified models, SVMJ and SVCMJ. The 95% confidence bands of the residuals for the three fitted models are reported in Panels A, B, and C of Figure 2, respectively. Similar as before, we find that the residuals of the true model follow closely $N(0,1)$. On

the other hand, the residuals of SVMJ and SVCMJ exhibit higher peaks, thinner wings, and heavier left and right tails than $N(0,1)$. This suggests that the data generated from VG are leptokurtic: There are too many observations near the mean and at extreme tails. We repeat the same analysis for SVLS by simulating data from SVLS and fitting the simulated data using the true model and SVMJ and SVCMJ. Panels D, E, and F of Figure 2 report the residuals of the three fitted models, respectively. Not surprisingly, the residuals of the true models follow $N(0,1)$ closely. Again the residuals of SVMJ and SVCMJ have much higher peaks than $N(0,1)$ and are left skewed. This again suggests that compared to SVVG, SVLS is more capable of generating negatively skewed data. The Kolmogorov-Smirnov test in Panel B of Table 2 rejects the null hypothesis of $N(0,1)$ for all 100 sets of estimated SVMJ and SVCMJ residuals, and all p-values are close to zero. Again, the Kolmogorov-Smirnov test fails to reject the null hypothesis of $N(0,1)$ for residuals obtained by fitting corresponding true models to simulated data.

Our simulation results show that finite-activity jumps cannot adequately approximate the infinite-activity jump behavior, even when finite-activity jump models are estimated using data simulated from infinite-activity jump models. In particular, we find that in all our simulations there are too many small movements in infinite-activity jumps that cannot be captured by AJD models. Our simulation results suggest that the infinitely many small Lévy jumps have the potential to capture asset price movements that are too big for Brownian motion to model and in the meantime are too small for compound Poisson process to capture.

4. Bayesian Inferences of Lévy Jump Models: Empirical Results

Despite their theoretical advantages, it is not immediately clear that our Lévy jump models can significantly outperform the sophisticated AJD models with stochastic volatility, leverage effect, and compound Poisson jumps in returns and volatility in empirical applications. In this section, we examine empirically the advantages of Lévy jump models over AJD models in modeling the return

dynamics of the S&P 500 index.

Table 3 provides summary statistics of the continuously compounded returns of the S&P 500 index (the log difference of index levels) between January 2, 1980, and December 29, 2000. The index is sampled at daily frequency, and in total we have about 5,000 observations. The mean annual rate of return of the S&P 500 is 11.995%, and the volatility is 16.565%. The summary statistics of the S&P 500 are very similar to that of EJP (2003) who consider daily S&P 500 returns from January 2, 1980, to December 31, 1999. Figure 3 plots the level and log difference of the S&P 500 index. The sample period covers some major events in the history of the US stock market, such as the stock market crash of 1987 and the long boom in the late 1990s. The S&P 500 index declined dramatically in 1987 (-25% on October 19). The volatility of the index was low in the mid-1990s and increased toward the end of the sample as the economic environment deteriorated.

We provide MCMC estimates of SVMJ, SVCMJ, SVVG, and SVLS models using daily S&P 500 returns. In our MCMC simulations, we discard the first 30,000 runs as "burn-in" period and use the last 20,000 iterations to estimate model parameters. Specifically, we take the means and the standard deviations of the posterior samples as parameter estimates and standard errors, respectively.

The parameter estimates of SVMJ (SVCMJ) in the second (third) column of Table 3 are quite similar to that of EJP (2003). The long-run mean return in SVMJ (0.0405*252= 10.206%) is close to the summary statistics of mean return (11.995%). The average annualized volatility of SVMJ ($\sqrt{252 \cdot \theta} = \sqrt{252 * 0.9012} = 15.07\%$) also is close to the summary statistics of volatility (16.565%). The estimates indicate strong negative correlations between instantaneous volatility and returns ($\rho = -0.5685$ for SVMJ and $\rho = -0.4656$ for SVCMJ), which are close to those obtained from option prices (Bakshi et al. 1997, Bates 2000, and Pan 2002). After allowing jumps in volatility, the estimated volatility process has much higher κ and lower θ. Jumps in volatility helps capturing large movements in volatility, and as a result the continuous part of the volatility process does not have

to be as high as before. In both SVMJ and SVCMJ, jumps in both returns and volatility happen infrequently; our estimates of λ_y suggest on average 0.0064*252= 1. 618 jumps per year. Jump sizes in returns tend to be negative and quite large compared to day-to-day movements in prices.

The model residuals ϵ_t^y should follow an $N(0,1)$ distribution if the model is correctly specified. For each of the last 100 of the 50,000 iterations in our MCMC algorithm, we calculate one set of residuals using the model parameters and volatility/jump variables in that iteration. Therefore, in total we have 100 sets of residuals, based on which we construct the confidence band of the kernel density estimator.

In Panel A of Figure 4, we compare the kernel density estimator of model residuals of SVMJ with $N(0,1)$.[6] Similar to our simulation evidence, we find that SVMJ performs quite poorly in capturing the small movements in index returns: The density of the residuals have a higher peak around zero and thinner wings than $N(0,1)$. Panel B of Figure 4 shows that SVCMJ is also severely misspecified. The model residuals exhibit a higher peak and thinner wings than $N(0,1)$, suggesting that there are still too many small movements in stock returns that cannot be captured by the model. To better understand the tail behavior of the two models, we provide QQ plots for the residuals of the two models in Panels C and D of Figure 4, respectively. It is clear that neither model can adequately capture the left tail of the data, although SVCMJ has a slightly better performance. It seems that the return data have heavier left tail than predicted by the two models.

In addition to graphical illustrations, we also formally test whether each of the 100 sets of residuals constructed above follow $N(0,1)$ using the Kolmogorov-Smirnov test. We report the average p-value and the percentage of rejections of the null hypothesis out of the 100 residuals for SVMJ and SVCMJ in Table 5. For SVMJ, the Kolmogorov-Smirnov test fails to reject the null hypothesis of $N(0,1)$ for

[6]It is interesting to see that the 95% confidence bands of model residuals estimated using actual data are much narrower than those using simulated data. This is because our simulations consider 100 different sample paths, which have much bigger variations than the actual data with only one sample path.

only one out of the 100 sets of residuals. The average p-value is as low as 0.014. The Kolmogorov-Smirnov test in Table 5 rejects the null hypothesis of $N(0,1)$ for all 100 sets of SVCMJ residuals and the average p-value is as low as 0.006. Collectively, the evidence in Figure 4 and Table 5 suggest that there are too many small jumps in stock returns that cannot be captured by SVMJ and SVCMJ.

Next we consider SVVG and SVLS, models with infinite-activity Lévy jumps. Parameter estimates of SVVG and SVLS are reported in columns four and five of Table 4, respectively. The estimates of the stochastic volatility parameters are very similar to that of SVMJ and SVCMJ. Panels A and B of Figure 5 report the kernel density estimators of the residuals of SVVG and SVLS, respectively. It is clear that the residuals of both SVVG and SVLS are very close to $N(0,1)$. In contrast to that of SVMJ and SVCMJ, the QQ plots of the residuals of SVVG and SVLS in Panels C and D of Figure 5 show that both models capture the tails of return distribution quite well. Consistent with these graphical illustrations, the Kolmogorov-Smirnov test in Table 5 rejects the null hypothesis of $N(0,1)$ for only 15 out of the 100 sets of SVVG residuals and the average p-value is about 0.23. Similarly, the Kolmogorov-Smirnov test rejects the null hypothesis of $N(0,1)$ for 36 out of the 100 sets of SVLS residuals and the average p-value is about 0.10. The above results, however, do not suggest that infinite-variation Lévy jumps in general do not perform as well as finite-variation Lévy jumps. This could simply be due to the fact that we restrict β to be -1 in the LS model.

Finally, we examine the filtered latent volatility and jump variables for all four models. In Figure 6, we see high volatilities around the stock market crash of 1987 and toward the end of the sample period in all four models. The dramatic increases in filtered volatilities in October 1987 suggest that jumps in volatility might still be needed to completely capture the volatility dynamics. Figure 7 presents filtered jump variables for all four models. We see significant downward jumps in returns around the stock market crash of 1987 in all four models. There are also a few large jumps in stochastic volatility which, by model design, coincide with jumps in returns in SVCMJ.

The filtered jump sizes and times of the four models in Figure 7 also clearly illustrate the differences between Lévy and compound Poisson jump models. While SVMJ and SVCMJ exhibit about 20 some relatively large jumps in the twenty-year sample period, there are many more jumps in SVVG and SVLS, which include both the large jumps identified in SVMJ and SVCMJ and many small jumps. Although jumps in SVVG and SVLS are very similar to each other, there are still some differences between the jumps in these two models.

The empirical analysis of the four models in this section clearly demonstrates the empirical relevance of the infinite-activity Lévy jumps in modeling the S&P 500 index returns. Even some of the most sophisticated AJD models, such as SVMJ and SVCMJ, still cannot completely capture the leptokurtosis in index returns. In particular, there are too many small movements in index returns that cannot be captured by the AJD models. The patterns of misspecifications of AJD models based on actual data are similar to that in our simulation studies. These results strongly suggest that the infinite-activity Lévy jumps capture some intrinsic features of the S&P 500 returns and naturally fill a gap in AJD models. These results, however, do not suggest that we should abandon existing AJD models. Instead they show that we can enrich existing AJD models by including Lévy jumps to capture certain features of the data that are difficult for AJD models to capture.

