

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**Bayesian Analysis of Stochastic Volatility Models with Levy
Jumps: Application to Risk Analysis**

Pawel J. Szerszen

2009-40

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

Bayesian Analysis of Stochastic Volatility Models with Lévy Jumps: Application to Risk Analysis

Pawel Szerszen*

Board of Governors of the Federal Reserve System

July 30, 2009

Abstract

In this paper I analyze a broad class of continuous-time jump diffusion models of asset returns. In the models, stochastic volatility can arise either from a diffusion part, or a jump part, or both. The jump component includes either compound Poisson or Lévy α -stable jumps. To be able to estimate the models with latent Lévy α -stable jumps, I construct a new Markov chain Monte Carlo algorithm. I estimate all model specifications with S&P500 daily returns. I find that models with Lévy α -stable jumps perform well in capturing return characteristics if diffusion is a source of stochastic volatility. Models with stochastic volatility from jumps and models with Poisson jumps cannot represent excess kurtosis and tails of return distribution. In density forecast and VaR analysis, the model with Lévy α -stable jumps and joint stochastic volatility performs the best among all other specifications, since both diffusion and infinite activity jump part provide information about latent volatility.

JEL classification: C1; C11; G1; G12

1 Introduction

In this paper I estimate a broad class of asset pricing models. I evaluate their performance with respect to goodness of fit, density forecast and Value at Risk (VaR) analysis. This is not trivial since there is a need for a balance between a level of model complexity - which always has a positive effect on the goodness of fit, and a possible extent of model overfitting - which decreases the forecasting power of the model. Specifically, I consider the family of continuous-time, time-changed jump diffusion models developed in Carr and Wu (2004). Stochastic volatility, or time-change, can arise either from a diffusion part, or a jump part, or both. The leverage effect is assumed to arise from the diffusion part if diffusion is a source of stochastic volatility. The jump component includes either finite activity compound Poisson or infinite activity Lévy α -stable jumps. I consider an estimation under the statistical measure, since it allows to perform density forecast and VaR analysis and use data on daily S&P 500 index returns. I choose this data for empirical study since it is a broad indicator of the equity market and it has been used in other comparable studies in

*The analysis and views presented here are solely mine and do not necessarily represent those of the Federal Reserve Board or its staff.

the literature. An important advantage of my empirical analysis is that I consider a large family of models. Therefore I can study in depth the marginal effects of different jump structures and source of stochastic volatility with respect to goodness of fit and density forecast performance.

My contribution to the literature is two-folded. First, I propose a Bayesian estimation method to estimate the general continuous-time, time-changed jump diffusion models with compound Poisson or, most importantly, infinite activity Lévy α -stable jumps. Second, I analyze the marginal contribution of jumps and volatility specifications in goodness of fit and density forecast. Intuitively, the more general jump structure with infinite activity should fit the data better than the finite activity compound Poisson jumps as found in Li, Wells and Yu (2008). However, it has not been studied in the literature, what the effect of infinite activity jumps is on the density forecast and VaR analysis. Moreover, it is important to address the choice of the source of stochastic volatility when we condition on the jump structure. How, if at all, the specification of stochastic volatility contributes to goodness of fit and density forecast?

I estimate my models by MCMC Bayesian methods and directly address the problem of parameter estimation in the presence of both latent volatility and latent jump sizes. The recent attempt to estimate models with latent Lévy α -stable jumps in returns by Li, Wells and Yu (2008) constitutes the foundation to solve this problem but it also introduces separability on the Markov chain state-space. I fill this gap in the literature by constructing an MCMC algorithm free of the separability flaw. The proposed algorithm is applicable in any stochastic volatility specification and is based on the Buckle's (1995) Bayesian method.

