Mathematical Finance, Vol. 21, No. 3 (July 2011), 383–422

MCMC ESTIMATION OF LÉVY JUMP MODELS USING STOCK AND OPTION PRICES

Cindy L. Yu

Iowa State University

Haitao Li

University of Michigan

Martin T. Wells

Cornell University

We examine the performances of several popular Lévy jump models and some of the most sophisticated affine jump-diffusion models in capturing the joint dynamics of stock and option prices. We develop efficient Markov chain Monte Carlo methods for estimating parameters and latent volatility/jump variables of the Lévy jump models using stock and option prices. We show that models with infinite-activity Lévy jumps in returns significantly outperform affine jump-diffusion models with compound Poisson jumps in returns and volatility in capturing both the physical and risk-neutral dynamics of the S&P 500 index. We also find that the variance gamma model of Madan, Carr, and Chang with stochastic volatility has the best performance among all the models we consider.

Key Words: Levy processes, variance gamma model, Markov Chain Monte Carlo, option pricing.

1. INTRODUCTION

Modeling the dynamics of stock returns is one of the most important issues in modern finance. A realistic model of return dynamics is essential for option pricing, portfolio analysis, and risk management. Although continuous-time models for return dynamics since Black and Scholes (1973) and Merton (1976) have mainly relied on Brownian motion and compound Poisson process as basic model building blocks, Lévy processes have become increasingly popular for modeling asset price dynamics in recent years.[1]

Martin Wells gratefully acknowledges the support of NSF Grant DMS 02-04252. We thank Yacine Aït-Sahalia for providing the data used in this study. We thank Yacine Aït-Sahalia, Antje Berndt, Peter Carr, Francois Derrien, Jefferson Duarte, Bjorn Eraker, Wayne Fuller, John Hull, Raymond Kan, Bob Jarrow, George Jiang, Dilip Madan (the editor), John Maheu, Nour Meddahi, Tom McCurdy, Ray Renken, Sidney Resnick, Ernst Schaumburg, Neil Shephard, George Tauchen, Liuren Wu, the associate editor, two anonymous referees, and seminar participants at Cornell University, Hong Kong University of Science and Technology, Iowa State University, Rice University, Virginia Commonwealth University, the University of Arizona, the University of Toronto, and the 17th Derivatives Conference at FDIC for helpful comments. We are responsible for any remaining errors.

Manuscript received June 2007; final revision received August 2009.

Address correspondence to Cindy L. Yu, Department of Statistics, Iowa State University, Ames, IA 50010; e-mail: cindyyu@iastate.edu.

[1] Prominent examples of Lévy models in the literature include the inverse Gaussian model of Barndorff-Nielsen (1998); the generalized hyperbolic class of Eberlein, Keller, and Prause (1998); the variance gamma

DOI: 10.1111/j.1467-9965.2010.00439.x

Lévy processes are continuous-time stochastic processes with stationary and independent increments. Although Brownian motion and compound Poisson process are two of the most well-known special cases of Lévy processes, there are many other members of the Lévy family that offer greater flexibility for modeling purposes. For example, Lévy processes allow nonnormal increments as compared to normal increments of Brownian motion. The jump component of a general Lévy process also is much more flexible than a compound Poisson process. In particular, 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.

There are concerns, however, that infinite-activity Lévy jumps, despite their *theoretical* appeals, may not have significant *empirical* advantages over some of the most flexible models of stock returns based on affine jump-diffusions (hereafter AJD) of Duffie, Pan, and Singleton (2000) (hereafter DPS). In AJD models, stock returns are driven by affine diffusions and compound Poisson processes. One of the most sophisticated AJD models for stock returns is the double-jump models of Eraker, Johannes, and Polson (2003) (hereafter EJP), which include not only stochastic volatility and leverage effect, but also compound Poisson jumps in both returns and volatility. The double-jump models capture important stylized behaviors of both returns and volatility of major U.S. stock indices. Therefore, it is not clear that infinite-activity Lévy jump models can significantly outperform the double-jump models in empirical applications. Unfortunately, there are no direct comparisons between Lévy jump models and the double-jump models of EJP (2003) in capturing the joint dynamics of stock and option prices in the current literature.[2]

Our paper addresses a basic and yet fundamental empirical issue in the current continuous-time finance literature: Can commonly used Lévy jump models outperform the most sophisticated AJD models in capturing the joint dynamics of spot and option prices? In particular, we consider models with stochastic volatility and jumps in returns that follow the variance gamma (VG) model of Madan, Carr, and Chang (1998) or the log stable (LS) model of Carr and Wu (2003), two of the most widely used Lévy processes in the current literature.[3] We also consider AJD models with stochastic volatility and compound Poisson jumps in returns or correlated compound Poisson jumps in both returns and volatility. The latter is the preferred model of EJP (2003).[4]

Statistical analysis of Lévy processes, however, can be difficult due to various reasons. First, the probability densities of most Lévy processes are not known in closed form and for certain processes, such as stable processes, not all moments exist. As a result, it is difficult to use either likelihood- or moment-based methods for estimation. Second, it is computationally demanding to deal with the high-dimensional latent volatility variables

model of Madan, Carr, and Chang (1998); the generalization of the variance gamma model in Carr et al. (2002); and the finite moment log-stable model of Carr and Wu (2003) among others. See also Wu (2006b) for an excellent review of the current literature on Lévy processes.

[2] Existing studies of Lévy processes using option prices, such as Huang and Wu (2003), do not compare the performances of Lévy jump models with that of the double-jump model.

[3] Earlier studies on variance gamma processes include Madan and Seneta (1990) and Madan and Milne (1991).

[4] We emphasize that the continuous part of the volatility process in both the AJD and Lévy jump models follows affine diffusion. Therefore, the main focus of our comparison is on the jump structures of the two classes of models. We refer to the two classes of models as the AJD and Lévy jump models mainly for ease of distinction.

typically included in some of the most sophisticated Lévy models.[5] Finally, attempts to include option prices in model estimation significantly increase the computational complexity because calculations of option prices involve numerical integrations.

We first develop efficient computational Bayesian Markov chain Monte Carlo (hereafter MCMC) methods for estimating the above Lévy and AJD models using both stock and option prices. Our focus on the joint dynamics makes it possible to estimate simultaneously the risk-neutral and physical dynamics of asset returns, as well as the market prices of risks that govern the change of measure process. Based on our MCMC methods, we obtain estimates of model parameters and posterior distributions of latent volatility/jump variables, which are important for understanding different aspects of model performance. Although Li, Wells, and Yu (2008) (hereafter LWY) have examined MCMC estimation of Lévy jump models using stock prices, the estimation problem becomes computationally much more challenging due to the inclusion of option prices. As a result, we rely on more sophisticated updating procedures to estimate many model parameters and latent variables.

Based on the new MCMC methods, we estimate the AJD and Lévy jump models using daily returns of the S&P 500 index and daily prices of a short-term ATM SPX option. We show that the Lévy jump models significantly outperform the preferred AJD model of EJP (2003) in capturing the joint dynamics of the spot and option prices of the S&P 500 index. For the physical dynamics, the infinite-activity Lévy jumps capture many small movements in index returns that cannot be captured by the AJD models. For the risk-neutral dynamics, the Lévy jump models have significantly smaller in-sample and out-of-sample option pricing errors than the preferred AJD model. We also find that the VG model Madan, Carr, and Chang (1998) with stochastic volatility has the best performance among all the models we consider.[6]

There are only a few other studies that estimate Lévy processes using spot and option prices jointly. Wu (2006a) introduces the so-called dampened power law to capture the tail behaviors of index returns under the physical and the risk-neutral measures. Bakshi and Wu (2005) estimate Lévy jump models using the spot and option prices of the Nasdaq 100 index during the Internet "bubble" period. Although Wu (2006a) and Bakshi and Wu (2005) use numerical likelihood method to estimate model parameters, the MCMC methods we adopt are particularly suitable to deal with the large number of latent volatility and jump variables. The Bayesian approach also makes it possible to study the impacts of priors and parameter uncertainties in applications such as hedging, portfolio selection, and VaR calculation involving Lévy processes. Consistent with the empirical focus of our study, we also adopt a different approach to the change of measure for Lévy processes from that of Wu (2006a) and Bakshi and Wu (2005). We require that jumps follow the same Lévy processes under the physical and the risk-neutral measures in order to have a fair comparison with AJD models in which jumps under both measures follow compound Poisson processes. Given this restriction, we obtain the Radon-Nikodym derivatives for VG and LS processes based on Sato's (1999) theorem. In contrast, Wu

[5] Stochastic volatility is essential for capturing empirical behaviors of stock returns, and existing studies, such as Carr et al. (2003) and Carr and Wu (2004), have used stochastic time change to generate stochastic volatility in Lévy processes.

[6] Aït-Sahalia and Jacod (2009) and Tauchen and Todorov (2008) provide nonparametric evidence that jump activity of asset prices is higher than that of the variance gamma process. However, their evidence is based on intraday stock prices and exchange rates, respectively, which might exhibit more active jumps that the daily data we use. Nevertheless, the two papers support the main point of our paper of using infinite-activity jumps in modeling asset prices.

(2006a) and Bakshi and Wu (2005) fix the form of the Radon-Nikodym derivative, which is defined by the so-called Esscher transform. Under this transform, jumps generally follow different Lévy processes under the two measures.[7]

The rest of the paper is organized as follows. In Section 2, we introduce the AJD and Lévy jump models and discuss the change of measure and option pricing under these models. In Section 3, we develop MCMC methods for estimating model parameters and latent variables of the Lévy jump models using spot and option prices. Section 4 contains empirical results using daily S&P 500 index returns and prices of SPX options. Section 5 concludes the paper. The appendix provides additional information on the four jump models we consider and detailed discussions of the MCMC methods.

2. AJD AND LÉVY JUMP MODELS FOR RETURN DYNAMICS

In this section, we introduce the AJD and Lévy jump models considered in our study. We also discuss the change of measure (between the physical and the risk-neutral measures) and option pricing under these models.

2.1. AJD and Lévy Jump Models for Return Dynamics

Suppose the uncertainty of the economy is described by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{F}_t\}$. We refer to $\mathbb{P}$ as the physical probability measure which represents the probability measure of the real world in which we reside. Let S_t be the price of a stock and Y_t be the continuously compounded return on the stock, that is, $Y_t = \log S_t$. We assume that the dynamics of Y_t are characterized by the following model:

$$dY_t = \mu\, dt + \sqrt{v_t}\, dW_t^{(1)}(\mathbb{P}) + dJ_t^y(\mathbb{P}), \tag{2.1}$$

$$dv_t = \kappa(\theta - v_t)\, dt + \sigma_v \sqrt{v_t}\big(\rho dW_t^{(1)}(\mathbb{P}) + \sqrt{1-\rho^2}\, dW_t^{(2)}(\mathbb{P})\big) + dJ_t^v(\mathbb{P}), \tag{2.2}$$

where μ measures the expected rate of return, v_t measures the instantaneous volatility of return, $W_t^{(1)}(\mathbb{P})$ and $W_t^{(2)}(\mathbb{P})$ are independent standard Brownian motions under $\mathbb{P}$, and $J_t^y(\mathbb{P})$ and $J_t^v(\mathbb{P})$ represent jumps in returns and volatility under $\mathbb{P}$, respectively.

The above model nests all the models considered in this paper. In particular, the continuous part of the instantaneous volatility of returns in all models follows the square-root process of Heston (1993): θ represents the long-run mean of v_t, κ is the speed of mean reversion, σ_v is the so-called volatility of volatility, and ρ measures the correlation between volatility and returns. Many studies have documented a strong negative correlation between volatility and returns, the so-called "leverage" effect, and the correlation coefficient ρ helps to capture this phenomenon. The main difference between AJD and Lévy jump models is the jump process. In AJD models, jumps follow compound Poisson processes, which are finite-activity jumps. In Lévy models, jumps are infinite-activity.

In the first AJD model, we consider, $J_t^v(\mathbb{P}) = 0$, and $J_t^y(\mathbb{P})$ follows a compound Poisson process with a constant jump intensity and jump sizes that follow a normal distribution:

[7] Other studies that estimate Lévy processes using underlying or option prices include Barndorff-Nielsen and Shephard (2004), Belomestny and Reiss (2006), Cont and Tankov (2004a), Griffin and Steel (2006), among others. Aït-Sahalia (2004) and Aït-Sahalia and Jacod (2008) provide theoretical analyses on statistical inferences of Lévy processes.

$$J_t^y(\mathbb{P}) = \sum_{n=1}^{N_t} \xi_n^y, \tag{2.3}$$

where $N_t \sim Poisson(\lambda t)$ and $\xi_n^y \sim N(\mu_y, \sigma_y^2)$. We refer to this model as the stochastic volatility Merton jump (hereafter SVMJ) model because the jump process was first introduced in Merton (1976).

The second AJD model we consider allows correlated jumps in both returns and volatility. The stochastic volatility correlated Merton jump (hereafter SVCMJ) model is the preferred model in EJP (2003) and Eraker (2004):

$$\begin{pmatrix} J_t^y(\mathbb{P}) \\ J_t^v(\mathbb{P}) \end{pmatrix} = \sum_{n=1}^{N_t} \begin{pmatrix} \xi_n^y \\ \xi_n^v \end{pmatrix}, \tag{2.4}$$

where $N_t \sim Poisson(\lambda t)$, $\xi_n^v \sim \exp(\mu_v)$, and $\xi_n^y \mid \xi_n^v \sim N(\mu_y + \rho_J \xi_n^v, \sigma_y^2)$. The above model is sometimes referred to as the double-jump model because of the jumps in both returns and volatility. As shown in EJP (2003), the negative jumps in returns, $J_t^y(\mathbb{P})$, help to capture the major crashes observed in the U.S. market; and the jumps in volatility, $J_t^v(\mathbb{P})$, help to model rapid increase in volatility that cannot be easily captured by the square-root process.

The two basic building blocks for AJD models, Brownian motion and compound Poisson process, are special cases of Lévy processes, which are continuous-time stochastic processes with stationary and independent increments. Formally, if X_t is 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 $X_t - X_s$ is independent of $\mathcal{F}_t$ and distributed as X_{t-s} for $0 \leq s < t$. Lévy processes are much more flexible than Brownian motion and compound Poisson process because they allow discontinuous sample paths, nonnormal increments, and more flexible jump structures that have (possibly) infinite arrival rates.[8]

Unlike finite-activity jump processes, an infinite-activity jump process allows an (possibly) infinite number of jumps within any finite time interval. 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 VG model of Madan, Carr, and Chang (1998) 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,

$$X_{VG}(t \mid \sigma, \gamma, \nu) = \gamma G_t^\nu + \sigma W(G_t^\nu), \tag{2.5}$$

where $W(t)$ is a standard Brownian motion and is independent of G_t^ν. The model in (2.1) and (2.2) reduces to the SVVG model, if $J_t^y(\mathbb{P}) = X_{VG}(t \mid \sigma, \gamma, \nu)$ and $J_t^v(\mathbb{P}) = 0$.

We choose the finite moment LS process of Carr and Wu (2003) as a representative of the infinite-activity and infinite-variation jump model in our analysis. The increments of the LS process follow an α-stable distribution. That is, for $t > s$,

$$X_{LS}(t \mid \alpha, \sigma) - X_{LS}(s \mid \alpha, \sigma) \sim S_\alpha\left(-1, \sigma^{\frac{1}{\alpha}}(t-s)^{\frac{1}{\alpha}}, 0\right), \tag{2.6}$$

[8] For more detailed discussions on Lévy processes, see Cont and Tankov (2004b).

where a generic α-stable distribution is 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 $\gamma \in \mathbb{R}$. The parameter α determines the shape of the distribution, while β determines the skewness of the distribution. Stable densities are supported on either $\mathbb{R}$ or $\mathbb{R}^+$. The latter situation occurs only when $\alpha < 1$ and $\beta = \pm 1$. Following Carr and Wu (2003), we set $\beta = -1$ to achieve finite moments for index levels under the risk-neutral measure (and thus finite option prices), and negative skewness in the return density, a feature that cannot be captured by either a Brownian motion or a symmetric Lévy α-stable process. We also restrict $\alpha \in (1, 2)$ so that the process has the support of the whole real line. The model in (2.1) and (2.2) reduces to the SVLS model, if $J_t^y(\mathbb{P}) = X_{LS}(t \mid \alpha, \sigma)$ and $J_t^v(\mathbb{P}) = 0$.

The two models, SVVG and SVLS, allow us to compare the performances of infinite-activity jumps in returns with that of compound Poisson jumps in both returns and volatility.

2.2. Change of Measure and Option Pricing for the AJD and Lévy Jump Models

Although equations (2.1) and (2.2) describe the AJD and Lévy jump models under the physical measure $\mathbb{P}$, for the purpose of option pricing, we also need return dynamics under the risk-neutral measure $\mathbb{Q}$. Thus, we need to consider the change of measure between $\mathbb{P}$ and $\mathbb{Q}$ for these models.