5. Conclusion

Infinite-activity Lévy jumps have become increasingly popular in the continuous-time finance literature as an alternative to Brownian motion and compound Poisson process for modeling asset price dynamics. Our paper contributes to the fast-growing literature on Lévy processes by developing a Bayesian MCMC-based approach for inferences of continuous-time models with stochastic volatility and Lévy jumps using discretely sampled data. Through simulation studies and empirical analyses, we show that (i) our MCMC methods provide accurate joint identification of diffusion, stochastic volatility, and Lévy jumps, (ii) affine jump-diffusion (AJD) models fail to adequately approximate the

behavior of infinite-activity Lévy jumps, and (iii) infinite-activity Lévy jumps are crucial for modeling the S&P 500 index returns. Even though we consider only VG and LS jumps, the combination of Aït-Sahalia's (2004) theoretical results and our simulation/empirical findings strongly suggest that the advantages of VG and LS over compound Poisson jumps are likely to hold for other infinite-activity Lévy jumps as well. The MCMC techniques developed in this paper are very general and can be applied to a wide range of interesting problems in finance. To study the impact of Lévy jumps on modeling interest rates and to estimate stock return dynamics using both the underlying and option prices are two topics that will be addressed in future research.

REFERENCES

Aït-Sahalia, Y., 2004, Disentangling diffusion from jumps, *Journal of Financial Economics* 74, 487-528.

Aït-Sahalia, Y. and J. Jacod, 2004, Fisher's information for discretely sampled Lévy processes, Working paper, Princeton University.

Anderson, T., L. Benzoni, and J. Lund, 2002, An empirical investigation of continuous-time equity return models, *Journal of Finance* 57, 1239-1284.

Bakshi, G., C. Cao, and Z. Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003-2049.

Bates, D., 1996, Jumps and stochastic volatility: Exchange rate processes implicit in Deutschemark options, *Review of Financial Studies* 9, 69-107.

Bates, D., 2000, Post-'87 crash fears in S&P 500 futures options, *Journal of Econometrics* 94, 181-238.

Barndorff-Nielsen, O., 1998, Processes of normal inverse Gaussian type, *Finance and Stochastics*, 41-68.

Barndorff-Nielsen, O. and N. Shephard, 2004, Impact of jumps on returns and realized variances: Econometric analysis of time-deformed Lévy processes, Working paper, Nuffield College, University of Oxford.

Bertoin, J., 1996, *Lévy Processes*. Cambridge University Press, Cambridge.

Buckle, D.J., 1995, Bayesian inference for stable distributions, *Journal of the American Statistical Association* 90, 605-613.

Carr, P. and L. Wu, 2003, The finite moment log stable process and option pricing, *Journal of Finance* 58, 753-777.

Carr, P. and L. Wu, 2004, Time-changed Lévy processes and option pricing, *Journal of Financial Economics* 71, 113-141.

Carr, P., H. Geman, D. Madan, and M. Yor, 2002, The fine structure of asset returns: An empirical investigation, *Journal of Business* 75, 305-332.

Chambers, J., C. Mallows, and B. Stuck, 1976, A method for simulating stable random variables, *Journal of the American Statistical Association* 71, 340-344.

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

Chib, S. and E. Greenberg, 1995, Understanding the Metropolis-Hastings algorithm, *The American Statistician* 49, 327-335.

Devroye, L., 1986, *Nonuniform Random Variate Generation*, New York: Springer-Verlag.

Duffie, D., J. Pan, and K. Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343-1376.

Eberlein, E., U. Keller, and K. Prause, 1998, New insights into smile, mispricing and Value at Risk: The hyperbolic model, *Journal of Business* 71, 371-405.

Elerian, O., S. Chib, and N. Shephard, 2001, Likelihood inference for discretely observed nonlinear diffusions, *Econometrica* 69, 959-994.

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

Eraker, B., 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* 59, 1367-1403

Eraker, B., M. Johannes, and N. Polson, 2003, The impact of jumps in equity index volatility and returns, *Journal of Finance* 58, 1269-1300.

Gilk, W. R., 1992, Derivative-Free Adaptive Rejection Sampling for Gibbs Sampling, *Bayesian Statistics 4*, Oxford University Press.

Hastings, W.K., 1970, Monte Carlo sampling methods using Markov chain and their applications, *Biometrika* 57, 97-109.

Heston, S., 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327-343.

Huang, J. and L. Wu, 2003, Specification analysis of option pricing models based on time-changed Lévy processes, *Journal of Finance* 59, 1405-1439.

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

Johannes, M. and N. Polson, 2003, MCMC methods for continuous-time financial econometrics, *Handbook of Financial Econometrics,* forthcoming.

Jones, C.S., 1998, Bayesian estimation of continuous-time finance models, Working paper, USC.

Jones, C.S., 2003a, The dynamics of stochastic volatility: Evidence from underlying and options markets, *Journal of Econometrics* 116, 181-224.

Jones, C.S., 2003b, Nonlinear mean reversion in the short-term interest rates, *Review of Financial Studies* 16, 793-843.

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

Madan, D., P. Carr, and E. Chang, 1998, The variance gamma process and option pricing, *European Finance Review* 2, 79-105.

Merton, R., 1976, Option pricing when the underlying stock returns are discontinuous, *Journal of Financial Economics* 3, 125-144.

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

Qiou, Z. and N. Ravishanker, 2004, Bayesian inference for vector ARMA models with stable innovations, Working paper, University of Connecticut.

Ripley, B.D., 1987, *Stochastic Simulation*, New York: John Wiley.

Robert, C. and G. Casella, 2004, *Monte Carlo Statistical Methods*, 2nd edition, New York: Springer.

Schaumburg, E., 2004, Maximum likelihood estimation of jump processes with applications to finance, Working paper, Northwestern University.

APPENDIX. Detailed Descriptions of MCMC Methods

A.1 Priors of Model Parameters

In this section, we provide the priors for parameters of all four models. To simplify our numerical simulations, we choose standard conjugate priors whenever possible to obtain known posterior distributions.

- **Priors for parameters common to four models**. We consider the following prior distributions: $\mu \sim N(0,1)$, $\kappa \sim N(0,1)$ (truncated at zero), $\theta \sim N(0,1)$ (truncated at zero). Following Jacquier, Polson, and Rossi (1994), we reparameterize (ρ, σ_v) as (ϕ_v, w_v), where $\phi_v = \sigma_v \rho$ and $w_v = \sigma_v^2(1-\rho^2)$, and choose the priors of the new parameters as $\phi_v | w_v \sim N\left(0, \frac{1}{2} w_v\right)$ and $w_v \sim IG(2, 200)$.

- **Priors for parameters common to SVMJ and SVCMJ.** The priors for σ_y and λ_y are exactly the same as that in EJP (2003): $\mu_y \sim N(0, 100)$, $\sigma_y^2 \sim IG(5, 1/20)$, and $\lambda_y \sim Beta(2, 40)$, where IG and $Beta$ represent the inverse gamma and beta distributions, respectively.

- **Priors for parameters unique to SVCMJ**. For μ_v and ρ_J, we choose standard conjugate priors, which are $IG(10, 1/10)$ and $N(0, 4)$, respectively.

- **Priors for parameters unique to SVVG**. We also choose standard conjugate priors for the three parameters unique to SVVG: $\gamma \sim N(0,1)$, $\sigma^2 \sim IG(2.5, 5)$, and $\nu \sim IG(10, 1/10)$.

- **Priors for parameters unique to SVLS**. For α and σ, we choose the following joint priors to obtain known posterior distributions of σ: $\alpha \sim Uniform(1,2)$ and $\sigma^{\frac{\alpha}{\alpha-1}} | \alpha \sim IG(2.5, 10)$.

The priors of most parameters are proper priors, pretty uninformative, and have been used in previous literature. In general, as the sample size becomes large, the information contained in the likelihood function dominates that in the priors. As a result, we find the results computed later seem to be relatively invariant to the choice of priors.

A.2 MCMC Methods for SVMJ

In this section, we provide detailed expressions of the posterior distributions of model parameters and latent variables for SVMJ as well as the updating algorithm used there. We first focus on the posteriors of eight parameters of SVMJ, $\mathbf{\Theta} = \{\mu, \kappa, \theta, \sigma_v, \rho, \lambda_y, \mu_y, \sigma_y\}$.

- **Posterior of** μ. The posterior of μ follows a normal distribution $\mu \sim N\left(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}}\right)$, where $\mathcal{W} = \frac{\Delta}{(1-\rho^2)}\sum_{t=0}^{T-1}\frac{1}{v_t} + \frac{1}{M^2}, \mathcal{S} = \frac{1}{(1-\rho^2)}\sum_{t=0}^{T-1}\frac{1}{v_t}(C_{t+1} - \rho\frac{D_{t+1}}{\sigma_v}) + \frac{m}{M^2}, C_{t+1} = Y_{t+1} - Y_t - N_{t+1}^y\xi_{t+1}^y$, $D_{t+1} = v_{t+1} + (\kappa\Delta - 1)\, v_t - \kappa\theta\Delta$, and m and M are the hyperparameters for the prior of μ and equal to 0 and 1, respectively.

- **Posterior for** μ_y. The posterior of μ_y follows a normal distribution $\mu_y \sim N\left(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}}\right)$, where $\mathcal{W} = \frac{T}{\sigma_y^2} + \frac{1}{M^2}, \mathcal{S} = \frac{\sum_{t=0}^{T-1}\xi_{t+1}^y}{\sigma_y^2} + \frac{m}{M^2}$, and m and M are the hyperparameters for the prior of μ_y and equal to 0 and 10, respectively.

- **Posterior for** σ_y. The posterior of σ_y follows an inverse gamma distribution

$$\sigma_y^2 \sim IG\left(\frac{T}{2} + m, \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1}(\xi_{t+1}^y - \mu_y)^2 + \frac{1}{M}}\right),$$

  where m and M are the hyperparameters for the prior of σ_y and equal to 5 and 1/20, respectively.