In my empirical analysis of the S&P 500 returns, I find that the models with Lévy α -stable jumps in returns are able to represent well excess kurtosis and skewness of return distribution, if diffusion is included as a source of stochastic volatility. Lévy α -stable jumps dominate Poisson jumps specifications with respect to goodness of fit analysis, since the latter are only suited to fit big jumps. Most importantly, models with stochastic volatility coming only from pure jumps do not fit the asset returns well. Nevertheless, based only on goodness of fit measure, one cannot in a decisive way point out if there is a need for the jump component as the second source of stochastic volatility. This conclusion holds for the models with all considered jump structures including infinite activity Lévy α -stable jumps.

The density forecast and VaR analysis shed new light on the application of continuous-time jump diffusion models of asset returns. I find that correct specification of the source of stochastic volatility is of fundamental importance in the density forecast and VaR analysis. The performance of the compound Poisson jump models do not significantly change with the addition of the jump component to the diffusion as the source of stochastic volatility. On the contrary, models with Lévy α -stable jumps improve in the density forecast and VaR performance with the inclusion of both sources of stochastic volatility, thus dominating all other model specifications. The joint stochastic volatility enables us to extract information about latent volatility from both diffusion and jumps, where the jumps are more informative with its infinite activity property. However, one cannot go further and exclude the diffusion from the source of stochastic volatility. This conclusion does not depend on the jump structure and agrees with the goodness of fit analysis.

The most difficult problem that arises in the density forecast analysis involves approximation of the filtering density. I follow the auxiliary particle filter approach, as developed in Pitt and Shephard (1999), and modify it to allow for the new features of my model. Durham (2006) extends the basic particle filter for models with leverage effect but does not include jumps in returns. Moreover, he works with particle filter and does not apply auxiliary particle filter involving index

parameter draws. Johannes, Polson and Stroud (2008) offer further refinements to the auxiliary particle filter algorithm for models with jumps and stochastic volatility. However, their algorithm cannot be applied to specifications with Lévy α -stable jumps. I refine auxiliary particle filter to study jump-diffusion models with leverage effect. Moreover, I allow for various sources of stochastic volatility and most importantly for Lévy α -stable jumps in returns.

My model specifications are not new and are based on the continuous-time, time-changed jump diffusion framework, which is the direct outcome of the evolution in the asset pricing literature that started with Black and Scholes (1973). However, their model produces disappointing results both in fitting time-series of returns and cross section of option prices, since it lacks the ability to represent non-normality of asset returns. In recent asset pricing literature, stochastic volatility and jumps are found to be important, allowing to represent skewness and excess kurtosis both in unconditional and conditional return distribution. Merton (1976) was the first to consider jump-diffusion models. Heston (1993) assumed volatility to be stochastic and followed the square-root Cox, Ingersoll and Ross (1985) (CIR) specification, while Jacquier, Polson and Rossi (1994) (JPR) assumed log-volatility specification. In this paper I follow JPR specification, since it does not require additional constraints on the parameters, satisfies non-negativity after discretization and allows for convenient interpretation of the parameters in the models with joint stochastic volatility.

Das and Sundaram (1999) found that jumps and stochastic volatility have different effects on the conditional asset return distribution and hence they play complimentary role in the option pricing literature. The generalized version of the model with both stochastic volatility and jumps required different techniques of estimation under statistical measure, where estimation problems arise from the unobservable stochastic volatility. This was partially resolved with development of efficient method of moments (EMM) estimation of Gallant and Tauchen (1996) and Bayesian Markov chain Monte Carlo (MCMC) methods. However, the class of models with Lévy α -stable jumps in returns and the class of models with various sources of stochastic volatility lack a robust estimation method under the statistical measure. I construct a new MCMC method to estimate these models.