The change of measure for Brownian motion is well understood in the literature. Following the standard practice of Pan (2002), we assume that the market prices of risks of Brownian shocks to returns and volatility are

$$\gamma_t^{(1)} = \eta^s \sqrt{v_t}, \ \gamma_t^{(2)} = -\frac{1}{\sqrt{1-\rho^2}} \left(\rho \eta^s + \frac{\eta^v}{\sigma_v} \right) \sqrt{v_t}, \tag{2.7}$$

respectively. Thus, the change of measure for the two Brownian motions is

$$dW_t^{(1)}(\mathbb{Q}) = dW_t^{(1)}(\mathbb{P}) + \gamma_t^{(1)}\, dt, \quad dW_t^{(2)}(\mathbb{Q}) = dW_t^{(2)}(\mathbb{P}) + \gamma_t^{(2)}\, dt, \tag{2.8}$$

where $W_t^{(1)}(\mathbb{Q})$ and $W_t^{(2)}(\mathbb{Q})$ are independent standard Brownian motions under $\mathbb{Q}$.

Although the change of measure for Brownian motion only involves changing the drift term, the change of measure for Lévy processes is much more complicated. The important result of Sato (1999) (given in the Appendix) provides the theoretical foundation for the change of measure of Lévy processes considered in this paper. To apply Sato's (1999) general theorem to our setting, some restrictions on model structures have to be imposed.

Under AJD models, jumps under both $\mathbb{P}$ and $\mathbb{Q}$ follow the same compound Poisson processes with different parameters. To have a fair comparison with AJD models, we restrict Lévy jumps under $\mathbb{P}$ and $\mathbb{Q}$ to follow the same Lévy process. That is, if the Lévy jump under $\mathbb{P}$ is VG (LS), then the Lévy jump under $\mathbb{Q}$ has to be VG (LS) as well, although with possible different parameters. Under this restriction, the Radon-Nikodym derivative between $\mathbb{P}$ and $\mathbb{Q}$ generally will be different from that of Wu (2006a) and Bakshi and Wu (2005). Based on the general result of Sato (1999) and our specific model restriction, we obtain the following results on the change of measure for the four jump processes considered in our paper.

PROPOSITION 2.1. *The parameters of the following four jump processes under measures* $\mathbb{P}$ *and* $\mathbb{Q}$ *must satisfy the following restrictions:*

- *All parameters of MJ,* $(\lambda, \mu_y, \sigma_y)$*, can change freely between* $\mathbb{P}$ *and* $\mathbb{Q}$*;*
- *All parameters of CMJ,* $(\lambda, \mu_y, \sigma_y, \rho_J, \mu_v)$*, can change freely between* $\mathbb{P}$ *and* $\mathbb{Q}$*;*
- *Among the parameters of VG,* (ν, γ, σ)*,* γ *and* σ *can change freely between* $\mathbb{P}$ *and* $\mathbb{Q}$*,* while ν *has to be the same under* $\mathbb{P}$ *and* $\mathbb{Q}$*;*
- *None of the parameters of a Lévy* α*-stable process,* $(\alpha, \beta, \sigma, \gamma)$*, can change between* $\mathbb{P}$ *and* $\mathbb{Q}$*.*[9]

The above results impose restrictions on the physical and the risk-neutral parameters of the four jump processes. For MJ and CMJ, all parameters can take different values under the physical and the risk-neutral measures. Previous studies, such as Pan (2002) and Eraker (2004), show that allowing all the parameters to change between measures makes econometric identification difficult. As a result, they only allow the mean jump size μ_y to be different between $\mathbb{P}$ and $\mathbb{Q}$. To compare our results with existing studies, we follow the same approach. As a result, the parameters of MJ and CMJ under both measures are $(\lambda, \sigma_y, \mu_y, \mu_y^{\mathbb{Q}})$ and $(\lambda, \sigma_y, \sigma_v, \rho_J, \mu_y, \mu_y^{\mathbb{Q}})$, respectively. The parameters of VG and LS under both measures are $(\nu, \gamma, \sigma, \gamma^{\mathbb{Q}}, \sigma^{\mathbb{Q}})$ and (α, σ), respectively.

If the Lévy measures of the four jump processes under $\mathbb{P}$ and $\mathbb{Q}$ satisfy the restrictions in Proposition 2.1, then the Radon-Nikodym derivatives of these processes are given as e^{U_t}, where U_t is defined as in the second part of Sato's (1999) theorem.[10] Combining this with the change of measure for the two Brownian motions, we obtain the Radon-Nikodym derivatives for the AJD and Lévy jump models:

$$\begin{aligned} \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_t = \exp\Bigg\{ & -\int_0^t \gamma_s^{(1)}\, dW_s^{(1)}(\mathbb{P}) - \int_0^t \gamma_s^{(2)}\, dW_s^{(2)}(\mathbb{P}) \\ & -\frac{1}{2}\left(\int_0^t \gamma_s^{(1)2}\, ds + \int_0^t \gamma_s^{(2)2}\, ds\right)\Bigg\} \exp U_t. \end{aligned} \tag{2.9}$$

This naturally leads to the risk-neutral return dynamics of all four models we consider

$$dY_t = \left[r_t - \frac{1}{2}v_t + \psi_J^{\mathbb{Q}}(-i)\right]dt + \sqrt{v_t}\, dW_t^{(1)}(\mathbb{Q}) + dJ_t^y(\mathbb{Q}), \tag{2.10}$$

$$dv_t = [\kappa(\theta - v_t) + \eta^v v_t]dt + \sigma_v\sqrt{v_t}\big(\rho dW_t^{(1)}(\mathbb{Q}) + \sqrt{1-\rho^2}\, dW_t^{(2)}(\mathbb{Q})\big) + dJ_t^v(\mathbb{Q}), \tag{2.11}$$

where $J_t^v(\mathbb{Q}) = 0$ for SVMJ, SVVG, and SVLS. The drift term of the return process under $\mathbb{Q}$ has three components: the risk-free interest rate r_t, the Ito adjustment for log price $-\frac{1}{2}v_t$, and the jump compensator in returns $\psi_J^{\mathbb{Q}}(-i)$ under $\mathbb{Q}$. Consequently, the drift term of the return process under $\mathbb{P}$ equals $\mu = r_t - \frac{1}{2}v_t + \psi_J^{\mathbb{Q}}(-i) + \eta^s v_t$.[11]

Option prices are determined by the risk-neutral dynamics of stock returns. Carr and Wu (2004) show that Lévy processes are as tractable as AJD models for the purpose of option pricing: The risk-neutral dynamics in (2.10) and (2.11) lead to closed-form

[9] The proof of the Proposition involves straightforward verification of the conditions of Sato's theorem for each of the jump processes and is available from the authors upon request.

[10] Because of the restriction that jumps under $\mathbb{P}$ and $\mathbb{Q}$ have to follow the same process (with different parameters), the four jump models have different U_ts as well. Therefore, the Lévy jump models differ from the AJD models not only in jump structures but also in the Radon-Nikodym derivatives of the jump processes.

[11] The explicit expressions of $\psi_J^{\mathbb{Q}}(\cdot)$ of the four jump processes are given in the Appendix.

solution to the characteristic function of the log stock price under $\mathbb{Q}$. That is, when interest rate is constant,

$$\begin{aligned}\phi_t(u) &= E_0^{\mathbb{Q}}\big[e^{iuY_t}\big] = E_0^{\mathbb{Q}}\big[e^{iuY_0+iu(r+\psi_J^{\mathbb{Q}}(-i))t+iu(\int_0^t \sqrt{v_t}\,dW_s^{(1)}(\mathbb{Q})-\frac{1}{2}\int_0^t v_s\,ds)+iuJ_t^y}\big]\\ &= e^{iuY_0+iu(r+\psi_J^{\mathbb{Q}}(-i))t}E_0^{\mathbb{Q}}\big[e^{iuJ_t^y}\big]E_0^{\mathbb{Q}}\big[e^{iu(\int_0^t \sqrt{v_t}\,dW_s^{(1)}(\mathbb{Q})-\frac{1}{2}\int_0^t v_s\,ds)}\big]\\ &= e^{iuY_0+iu(r+\psi_J^{\mathbb{Q}}(-i))t}e^{-t\psi_J^{\mathbb{Q}}(u)}e^{-b(t)v_0-c(t)},\end{aligned}$$

where

$$b(t) = \frac{(iu+u^2)(1-e^{-\delta t})}{(\delta+\kappa^M)+(\delta-\kappa^M)e^{-\delta t}},\quad c(t) = \frac{\kappa\theta}{\sigma_v^2}\left[2\ln\frac{2\delta-(\delta-\kappa^M)(1-e^{-\delta t})}{2\delta}+(\delta-\kappa^M)t\right],$$

$$\kappa^M = \kappa-\eta^v-iu\sigma_v\rho,\quad \delta = \sqrt{(\kappa^M)^2+(iu+u^2)\sigma_v^2},\quad \text{and}\quad Y_0 = \log(S_0).$$

The closed-form expression of the characteristic function of the log stock price naturally leads to closed-form expression of the Fourier transform of option prices. Consequently, option price can be solved using the Fourier inversion formula. The time-0 price of a European call option with time-to-maturity of τ and strike price of K equals

$$F(Y_0, v_0, \tau, K) = E_0^{\mathbb{Q}}[e^{-r\tau}(S_\tau-K)^+] = \frac{e^{-r\tau}}{\pi}\times\mathrm{Re}\left(\int_0^\infty e^{-ix\log(K)}\frac{\phi_\tau(x-i)}{-x^2+ix}\,dx\right).$$

In addition to the contractual terms of the option, the option price also depends on the current levels of the stock price (Y_0) and the instantaneous stochastic volatility (v_0).

3. MCMC ESTIMATION OF LÉVY JUMP MODELS USING SPOT AND OPTION PRICES

In this section, we discuss Bayesian MCMC estimation of Lévy jump models using spot and option prices. We first summarize the specifications of all models considered in our analysis. Then we discuss the statistical methods used for model estimation and comparison.

3.1. Summary of Model Specifications

In our joint estimation of Lévy jump models, we use daily returns on the S&P 500 index and daily prices of a short-term ATM SPX option. Let $C(t, \tau, K)$ be the market price at t of the option with time-to-maturity τ and strike price K, and $F(t, \tau, K, Y_t, v_t; \mathbf{\Theta})$ be the theoretical price of the same option in a given model where the log stock price equals Y_t, the instantaneous volatility equals v_t, and the vector of model parameters is denoted as $\mathbf{\Theta}$. We assume that the market price of the option equals its theoretical price plus some random noises:

$$C(t, \tau, K) = F(t, \tau, K, Y_t, v_t; \mathbf{\Theta}) + \varpi_t^c,$$

where $\varpi_t^c \sim N(\rho_c\varpi_{t-1}^c, \sigma_c^2)$. The first-order autocorrelation in option pricing errors also has been considered in Eraker (2004) and captures the phenomenon that if option pricing error is high on one day, it is likely to be high on the next day.

We consider first-order Euler discretization of the continuous-time models at daily frequency. Simulation studies in EJP (2003) and LWY (2008) show that the bias introduced by daily discretization is very small. Therefore, the joint dynamics of the daily spot and the option prices under the four models we consider are summarized by the following system of equations:

$$
\begin{cases}
C_{t+1} - F_{t+1} = \rho_c(C_t - F_t) + \sigma_c \epsilon_t^c, \\
Y_{t+1} = Y_t + \mu\Delta + \sqrt{v_t \Delta}\epsilon_{t+1}^y + J_{t+1}^y, \\
v_{t+1} = v_t + \kappa(\theta - v_t)\Delta + \sigma_v\sqrt{v_t\Delta}\epsilon_{t+1}^v + J_{t+1}^v,
\end{cases} \tag{3.1}
$$

where $\Delta = \frac{1}{252}$, $\mu = r_t - \frac{1}{2}v_t + \psi_J^{\mathbb{Q}}(-i) + \eta^s v_t$, $\epsilon_t^c, \epsilon_{t+1}^y$, and $\epsilon_{t+1}^v \sim N(0,1)$, $\mathrm{corr}(\epsilon_{t+1}^y, \epsilon_{t+1}^v) = \rho$, and ϵ_t^c is independent of ϵ_{t+1}^y and ϵ_{t+1}^v.

Specializing (12) to each of the four models, we have the following exact specifications of each model.

- **SVMJ.** In this model, $J_{t+1}^y = \xi_{t+1}^y N_{t+1}^y$, $P(N_{t+1}^y = 1) = \lambda\Delta$, $\xi_{t+1}^y \sim N(\mu_y, \sigma_y^2)$, and $J_{t+1}^v = 0$ for all t. We have observations $(Y_t, C_t)_{t=0}^T$; latent volatility variables $(v_t)_{t=0}^T$, jump times $(N_t^y)_{t=1}^T$, and jump sizes $(\xi_t^y)_{t=1}^T$; and parameters $\mathbf{\Theta} = \{(\kappa, \theta, \sigma_v, \rho, \mu_y, \sigma_y, \lambda), (\mu_y^{\mathbb{Q}}), (\eta^s, \eta^v), (\rho_c, \sigma_c)\}$, where the first group of parameters is either common to both measures or unique to the physical measure, the second one is unique to the risk-neutral measure, the third one represents the market prices of return and volatility risks, and the last one represents option pricing errors.
- **SVCMJ.** In this model, $J_{t+1}^y = \xi_{t+1}^y N_{t+1}$, $J_{t+1}^v = \xi_{t+1}^v N_{t+1}$, $P(N_{t+1} = 1) = \lambda\Delta$, $\xi_{t+1}^v \sim \exp(\mu_v)$, and $\xi_{t+1}^y \mid \xi_{t+1}^v \sim N(\mu_y + \rho_J \xi_{t+1}^v, \sigma_y^2)$. We have observations $(Y_t, C_t)_{t=0}^T$; latent volatility variables $(v_t)_{t=0}^T$, jump times $(N_t)_{t=1}^T$, and jump sizes $(\xi_t^v)_{t=1}^T$ and $(\xi_t^y)_{t=1}^T$; and parameters $\mathbf{\Theta} = \{(\kappa, \theta, \sigma_v, \rho, \mu_y, \sigma_y, \lambda, \rho_J, \mu_v), (\mu_y^{\mathbb{Q}}), (\eta^s, \eta^v), (\rho_c, \sigma_c)\}$, where the first group of parameters is either common to both measures or unique to the physical measure, the second one is unique to the risk-neutral measure, the third one represents the market prices of return and volatility risks, and the last one represents option pricing errors.
- **SVVG.** In this model, $J_{t+1}^v = 0$ for all t, and J_{t+1}^y follows a VG process whose discretized version is

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

  where $\epsilon_{t+1}^J \sim N(0,1)$ and $G_{t+1} \sim \Gamma(\frac{\Delta}{\nu}, \nu)$.$\epsilon_{t+1}^J$ and G_{t+1} are independent of each other and are independent of ϵ_{t+1}^y and ϵ_{t+1}^v. The parametrization of the Gamma distribution, $\Gamma(\alpha, \beta)$, used in this paper has density form $\frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$. We have observations $(Y_t, C_t)_{t=0}^T$; latent volatility variables $(v_t)_{t=0}^T$, jump times/sizes $(J_t^y)_{t=1}^T$, and time-change variables $(G_t)_{t=1}^T$; and parameters $\mathbf{\Theta} = \{(\kappa, \theta, \sigma_v, \rho, \nu, \gamma, \sigma), (\gamma^{\mathbb{Q}}, \sigma^{\mathbb{Q}}), (\eta^s, \eta^v), (\rho_c, \sigma_c)\}$, where the first group of parameters is either common to both measures or unique to the physical measure, the second one is unique to the risk-neutral measure, the third one represents the market prices of return and volatility risks, and the last one represents option pricing errors.
- **SVLS.** In this model, $J_{t+1}^v = 0$ for all t. The jump size J_{t+1}^y, independent of ϵ_{t+1}^y and ϵ_{t+1}^v, follows a stable distribution with shape parameter α, skewness parameter -1, zero drift, and scale parameter $\sigma\Delta^{\frac{1}{\alpha}}$. That is, $J_{t+1}^y \sim S_\alpha(-1, \sigma\Delta^{\frac{1}{\alpha}}, 0)$. We have observations $(Y_t, C_t)_{t=1}^T$; latent volatility variables $(v_t)_{t=0}^T$, and jump times/sizes $(J_t^y)_{t=1}^T$; and parameters $\mathbf{\Theta} = \{(\kappa, \theta, \sigma_v, \rho, \alpha, \sigma), (\eta^s, \eta^v), (\rho_c, \sigma_c)\}$, where the first group of

parameters is either common to both measures or unique to the physical measure, the second one represents the market prices of return and volatility risks, and the last one represents option pricing errors.