- **Posterior for** λ_y. The posterior of λ_y follows a beta distribution

$$\lambda_y \sim Beta\left(\sum_{t=0}^{T-1} N_{t+1}^y + m, T - \sum_{t=0}^{T-1} N_{t+1}^y + M\right),$$

  where m and M are the hyperparameters for the prior of λ and equal to 2 and 40, respectively.

- **Posterior for** θ. The posterior of θ follows a truncated normal distribution $\theta \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})\, 1_{\theta>0}$, where $\mathcal{W} = \frac{\kappa^2\Delta}{\sigma_v^2(1-\rho^2)}\sum_{t=0}^{T-1}\frac{1}{v_t} + \frac{1}{M^2}, \mathcal{S} = \frac{\kappa}{(1-\rho^2)\sigma_v}\sum_{t=0}^{T-1}(\frac{D_{t+1}/\sigma_v - \rho C_{t+1}}{v_t}) + \frac{m}{M^2}, C_{t+1} = Y_{t+1} - Y_t - \mu\Delta - N_{t+1}^y\xi_{t+1}^y$, $D_{t+1} = v_{t+1} + (\kappa\Delta - 1)v_t$, and m and M are the hyperparameters for the prior of θ and equal to 0 and 1 respectively. We follow the steps below to sample from the

truncated normal distribution. For the $(g+1)$-th iteration, (1) Generate u from $uniform(0,1)$; (2) Get $\theta^{(g+1)} = F^{-1}((1-F(0))u + F(0))$ where $F(x)$ is cumulative distribution function (cdf) of $N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$ and $F^{-1}(x)$ is inverse cdf of $N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$.

- **Posterior for κ.** The posterior of κ follows a truncated normal distribution $\kappa \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})\, 1_{\kappa>0}$, where $\mathcal{W} = \frac{\Delta}{(1-\rho^2)\sigma_v^2} \sum_{t=0}^{T-1} \frac{(\theta - v_t)^2}{v_t} + \frac{1}{M^2}$, $\mathcal{S} = \frac{1}{\sigma_v(1-\rho^2)} \sum_{t=0}^{T-1} \frac{(\theta - v_t)(\frac{D_{t+1}}{\sigma_v} - \rho C_{t+1})}{v_t} + \frac{m}{M^2}, C_{t+1} = Y_{t+1} - Y_t - \mu\Delta - N_{t+1}^y \xi_{t+1}^y$, $D_{t+1} = v_{t+1} - v_t$, and m and M are the hyperparameters for the prior of κ and are equal to 0 and 1, respectively. Since the posterior is a truncated normal distribution, we follow the same procedure as for θ in generating posterior samples for κ.

- **Posteriors for σ_v and ρ.** Following Jacquier, Polson, and Rossi (1994), we transform (ρ, σ_v) to (ϕ_v, w_v) where $\phi_v = \sigma_v \rho$ and $w_v = \sigma_v^2(1-\rho^2)$. This transformation is motivated by the observation that the volatility innovation can be written as $\sigma_v \rho \epsilon_t + \sigma_v \sqrt{1-\rho^2}\eta_t$ where ϵ_t and η_t are two independent $N(0,1)$. The priors $\phi_v | w_v \sim N(0, \frac{1}{2} w_{v)}$ and $w_v \sim IG(2, 200)$ induce a diffuse distribution on ρ, ruling out very large correlations. Given this re-parameterization, the joint posteriors of (ϕ_v, w_v) are the conjugate of the priors:

$$w_v \sim IG(\frac{T}{2} + m, \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1} D_{t+1}^2 + \frac{1}{M} - \frac{\mathcal{S}^2}{2\mathcal{W}}}) \text{ and } \phi_v | w_v \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{w_v}{\mathcal{W}}),$$

where $\mathcal{W} = \sum_{t=0}^{T-1} C_{t+1}^2 + 2$, $\mathcal{S} = \sum_{t=0}^{T-1} C_{t+1} D_{t+1}, C_{t+1} = \left(Y_{t+1} - Y_t - \mu\Delta - N_{t+1}^y \xi_{t+1}^y\right) / \sqrt{v_t \Delta}$, $D_{t+1} = \left(v_{t+1} - v_t - \kappa\left(\theta - v_t\right)\Delta\right) / \sqrt{v_t \Delta}$, and m and M are the hyperparameters of w_v and are equal to 2 and 200, respectively. After drawing (ϕ_v, w_v) from their joint posteriors, we obtain draws of (ρ, σ_v) as $\sigma_v^2 = w_v + \phi_v^2$, $\rho = \frac{\phi_v}{\sigma_v}$. The draws from joint posteriors of (ϕ_v, w_v) simplify our analysis, because the individual posteriors of ρ and σ_v are not known distributions and Metropolis type of methods have to be used.

Next we consider the posteriors of latent variables of jump size ξ_t^y, jump time N_t^y, and stochastic volatility v_t.

- **Posterior for** ξ_{t+1}^y. The posterior of ξ_{t+1}^y follows a normal distribution $\xi_{t+1}^y \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$,where $\mathcal{W} = \frac{N_{t+1}^{y2}}{(1-\rho^2)v_t\Delta} + \frac{1}{\sigma_y^2}, \mathcal{S} = \frac{N_{t+1}^{y}}{(1-\rho^2)v_t\Delta}(C_{t+1} - \rho D_{t+1}/\sigma_v) + \frac{\mu_y}{\sigma_y^2}$, $C_{t+1} = Y_{t+1} - Y_t - \mu\Delta$, and $D_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta$.

- **Posterior for** N_{t+1}^y. The posterior of N_{t+1}^y follows $N_{t+1}^y \sim Bernoulli(\frac{\alpha_1}{\alpha_1+\alpha_2})$, where $\alpha_1 = \exp\left\{-\frac{1}{2(1-\rho^2)}[A_1^2 - 2\rho A_1 B]\right\}\lambda_y$, $\alpha_2 = \exp\left\{-\frac{1}{2(1-\rho^2)}[A_2^2 - 2\rho A_2 B]\right\}(1-\lambda_y)$,
$A_1 = \left(Y_{t+1} - Y_t - \mu\Delta - \xi_{t+1}^y\right)/\sqrt{v_t\Delta}$, $A_2 = \left(Y_{t+1} - Y_t - \mu\Delta\right)/\sqrt{v_t\Delta}$, and
$B = \left(v_{t+1} - v_t - \kappa(\theta - v_t)\Delta\right)/(\sigma_v\sqrt{v_t\Delta})$.

- **Posterior for** v_{t+1}. For $0 < t+1 < T$, the posterior of v_{t+1} equals

$$p(v_{t+1}|\cdot) \propto \exp\left\{-\frac{\left[-2\rho\epsilon_{t+1}^y\epsilon_{t+1}^v + \left(\epsilon_{t+1}^v\right)^2\right]}{2(1-\rho^2)}\right\} \times \frac{1}{v_{t+1}} \times \exp\left\{-\frac{\left[(\epsilon_{t+2}^y)^2 - 2\rho\epsilon_{t+2}^y\epsilon_{t+2}^v + (\epsilon_{t+2}^v)^2\right]}{2(1-\rho^2)}\right\},$$

where $\epsilon_{t+1}^y = \left(Y_{t+1} - Y_t - \mu\Delta - N_{t+1}^y\xi_{t+1}^y\right)/\sqrt{v_t\Delta}$, and $\epsilon_{t+1}^v = \left(v_{t+1} - v_t - \kappa(\theta - v_t)\Delta\right)/(\sigma_v\sqrt{v_t\Delta})$. For $t+1 = T$, the above posterior only has the first exponential part because v_T depends only on v_{T-1}. Similarly, the posterior of $p(v_0|\cdot)$ depends on $\frac{1}{v_0}$ and the second exponential part.

The posterior distribution of v_t is very complicated and difficult to simulate from. After considering a variety of updating methods, we choose the Adaptive Rejection Metropolitan Sampling (ARMS) method of Gilks, Best, and Tan (1995) to update volatility variables one at a time in our estimation of all four models. ARMS is a generalization of the Adaptive Rejection Sampling (ARS) method of Gilks (1992), which is very efficient for sampling from posterior densities that are log-concave. ARS works by constructing an envelope function of the log of the target density, which is then used in rejection sampling (see, for example, Ripley, 1987). Whenever a point is rejected by ARS, the envelope is updated to correspond more closely to the true log density, thereby reducing the chance of rejecting subsequent points. To accommodate densities that are not log concave, ARMS performs a Metropolis step on each point accepted at an ARS rejection step. In the Metropolis step, the new point is weighed against the previous point sampled. If the new point is rejected, the

previous point is retained as the new point. The procedure returns samples from the exact target density, regardless of the degree of complexity of the log density (See Robert and Casella (2004) for more detailed discussions of the method). Our simulation studies have shown that ARMS has excellent performance in updating volatility variables. We also update ν and G_t using ARMS and obtain excellent results.

Other methods we consider for updating volatility variables include Random Walk Metropolis, Accept-Reject method of Tierney (1994), and Kalman filter block updating of Carter and John (1994) and Kim, Shephard, and Chib (1998). We find that Random Walk Metropolis is very sensitive to initial values of volatility and the variance of the independent error term used to generate the candidate draws. We also find that it is difficult to obtain good blanket proposal densities for the posterior densities of volatility variables and as a result the Accept-Reject method tend to converge very slowly. It is also quite difficult to apply the Kalman filter block updating method to the square-root volatility process which is neither linear nor Gaussian and is negatively correlated with returns.