3.2. MCMC Methods

Estimation of Lévy processes is generally very difficult for several reasons. First, the probability densities for most Lévy processes are not known in closed form, and for certain Lévy processes higher moments of asset returns do not even exist. Second, the high dimensionality of latent variables, such as stochastic volatility, jump sizes, and jump times, significantly complicates the estimation. Computationally it is very demanding to integrate out the large number of latent variables when implementing either likelihood or moment-based approaches. The inclusion of option prices significantly increases the computational complexity because certain parameters enter into the option pricing formulae nonlinearly, and the computation of option prices involves numerical integrations.

LWY (2008) have developed efficient Bayesian MCMC methods for estimating Lévy processes using only the spot price.[12] We extend their methods to estimate the physical and risk-neutral dynamics of Lévy processes jointly using spot and option prices. The main difference here is that we need to rely on more sophisticated updating procedures for many model parameters and latent variables due to the nonlinear option pricing formula involved.

Because MCMC analysis of SVMJ and SVCMJ has been considered in previous studies, such as EJP (2003) and Eraker (2004), 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.

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

$$\begin{aligned} p(\mathbf{\Theta}, V, \mathbf{J}, \mathbf{G} \mid \mathbf{Y}, \mathbf{C}) &\propto p(\mathbf{Y}, \mathbf{C}, \mathbf{V}, \mathbf{J}, \mathbf{G}, \mathbf{\Theta}) \\ &= p(\mathbf{C} \mid \mathbf{Y}, \mathbf{V}, \mathbf{\Theta}) p(\mathbf{Y}, \mathbf{V} \mid \mathbf{J}, \Theta) p(\mathbf{J} \mid \mathbf{G}, \mathbf{\Theta}) p(\mathbf{G} \mid \mathbf{\Theta}) p(\mathbf{\Theta}). \end{aligned}$$

Given the assumed option price dynamics, we have

$$p(\mathbf{C} \mid \mathbf{Y}, \mathbf{V}, \Theta) = \prod_{t=0}^{T-1} \frac{1}{\sqrt{2\pi}\sigma_c} \exp\left\{ -\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2} \right\}.$$

Conditioning on v_t and J_{t+1}^y, $Y_{t+1} - Y_t$ and $v_{t+1} - v_t$ follow a bivariate normal distribution

[12] Earlier studies, such as Jacquier, Polson, and Rossi (1994), Kim, Shephard, and Chib (1998), Chib, Nardari, and Shephard (2002), and Maheu and McCurdy (2004) among others, apply MCMC methods to estimate discrete-time stochastic volatility models.

$$\begin{pmatrix} Y_{t+1} - Y_t \\ v_{t+1} - v_t \end{pmatrix} \Bigg| v_t, J^y_{t+1} \sim N\left(\begin{pmatrix} \mu\Delta + J^y_{t+1} \\ \kappa(\theta - v_t)\Delta \end{pmatrix}, v_t\Delta \begin{pmatrix} 1 & \rho\sigma_v \\ \rho\sigma_v & \sigma_v^2 \end{pmatrix} \right),$$

$$J^y_{t+1} \mid G_{t+1}, \mathbf{\Theta} \sim N(\gamma G_{t+1}, \sigma^2 G_{t+1}) \quad \text{and} \quad G_{t+1} \mid \mathbf{\Theta} \sim \Gamma\left(\frac{\Delta}{\nu}, \nu\right).$$

Therefore, the joint posterior distribution of parameters and latent variables is given as

$$\begin{aligned} p(\mathbf{\Theta}, V, \mathbf{J}, \mathbf{G} \mid \mathbf{Y}, \mathbf{C}) \propto & \prod_{t=0}^{T-1} \frac{1}{\sqrt{2\pi}\sigma_c} \exp\left\{ -\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2} \right\} \\ & \times \prod_{t=0}^{T-1} \frac{1}{\sigma_v v_t \Delta \sqrt{1-\rho^2}} \\ & \times \exp\left\{ -\frac{1}{2(1-\rho^2)} \left(\left(\epsilon^y_{t+1}\right)^2 - 2\rho\epsilon^y_{t+1}\epsilon^v_{t+1} + \left(\epsilon^v_{t+1}\right)^2 \right) \right\} \\ & \times \prod_{t=0}^{T-1} \frac{1}{\sigma\sqrt{G_{t+1}}} \exp\left\{ -\frac{(J_{t+1} - \gamma G_{t+1})^2}{2\sigma^2 G_{t+1}} \right\} \\ & \times \prod_{t=0}^{T-1} \frac{1}{\nu^{\frac{\Delta}{\nu}} \Gamma\left(\frac{\Delta}{\nu}\right)} G_{t+1}^{\frac{\Delta}{\nu}-1} \exp\left\{ -\frac{G_{t+1}}{\nu} \right\} \times p(\mathbf{\Theta}), \end{aligned}$$

where $\epsilon^y_{t+1} = (Y_{t+1} - Y_t - \mu\Delta - J^y_{t+1})/\sqrt{v_t\Delta}$ and $\epsilon^v_{t+1} = (v_{t+1} - v_t - \kappa(\theta - v_t)\Delta)/(\sigma_v\sqrt{v_t\Delta})$.

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

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

$$S_{t+1} \sim S_\alpha\left(-1, \sigma\Delta^{\frac{1}{\alpha}}, 0\right).$$

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}, $p(S_{t+1} \mid \mathbf{\Theta})$, 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 $p(Y_{t+1}, v_{t+1}, S_{t+1} \mid \mathbf{\Theta}) = p(Y_{t+1}, v_{t+1} \mid S_{t+1}, \mathbf{\Theta}) p(S_{t+1} \mid \mathbf{\Theta})$. Consequently, it is difficult to obtain the joint posterior distribution for SVLS.

Buckle (1995) provides a representation of a 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 option prices as $\mathbf{C} = \{C_t\}_{t=0}^T$, 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

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

where $\epsilon_{t+1}^y = (Y_{t+1} - Y_t - \mu\Delta - S_{t+1})/\sqrt{v_t \Delta}$, $\epsilon_{t+1}^v = (v_{t+1} - v_t - \kappa(\theta - v_t)\Delta)/(\sigma_v \sqrt{v_t \Delta})$, $l_\alpha = \frac{\alpha-2}{2\alpha}$, and $t_\alpha(U_{t+1}) = (\frac{sin[\pi\alpha U_{t+1} + \frac{(2-\alpha)\pi}{2}]}{cos[\pi U_{t+1}]})(\frac{cos[\pi U_{t+1}]}{cos[\pi(\alpha-1)U_{t+1} + \frac{(2-\alpha)\pi}{2}]})^{(\alpha-1)/\alpha}$. We obtain joint posterior samples of $\boldsymbol{\Theta}$, V, $\mathbf{S}$, and $\mathbf{U}$ by simulating from the above joint posterior density. We then marginalize $\mathbf{U}$ out to obtain the samples for $\boldsymbol{\Theta}$, V, and $\mathbf{S}$. That is, we simply throw away the observations of $\mathbf{U}$ and retain the observations of $\boldsymbol{\Theta}$, 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 $p(\Theta_i \mid \boldsymbol{\Theta}_{-i}, \mathbf{J}, \mathbf{G}, \mathbf{V}, \mathbf{Y}, \mathbf{C})$ for $i = 1, \ldots, k$, where Θ_i is the ith element of $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k)$, the posterior distribution for jump times $p(J_t^y \mid \Theta, \mathbf{G}, \mathbf{V}, \mathbf{Y}, \mathbf{C})$, jump sizes $p(G_t \mid \Theta, \mathbf{J}, \mathbf{V}, \mathbf{Y}, \mathbf{C})$, and latent volatility variables $p(v_t \mid v_{t+1}, v_{t-1}, \boldsymbol{\Theta}, J, G, Y, C)$, 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.

In an interesting paper, Griffin and Steel (2006) have developed MCMC methods for estimating an Ornstein–Uhlenbeck (OU) volatility process driven by a positive Lévy process without Gaussian component. They rely on a series representation of Lévy processes for drawing latent volatility variables. Their approach requires the inverse tail mass function of Lévy process to be known analytically.

3.3. Model Diagnostics and Comparisons

The posterior estimates of model parameters and latent state variables allow us to examine the performances of all four models in capturing the joint dynamics of spot and option prices.

One way to gauge the performances of each model in capturing the spot price is to test whether the standardized model residuals of both returns and volatility follow an $N(0,1)$ distribution as in EJP (2003) and LWY (2008). For example, for SVLS, if the model is correctly specified, then

$$\frac{Y_{t+1}-Y_t-\mu\Delta-S_{t+1}}{\sqrt{v_t\Delta}}=\epsilon^y_{t+1}\sim N(0,1)\quad\text{and}$$
$$\frac{v_{t+1}-v_t-\kappa(\theta-v_t)\Delta}{\sigma_v\sqrt{v_t\Delta}}=\epsilon^v_{t+1}\sim N(0,1).$$

Deviations of ϵ^y_{t+1} and ϵ^v_{t+1} from $N(0,1)$ can reveal rich information on potential sources of model misspecifications.

To compare the performances of different models in capturing the risk-neutral dynamics, we test whether one model has significantly smaller option pricing errors than another. For this purpose, we adopt an approach developed by Diebold and Mariano (1995) (hereafter DM) in time series forecasting literature. Consider two models whose associated true daily option pricing errors (calculated at true model parameters) are $\{e_1(t)\}_{t=1}^T$ and$\{e_2(t)\}_{t=1}^T$, respectively. The null hypothesis that the two models have the same squared pricing errors is $E[e_1^2(t)]=E[e_2^2(t)]$, or $E[d(t)]=0$, where $d(t)=e_1^2(t)-e_2^2(t)$. DM (1995) show that if $\{d(t)\}_{t=1}^T$ is covariance stationary and short memory, then

$$\sqrt{T}(\overline{d}-\mu_d)\sim N(0,2\pi f_d(0)),$$

where $\overline{d}=\frac{1}{T}\sum_{t=1}^T[\varepsilon_1^2(t)-\varepsilon_2^2(t)]$, $f_d(0)=\frac{1}{2\pi}\sum_{q=-\infty}^{\infty}\gamma_d(q)$, and $\gamma_d(q)=E[(d_t-\mu_d)(d_{t-q}-\mu_d)]$. In large samples, $\overline{d}$ is approximately normally distributed with mean μ_d and variance $2\pi f_d(0)/T$. Thus, under the null hypothesis of equal squared pricing errors, the following DM statistic

$$DM=\frac{\overline{d}}{\sqrt{2\pi\widehat{f_d}(0)/T}}$$

is distributed asymptotically as $N(0,1)$, where $\widehat{f_d}(0)$ is a consistent estimator of $f_d(0)$.[13]

In empirical analysis, however, we do not observe the true pricing errors $\{e_1(t)\}_{t=1}^T$ and$\{e_2(t)\}_{t=1}^T$. Instead we only observe the estimated pricing errors (calculated at estimated model parameters) $\{\hat{e}_1(t)\}_{t=1}^T$ and $\{\hat{e}_2(t)\}_{t=1}^T$. Because of parameter estimation uncertainty, $E[\hat{e}_i^2(t)]\neq E[e_i^2(t)]$, for $i=1,2$. To address this issue, we use modified pricing errors $\sqrt{\frac{T}{T-k_i}}\hat{e}_i(t)$ in our implementation of the DM test, where k_i represents the number of parameters for model i. This approach is based on the fact that in both linear and nonlinear regressions, $\frac{1}{T-k_i}\sum_{t=1}^T\hat{e}_i^2(t)$ is an unbiased estimator of $E(e_i^2(t))$ as $T\to\infty$, for $i=1,2$. Our approach not only takes into account parameter estimation uncertainty but also penalizes more complex models with a larger number of parameters.[14]

[13] We estimate the variance of the test statistic using the Bartlett estimate of Newey and West (1987) with a lag order of 50.

[14] We thank Wayne Fuller for suggesting this approach. For references, see Casella and Berger (2001) and Gallant (1987). We acknowledge that our argument for the modified DM statistic is mainly heuristic and may not completely address the problem of parameter estimation uncertainty.

To compare the overall performances of the two models, we use the DM statistic to measure whether one model has significantly smaller squared option pricing errors than another. We also use the DM statistic to measure whether one model has smaller squared pricing errors than another for options in a specific moneyness and maturity group.

4. EMPIRICAL RESULTS

In this section, we provide empirical analysis of the four models (SVMJ, SVCMJ, SVVG, and SVLS) using the spot and option prices of the S&P 500 index. We first introduce the data used in our analysis. We then examine the performances of the four models based on their (i) posterior means of model parameters and latent volatility/jump variables; (ii) empirical fits of the spot price; and (iii) in-sample and out-of-sample option pricing errors.

4.1. The Data

We use the same data as that in Aıt-Sahalia and Lo (1998), which include daily spot and option prices of the S&P 500 index between January 4, 1993 and December 31, 1993. Aıt-Sahalia and Lo (1998) take the midpoint of the bid and ask prices of each option as observed market price and eliminate observations with time-to-maturity less than 1 day, implied volatility greater than 70%, and price less than $\frac{1}{8}$. To deal with potential nonsynchronous trading and unobservable dividend yield, they back out the futures price of the underlying index at the time the option prices are observed. They obtain prices of calls and puts that have the same time-to-maturity and strike price and are closest to the money. Using put-call parity, they solve for the futures price at that certain maturity, which then can be used to back out the implied dividend yield via the cost-of-carry relation.[15]

Our estimation uses daily returns of the S&P 500 index and daily prices of a short-term ATM SPX option that we choose for each day.[16] We require that the option has a time-to-maturity between 20 and 50 days and is closest to the money, that is, its strike to spot price ratio is closest to one.[17] On a few days without such options, we use an option whose time-to-maturity is closest to 20 days. Table 4.1 provides summary statistics on the data used directly in our estimation. During 1993, the mean and standard deviation of annualized continuously compounded daily returns of the index are 7.36% and 8.94%, respectively. Index returns exhibit slight negative skewness and high kurtosis. The mean and median time-to-maturities of the short-term options are 34 and 35 days, respectively, whereas the shortest and longest time-to-maturities are 16 and 50 days, respectively. The price of the options has a mean of \$7.14 and a range between \$3.44 and \$10.72. The implied volatility has a mean of 9.2% and a range between 6.7% and 12.23%. The ratio between the strike and the spot price of the short-term option is very close to 1. Aıt-Sahalia and Lo (1998) note that the short-term interest rates exhibit little variation during 1993, ranging from 2.85% to 3.21%. As a result, we assume constant interest rate in our estimation and use the prevailing interest rate each day in our pricing formula.

[15] See Aït-Sahalia and Lo (1998) for more detailed descriptions of the data set.

[16] Short-term ATM options are among the most liquid options and should have the most efficient prices in the market.

[17] Because the time-to-maturity of an option changes daily, we have to use different options on different days in our estimation.

TABLE 4.1
Summary Statistics of Spot and Option Prices of the S&P 500 Index

	Mean	Variance	Skewness	Kurtosis	Min	Max
Panel A. Summary statistics of continuously compounded daily returns of the S&P 500 index between January 4, 1993 and December 31, 1993.						
S&P 500	0.000292	0.0000316	−0.0332	5.5602	−0.0256	0.0223

	Mean	Median	Std. Dev.	Min	Max
Panel B. Summary statistics for the short-term ATM SPX option used in model estimation between January 4, 1993 and December 31, 1993.					
Time-to-maturity	34	35	9.24	16	50
Option price	7.14	7.25	1.61	3.44	10.72
Implied volatility	0.092	0.0914	0.0095	0.0679	0.1223
Strike price	449.8207	450	10.3086	425	450
Spot price	450.0755	448.394	10.1711	427.0155	470.0928
Moneyness (strike/spot)	0.9994	0.9994	0.003	0.9946	1.0055

Note: This table provides summary statistics of spot and option prices of the S&P 500 index between January 4, 1993 and December 31, 1993. Panel A reports summary statistics of continuously compounded daily returns of the S&P 500 index during the sample period. Panel B reports summary statistics on time-to-maturity, price, implied volatility, strike price, spot price, and moneyness (strike/spot) of the short-term ATM SPX option used in model estimation. We restrict the time-to-maturity of the option to be between 20 and 50 days. On a few days without such options, we use an option whose time-to-maturity is closest to 20 days. Because the time-to-maturity of an option changes daily, in general we have to use different options on different dates.

Figure 4.1 provides time series plots of the level and log change of the S&P 500 index, and the implied volatility of the short-term SPX options. The level of the index has increased steadily during 1993, with occasional relatively large negative returns, although none is as large as that of several major stock market crashes in other periods. The implied volatility fluctuates between 5% and 15% during 1993 with strong mean reversion.

4.2. Estimates of Model Parameters and Latent Volatility/Jump Variables

Table 4.2 reports posterior estimates of (i) model parameters under both the physical and the risk-neutral measures; (ii) market prices of risks for the two Brownian shocks (η^v and η^s); and (iii) parameters describing option pricing errors (ρ_c and σ_c). Figures 4.2 and 4.3 provide time series plots of the filtered volatility and jump variables for the four models, respectively. The estimates of model parameters and latent variables reveal both similarities and differences among the four models.

Consistent with existing studies, all four models exhibit strong negative correlations between volatility and returns: The estimates of ρ for the four models range from −0.56 to

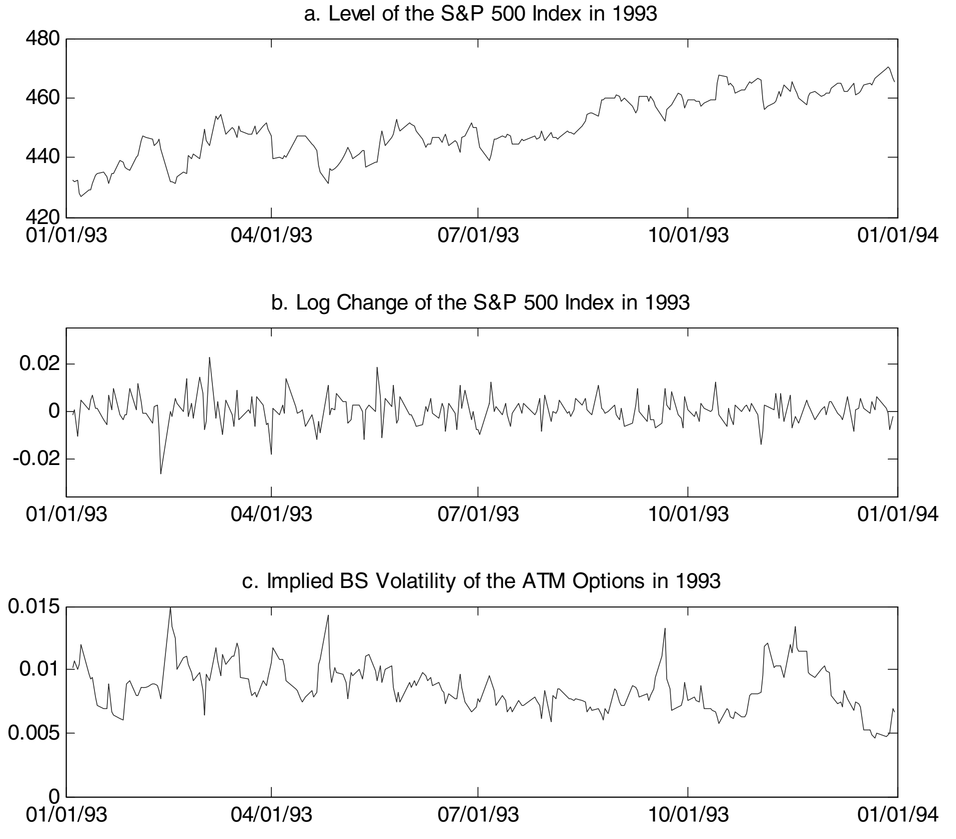

FIGURE 4.1. Level and log change of the S&P 500 index, and implied volatility of the short-term ATM SPX options used in model estimation between January 4, 1993 and December 31, 1993.

-0.82. The four models share similar estimates of the long-run mean (θ) of the volatility processes.[18] The estimates of the market prices of return and volatility risks are very similar across the four models and are similar to those in previous studies. For example, the estimates of η^s (η^v) in the four models are between 3.5 and 4.4 (2.9 and 4.8), whereas the estimate of η^s (η^v) in Pan (2002) equals 3.6 (3.1). The four models also share similar estimates of parameters describing option pricing errors (ρ_c and σ_c). In particular, the estimates of ρ_c in the four models are about 0.90, confirming that there is indeed strong autocorrelation in option pricing errors.

The four models also differ from each other in important ways. For example, the volatility process of SVVG has the strongest mean-reversion (κ) and the highest volatility of volatility (σ_v) among the four models.[19] The filtered volatility variables of the four models in Figure 4.2 confirm this fact and show that the other three models have much

[18] Because of jumps in volatility in SVCMJ, the long-run mean of volatility in this model should include the impact of jumps.

[19] The estimates of κ in this paper differ from that in LWY (2008) in magnitude mainly because we use a different scale on observables in our estimation. While LWY (2008) consider index returns in percentages, we express index returns in decimal points.

TABLE 4.2
Parameter Estimates of AJD and Lévy Jump Models

	SVMJ	SVCMJ	SVVG	SVLS
κ	2.6387	3.3627	15.778	6.2792
	(0.544)	(0.6452)	(1.3706)	(0.467)
θ	0.0049	0.0076	0.0060	0.0055
	(0.0030)	(0.0022)	(0.0011)	(0.0017)
σ_v	0.1198	0.1676	0.3043	0.1852
	(0.0116)	(0.0179)	(0.0315)	(0.0268)
ρ	−0.7014	−0.7786	−0.8167	−0.5619
	(0.0163)	(0.0324)	(0.0511)	(0.0746)
η^v	3.0526	1.1074	4.7128	2.9419
	(0.8005)	(0.5933)	(2.2753)	(1.336)
η^s	3.7020	4.3586	4.328	3.5962
	(2.7850)	(2.499)	(3.046)	(1.784)
ρ_c	0.8952	0.8665	0.895	0.9023
	(0.0557)	(0.0495)	(0.0584)	(0.0660)
σ_c	0.2039	0.2257	0.1869	0.2666
	(0.0275)	(0.0216)	(0.0189)	(0.0256)
λ	0.0103	0.0048	-	-
	(0.0216)	(0.0040)		
${\mu_y}^P$	0.0150	−0.03376	-	-
	(0.0108)	(0.0108)		
${\mu_y}^Q$	−0.3091	−0.3414	-	-
	(0.1294)	(0.0892)		
σ_y	0.0107	0.0103	-	-
	(0.0064)	(0.0063)		
μ_v	-	0.00849	-	-
		(0.0075)		
ρ_J	-	−0.0038	-	-
		(0.00492)		
ν	-	-	0.0142	-
			(0.0017)	
γ^P	-	-	0.0256	-
			(0.0315)	
σ^P	-	-	0.0462	-
			(0.0070)	
γ^Q	-	-	0.0030	-
			(0.0056)	
σ^Q	-	-	0.0412	-
			(0.0150)	
α	-	-	-	1.846
				(0.0012)
σ	-	-	-	0.0352
				(0.0014)

This table reports posterior estimates of model parameters of AJD and Lévy jump models using daily returns on the S&P 500 index and daily prices of a short-term ATM SPX option between January 4, 1993 and December 31, 1993. We discard the first 10,000 runs as "burn-in" period and use the last 90,000 iterations in MCMC simulations to estimate model parameters. Specifically, we take the mean of the posterior distribution as parameter estimate and the standard deviation of the posterior as standard error.

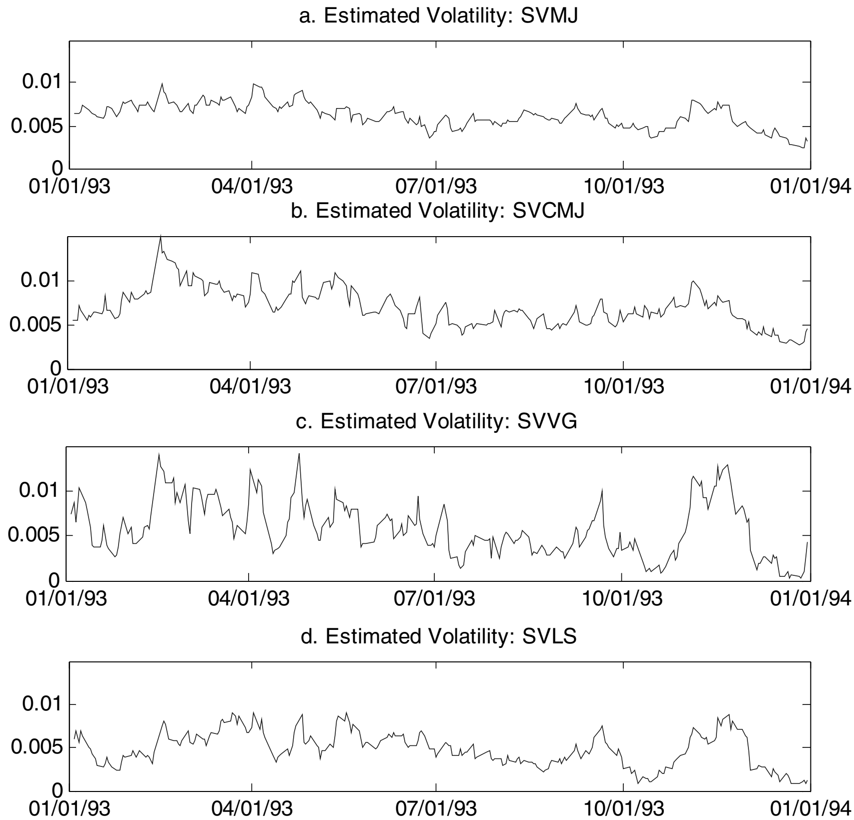

FIGURE 4.2. Estimated volatility variables of SVMJ, SVCMJ, SVVG, and SVLS using daily returns of the S&P 500 index and daily prices of the short-term ATM SPX options between January 4, 1993 and December 31, 1993.

smoother volatility factors. Interestingly, the filtered volatility variables of SVVG mimic the behavior of the implied volatilities of the short-term SPX options (shown in Figure 4.1) much more closely than that of the other three models.[20]

The AJD and Lévy jump models exhibit dramatically different jump behaviors. The estimated jump intensities (λ) for SVMJ and SVCMJ suggest that on average there are about one to two jumps per year. While the mean jump sizes under $\mathbb{P}\left(\mu_y^{\mathbb{P}}\right)$ in the two models are close to zero, the mean jump sizes under $\mathbb{Q}\left(\mu_y^{\mathbb{Q}}\right)$ are much more negative. The filtered jump sizes and times of the two models in Figure 4.3 also show that there are a few large jumps in returns (and volatility) in SVMJ (SVCMJ). On the other hand, Figure 4.3 shows that in addition to several large jumps, SVVG and SVLS also exhibit many frequent small jumps in returns. Hence, VG and LS have the advantage over MJ and CMJ in capturing both the infrequent large jumps as well as the frequent small jumps

[20] The simulation evidence in LWY (2008) shows that the MCMC methods can estimate the AJD and Lévy jump models very accurately using return data alone. The inclusion of option prices should make it even easier to identify model parameters. Therefore, the differences in parameter estimates across different models are an indication of model misspecification. Parameters of a misspecified model may have to take unreasonable values to capture certain features of the data that the model inherently cannot capture. This in turn could lead to large option pricing errors.

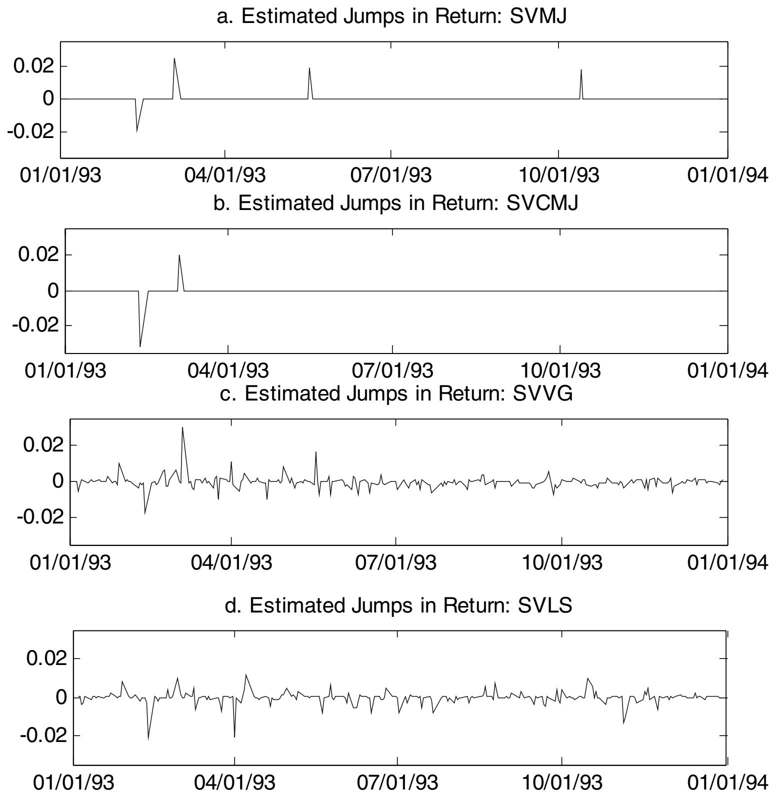

FIGURE 4.3. Estimated jumps in returns of SVMJ, SVCMJ, SVVG, and SVLS using daily returns of the S&P 500 index and daily prices of the short-term ATM SPX options between January 4, 1993 and December 31, 1993.

in returns. The risk-neutral jump distribution of VG is less positively skewed than its physical jump distribution, suggesting that jumps are less positive under $\mathbb{Q}$ than under $\mathbb{P}$. This fact suggests that LS is likely to underperform VG in modeling the joint dynamics of index returns because its parameters are restricted to be the same under both measures. The estimated jump risk premium in index returns is given by $\psi_J^{\mathbb{Q}}(-i) - \psi_J^{\mathbb{P}}(-i)$ for each model. The jump risk premiums for SVMJ and SVCMJ are 0.29% and 0.12%, respectively. The jump risk premium for SVVG is much higher at 2.28%, and by definition the jump risk premium for SVLS is zero.

4.3. Performances in Modeling the Spot Price

In this section, we examine the performances of the four models in capturing the physical dynamics of the S&P 500 index. Based on estimated model parameters and latent volatility/jump variables, we calculate the standardized residuals for both returns and volatility, ϵ_{t+1}^{y} and ϵ_{t+1}^{v}. If a given model is correctly specified, then the distributions of both residuals should be close to $N(0, 1)$.

Figure 4.4 (5) plots kernel density estimators of $\epsilon_{t+1}^{y}(\epsilon_{t+1}^{v})$ of each of the four models and the density function of $N(0, 1)$. For both SVMJ and SVCMJ, ϵ_{t+1}^{y} and ϵ_{t+1}^{v} exhibit clear

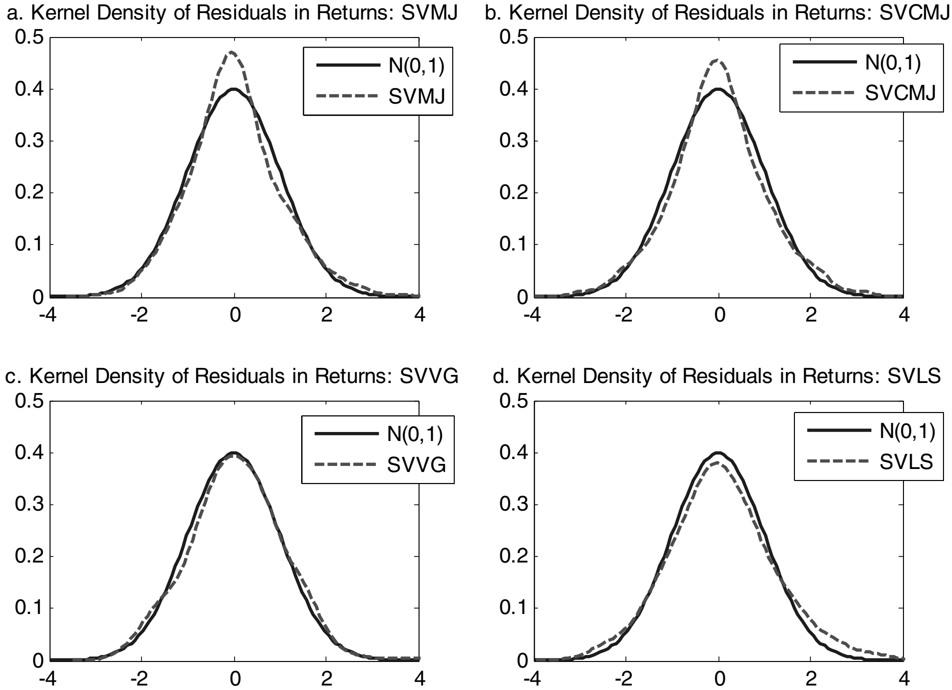

FIGURE 4.4. Kernel densities of standardized model residuals of returns of SVMJ, SVCMJ, SVVG, and SVLS, which are estimated using daily returns of the S&P 500 index and daily prices of the short-term ATM SPX options between January 4, 1993 and December 31, 1993.

deviations from standard normal: There is a high peak at the center of the distributions of both residuals, suggesting that the two models fail to capture the many small movements in both returns and volatility. On the other hand, the distributions of ϵ_{t+1}^{y} and ϵ_{t+1}^{v} of the two Lévy jump models are much closer to standard normal. The residuals of SVVG are closer to standard normal than that of SVLS. The fact that none of the parameters of LS can change between $\mathbb{P}$ and $\mathbb{Q}$ limits its ability in capturing the joint dynamics of index returns.