A.3 MCMC Methods for SVCMJ

The common parameters and latent variables between SVMJ and SVCMJ have similar posterior distributions. So in this section, we focus on the posterior distributions of the additional parameters and latent variables that are unique to SVCMJ.

- **Posterior for** μ_v. The posterior of μ_v follows an inverse gamma distribution

$$\mu_v \sim IG\left(T+m, \frac{1}{\sum_{t=0}^{T-1}\xi_{t+1}^v + \frac{1}{M}}\right),$$

  where m and M are the hyperparameters of the prior of μ_v and equal to 10 and 1/10, respectively.

- **Posterior for** ρ_J. The posterior for ρ_J follows a normal distribution $\rho_J \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$,where $\mathcal{W} = \frac{\sum_{t=0}^{T-1}(\xi_{t+1}^v)^2}{\sigma_y^2} + \frac{1}{M^2}$, $\mathcal{S} = \frac{\sum_{t=0}^{T-1}\xi_{t+1}^v C_{t+1}}{\sigma_y^2} + \frac{m}{M^2}$, $C_{t+1} = \xi_{t+1}^y - \mu_y$, and m and M are the hyperparameters of the prior of ρ_J and equal to 0 and 2, respectively.

- **Posteriors for ξ^y_{t+1} and ξ^v_{t+1}.** Since ξ^y and ξ^v are correlated, it is more efficient to draw $(\xi^y_{t+1},\ \xi^v_{t+1})$ from their joint posterior distribution, $\xi^v_{t+1} \sim N(\frac{\mathcal{S}_2}{\mathcal{W}_2}, \frac{1}{\mathcal{W}_2})\ 1_{\xi^v_{t+1}>0}$, $\xi^y_{t+1}|\xi^v_{t+1} \sim N(\frac{\mathcal{S}_1}{\mathcal{W}_1}, \frac{1}{\mathcal{W}_1})$,where $\mathcal{W}_2 = \frac{N^2_{t+1}}{(1-\rho^2)\sigma_v^2 v_t \Delta} + \frac{\rho_J^2}{\sigma_y^2} - \frac{B^2}{\mathcal{W}_1}$, $\mathcal{S}_2 = \frac{N_{t+1}}{(1-\rho^2)\sigma_v v_t \Delta}(-\rho C_{t+1} + \frac{D_{t+1}}{\sigma_v}) - \frac{\mu_y \rho_J}{\sigma_y^2} - \frac{1}{\mu_v} + \frac{AB}{\mathcal{W}_1}$, $\mathcal{W}_1 = \frac{N^2_{t+1}}{(1-\rho^2)v_t\Delta} + \frac{1}{\sigma_y^2}$, $\mathcal{S}_1 = A + B\xi^v_{t+1}$, $A = \frac{N_{t+1}}{(1-\rho^2)v_t\Delta}(C_{t+1} - \frac{\rho D_{t+1}}{\sigma_v}) + \frac{\mu_y}{\sigma_y^2}$, $B = \frac{\rho N^2_{t+1}}{(1-\rho^2)\sigma_v v_t \Delta} + \frac{\rho_J}{\sigma_y^2}$, and $C_{t+1} = Y_{t+1} - Y_t - \mu\Delta, D_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta$. The posterior of ξ^v_t is a truncated normal distribution, because ξ^v_t is constrained to be positive to ensure that volatility process does not turn negative. We follow the same algorithm discussed before to simulate jumps in volatility from its truncated normal posterior.

A.4 MCMC Methods for SVVG

The common parameters and latent variables between SVMJ and SVVG have similar posterior distributions. So in this section we focus on the posterior distributions of the parameters and latent variables that are unique to SVVG.

- **Posterior for γ.** The posterior of γ is $\gamma \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{1}{\sigma^2}\sum_{t=0}^{T-1} G_{t+1} + \frac{1}{M^2}$, $\mathcal{S} = \frac{1}{\sigma^2}\sum_{t=0}^{T-1} J_{t+1} + \frac{m}{M^2}$, and m and M are the hyperparameters of the prior of γ and equal to 0 and 1, respectively.

- **Posterior for σ.** The posterior of σ is $\sigma^2 \sim IG(\frac{T}{2} + m, \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1}\frac{(J_{t+1} - \theta G_{t+1})^2}{G_{t+1}} + \frac{1}{M}})$, where m and M are the hyperparameters of the prior of σ and equal to 2.5 and 5, respectively.

- **Posterior for ν.** The posterior of ν is

$$p(\nu|\cdot) \propto \left(\frac{1}{\nu^{\frac{\Delta}{\nu}}\Gamma(\frac{\Delta}{\nu})}\right)^T \left(\prod_{t=0}^{T-1} G_t\right)^{\frac{\Delta}{\nu}} \exp\left\{-\frac{1}{\nu}(\sum_{t=0}^{T-1} G_t + \frac{1}{M})\right\}(\frac{1}{\nu})^{m+1},$$

  where m and M are the hyperparameters of the prior of ν and equal to 10 and 1/10, respectively. This posterior distribution is quite complicated, and our simulation studies show that ARMS has a very good performance in drawing from this distribution.

- **Posterior for** J_{t+1}. The posterior of J_{t+1} follows a normal distribution $J_{t+1} \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{1}{(1-\rho^2)v_t\Delta} + \frac{1}{\sigma^2 G_{t+1}}$, $\mathcal{S} = \frac{1}{(1-\rho^2)v_t\Delta}(C_{t+1} - \frac{\rho D_{t+1}}{\sigma_v}) + \frac{\gamma}{\sigma^2}$, $C_{t+1} = Y_{t+1} - Y_t - \mu\Delta$, and $D_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta$.

- **Posterior for** G_{t+1}. The posterior of G_{t+1} is

$$p(G_{t+1}|\cdot) \propto G_{t+1}^{\frac{\Delta}{\nu}-\frac{3}{2}} \exp\left\{-\frac{J_t^2}{2\sigma^2}\frac{1}{G_{t+1}}\right\} \exp\left\{-\left(\frac{\gamma^2}{2\sigma^2} + \frac{1}{\nu}\right) G_{t+1}\right\}.$$

  Again ARMS has a very good performance in updating this posterior distribution. In our ARMS algorithm, we set the left bound to zero, because G_{t+1} follows a Gamma process.

A.5 MCMC Methods for SVLS

The common parameters and latent variables between SVMJ and SVLS have similar posterior distributions. So in this section we focus on the posterior distributions of the parameters and latent variables that are unique to SVLS.

- **Posterior for** σ. We choose the prior of σ conditioning on α is $\sigma^{\frac{\alpha}{\alpha-1}} \sim IG(m, M)$, then the posterior of σ is

$$p(\sigma|\cdot) \propto \left[\left(\frac{1}{\sigma}\right)^{\frac{\alpha}{\alpha-1}}\right]^{(T+m+1)} \exp\left\{-\left(\frac{1}{\sigma}\right)^{\frac{\alpha}{\alpha-1}}\left(\sum_{t=0}^{T-1}|\frac{S_{t+1}}{\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}} + \frac{1}{M}\right)\right\},$$

or alternative $\sigma^{\frac{\alpha}{\alpha-1}} \sim IG(T+m, \frac{1}{\sum_{t=0}^{T-1}|\frac{S_{t+1}}{\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}} + \frac{1}{M}})$, where m and M are the hyperparameters of the prior of σ and equal to 2.5 and 10, respectively. In each iteration, a sample τ is generated from the above inverse gamma distribution, then the current draw of σ is $\tau^{\frac{\alpha-1}{\alpha}}$ which depends on the most updated sample of α.

- **Posterior for** α. The posterior of α is

$$\begin{aligned} p(\alpha|\cdot) \quad \propto \quad & (\tfrac{\alpha}{\alpha-1})^T \exp\left\{-\textstyle\sum_{t=0}^{T-1}|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}}\right\} \times \prod_{t=0}^{T-1}|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}} \\ & \times \left[(\tfrac{1}{\sigma})^{\frac{\alpha}{\alpha-1}}\right]^{m+1} exp\{-(\tfrac{1}{\sigma})^{\frac{\alpha}{\alpha-1}}\tfrac{1}{M}\} \times \mathbf{1}(\alpha)_{\alpha\in[1.01,2]}, \end{aligned}$$

where m and M are the hyperparameters of the prior of σ and equal to 2.5 and 10, respectively. As pointed out by Buckle (1995), we tend to have computer overflow problems when α is very close to 1 because of the term $(\frac{\alpha}{\alpha-1})^T$ in all the conditional posterior densities. As a result, we choose a uniform prior of α over $[1.01, 2]$ in our implementation of the MCMC methods. It is notoriously difficult to estimate the shape parameter of a stable distribution since the complete conditional distribution for α does not have a standard form. Motivated by the idea in Qiou and Ravishanker (2004), we use the Metropolis-Hastings Algorithm with a linearly transformed Beta distribution as the proposal density. This is mainly because α is bounded from both above and below and its density appears to be unimodal. We choose the parameters of the proposal beta density, a and b, such that the previous draw $\alpha^{(g)}$ is the mode of this density and $a + b = 5log(T)$, a constant suggested by Buckle (1995). Define

$$g(\alpha|a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}(\frac{\alpha - 1.01}{0.99})^{\alpha-1}(\frac{2-\alpha}{0.99})^{b-1}.$$

Then, the algorithm works in the following way:

1. Calculate
$$\begin{cases} a_1 = (\frac{\alpha^{(g)}-1.01}{0.99})(5log(T) - 2) + 1 \\ b_1 = 5log(T) - a_1 \end{cases}$$
and then draw τ from $Beta(a_1, b_1)$ and set $\alpha^{(g+1)} = 0.99\tau + 1.01$;

2. Calculate
$$\begin{cases} a_2 = (\frac{\alpha^{(g+1)}-1.01}{0.99})(5log(T) - 2) + 1 \\ b_2 = 5log(T) - a_2 \end{cases};$$

3. Draw u from $Uniform(0, 1)$;

4. Accept $\alpha^{(g+1)}$ if $u > \min(\frac{p(\alpha^{(g+1)})}{p(\alpha^{(g)})} \times \frac{g(\alpha^{(g)}|a_2,b_2)}{g(\alpha^{(g+1)}|a_1,b_1)}, 1)$, otherwise keep the previous draw.

- **Posterior for** S_{t+1}. The posterior of S_{t+1} is

$$p(S_{t+1}|\cdot) \propto \exp\left\{-\frac{S_{t+1}}{2(1-\rho^2)v_t\Delta}[S_{t+1} - 2(C_{t+1} - \frac{\rho}{\sigma_v}D_{t+1})]\right\} \times \exp\left\{-|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}}\right\}|S_{t+1}|^{\frac{1}{\alpha-1}},$$

where $C_{t+1} = Y_{t+1} - Y_t - \mu\Delta$ and $D_{t+1} = v_{t+1} - vt - \kappa(\theta - v_t)\Delta$. Simple algebra shows this posterior is log-concave. So it is very efficient to use the ARS algorithm of Gilks (1992) to sample from this posterior distribution.

- **Posterior for** U_{t+1}. The posterior of U_{t+1} is

$$p(U_{t+1}|\cdot) \propto \underbrace{\exp\left\{-|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}} + 1\right\}|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}}}_{g(U_{t+1})}$$

$$\times[\mathbf{1}_{S_{t+1}\in(-\infty,0)\cap U_{t+1}\in(-\frac{1}{2},l_\alpha)} + \mathbf{1}_{S_{t+1}\in(0,\infty)\cap U_{t+1}\in(l_\alpha,\frac{1}{2})}].$$

Due to the monotonicity of $t_\alpha(U_{t+1})$, we know that $p(U_{t+1}|\cdot)$ has a global maximum which equals 1 at $t_\alpha(U_{t+1}) = \frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}}}$. The knowledge of this maximum makes the Rejection algorithm of Devroye (1986) or Ripley (1987) a suitable method to sample from $p(U_{t+1}|\cdot)$. This algorithm works in the following way:

1. Draw
$$U_{t+1}^{(g+1)} \longleftarrow \begin{cases} Uniform(-\frac{1}{2}, l_\alpha) \text{ if } S_{t+1} < 0 \\ Uniform(l_\alpha, \frac{1}{2}) \text{ if } S_{t+1} > 0 \end{cases};$$
2. Draw u from $Uniform(0,1)$;
3. Accept U_{t+1} if $u < g(U_{t+1}^{(g+1)})$, otherwise return to 1.

Table 1. Simulation Results on the Accuracy of MCMC Estimators of Model Parameters

This table reports simulation results on the accuracy of MCMC estimators of parameters of SVMJ, SVCMJ, SVVG, and SVLS models. We simulate 100 sample paths of data with length of 20 years of daily data from each of the four models. Then we estimate parameters of each model for each sample path. In each panel below, we report the true model parameters used in simulation, the mean of the estimated parameters from the 100 sample paths, and the dispersion of such estimates measured by RMSE. We use 50,000 iterations in our MCMC estimation.

Panel A. SVMJ Model

	μ	κ	θ	σ_v	ρ	μ_y	σ_y	λ_y
True	0.05	0.015	0.8	0.1	-0.4	-3.0	3.5	0.015
Mean	0.0509	0.0155	0.8088	0.1000	-0.3959	-3.0328	3.2446	0.0157
RMSE	0.0103	0.0024	0.0822	0.0021	0.0228	0.6964	0.4330	0.0029

Panel B. SVCMJ Model

	μ	κ	θ	σ_v	ρ	μ_y	σ_y	λ	μ_v	ρ_J
True	0.05	0.015	0.8	0.1	-0.4	-3.0	3.5	0.015	1.0	-0.4
Mean	0.0497	0.0152	0.7929	0.0983	-0.4094	-3.0657	3.2003	0.0162	0.9966	-0.3734
RMSE	0.0141	0.0018	0.1045	0.0064	0.0583	0.8840	0.4413	0.0031	0.1378	0.5086

Panel C. SVVG Model

	μ	κ	θ	σ_v	ρ	γ	σ	ν
True	0.05	0.015	0.8	0.1	-0.4	-0.01	0.4	3.0
Mean	0.0487	0.0159	0.8073	0.0981	-0.3796	-0.0115	0.3808	3.0802
RMSE	0.0212	0.0033	0.0801	0.008	0.0713	0.0229	0.0341	0.1176

Panel D. SVLS Model

	μ	κ	θ	σ_v	ρ	A	σ
True	0.05	0.015	0.8	0.1	-0.4	1.8	0.3
Mean	0.0491	0.0154	0.8153	0.099	-0.3975	1.807	0.2994
RMSE	0.0114	0.0029	0.0861	0.006	0.0506	0.0073	0.0019

Table 2. Kolmogorov-Smirnov Goodness-of-Fit Test for Simulation Analysis

Panel A. Simulation Results for Pure Jump Models

We simulate 100 sample paths of 20 years of daily data from Cauchy, VG, and LS models and fit each sample path using the corresponding true model and the Merton Jump (MJ) model. For each sample path, we obtain one set of model residuals ε_t^y based on the estimated model parameters and latent variables. Then we test whether each of the 100 sets of residuals follow N(0,1) using the Kolmogorov-Smirnov test at the 5% critical value. We report the percentage of the 100 sets of residuals for which the null hypothesis of N(0,1) is rejected. We also report the average p-values of the 100 sets of residuals.

	Percentage of rejection	Mean p-values
Cauchy fitted by Cauchy	7%	0.2961
Cauchy fitted by MJ	100%	~0.0
VG fitted by VG	20%	0.1616
VG fitted by MJ	100%	~0.0
LS fitted by LS	8%	0.4146
LS fitted by MJ	100%	~0.0

Panel B. Simulation Results for Stochastic Volatility and Jump Models

We simulate 100 sample paths of 20 years of daily data from SVVG and SVLS models and fitted each sample path using the corresponding true model and SVMJ and SVCMJ. For each sample path, we obtain one set of model residuals ε_t^y based on the estimated model parameters and latent variables. Then we test whether each of the 100 sets of residuals follow N(0,1) using the Kolmogorov-Smirnov test at the 5% critical value. We report the percentage of the 100 sets of residuals for which the null hypothesis of N(0,1) is rejected. We also report the average p-values of the 100 sets of residuals.

	Percentage of rejection	Mean p-values
SVVG fitted by SVVG	24%	0.346
SVVG fitted by SVMJ	100%	~0.0
SVVG fitted by SVCMJ	100%	~0.0
SVLS fitted by SVLS	17%	0.2528
SVLS fitted by SVMJ	100%	~0.0
SVLS fitted by SVCMJ	100%	~0.0

Table 3. Summary of Continuously Compounded Returns of S&P 500 Index

This table provides summary statistics of continuously compounded returns of S&P 500 index from January 2, 1980, to December 29, 2000. The continuously compounded returns are calculated as the log difference of S&P 500 index levels.

	Mean	Volatility	Skewness	Kurtosis	Min	Max
S&P 500	0.0476	1.0435	-2.3584	55.6080	-22.8997	8.7089

Table 4. MCMC Estimates of Model Parameters

This table provides MCMC estimates of model parameters using S&P 500 index returns from 01/02/1980 to 12/29/2000. Parameter estimates and standard errors (in parentheses) are the mean and standard deviation of posterior distributions of model parameters, respectively. In MCMC simulation, we discard the first 30,000 simulations as burn-in period and use the next 20,000 simulations in estimation.

	SVMJ	SVCMJ	SVVG	SVLS
μ	0.0405	0.0447	0.0753	0.0731
	(0.0099)	(0.0107)	(0.0102)	(0.0085)
κ	0.0150	0.0233	0.0143	0.0162
	(0.0022)	(0.0044)	(0.0023)	(0.0022)
θ	0.9012	0.7351	0.8427	0.8465
	(0.0924)	(0.0604)	(0.1004)	(0.0927)
σ_v	0.1053	0.1085	0.1101	0.1170
	(0.0011)	(0.0088)	(0.0030)	(0.0011)
ρ	-0.5685	-0.4656	-0.5839	-0.6001
	(0.0136)	(0.0565)	(0.0230)	(0.0080)
μ_y	-1.3881	-1.5377	--	--
	(1.0720)	(1.2142)		
σ_y	5.9309	6.1054	--	--
	(0.770)	(0.7924)		
λ_y	0.0064	0.0064	--	--
	(0.0011)	(0.0011)		
μ_v	--	0.5188	--	--
		(0.1548)		
ρ_J	--	0.0209	--	--
		(0.9924)		
γ	--	-	-0.0450	--
			(0.0072)	
σ	--	--	0.3863	0.2787
			(0.0202)	(0.0064)
ν	--	--	6.4858	--
			(0.0965)	
α	--	--	--	1.7629
				(0.0004)

Table 5. Kolmogorov-Smirnov Goodness-of-Fit Test for Empirical Analysis

In our estimates of SVMJ, SVCMJ, SVVG, and SVLS models using the S&P 500 index returns, we conduct 50,000 iterations in our MCMC algorithms. For each of the final 100 iterations, we obtain one set of model residuals ε_t^y based on the estimated model parameters and latent variables of that iteration. Then we test whether each of the 100 sets of residuals follow N(0,1) using the Kolmogorov-Smirnov test at the 5% critical value. We report the percentage of the 100 sets of residuals for which the null hypothesis of N(0,1) is rejected. We also report the average p-values of the 100 sets of residuals.