In addition to graphical illustrations, we also formally test whether ϵ_{t+1}^{y} and ϵ_{t+1}^{v} follow $N(0, 1)$ using the Kolmogorov–Smirnov (KS hereafter) test. For each set of the residuals, the KS test compares the empirical cumulative distribution function (CDF) with the CDF of $N(0, 1)$ and rejects the null hypothesis if the maximum distance between the two CDFs is too big. The KS tests in Table 4.3 reject the null hypothesis that ϵ_{t+1}^{y} and ϵ_{t+1}^{v} of SVMJ and SVCMJ follow a standard normal distribution. The p-values are between 3% and 4% for most cases, except that the p-value equals 5.37% for ϵ_{t+1}^{v} of SVCMJ. This suggests that including MJ jumps in volatility improves the modeling of the volatility process. Consistent with Figures 4.4 and 4.5, the KS test fails to reject the null hypothesis that ϵ_{t+1}^{y} and ϵ_{t+1}^{v} of the two Lévy models follow a standard normal distribution (p-values range from 25% to 38% for the two residuals under both models).

The above findings are consistent with the simulation and empirical evidence of LWY (2008). In particular, LWY (2008) fit SVMJ and SVCMJ to return data simulated from SVVG and SVLS using similar MCMC methods developed for return data only. They

TABLE 4.3
Kolmogorov–Smirnov Goodness-of-Fit Test of Model Residuals

	Return residuals				Volatility residuals			
	SVMJ	SVCMJ	SVVG	SVLS	SVMJ	SVCMJ	SVVG	SVLS
KS statistics	0.096	0.0934	0.0619	0.0695	0.0950	0.0893	0.0642	0.0592
P-values	0.0317	0.041	0.3246	0.2531	0.0305	0.0537	0.2902	0.3797

This table provides Kolmogorov–Smirnov (KS) tests of the hypotheses that the standardized model residuals of returns and volatility of each of the four models follow N(0,1). We report the KS statistics and their corresponding *p*-values for both residuals of all four models.

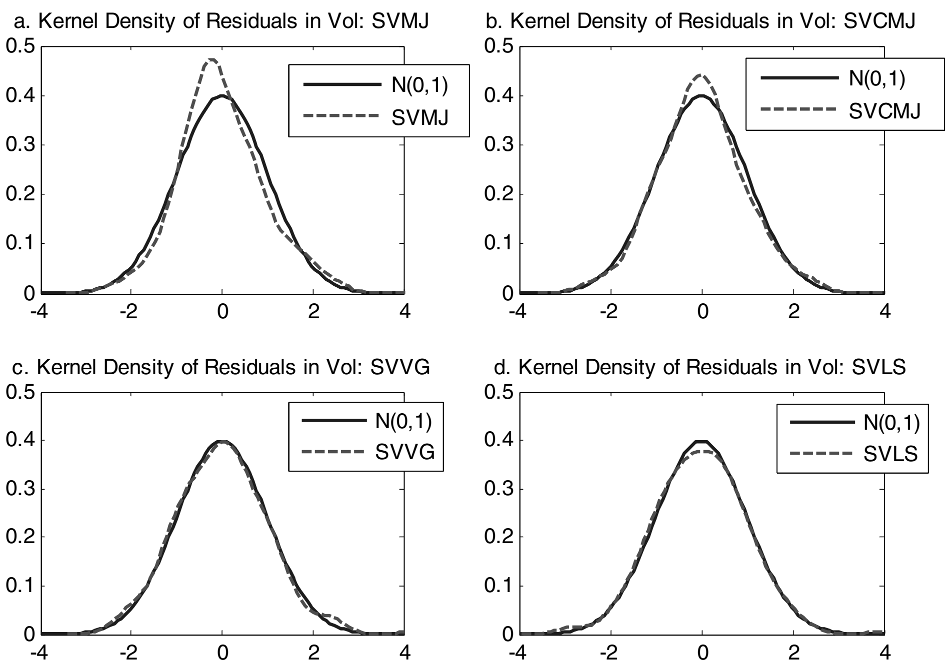

FIGURE 4.5. Kernel densities of standardized model residuals of volatility of SVMJ, SVCMJ, SVVG, and SVLS, which are estimated using daily returns of the S&P 500 index and daily prices of the short-term ATM SPX options between January 4, 1993 and December 31, 1993.

also estimate the four models using daily returns of the S&P 500 index between January 1980 and December 2000. The deviations from $N(0, 1)$ of return residuals $\left(\epsilon_{t+1}^{y}\right)$ from simulated and actual data for both SVMJ and SVCMJ documented in LWY (2008) are very similar to that observed in Figure 4.4. These results show that the parametric specifications of existing AJD models are not flexible enough to capture the many small movements in index returns. In contrast, because VG and LS can generate both large and small jumps, they can capture those movements that are too big for the diffusion part and too small for MJ/CMJ in the AJD models.[21]

4.4. Performances in Modeling Option Prices

There is no guarantee that a model that captures the physical dynamics better also can fit option prices better. For example, Eraker (2004) shows that while the double-jump model of EJP (2003) captures index returns better than SVMJ, it does not have significantly smaller option pricing errors. In this section, we address the basic question whether the Lévy jump models we consider can capture the joint dynamics of the S&P 500 index returns better than the AJD models.

Panel A of Table 4.4 reports the time series mean of daily absolute and percentage pricing errors of the short-term ATM SPX options used in model estimation for the four models.[22] We find similar pricing errors for SVMJ and SVCMJ: The mean absolute pricing errors of the two models are about 44 cents (the mean option price is $7.14); and the mean percentage pricing errors of the two models are about 6.3%, which is bigger than the percentage bid-ask spread of the option. On the other hand, the mean absolute pricing errors of SVVG and SVLS are about 16 and 24 cents, respectively, and the mean percentage pricing errors are about 2.4% and 3.6%, respectively. Consistent with the results of Eraker (2004), the DM statistics in Panel B of Table 4.4 show that the squared pricing errors of SVMJ and SVCMJ are not significantly different from each other. In contrast, SVVG and SVLS have significantly smaller squared pricing errors than SVMJ and SVCMJ, and SVVG has significantly smaller squared pricing errors than SVLS. The time series plots of the daily absolute pricing errors of the four models in Figure 4.6 show that SVVG and SVLS have smaller absolute pricing errors than SVMJ and SVCMJ during most of the sample period. In particular, SVVG has almost uniformly smaller in-sample option pricing errors than the AJD models. SVLS has somewhat worse performances than SVVG.[23] Panel C of Table 4.4 shows that the KS test fails to reject the null hypothesis at the 5% level that the option pricing errors ϵ_t^c follow $N(0, 1)$ for all models, confirming our econometric specification of option pricing errors.

In addition to the short-term ATM SPX options used in estimation, we also examine the performances of the four models in pricing 12,725 other options in the data set.[24] Because these options have not been used in model estimation, they provide evidence on the out-of-sample performances of the four models in option pricing. We divide all options into six moneyness groups, from deep in-the-money (ITM) to deep out-of-the-money (OTM) options, and five maturity groups, with time-to-maturities from less than 1 month to longer than 6 months. The majority of these options are ITM options with time-to-maturities between 1 and 6 months, and we do not observe many short-term deep

[21] We emphasize that although compound Poisson processes can approximate an infinite activity Lévy processes with arbitrary precision, such approximation would require a much richer specification of compound Poisson processes than those in the current AJD literature. The basic point of our paper is that the parametric specifications of compound Poisson processes in the current AJD literature are not as flexible as the Lévy jump models in capturing return dynamics.

[22] Absolute pricing error of an option is the absolute value of the difference between model and market prices of the option, and percentage pricing error of an option is the absolute pricing error divided by the market price of the option.

[23] We obtain very similar results using both absolute and percentage pricing errors. For the rest of the paper, we only report results based on absolute pricing errors.

[24] We eliminate options with prices that are less than one dollar.

TABLE 4.4
In-Sample Performances in Option Pricing

	SVMJ	SVCMJ	SVVG	SVLS
Panel A. Time series mean and standard deviation (in parentheses) of the absolute and percentage pricing errors of the short-term ATM options used in model estimation.				
Absolute (in dollars)	0.44	0.44	0.16	0.24
	(0.2913)	(0.3268)	(0.1189)	(0.1890)
Percentage	0.0629	0.0634	0.024	0.0361
	(0.0419)	(0.0467)	(0.0186)	(0.0329)

	VG-MJ	LS-MJ	VG-LS	VG-CMJ	LS-CMJ	MJ-CMJ
Panel B. Diebold–Mariano (DM) statistics for in-sample option pricing errors. The DM statistics measure whether the first model has significantly smaller squared pricing errors than the second model in each of the six pairs of models in the first row. Bold entries mean that the difference is significant at the 5% level for one-sided test. To save space, we omit “SV” in the names of all four models.						
DM Stats	**−2.2194**	**−1.9954**	**−2.1095**	**−2.1767**	**−1.9255**	−0.5283

	SVMJ	SVCMJ	SVVG	SVLS
Panel C. Kolmogorov–Smirnov test of the hypotheses that the standardized option pricing errors of each of the four models follow $N(0,1)$. We report the KS statistics and their corresponding p-values for each model.				
KS statistics	0.0846	0.0794	0.0800	0.0765
P-values	0.0525	0.0812	0.0771	0.1022

This table provides summary information on the in-sample performances of the four models in pricing the short-term ATM options used in model estimation. Absolute pricing error is defined as the absolute value of the difference between model and market prices of an option. Percentage pricing error is defined as the absolute pricing error of an option divided by the market price of the option.

OTM options. Based on the estimated model parameters and latent volatility variables, we calculate the theoretical price of each of these options under each model. Then based on options that are available on each day, we obtain daily arithmetic weighted average of absolute and percentage pricing errors for (i) all options; (ii) options within each of the moneyness groups (options across all maturities that belong to a certain moneyness group) or each of the maturity groups (options across all moneyness that belong to a certain maturity group); and (iii) options within each individual moneyness/maturity group.

We first examine the overall performances of the four models by focusing on the average pricing errors of the 12,725 out-of-sample options. The time series mean of daily weighted average of the absolute pricing errors of all options are reported in the last four

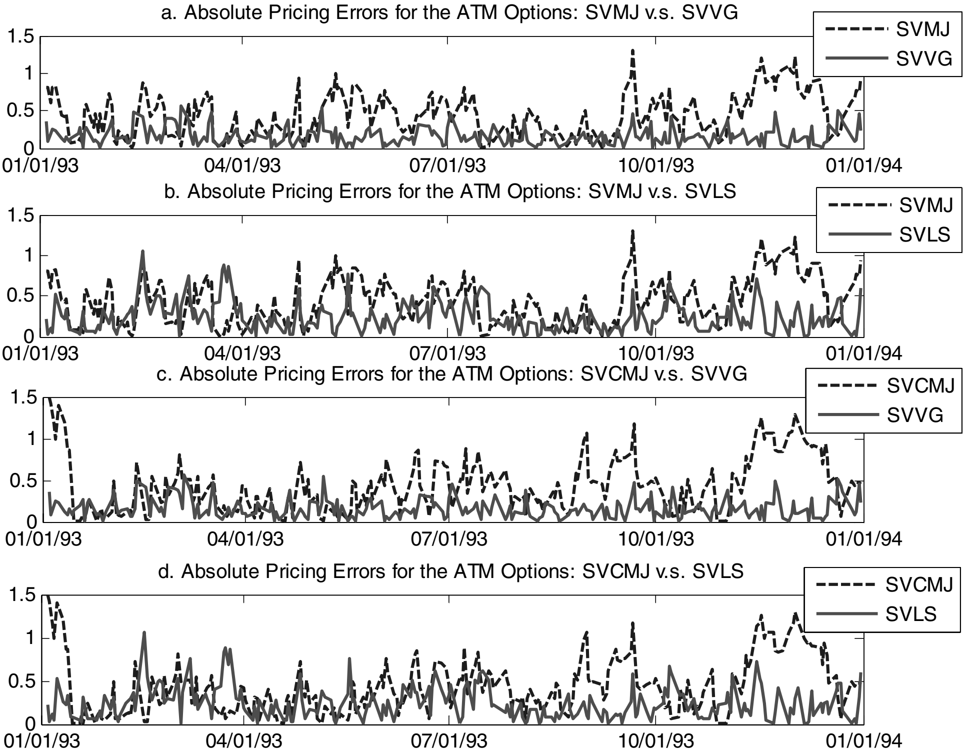

FIGURE 4.6. In-sample absolute option pricing errors of SVMJ, SVCMJ, SVVG, and SVLS, which are estimated using daily returns of the S&P 500 index and daily prices of the short-term ATM SPX options between January 4, 1993 and December 31, 1993.

rows of the last column in Panel A of Table 4.5. We see clearly that SVCMJ has smaller absolute pricing errors than SVMJ, and SVVG and SVLS have smaller absolute pricing errors than SVMJ and SVCMJ. The DM statistics for pair-wise comparisons of the four models based on the squared pricing errors of all options are reported in the last six rows of the last column in Panel B of Table 4.5. SVCMJ has significantly smaller squared pricing errors than SVMJ. SVVG has significantly smaller squared pricing errors than both SVMJ and SVCMJ. SVLS has somewhat worse performances than SVVG. Figure 4.7 provides time series plots of daily weighted average of the absolute pricing errors of all options for the four models during our sample period. Consistent with the DM statistics, we find that SVVG and SVLS have smaller absolute pricing errors than SVMJ and SVCMJ during most of the sample period.

Next we examine the performances of the four models in pricing options grouped by time-to-maturity. The time series mean of daily weighted average of the absolute pricing errors of options in each of the five maturity groups are reported in the last column in Panel A of Table 4.5. The DM statistics for pair-wise comparisons of the four models based on the squared pricing errors of options in the five maturity groups are reported in the last column in Panel B of Table 4.5. We find similar patterns in model performances for options in each maturity group as that for all options. For example, we find that SVVG has significantly smaller squared pricing errors than SVMJ and SVCMJ for most maturity groups.

Finally, we examine the performances of the four models in pricing options grouped by moneyness. The time series mean of daily weighted average of the absolute pricing errors of options in each of the six moneyness groups are reported in the last four rows in Panel A of Table 4.5. The DM statistics for pair-wise comparisons of the four models based on the squared pricing errors of options in the six moneyness groups are

TABLE 4.5
Out-of-Sample Performances in Option Pricing

		<0.93	0.93–0.97	0.97–1.0	1.0–1.03	1.03–1.07	>1.07	All
Panel A. Time series mean of daily weighted average of absolute pricing errors (in dollar) of out-of-sample options in each moneyness/maturity group.								
<1 m	No.	410	731	650	387	9	0	2187
	SVMJ	0.2265	0.3663	0.4277	0.3347	0.6449	N/A	0.3410
	SVCMJ	0.2148	0.3518	0.3867	0.3025	0.3220	N/A	0.3172
	SVVG	0.2319	0.3061	0.2399	0.2234	0.2404	N/A	0.2553
	SVLS	0.1779	0.2817	0.2760	0.2580	0.2277	N/A	0.2524
1–2 m	No.	694	896	679	676	306	0	3251
	SVMJ	0.5133	0.8113	0.7902	0.5226	0.3700	N/A	0.6371
	SVCMJ	0.4575	0.6893	0.6252	0.4697	0.3393	N/A	0.5400
	SVVG	0.4667	0.5653	0.3045	0.2467	0.2835	N/A	0.3915
	SVLS	0.3996	0.5682	0.3721	0.3444	0.4344	N/A	0.4297
2–3-m	No.	605	693	611	612	613	16	3150
	SVMJ	0.7937	1.3026	1.2660	0.8602	0.4593	0.4491	0.9452
	SVCMJ	0.6335	0.9732	0.9043	0.7091	0.4726	0.3407	0.7250
	SVVG	0.6639	0.8261	0.4889	0.3119	0.4252	0.1543	0.5286
	SVLS	0.5467	0.7267	0.4215	0.4468	0.6531	0.5792	0.5527
3–6 m	No.	941	415	334	328	370	170	2558
	SVMJ	1.1953	1.8914	1.9260	1.5650	0.8238	0.4239	1.3352
	SVCMJ	0.8150	1.2498	1.3257	1.2193	0.8240	0.5618	0.9805
	SVVG	0.9454	1.1721	0.8257	0.4151	0.4524	0.5474	0.7876
	SVLS	0.6982	0.8231	0.4669	0.3818	0.8369	0.9546	0.6700
>6 m	No.	696	170	128	120	154	311	1579
	SVMJ	1.8625	3.2767	3.0434	3.2549	2.4035	1.0897	2.1051
	SVCMJ	0.9751	1.7010	1.7761	1.8833	1.6684	1.0344	1.2712
	SVVG	1.4383	2.1033	1.4651	1.4414	0.6432	0.4837	1.2285
	SVLS	0.8029	0.9653	0.4122	0.4376	0.8658	1.3964	0.8610
All	No.	3346	2905	2402	2123	1452	497	12725
	SVMJ	0.8482	1.0786	1.0641	0.8752	0.6895	0.7868	0.9296
	SVCMJ	0.5997	0.7877	0.7595	0.6776	0.6109	0.7943	0.6832
	SVVG	0.6953	0.7172	0.4541	0.3392	0.3832	0.4422	0.5444
	SVLS	0.5095	0.5792	0.3702	0.3768	0.6378	1.1006	0.5093

(continued)

reported in the last six rows in Panel B of Table 4.5. We find that SVVG and SVLS have significantly smaller squared pricing errors than SVMJ and SVCMJ for most ITM and slightly OTM ($1.0 < K/S < 1.03$) options. While SVVG has smaller squared pricing errors than SVMJ and SVCMJ for all deep OTM options ($K/S > 1.03$) , the differences are statistically significant only for absolute pricing errors. SVVG tends to have larger

TABLE 4.5
Continued

Panel B. Diebold–Mariano statistics for out-of-sample squared absolute option pricing errors. The DM statistics provide pair-wise comparison of the four models by testing whether one model has significantly smaller average squared pricing errors for all options in a moneyness/maturity group than another model. Bold entries mean that the difference is significant at the 5% level for one-sided test. To save space, we omit "SV" in the names of all four models.