	Percentage of rejection	Mean p-values
SVMJ	99%	0.014
SVCMJ	100%	0.0060
SVVG	15%	0.23
SVLS	36%	0.098

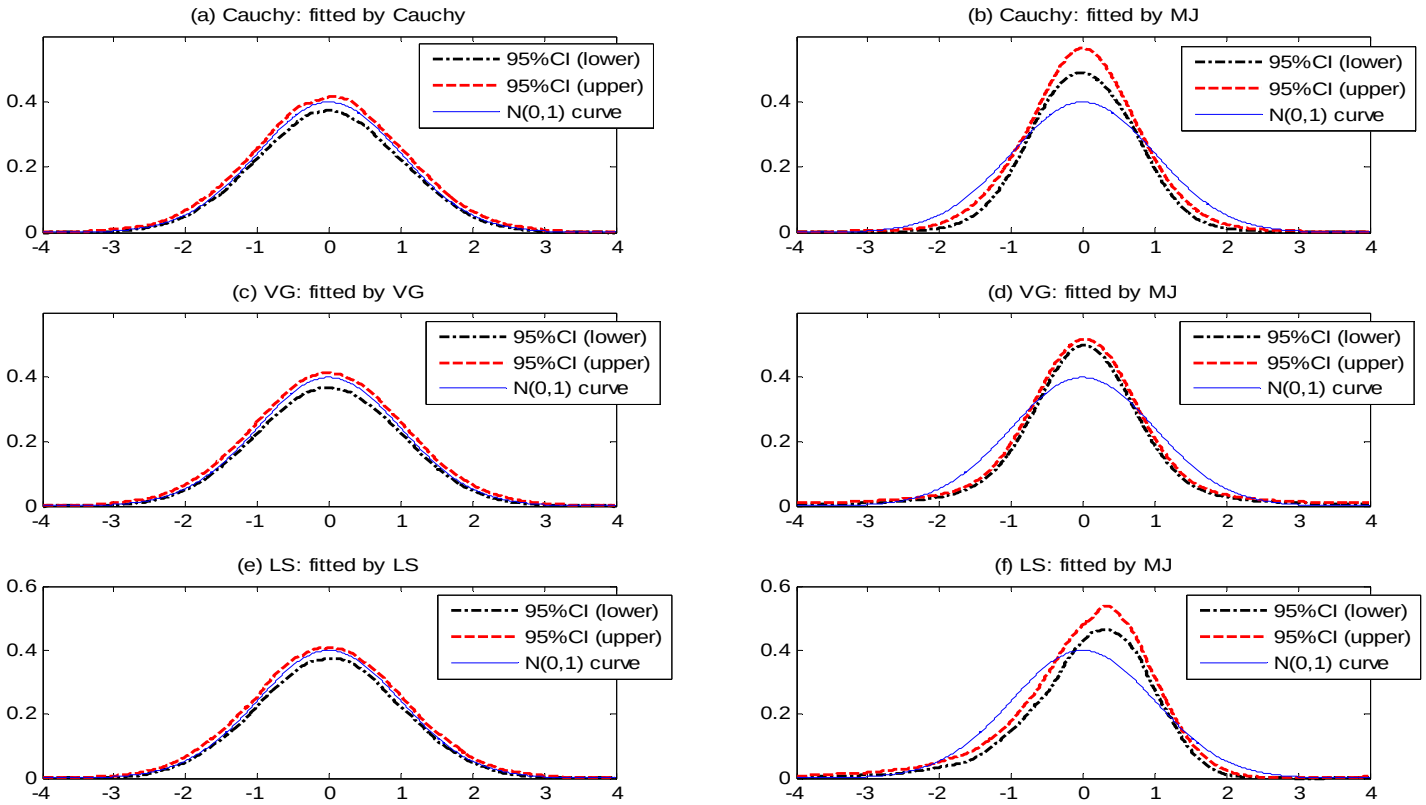

Figure 1. Residuals of pure Levy jump models (Cauchy, VG, and LS) and MJ models estimated using data simulated from corresponding Levy jump models. Panels A and B report the 95% confidence bands of kernel density estimators of model residuals obtained from fitting Cauchy and MJ models to the 100 sample paths simulated from Cauchy process, respectively. Panels C and D report the 95% confidence bands of kernel density estimators of model residuals obtained from fitting VG and MJ models to the 100 sample paths simulated from VG process, respectively. Panels E and F report the 95% confidence bands of kernel density estimators of model residuals obtained from fitting LS and MJ models to the 100 sample paths simulated from LS process, respectively.

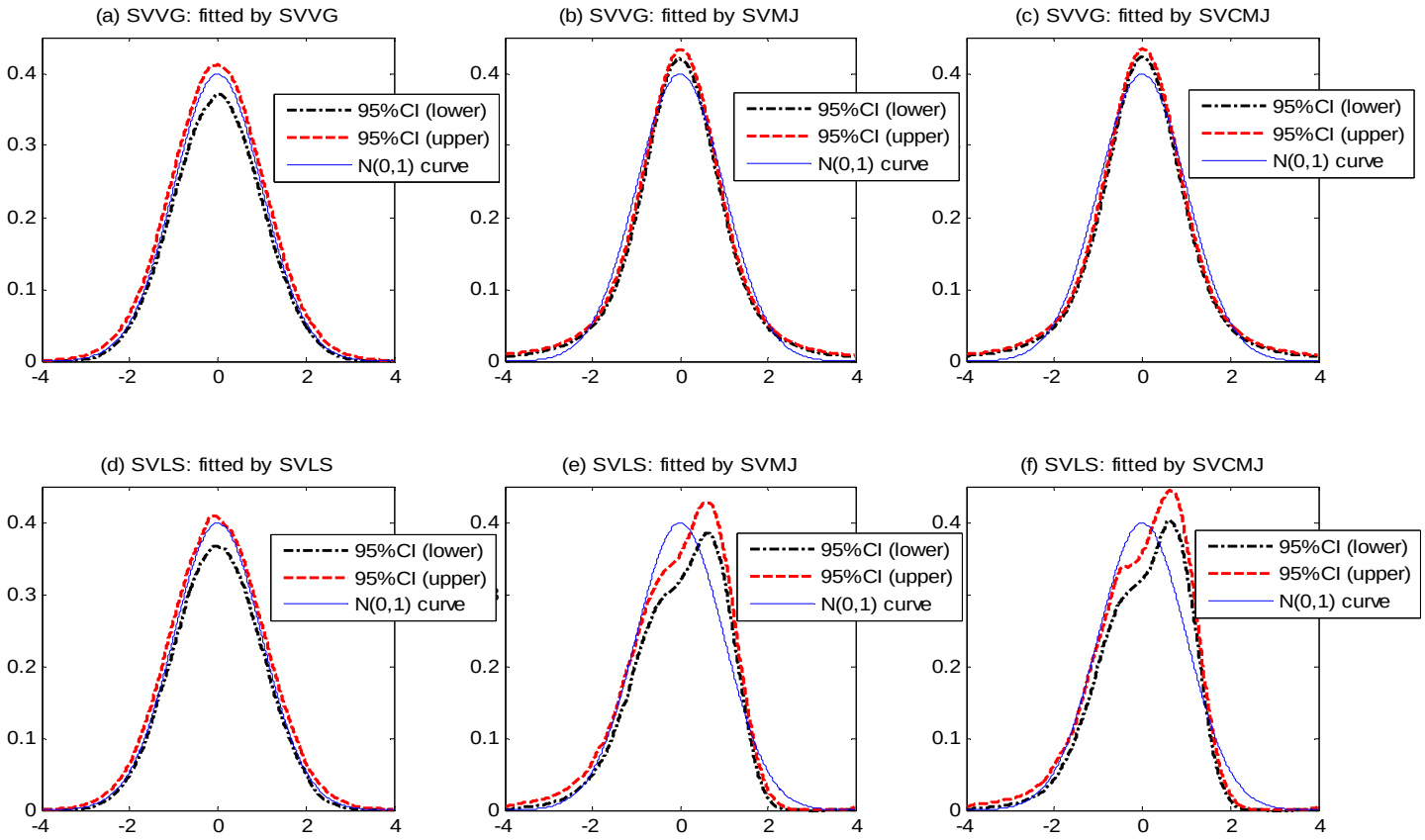

Figure 2. Residuals of models with stochastic volatility, Levy, and Merton jumps estimated using data simulated from models with stochastic volatility and Levy jumps. Panels A , B, and C report the 95% confidence bands of kernel density estimators of model residuals obtained from fitting SVVG, SVMJ, and SVCMJ models to the 100 sample paths simulated from SVVG process, respectively. Panels D , E, and F report the 95% confidence bands of kernel density estimators of model residuals obtained from fitting SVLS, SVMJ, and SVCMJ models to the 100 sample paths simulated from SVLS process, respectively.

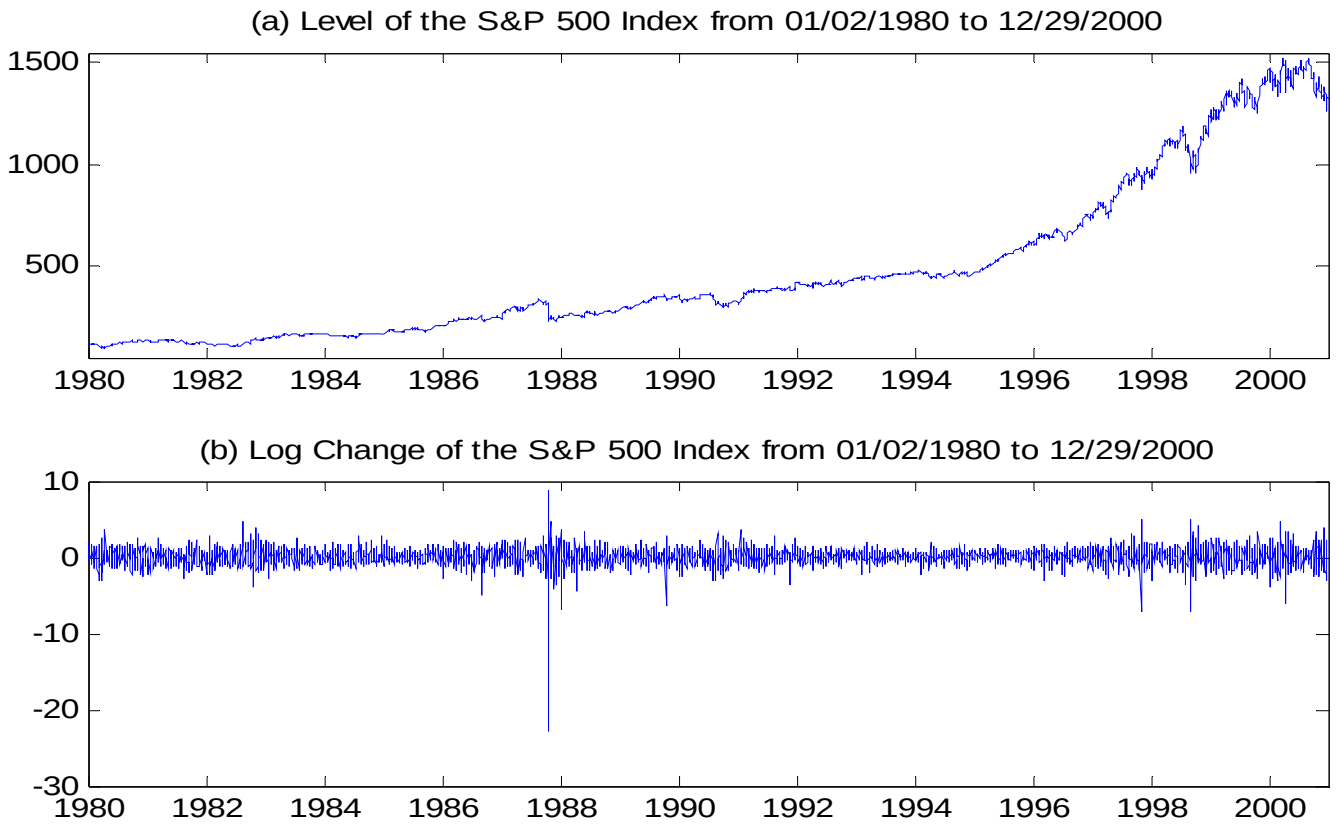

Figure 3. Level and log changes of the S&P 500 index from January 2, 1980, to December 29, 2000.

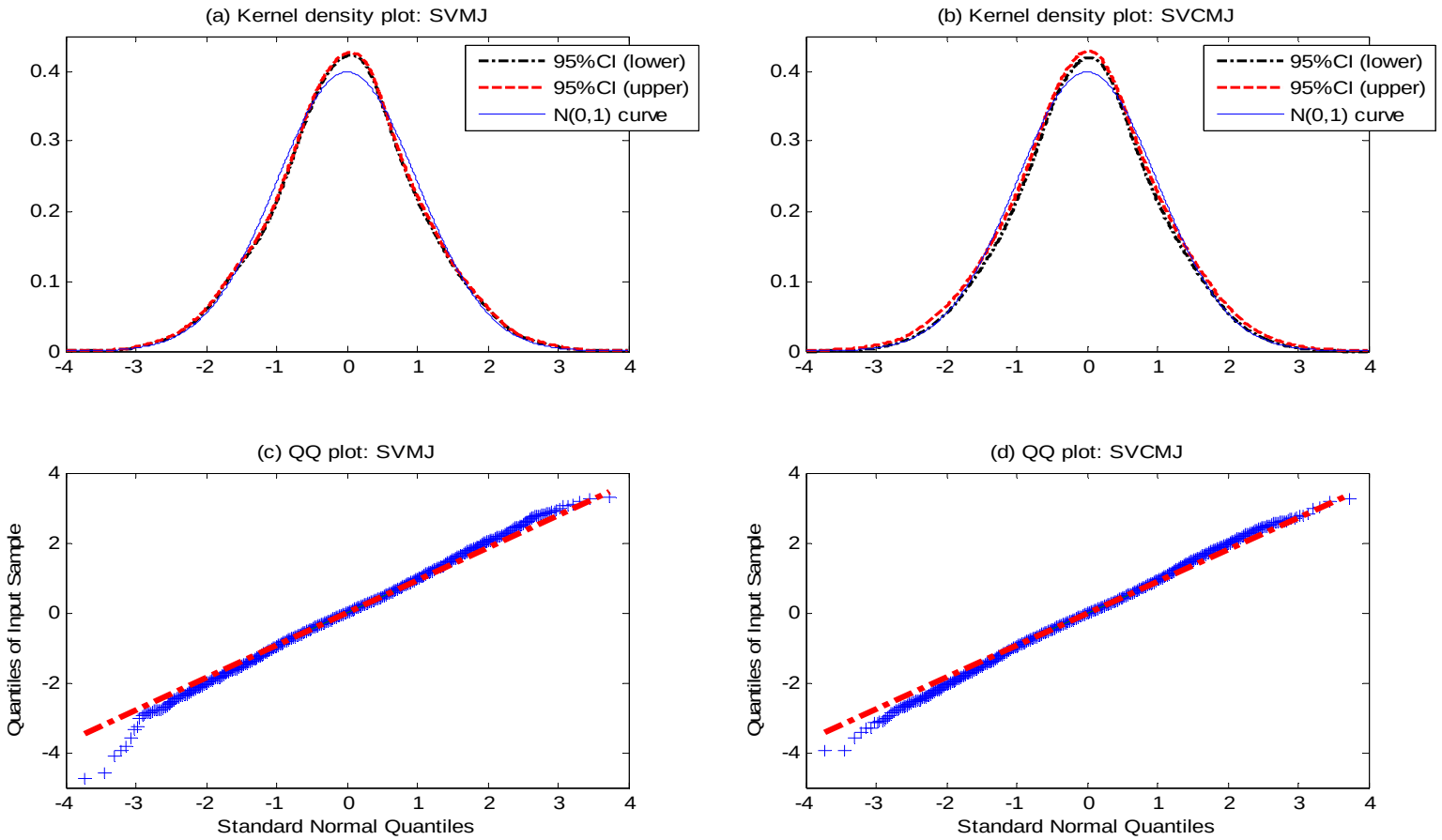

Figure 4. Kernel density and QQ plots of the residuals of SVMJ and SVCMJ estimated using daily returns of the S&P 500 index between January 2, 1980, and December 29, 2000. Panels A and B report the 95% confidence bands of kernel density estimators of the residuals of the estimated SVMJ and SVCMJ models, respectively. Panels C and D report the QQ plots of the residuals of the estimated SVMJ and SVCMJ models, respectively.