		<0.93	0.93–0.97	0.97–1.0	1.0–1.03	1.03–1.07	>1.07	All
<1 m	CMJ-MJ	**−1.8760**	−1.2468	−1.3720	−1.3913	−1.1857	N/A	−1.4356
	VG-MJ	−0.9924	**−2.0587**	**−2.2073**	**−2.1726**	−1.1646	N/A	**−2.2196**
	LS-MJ	**−1.9788**	**−2.1380**	**−2.1593**	**−2.0294**	−1.1382	N/A	**−2.2431**
	VG-CMJ	1.8267	**−1.8767**	**−2.0081**	**−2.0682**	−0.8847	N/A	**−2.1357**
	LS-CMJ	**−1.9612**	**−1.8109**	**−1.7500**	−1.2245	−0.6546	N/A	**−1.8931**
	VG-LS	2.0098	0.3429	−1.6095	−1.5835	−0.0798	N/A	−1.0260
1–2 m	CMJ-MJ	**−1.8547**	**−1.7263**	−1.4837	−0.9065	−0.4881	N/A	−1.4702
	VG-MJ	**−2.0359**	**−2.3167**	**−2.3194**	**−2.1112**	−1.0077	N/A	**−2.2518**
	LS-MJ	**−2.2008**	**−2.3157**	**−2.2948**	**−1.9670**	1.1143	N/A	**−2.2530**
	VG-CMJ	0.3630	**−1.9169**	**−2.1336**	**−2.2551**	−0.9283	N/A	**−2.1883**
	LS-CMJ	**−2.2439**	**−1.7375**	**−1.9642**	−1.3159	1.3400	N/A	**−1.7832**
	VG-LS	2.2363	−0.3727	**−1.6809**	−1.6245	**−1.9018**	N/A	**−1.8366**
2–3 m	CMJ-MJ	**−1.8094**	**−1.9877**	**−1.7572**	−1.0595	−0.0972	−1.0426	**−1.7109**
	VG-MJ	**−1.9395**	**−2.3205**	**−2.3510**	**−2.2343**	−0.6682	−1.1401	**−2.2865**
	LS-MJ	**−1.9849**	**−2.3021**	**−2.3219**	**−2.1415**	1.2392	1.1848	**−2.2552**
	VG-CMJ	0.9804	**−1.6520**	**−2.1103**	**−2.2833**	−1.0680	−1.2298	**−2.1052**
	LS-CMJ	**−2.0228**	**−2.0675**	**−2.2546**	**−2.0913**	1.6694	1.2528	**−2.1740**
	VG-LS	2.0255	1.8253	1.3916	**−1.7197**	**−2.1155**	−1.2681	−0.5989
3–6 m	CMJ-MJ	**−1.7672**	−1.6398	**−1.6616**	−1.4190	−0.4070	1.4619	**−1.6688**
	VG-MJ	**−1.8426**	**−1.7313**	**−1.8234**	**−1.8212**	**−1.7285**	0.9065	**−1.8327**
	LS-MJ	**−1.8208**	**−1.7195**	**−1.8171**	**−1.8127**	−1.0470	1.3944	**−1.8220**
	VG-CMJ	1.4865	−0.9991	**−1.7170**	**−1.7974**	**−1.6710**	0.1463	**−1.6100**
	LS-CMJ	**−1.8087**	**−1.7686**	**−1.8105**	**−1.8049**	−0.5355	1.2670	−1.7987
	VG-LS	1.7821	1.6601	1.7202	0.5454	−1.6131	−1.5439	1.6706
>6 m	CMJ-MJ	**−1.7002**	−1.6398	−1.4476	−1.4017	−1.3692	−0.6747	**−1.6748**
	VG-MJ	**−1.7200**	**−1.6758**	−1.4735	−1.4841	−1.5470	−1.6319	**−1.7421**
	LS-MJ	**−1.7155**	**−1.6698**	−1.4795	−1.4664	−1.5237	0.9748	**−1.7456**
	VG-CMJ	1.6634	1.4437	−1.3809	−1.4402	−1.5931	**−1.6835**	−0.1868
	LS-CMJ	**−1.7297**	**−1.7288**	−1.4843	−1.5707	−1.4867	1.0068	**−1.7471**
	VG-LS	1.7090	1.6575	1.4961	1.3922	−0.9588	−1.5663	1.7428

(continued)

TABLE 4.5
Continued

		<0.93	0.93–0.97	0.97–1.0	1.0–1.03	1.03–1.07	>1.07	All
All	CMJ-MJ	**−1.9279**	**−2.0100**	**−2.1283**	**−1.7568**	−1.3439	−0.3758	**−2.0548**
	VG-MJ	**−2.0673**	**−2.1634**	**−2.3182**	**−2.1807**	**−1.9363**	**−1.7026**	**−2.2968**
	LS-MJ	**−2.0069**	**−2.0929**	**−2.2794**	**−2.0398**	−1.4463	1.3682	**−2.2150**
	VG-CMJ	1.6460	−1.0717	**−2.0806**	**−2.2271**	**−2.0860**	**−1.7240**	**−1.9971**
	LS-CMJ	**−2.1797**	**−2.0070**	**−2.0740**	**−1.9737**	−0.6785	1.2596	**−2.0679**
	VG-LS	1.9318	1.8282	1.7840	0.0511	**−2.1586**	−1.6360	1.4424

This table reports the out-of-sample performances of the four models in option pricing. Based on the estimates of model parameters and latent volatility variables using the spot and option prices of the S&P 500 index, we obtain the theoretical price of each option that is not used in model estimation (12,725 in total) under each of the four models. We divide these options into six moneyness (defined as the ratio between strike and spot prices) and five maturity groups. The numbers of options belonging to each moneyness/maturity group during the entire sample also are reported. Based on options that are available on each day, we obtain daily arithmetic weighted average of the absolute pricing errors of options within each moneyness/maturity group. Then we obtain the time series means of the daily pricing errors over the sample period for each option group. Absolute pricing error is defined as the absolute value of the difference between model and market prices of an option.

(smaller) pricing errors than SVLS for ITM (OTM) options. We obtain similar findings for moneyness groups with different time-to-maturities, although the advantages of the Lévy jump models over the AJD models become less significant for options with longer time-to-maturities.

The analysis in this section clearly demonstrates the advantages of the Lévy jump models over the AJD models in modeling the joint dynamics of the spot and option prices of the S&P 500 index. The VG and LS models capture the many small movements in index returns that cannot be captured by the AJD models. The Lévy jump models also have significantly smaller in-sample and out-of-sample option pricing errors than the AJD models. Among all the models we consider, we find that the VG model of Madan, Carr, and Chang (1998) with stochastic volatility has the best performance in modeling the risk-neutral and physical dynamics of the S&P 500 index returns. We emphasize that the superior performances of the Lévy jump models are obtained under the restriction that jumps under the physical and the risk-neutral measures must follow the same Lévy process. If we allow jumps to follow different Lévy processes under the two measures, Lévy jump models are likely to have even better performances in capturing the joint dynamics of index returns. Therefore, our analysis points out the great potentials of Lévy processes for continuous-time finance modeling and strongly suggests that we can enrich existing AJD models by incorporating infinite-activity Lévy jumps.

5. CONCLUSION

In this paper, we address a basic and yet fundamental empirical issue in the current continuous-time finance literature: Whether newly developed Lévy jump models can

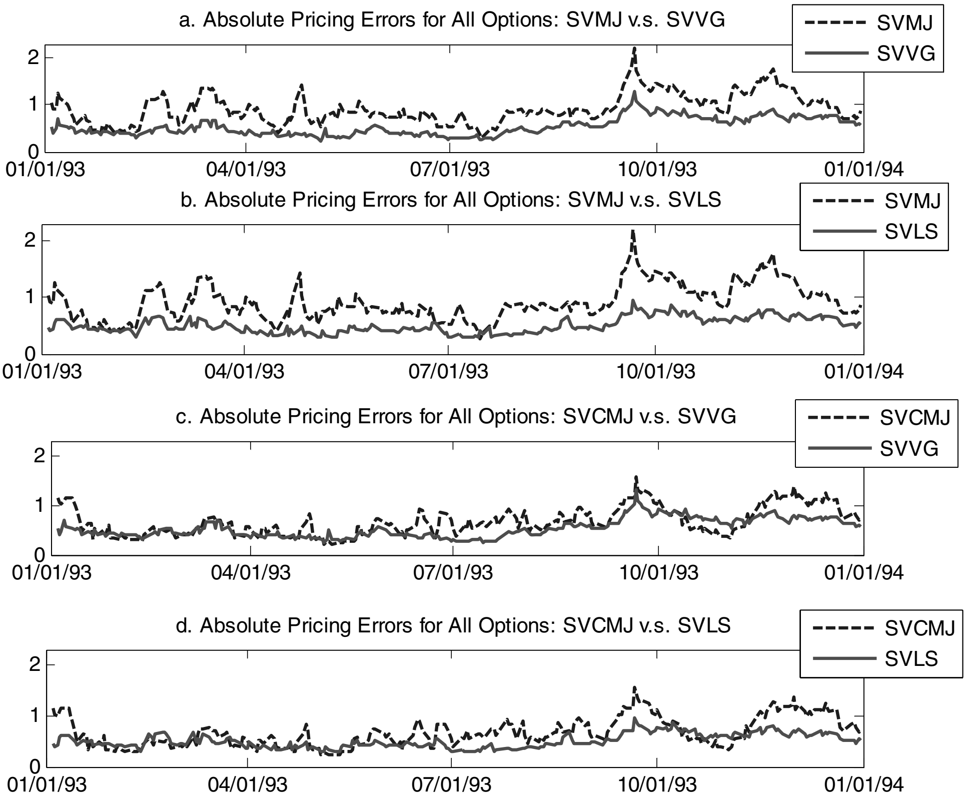

FIGURE 4.7. Average absolute pricing errors for all out-of-sample options of SVMJ, SVCMJ, SVVG, and SVLS, which are estimated using daily returns of the S&P 500 index and daily prices of the short-term ATM SPX options between January 4, 1993 and December 31, 1993.

outperform some of the most sophisticated AJD models in capturing the joint dynamics of stock and option prices. We develop efficient MCMC methods to estimate the posterior distributions of parameters and latent volatility/jump variables of the Lévy jump models using stock and option prices. We show that models with infinite-activity Lévy jumps in returns significantly outperform the AJD models with compound Poisson jumps in returns and volatility in capturing the joint dynamics of the spot and option prices of the S&P 500 index. We also find that the variance gamma model of Madan, Carr, and Chang (1998) with stochastic volatility has the best performance among all the models we consider.

Our results suggest that existing AJD models need much more sophisticated parametric specifications of the compound Poisson process to capture the data. The more general Lévy processes we consider provide a convenient alternative and are suggestive of presence of infinite activity jumps in asset prices. Our results are consistent with the nonparametric evidence of Aït-Sahalia and Jacod (2009) and Tauchen and Todorov (2008) and suggest that incorporating infinite-activity Lévy jumps into existing AJD models can substantially increase the flexibility of AJD models without sacrificing their tractability.

APPENDIX

In this appendix, we first provide more detailed information on the finite- and infinite-activity jump processes considered in the paper. Then we provide the details of the MCMC methods for estimating SVMJ, SVCMJ, SVVG, and SVLS models.

A.1. Characteristic Component, Lévy Measure and Drift for MJ, CMJ, VG, and LS

In this section, we provide analytical expressions of the characteristic component, Lévy measure and drift for MJ, CMJ, VG, and LS, which have been used in the paper. To emphasize the generality of these results, we omit dependence of model parameters on probability measures.

MJ:

$$\psi_J(u) = \lambda\left(1 - e^{iu\mu_y - \frac{1}{2}\sigma_y^2 u^2}\right), \quad \pi(x) = \frac{\lambda}{\sqrt{2\pi}\sigma_y} e^{-\frac{(x-\mu_y)^2}{2\sigma_y^2}}, \quad \bar{\mu} = \int_{|x|\leq 1} x\pi(dx).$$

CMJ:

$$\psi_J(u) = \lambda\left(1 - \frac{e^{iu_1\mu_y - \frac{1}{2}\sigma_y^2 u_1^2}}{1 - iu_1\mu_v\rho_J - iu_2\mu_v}\right), \quad \pi(x) = \frac{\lambda}{\mu_v\sqrt{2\pi}\sigma_y} e^{-\frac{x_2}{\mu_v} - \frac{(x_1-\mu_y-\rho_J x_2)^2}{2\sigma_y^2}},$$

$$\bar{\mu} = \int_{|x|\leq 1} x\pi(dx).$$

VG:

$$\psi_J(u) = \frac{\log\left(1 - iu\gamma\nu + \frac{\sigma^2\nu u^2}{2}\right)}{\nu},$$

$$\pi(x) = \begin{cases} \frac{1}{\nu}\frac{\exp(-Mx)}{x} x > 0 \\ \frac{1}{\nu}\frac{\exp(-G\mid x\mid)}{\mid x\mid} x < 0 \end{cases}, \quad \bar{\mu} = \int_{|x|\leq 1} x\pi(dx),$$

where

$$M = \left(\sqrt{\frac{1}{4}\gamma^2\nu^2 + \frac{1}{2}\sigma^2\nu} + \frac{1}{2}\gamma\nu\right)^{-1} \quad \text{and} \quad G = \left(\sqrt{\frac{1}{4}\gamma^2\nu^2 + \frac{1}{2}\sigma^2\nu} - \frac{1}{2}\gamma\nu\right)^{-1}.$$

If $\gamma = 0$, then the jump structure is symmetric around zero, and $M = G$.

Lévy α-stable Process: Suppose $X_1 \sim S_\alpha(\beta, \sigma, \gamma)$, which reduces to Log-Stable process if $\beta = -1$ and $\gamma = 0$, then

$$\psi_J(u) = \sigma^\alpha |u|^\alpha (1 - i\beta \text{sign}(u) \tan\left(\frac{\pi\alpha}{2}\right) + i\gamma u,$$

$$\pi(x) = \begin{cases} c_1 \frac{1}{x^{1+\alpha}} x > 0 \\ c_2 \frac{1}{|x|^{1+\alpha}} x < 0 \end{cases}, \quad \bar{\mu} = \gamma - \int 1_{|x|>1} x\pi(dx),$$

where $c_1 = \frac{\sigma^\alpha(1+\beta)}{2}$ and $c_2 = \frac{\sigma^\alpha(1-\beta)}{2}$. In the LS model, c_1 becomes 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 pth moment for $p > \alpha$.

A.2. Change of Measure for Lévy Processes

Theorem (Sato 1999). Let $(X_t^{\mathbb{P}}, \mathbb{P})$ and $(X_t^{\mathbb{Q}}, \mathbb{Q})$ be two Lévy processes on $\mathbb{R}$ with corresponding characteristic triplets $(\bar{\mu}_{\mathbb{P}}, \bar{\sigma}_{\mathbb{P}}^2, \pi_{\mathbb{P}}(dx))$ and $(\bar{\mu}_{\mathbb{Q}}, \bar{\sigma}_{\mathbb{Q}}^2, \pi_{\mathbb{Q}}(dx))$, and $\phi(x) = \log(\frac{\pi_{\mathbb{Q}}(x)}{\pi_{\mathbb{P}}(x)})$. Sato (1999) shows that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent for all t if and only if the following conditions are satisfied: (i) $\bar{\sigma}_{\mathbb{P}} = \bar{\sigma}_{\mathbb{Q}}$; (ii) The Lévy measures are equivalent with $\int_{-\infty}^{\infty}(e^{\phi(x)/2} - 1)^2 \pi_{\mathbb{P}}(dx) < \infty$; and (iii) If $\bar{\sigma}_{\mathbb{P}} = 0$, then we must in addition have $\bar{\mu}_{\mathbb{Q}} - \bar{\mu}_{\mathbb{P}} = \int_{-1}^{1} x(\pi_{\mathbb{Q}}(x) - \pi_{\mathbb{P}}(x))\, dx$. And the Radon-Nikodym derivative equals e^{U_t}, where U_t is a Lévy process with characteristic triplet $(\bar{\mu}_u, \bar{\sigma}_u^2, \pi_u(dx))$: (i) $\bar{\sigma}_u^2 = 0$; (ii) $\bar{\mu}_u = -\int_{-\infty}^{\infty}(e^y - 1 - y_{|y|\leq 1})(\pi_{\mathbb{P}}\phi^{-1})\, dy$; and (iii) $\pi_u = \pi_{\mathbb{P}}\phi^{-1}$.

This theorem provides the necessary and sufficient conditions for two probability measures of Lévy processes to be equivalent. The first condition requires that the change of measure does not affect the volatility of the Brownian part of a Lévy process, which is similar to the change of measure for Brownian motions. The second condition requires the Hellinger distance between the two Lévy measures to be finite. That is, for the two probability measures to be equivalent, the jump structures of the two Lévy processes cannot be too different from each other. The third condition imposes restriction between the drift terms and the Lévy measures of the two Lévy processes.

A.3. Priors for Model Parameters

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

- **Priors for parameters common to four models**. We consider the following prior distributions: $\kappa \sim N(0, 1)$ (truncated at zero), $\theta \sim N(0, 1)$ (truncated at zero), $\rho \sim Uniform(0, 1)$, $\sigma_v \sim \frac{1}{\sigma_v}$, $\eta^s \sim N(0, 1)$, $\eta^v \sim N(0, 1)$, $\rho_c \sim N(0, 1)$, and $\sigma_c \sim \frac{1}{\sigma_c}$.
- **Priors for parameters common to SVMJ and SVCMJ.** For $\mu_y^{\mathbb{P}}$ and $\mu_y^{\mathbb{Q}}$, we choose standard conjugate priors: $\mu_y^{\mathbb{P}} \sim N(0, 1)$ and $\mu_y^{\mathbb{Q}} \sim N(0, 1)$. We choose flat priors for σ_y and λ: $\sigma_y \sim \frac{1}{\sigma_y}$ and $\lambda \sim Uniform(0, 1)$.
- **Priors for parameters unique to SVCMJ**. For μ_v and ρ_J, we choose the following priors: $\mu_v \sim \frac{1}{\mu_v}$ and $\rho_J \sim N(0, 1)$.
- **Priors for parameters unique to SVVG**. We choose the following priors for the five parameters that are unique to SVVG: $\gamma^{\mathbb{P}} \sim N(0, 1)$, $\gamma^{\mathbb{Q}} \sim N(0, 1)$, $\nu \sim \frac{1}{\nu}$, $\sigma^{\mathbb{P}} \sim \frac{1}{\sigma^{\mathbb{P}}}$, and $\sigma^{\mathbb{Q}} \sim \frac{1}{\sigma^{\mathbb{Q}}}$.
- **Priors for parameters unique to SVLS**. For α and σ, we choose the following joint priors: $\alpha \sim Uniform(1, 2)$ and $\sigma \sim \frac{1}{\sigma}$.

Although we choose flat priors for the variance parameters, the priors of most other parameters are proper priors, pretty uninformative, and have been used in previous studies. 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.4. MCMC Methods for SVMJ

In this section, we discuss the updating algorithms and the posterior distributions of model parameters and latent variables for SVMJ. Compared to that of LWY (2008), which only uses stock prices, the posterior likelihood here always has an additional component, which is the likelihood of option pricing errors. Because the computation of option price involves numerical integration, the parameters that appear in the option pricing formula usually do not have known posterior distributions. To overcome this difficulty, we adopt the method of Damine, Wakefield, and Walker (1999) (hereafter DWW) to update these parameters. Parameters that are not involved in the option pricing formula usually have standard known posterior distributions, from which we draw posterior samples. In this and the following sections, we discuss the updating methods, first for parameters that appear in the option pricing formula, then for the rest.

- **Posterior for** κ. The posterior of κ is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1}-F_{t+1})-\rho_c(C_t-F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\kappa)} \times N\left(\frac{\mathcal{S}}{\mathcal{W}}, \sqrt{\frac{1}{\mathcal{W}}}\right) 1_{\kappa>0}$$

where $\mathcal{W} = \frac{\Delta}{(1-\rho^2)\sigma_v^2} \sum_{t=0}^{T-1} \frac{(\theta-v_t)^2}{v_t} + 1$, $\mathcal{S} = \frac{1}{\sigma_v(1-\rho^2)} \sum_{t=0}^{T-1} \frac{(\theta-v_t)(\frac{B_{t+1}}{\sigma_v} - \rho A_{t+1})}{v_t}$, $A_{t+1} = Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta - N_{t+1}^y \xi_{t+1}^y$, and $B_{t+1} = v_{t+1} - v_t$. We denote the first term as $l(\kappa)$, omitting dependence on other parameters to simplify notation. Its calculation involves numerical integration because of the option pricing formula involved. The second term in the posterior is a truncated normal distribution. This combination motivates us to use the DWW method. Specifically, for a given previous draw, $\kappa^{(g)}$, the algorithm for $(g+1)$-th iteration is:
1. Draw $\kappa^{(g+1)}$ from $N(\frac{\mathcal{S}}{\mathcal{W}}, \sqrt{\frac{1}{\mathcal{W}}})1_{\kappa>0}$;
2. Draw an auxiliary variable u from $Uniform(0, l(\kappa^{(g)}))$;
3. Accept $\kappa^{(g+1)}$ if $l(\kappa^{(g+1)}) > u$; otherwise, keep $\kappa^{(g)}$.
- **Posterior for** θ. Similarly, the posterior of θ is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1}-F_{t+1})-\rho_c(C_t-F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\theta)} \times N\left(\frac{\mathcal{S}}{\mathcal{W}}, \sqrt{\frac{1}{\mathcal{W}}}\right) 1_{\theta>0}$$

where $\mathcal{W} = \frac{\kappa^2\Delta}{\sigma_v^2(1-\rho^2)} \sum_{t=0}^{T-1} \frac{1}{v_t} + 1$, $\mathcal{S} = \frac{\kappa}{(1-\rho^2)\sigma_v} \sum_{t=0}^{T-1} (\frac{B_{t+1}/\sigma_v - \rho A_{t+1}}{v_t})$, $A_{t+1} = Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta - N_{t+1}^y \xi_{t+1}^y$, and $B_{t+1} = v_{t+1} + (\kappa\Delta - 1)v_t$. Again we use the DWW method and the updating algorithm is the same as that for κ.

- **Posterior for** σ_v. The posterior of σ_v is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1}-F_{t+1})-\rho_c(C_t-F_t)]^2}{2\sigma_c^2}\right) \times \exp\left(\frac{\rho}{1-\rho^2}\left(\sum_{t=0}^{T-1} A_{t+1}B_{t+1}\right)\frac{1}{\sigma_v}\right)}_{:=l(\sigma_v)}$$

$$\times \left(\frac{1}{\sigma_v^2}\right)^{\frac{T}{2}+\frac{1}{2}} \exp\left(-\frac{\sum_{t=0}^{T-1} B_{t+1}^2}{2(1-\rho^2)}\frac{1}{\sigma_v^2}\right)$$

where $A_{t+1} = \frac{Y_{t+1}-Y_t-(r_t-\frac{1}{2}v_t+\psi_J(-i)+\eta^s v_t)\Delta - N_{t+1}^y \xi_{t+1}^y}{\sqrt{v_t\Delta}}$ and $B_{t+1} = \frac{v_{t+1}-v_t-\kappa(\theta-v_t)\Delta}{\sqrt{v_t\Delta}}$. The algorithm is similar to that for κ:
1. Draw $\frac{1}{(\sigma_v^{(g+1)})^2}$ from $\Gamma(\frac{T}{2}+\frac{3}{2}, (\frac{\sum_{t=0}^{T-1} B_{t+1}^2}{2(1-\rho^2)})^{-1})$;
2. Draw an auxiliary variable u from $Uniform(0, l(\sigma_v^{(g)}))$;
3. Accept $\sigma_v^{(g+1)}$ if $l(\sigma_v^{(g+1)}) > u$; otherwise, keep $\sigma_v^{(g)}$.

- **Posterior for** ρ. The posterior of ρ is proportional to the function $\pi(\rho)$

$$\propto \pi(\rho) := \prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1}-F_{t+1})-\rho_c(C_t-F_t)]^2}{2\sigma_c^2}\right)$$

$$\times (1-\rho^2)^{-\frac{T}{2}} \exp\left(-12(1-\rho^2)\sum_{t=0}^{T-1}\left(A_{t+1}^2+B_{t+1}^2\right) + \frac{\rho}{(1-\rho^2)}\sum_{t=0}^{T-1} A_{t+1}B_{t+1}\right)$$

where $A_{t+1} = \frac{Y_{t+1}-Y_t-(r_t-\frac{1}{2}v_t+\psi_J(-i)+\eta^s v_t)\Delta - N_{t+1}^y \xi_{t+1}^y}{\sqrt{v_t\Delta}}$ and $B_{t+1} = \frac{v_{t+1}-v_t-\kappa(\theta-v_t)\Delta}{\sigma_v\sqrt{v_t\Delta}}$. It is well known that the sampling distribution of Pearson's correlation is negatively skewed and the so-called "Fisher's Z transformation" converts Pearson's correlation to a normally distributed variable. Motivated by Fisher's idea, we develop the following algorithm:
1. Draw $\frac{1}{2}\log\frac{1+\rho^{(g+1)}}{1-\rho^{(g+1)}}$ from $N(\frac{1}{2}\log\frac{1+\rho_r}{1-\rho_r}, \frac{1}{T-3})$, where $\rho_r = \text{Corr}(\mathbf{A}, \mathbf{B})$, $\mathbf{A} = \{A_{t+1}\}_{t=0}^{T-1}$, $\mathbf{B} = \{B_{t+1}\}_{t=0}^{T-1}$, and $g(\rho_r) = \frac{1}{2}\log\frac{1+\rho_r}{1-\rho_r}$ is the formula of Fisher's Z transformation.
2. Accept $\rho^{(g+1)}$ with probability

$$\min\left(\frac{\pi(\rho^{(g+1)})}{\pi(\rho^{(g)})} \times \frac{\exp\left(-\frac{(g(\rho^{(g)})-g(\rho_r))^2}{\frac{2}{T-3}}\right)}{\exp\left(-\frac{(g(\rho^{(g+1)})-g(\rho_r))^2}{\frac{2}{T-3}}\right)}, 1\right).$$

By removing the skewness of the distribution for the candidate draw, our algorithm converges more quickly than the one without the transformation.

- **Posteriors for** η^v **and** $\mu_y^{\mathbb{Q}}$. Because the updating methods and the posteriors of η^v and $\mu_y^{\mathbb{Q}}$ are the same, we focus our discussion on η^v. The posterior of η^v is proportional to

$$\propto \pi(\eta^v) := \prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right) \times \exp\left(-\frac{(\eta^v)^2}{2}\right).$$

We update the parameter using the Metropolis–Hasting algorithm. A normal distribution centered at the previous draw with constant variance 1 is used as the proposal distribution for the candidate draw, which is accepted with the probability $\min(\frac{\pi(\eta^{v(g+1)})}{\pi(\eta^{v(g)})}, 1)$.

- **Posterior for** σ_y. The posterior of σ_y is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\sigma_y)}$$
$$\times \left(\frac{1}{\sigma_y^2}\right)^{\frac{T}{2}+\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_{t=0}^{t-1}(\xi_{t+1}^y - \mu_y^{\mathbb{P}})^2 \frac{1}{\sigma_y^2}\right).$$

We use the DWW method to update the parameter:
1. Draw $\frac{1}{(\sigma_y^{(g+1)})^2}$ from $\Gamma(\frac{T}{2} + \frac{3}{2}, \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1}(\xi_{t+1}^y - \mu_y^{\mathbb{P}})^2})$;
2. Draw an auxiliary variable u from $Uniform(0, l(\sigma_y^{(g)}))$;
3. Accept $\sigma_y^{(g+1)}$ if $l(\sigma_y^{(g+1)}) > u$; otherwise, keep $\sigma_y^{(g)}$.

- **Posterior for** λ. The posterior of λ is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\lambda)} \times \lambda^{\sum_{t=0}^{T-1} N_{t+1}}(1-\lambda)^{T-\sum_{t=0}^{T-1} N_{t+1}}.$$

The DWW method is used and the proposal distribution for the candidate draw is $Beta(\sum_{t=0}^{T-1} N_{t+1} + 1, T - \sum_{t=0}^{T-1} N_{t+1} + 1)$. The algorithm is skipped.

For parameters that do not appear in the option pricing formula, that is, $(\eta^s, \mu_y^{\mathbb{P}}, \rho_c, \sigma_c)$, we obtain known posterior distributions.

- **Posterior for** η^s. The posterior of η^s follows a normal distribution $\eta^s \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{\Delta}{(1-\rho^2)}\sum_{t=0}^{T-1} v_t + 1$, $\mathcal{S} = \frac{1}{(1-\rho^2)}\sum_{t=0}^{T-1}(A_{t+1} - \frac{\rho}{\sigma_v}B_{t+1})$, $A_{t+1} = Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta - N_{t+1}^y \xi_{t+1}^y$, and $B_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta$.
- **Posterior for** $\mu_y^{\mathbb{P}}$. The posterior of $\mu_y^{\mathbb{P}}$ follows a normal distribution $\mu_y^{\mathbb{P}} \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{T}{\sigma_y^2} + 1$, and $\mathcal{S} = \frac{\xi_{t=0}^{T-1}\xi_{t+1}}{\sigma_y^2}$.
- **Posterior for** ρ_c. The posterior of ρ_c follows a normal distribution $\rho_c \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{\sum_{t=0}^{T-1} A_t^2}{\sigma_c^2} + 1$, $\mathcal{S} = \frac{\sum_{t=0}^{T-1} A_t A_{t+1}}{\sigma_c^2}$, and $A_{t+1} = C_{t+1} - F_{t+1}$.
- **Posterior for** σ_c. The posterior of σ_c follows a gamma distribution $\frac{1}{\sigma_c^2} \sim \Gamma(\frac{T}{2} + \frac{3}{2}, \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1}(A_{t+1} - \rho_c A_t)^2})$, where $A_{t+1} = C_{t+1} - F_{t+1}$.

Next we consider the posteriors of latent jump and volatility variables.

- **Posterior for** ξ_{t+1}^{y}. The posterior of ξ_{t+1}^{y} is $\xi_{t+1}^{y} \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{N_{t+1}^2}{(1-\rho^2)v_t\Delta} + \frac{1}{\sigma_y^2}$, $\mathcal{S} = \frac{N_{t+1}}{(1-\rho^2)v_t\Delta}(A_{t+1} - \rho B_{t+1}/\sigma_v) + \frac{\mu_y}{\sigma_y^2}$, $A_{t+1} = Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta$, and $B_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta$.
- **Posterior for** N_{t+1}. The posterior of N_{t+1} is $N_{t+1} \sim \text{Bernoulli}(\frac{\alpha_1}{\alpha_1+\alpha_2})$, where $\alpha_1 = e^{-\frac{1}{2(1-\rho^2)}[A_1^2 - 2\rho A_1 B]}\lambda$, $\alpha_2 = e^{-\frac{1}{2(1-\rho^2)}[A_2^2 - 2\rho A_2 B]}(1-\lambda)$, $A_1 = (Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta - \xi_{t+1}^{y})/\sqrt{v_t\Delta}$, $A_2 = (Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta)/\sqrt{v_t\Delta}$, and $B = (v_{t+1} - v_t - \kappa(\theta - v_t)\Delta)/(\sigma_v\sqrt{v_t\Delta})$.
- **Posterior for** v_{t+1}. For $0 < t+1 < T$, the posterior of v_{t+1} is proportional to

$$
\begin{aligned}
&\propto \exp\left(-\frac{1}{2\sigma_c^2}[(C_{t+1} - F_{t+1})^2 - 2\rho_c(C_{t+1} - F_{t+1})(C_t - F_t)]\right)\\
&\quad\times \exp\left(-\frac{1}{2\sigma_c^2}\left[\rho_c^2(C_{t+1} - F_{t+1})^2 - 2\rho_c(C_{t+2} - F_{t+2})(C_{t+1} - F_{t+1})\right]\right)\\
&\quad\times \exp\left\{-\frac{\left[-2\rho\epsilon_{t+1}^{y}\epsilon_{t+1}^{v} + \left(\epsilon_{t+1}^{v}\right)^2\right]}{2(1-\rho^2)}\right\} \times \frac{1}{v_{t+1}}\\
&\quad\times \exp\left\{-\frac{\left[\left(\epsilon_{t+2}^{y}\right)^2 - 2\rho\epsilon_{t+2}^{y}\epsilon_{t+2}^{v} + \left(\epsilon_{t+2}^{v}\right)^2\right]}{2(1-\rho^2)}\right\},
\end{aligned}
$$

  where $\epsilon_{t+1}^{y} = (Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta - N_{t+1}^{y}\xi_{t+1}^{y})/\sqrt{v_t\Delta}$, and $\epsilon_{t+1}^{v} = (v_{t+1} - v_t - \kappa(\theta - v_t)\Delta)/(\sigma_v\sqrt{v_t\Delta})$. And the posterior for v_t when $t = 0$ and $t = T$ can be derived in the similar way. We use the traditional Metropolis-Hasting method to update v_t, and use the Student-t distribution with a degree of freedom of 6 as the proposal distribution.