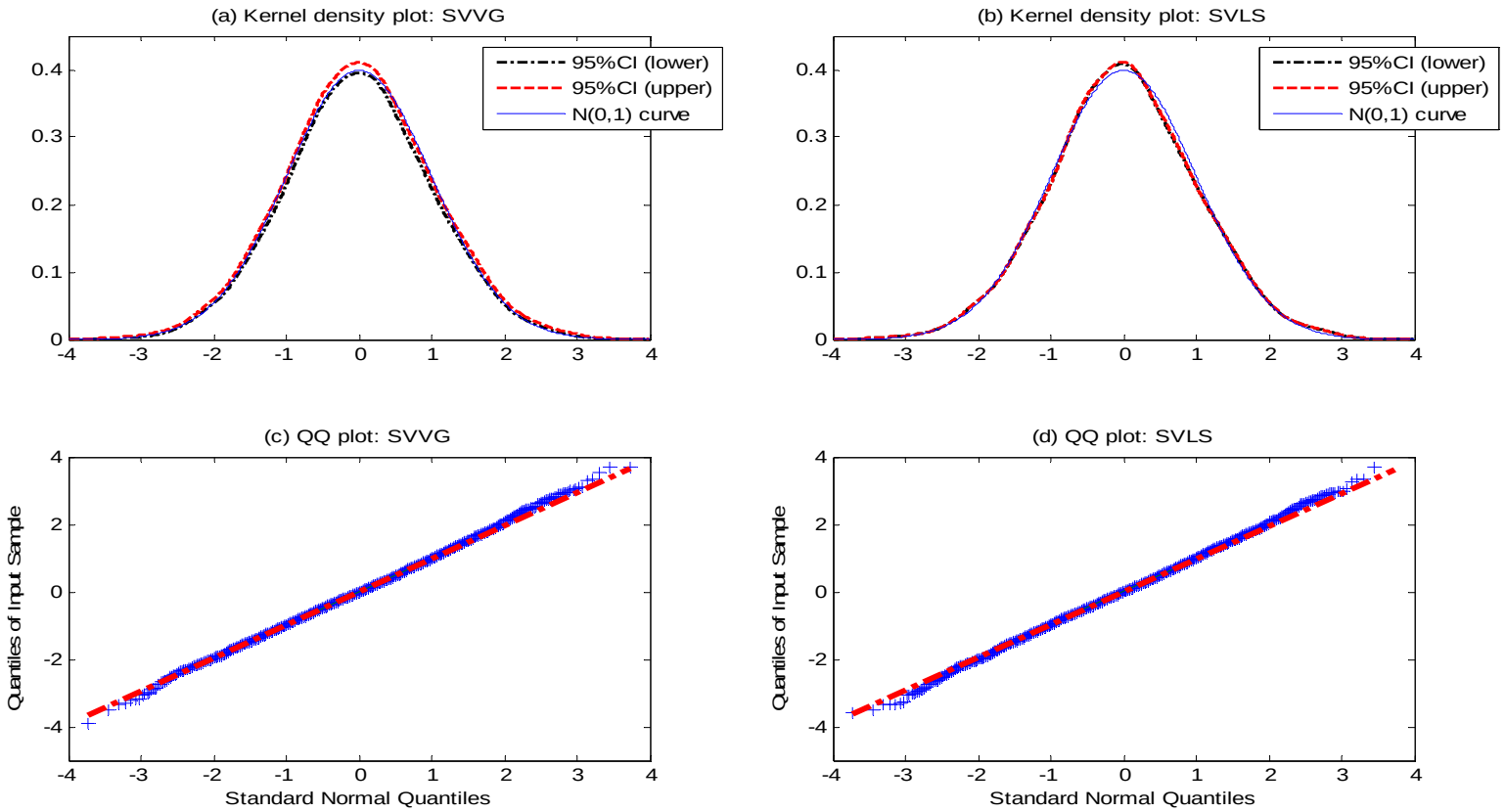

Figure 5. Kernel density and QQ plots of the residuals of SVVG and SVLS estimated using daily returns of the S&P 500 index between January 2, 1980, and December 29, 2000. Panels A and B report the 95% confidence bands of kernel density estimators of the residuals of the estimated SVVG and SVLS models, respectively. Panels C and D report the QQ plots of the residuals of the estimated SVVG and SVLS models, respectively.

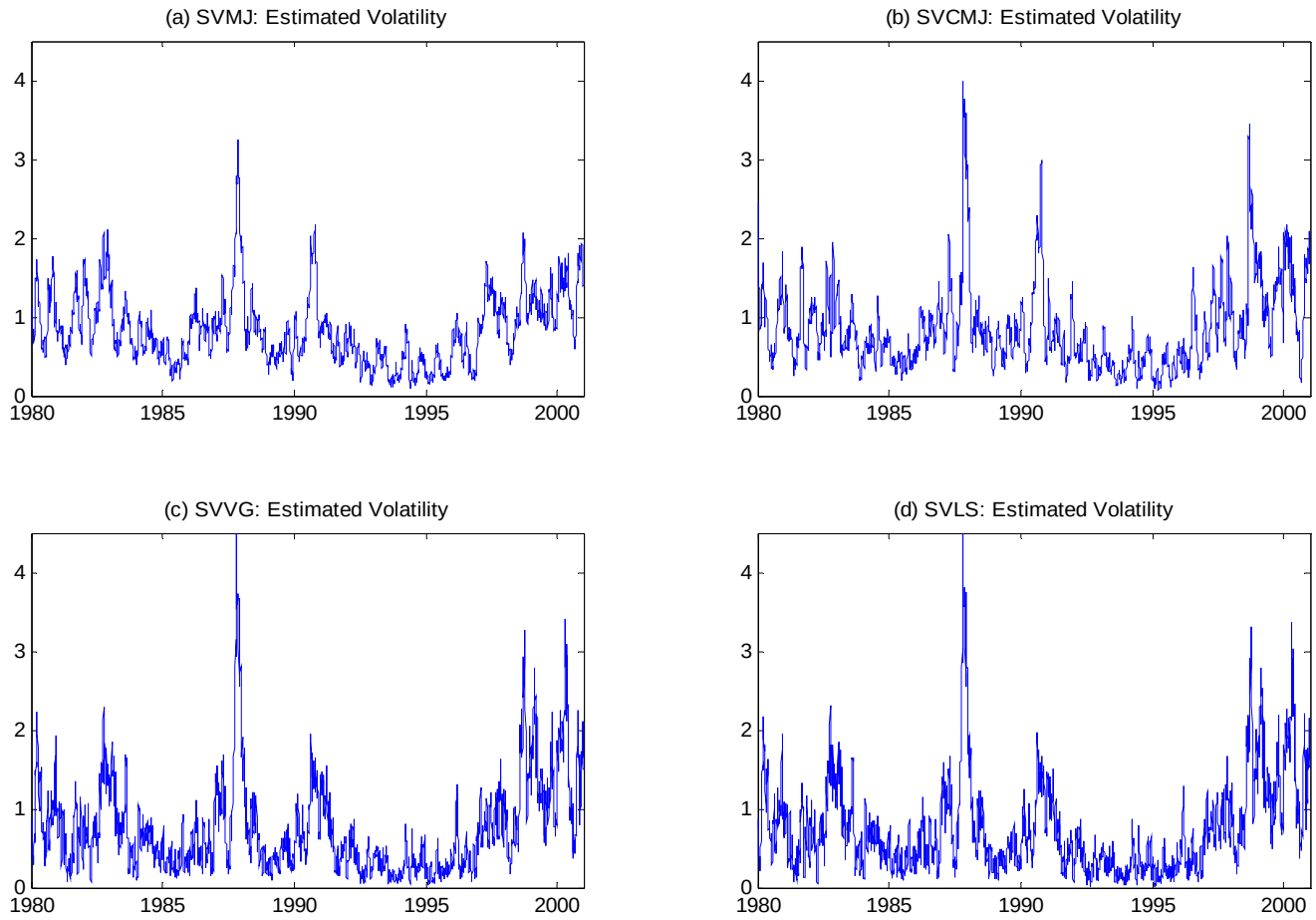

Figure 6. Filtered volatility variables of SVMJ, SVCMJ, SVVG, and SVLS models using daily returns of the S&P 500 index between January 2, 1980, and December 29, 2000.

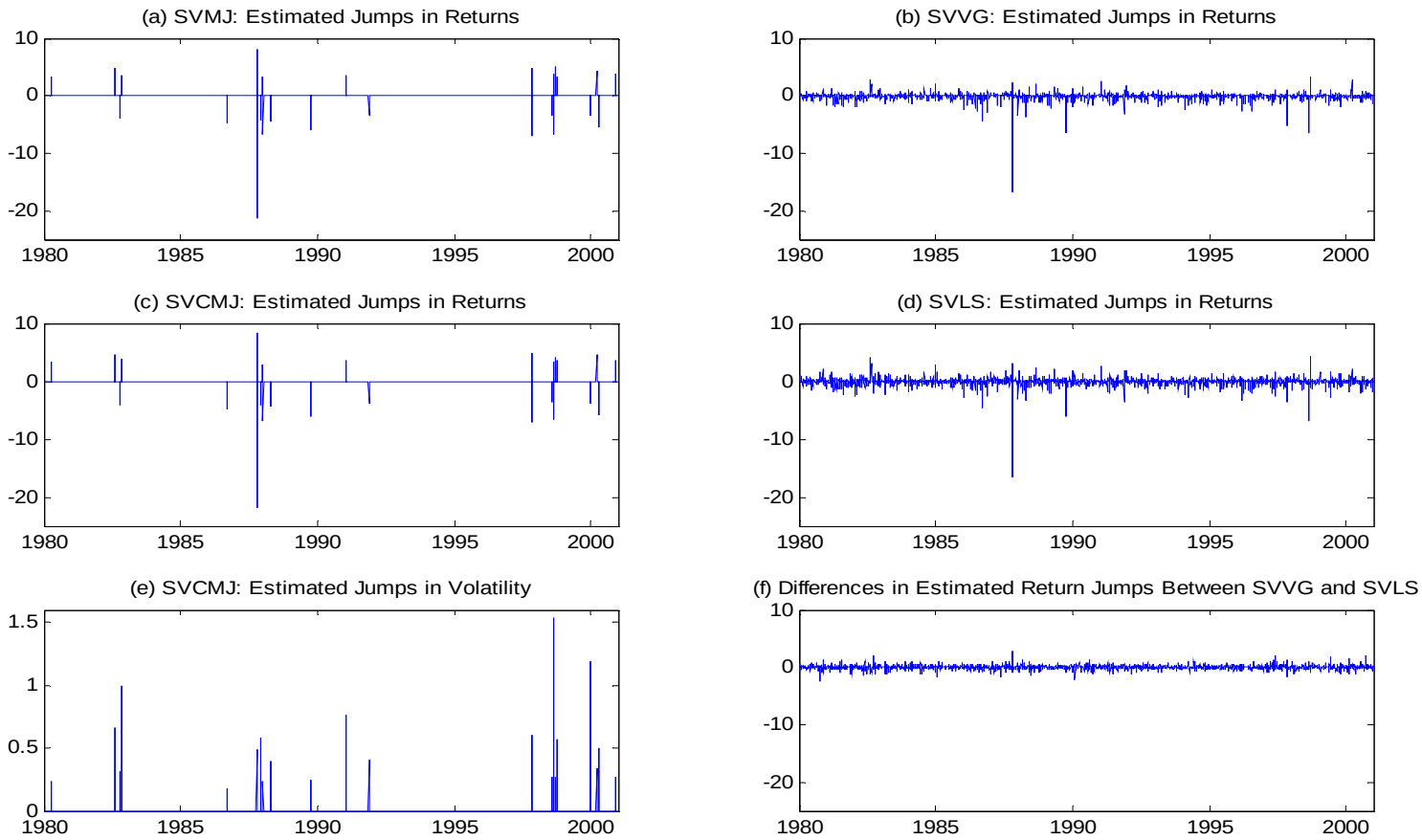

Figure 7. Filtered jump variables of SVMJ, SVCMJ, SVVG, and SVLS models using daily returns of the S&P 500 index between January 2, 1980, and December 29, 2000.