A.5. 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 parameters and latent variables that are unique to SVCMJ.

- **Posterior for** μ_v. The posterior of μ_v is proportional to

$$
\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\mu_v)} \times \left(\frac{1}{\mu_v}\right)^{T+1} \exp\left(-\frac{1}{\mu_v}\sum_{t=0}^{T-1}\xi_{t+1}^{v}\right).
$$

  The DWW method is used and the proposal distribution for the candidate draw is $IG(T+2, \frac{1}{\sum_{t=0}^{T-1}\xi_{t+1}^{v}})$.
- **Posterior for** ρ_J. The posterior of ρ_J is proportional to

$$
\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\rho_J)} \times N\left(\frac{\mathcal{S}}{\mathcal{W}}, \sqrt{\frac{1}{\mathcal{W}}}\right),
$$

where $\mathcal{W} = \frac{\sum_{t=0}^{T-1}(\xi_{t+1}^{v})^2}{\sigma_y^2} + 1, \mathcal{S} = \frac{\sum_{t=0}^{T-1}\xi_{t+1}^{v} A_{t+1}}{\sigma_y^2}$, and $A_{t+1} = \xi_{t+1}^{y} - \mu_y^{\mathbb{P}}$. The DWW method is used and the proposal distribution for the candidate draw is $N(\frac{\mathcal{S}}{\mathcal{W}}, \sqrt{\frac{1}{\mathcal{W}}})$.

- **Posterior for** ξ_{t+1}^{v}. The posterior of ξ_{t+1}^{v} follows a normal distribution $\xi_{t+1}^{v} \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})1_{\xi_{t+1}^{v}>0}$, where $\mathcal{W} = \frac{N_{t+1}^2}{(1-\rho^2)v_t\Delta} + \frac{\rho_J^2}{\sigma_y^2}, \mathcal{S} = \frac{N_{t+1}}{(1-\rho^2)v_t\Delta}(-\rho A_{t+1} + \frac{B_{t+1}\sigma_v}{+}\frac{\xi_{t+1}^{y}-\mu_y^{\mathbb{P}})\rho_J}{\sigma_y^2} - \frac{1}{\mu_v}$, $A_{t+1} = Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta$, and $B_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta - N_{t+1}\xi_{t+1}^{v}$.

A.6. 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** ν. The posterior of ν is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)\left(\frac{1}{\nu^{\frac{\Delta}{\nu}}\Gamma\left(\frac{\Delta}{\nu}\right)}\right)^T\left(\prod_{t=0}^{T-1} G_t\right)^{\frac{\Delta}{\nu}-1}}_{:=l(\nu)}$$

$$\times \exp\left\{-\frac{1}{\nu}\left(\sum_{t=0}^{T-1} G_t\right)\right\}\frac{1}{\nu}.$$

The DWW method is used and the proposal distribution for the candidate draw is $IG(2, \frac{1}{\sum_{t=0}^{T-1} G_{t+1}})$.

- **Posteriors for** $\gamma^{\mathbb{Q}}$ **and** $\sigma^{\mathbb{Q}}$. The algorithms for updating $\gamma^{\mathbb{Q}}$ and $\sigma^{\mathbb{Q}}$ are the same as that for η^{v} in SVMJ, except that the candidate draw for $\sigma^{\mathbb{Q}}$ needs to be truncated at zero because it has to be a positive number.
- **Posterior for** $\gamma^{\mathbb{P}}$. The posterior of $\gamma^{\mathbb{P}}$ is $\gamma^{\mathbb{P}} \sim N(\frac{\mathcal{S}}{\mathcal{W}}, \frac{1}{\mathcal{W}})$, where $\mathcal{W} = \frac{1}{(\sigma^{\mathbb{P}})^2}\sum_{t=0}^{T-1} G_{t+1} + 1$, and $\mathcal{S} = \frac{1}{(\sigma^{\mathbb{P}})^2}\sum_{t=0}^{T-1} J_{t+1}$.
- **Posterior for** $\sigma^{\mathbb{P}}$. The posterior of $\sigma^{\mathbb{P}}$ is $(\sigma^{\mathbb{P}})^2 \sim IG(\frac{T}{2} + \frac{3}{2}, \frac{1}{\frac{1}{2}\sum_{t=0}^{T-1}\frac{(J_{t+1}-\gamma^{\mathbb{P}}G_{t+1})^2}{G_{t+1}}})$.
- **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 $\mathcal{W} = \frac{1}{(1-\rho^2)v_t\Delta} + \frac{1}{(\sigma^{\mathbb{P}})^2 G_{t+1}}, \mathcal{S} = \frac{1}{(1-\rho^2)v_t\Delta}(A_{t+1} - \frac{\rho B_{t+1}}{\sigma_v}) + \frac{\gamma^{\mathbb{P}}}{(\sigma^{\mathbb{P}})^2}$, $A_{t+1} = Y_{t+1} - Y_t - (r_t - \frac{1}{2}v_t + \psi_J(-i) + \eta^s v_t)\Delta$, and $B_{t+1} = v_{t+1} - v_t - \kappa(\theta - v_t)\Delta$.
- **Posterior for** G_{t+1}. The posterior of G_{t+1} is proportional to

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

The posterior distribution of G_{t+1} is nonstandard 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, e.g., 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 G_t.

A.7. 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** α. The posterior of α is proportional to

$$\pi(\alpha) \propto \prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)$$
$$\times \left(\frac{\alpha}{\alpha-1}\right)^T \exp\left\{-\sum_{t=0}^{T-1}\left|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})}\right|^{\frac{\alpha}{\alpha-1}}\right\}$$
$$\times \prod_{t=0}^{T-1}\left|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})}\right|^{\frac{\alpha}{\alpha-1}} \times \left[\left(\frac{1}{\sigma}\right)^{\frac{\alpha}{\alpha-1}}\right]^{m+1} \exp\left\{-\left(\frac{1}{\sigma}\right)^{\frac{\alpha}{\alpha-1}} \frac{1}{M}\right\} \times \mathbf{1}(\alpha)_{\alpha\in[1.01,2]},$$

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 because 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 $a + b = 5\log(T)$, a constant suggested by Buckle (1995). Define

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

Then, the algorithm works in the following way:

1. Calculate

$$\begin{cases} a_1 = \left(\frac{\alpha^{(g)} - 1.01}{0.99}\right)(5\log(T) - 2) + 1 \\ b_1 = 5\log(T) - a_1 \end{cases}$$

and then draw τ from $Beta(a_1, b_1)$ and set $\alpha^{(g+1)} = 0.99\tau + 1.01$;

2. Calculate

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

3. Draw u from $Uniform(0, 1)$;
4. Accept $\alpha^{(g+1)}$ if $u > \min(\frac{p(\alpha^{(g+1)})}{p(\alpha^{(g)})} \times \frac{g(\alpha^{(g)} \mid a_2, b_2)}{g(\alpha^{(g+1)} \mid a_1, b_1)}, 1)$, otherwise keep the previous draw.

- **Posterior for** σ. The posterior of σ is proportional to

$$\propto \underbrace{\prod_{t=0}^{T-1} \exp\left(-\frac{[(C_{t+1} - F_{t+1}) - \rho_c(C_t - F_t)]^2}{2\sigma_c^2}\right)}_{:=l(\sigma)}$$

$$\times \left[\left(\frac{1}{\sigma}\right)^{\frac{\alpha}{\alpha-1}} T + 1\right] \exp\left\{-\left(\frac{1}{\sigma}\right)^{\frac{\alpha}{\alpha-1}} \left(\sum_{t=0}^{T-1} \left|\frac{S_{t+1}}{\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})}\right|^{\frac{\alpha}{\alpha-1}}\right)\right\}.$$

The DWW method is used with the following proposal distribution $(\frac{1}{\sigma})^{\frac{\alpha}{\alpha-1}} \sim \Gamma(T + \frac{\alpha-1}{\alpha} + 1, \frac{1}{\sum_{t=0}^{T-1} |\frac{S_{t+1}}{\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})}|^{\frac{\alpha}{\alpha-1}} + \frac{1}{M}})$.

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

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

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

$$p(U_{t+1} \mid \cdot) \propto \underbrace{\exp\left\{-\left|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})}\right|^{\frac{\alpha}{\alpha-1}} + 1\right\} \left|\frac{S_{t+1}}{\sigma\Delta^{\frac{1}{\alpha}} t_\alpha(U_{t+1})}\right|^{\frac{\alpha}{\alpha-1}}}_{g(U_{t+1})}$$
$$\times\left[\mathbf{1}_{S_{t+1}\in(-\infty,0)\cap U_{t+1}\in(-\frac{1}{2},l_\alpha)} + \mathbf{1}_{S_{t+1}\in(0,\infty)\cap U_{t+1}\in(l_\alpha,\frac{1}{2})}\right].$$

Due to the monotonicity of $t_\alpha(U_{t+1})$, we know that $p(U_{t+1} \mid \cdot)$ 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 $p(U_{t+1} \mid \cdot)$. This algorithm works in the following way:

1. Draw

$$U_{t+1}^{(g+1)} \longleftarrow \begin{cases} Uniform\left(-\frac{1}{2}, l_\alpha\right) & \text{if } S_{t+1} < 0 \\ Uniform\left(l_\alpha, \frac{1}{2}\right) & \text{if } S_{t+1} > 0 \end{cases};$$

2. Draw u from *Uniform*(0, 1);
3. Accept U_{t+1} if $u < g(U_{t+1}^{(g+1)})$, otherwise return to 1.

REFERENCES

AÏT-SAHALIA, Y. (2004): Disentangling Diffusion from Jumps, *J. Financ. Econ.* 74, 487–528.

AÏT-SAHALIA, Y., and J. JACOD (2008): Fisher's Information for Discretely Sampled Lévy Processes, *Econometrica*, 76, 727–761.

AÏT-SAHALIA, Y., and J. JACOD (2009): Estimating the Degree of Activity of Jumps in High Frequency Financial Data, *Ann. Stat.* 37, 2202–2244.

AÏT-SAHALIA, Y., and A. LO (1998): Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices, *J. Finance* 53, 499–547.

BAKSHI, G., and L. WU (2005): Investor Irrationality and the Nasdaq Bubble, Working Paper, Baruch College and University of Maryland.

BARNDORFF-NIELSEN, O. E. (1998): Processes of Normal Inverse Gaussian Type, *Finance Stoch.* 2, 41–68.

BARNDORFF-NIELSEN, O., and N. SHEPHARD (2004): Power and Bipower Variation with Stochastic Volatility and Jumps, *J. Financ. Econom.* 2, 1–48.

BELOMESTNY, D., and M. REISS (2006): Spectral Calibration of Exponential Lévy Models, *Finance Stoch.* 10, 449–474.

BLACK, F., and M. SCHOLES (1973): The Pricing of Options and Corporate Liabilities, *J. Polit. Econ.* 81, 637–654.

BUCKLE, D. J. (1995): Bayesian Inference for Stable Distributions, *J. Am. Stat. Assoc.* 90, 605–613.

CARR, P., and L. WU (2003): The Finite Moment Log Stable Process and Option Pricing, *J. Finance* 58, 753–777.

CARR, P., and L. WU (2004): Time-Changed Lévy Processes and Option Pricing, *J. Financ. Econ.* 71, 113–141.

CARR, P., H. GEMAN, D. MADAN, and M. YOR (2002): The Fine Structure of Asset Returns: An Empirical Investigation, *J. Business* 75, 305–332.

CARR, P., H. GEMAN, D. MADAN, and M. YOR (2003): Stochastic Volatility for Lévy Processes, *Math. Finance* 13, 345–382.

CASELLA, G., and R. BERGER (2001): *Statistical Inference*, Pacific Grove, CA: Duxbury Press.

CHIB, S., F. NARDARI, and N. SHEPHARD (2002): Markov Chain Monte Carlo Methods for Stochastic Volatility Models, *J. Econometr.* 108, 281–316.

CONT, R., and P. TANKOV (2004a): Non-Parametric Calibration of Jump-Diffusion Option Pricing Models, *J. Computat. Finance* 7, 1–49.

CONT, R., and P. TANKOV (2004b): *Financial Modelling with Jump Processes*, London: Chapman & Hall.

DAMINE, P., J. WAKEFIELD, and S. WALKER (1999): Gibbs Sampling for Bayesian Non-Conjugate and Hierarchical Models by Using Auxiliary Variables, *J. R. Stat. Soc.* 61, 331–344.

DEVROYE, L. (1986): *Nonuniform Random Variate Generation*, New York: Springer-Verlag.

DIEBOLD, F., and R. MARIANO (1995): Comparing Predictive Accuracy, *J. Business Econ. Stat.* 13, 253–265.

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, *J. Business* 71, 371–406.

ERAKER, B. (2004): Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *J. Finance* 59, 1367–1403.

ERAKER, B., M. JOHANNES, and N. POLSON (2003): The Impact of Jumps in Equity Index Volatility and Returns, *J. Finance* 58, 1269–1300.

GALLANT, R. (1987): *Nonlinear Statistical Models*, New York: John Wiley and Sons.

GILKS, W. (1992): Derivative-Free Adaptive Rejection Sampling for Gibbs Sampling, *Bayesian Statistics 4*, Oxford: Oxford University Press.

GILKS, W., N. BEST, and K. TAN (1995): Adaptive Rejection Metropolis Sampling, *Appl. Stat.* 44, 455–472.

GRIFFIN, J., and M. STEEL (2006): Inference with Non-Gaussian Ornstein-Uhlenbeck Processes for Stochastic Volatility, *J. Econometr.* 134, 605–644.

HESTON, S. (1993): A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Rev. Financ. Stud.* 6, 327–343.

HUANG, J., and L. WU (2003): Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes, *J. Finance* 59, 1405–1439.

JACQUIER, E., N. POLSON, and P. ROSSI (1994): Bayesian Analysis of Stochastic Volatility Models, *J. Business. Econ. Stat.* 12, 371–389.

KIM, S., N. SHEPHARD, and S. CHIB (1998): Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models, *Rev. Econ. Stud.* 65, 361–393.

LI, H., M. WELLS, and C. YU (2008): A Bayesian Analysis of Return Dynamics with Lévy Jumps, *Rev. Financ. Stud.* 21, 2345–2378.

MADAN, D., and E. SENETA (1990): The Variance Gamma (V.G.) Model for Share Market Returns, *J. Business* 63, 511–524.

MADAN, D., and F. MILNE (1991): Option Pricing with VG Martingale Components, *Math. Finance* 1, 39–55.

MADAN, D., P. CARR, and E. CHANG (1998): The Variance Gamma Process and Option Pricing, *Eur. Finance Rev.* 2, 79–105.

MAHEU, J., and T. MCCURDY (2004): News Arrival, Jump Dynamics and Volatility Components for Individual Stock Returns, *J. Finance* 59, 755–794.

MERTON, R. (1976): Option Pricing When the Underlying Stock Returns are Discontinuous, *J. Financ. Econ.* 3, 125–144.

NEWEY, W., and K. WEST (1987): A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

PAN, J. (2002): The Jump-Risk Premia Implicit in Options: Evidence From an Integrated Time-Series Study, *J. Financ. Econ.* 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 ed., New York: Springer.

SATO, K. (1999): *Lévy Processes and Infinitely Divisible Distributions*, Cambridge, UK: Cambridge University Press.

TAUCHEN, G., and V. TODOROV (2008): Activity Signature Functions for High-Frequency Data Analysis, Working Paper, Duke University.

WU, L. (2006a): Dampened Power Law: Reconciling the Tail Behavior of Financial Security Returns, *J. Business* 79, 1445–1474.

WU, L. (2006b): Modeling Financial Security Returns Using Lévy Processes, Working Paper, Baruch College.