The next generalization allowed for instantaneous correlation between increments of returns and volatility, the relation called leverage effect. Empirical results of Jacquier, Polson and Rossi (2004), Jones (2003), Andersen, Benzoni and Lund (2002) among others found the respective correlation to be significantly negative. The negative leverage effect has a deep intuitive explanation, since periods of high volatility on the market coincides more often with market crashes. The leverage effect helps in capturing the skewness of the stock returns and corrects estimates of parameters governing volatility as stated in Jacquier, Polson and Rossi (2004). Moreover, Jones (2003) included in one of his model specifications the leverage effect as a function of volatility. His findings suggest, that as volatility increases, the leverage effect is higher in magnitude. Hence, in periods of high volatility the probability of market crashes is higher than in periods with low volatility. Andersen, Benzoni and Lund (2002) estimated stochastic volatility models with compound Poisson jumps in returns and leverage effect under the statistical measure. They found that jumps, stochastic volatility and leverage effect are all important features of asset return models and generate skewness, excess kurtosis and conditional heteroscedasticity. Eraker, Johannes and Polson (2003) further extended the jump-diffusion model with stochastic volatility and studied jumps not only in returns but also in volatility. Although jumps in volatility are found to be an important feature in fitting the data on the 1987 crash, the discrete jumps in returns cannot be modeled successfully by jumps in volatility.

Since the arrival rate of Poisson jumps under the statistical measure was found to be small (about few jumps per year), the more subtle jumps cannot be modeled by rare and big compound Poisson jumps. This in turn is one of the main critiques of finite activity jumps in returns. The solution to this problem lies in the introduction of infinite activity Lévy jumps, that is the process with infinite number of "small" jumps in a finite time interval. The latest specifications include infinite activity jumps as in the case of variance-gamma (VG) model of Madan, Carr and Chang (1998) and CGMY class of models by Carr, Geman, Madan and Yor (2002). Li, Wells and Yu (2008) estimate the jump-diffusion model with VG jumps in returns and stochastic volatility from diffusion under the statistical measure and found its superior goodness of fit over the models with finite activity compound Poisson jumps. Lévy α -stable jumps, which are also of infinite activity, have already been studied in the literature under the risk-neutral measure in Huang and Wu (2004) and Carr and Wu (2003) but there has been so far no successful application of this jump structure under the statistical measure. A recent approach by Li, Wells and Yu (2008) introduces separability on the Markov chain state-space in the MCMC algorithm. I construct a robust MCMC algorithm to estimate models with Lévy α -stable jumps. In addition, I relax the assumption required in the option pricing approach that imposes maximum negative skewness on Lévy α -stable jumps, required to price options in a model with infinite second and higher moments. This allows modelling of the degree of skewness and the algorithm by Buckle (1995) is a suitable foundation to develop a method of estimation under the statistical measure. Finally, my analysis is based both on goodness of fit and density forecast. The latter is missing in the literature under the statistical measure for infinite activity jumps and hence I fill this gap in the literature. This lets us find how the models with infinite activity jumps perform in risk management. Finally, I allow for stochastic volatility to arise from diffusion, jumps or both and also look at its implications on the density forecast.

Another important issue in the literature has been the type of data used in the estimation. There are two general approaches to model asset returns. The first approach specifies models under the statistical measure, which allows for direct analysis of the return series and therefore density forecast and value at risk (VaR) analysis. The second approach uses options data and specifies models under the risk-neutral measure. There is also a way to utilize information from both worlds as in Chernov and Ghysels (2000) and Eraker (2004), however, it results in even further technical difficulties. Moreover, estimation under both the statistical and the risk-neutral measures requires definition of market risk premia, which can also be a potential source of misspecification as noted by Andersen, Benzoni and Lund (2002). Therefore I estimate the models under statistical measure which allows the study of density forecast and VaR analysis.

The rest of this paper is organized as follows, Section 2 introduces the concept of Lévy process and describes estimated model specifications, Section 3 describes MCMC estimation algorithm and the auxiliary particle filter, Section 4 gives a brief overview of the data used in the estimation and presents the results of the estimation with goodness of fit, density forecast and VaR performance analysis, and Section 5 concludes. The tables are presented at the end of the paper.

2 Model Specifications

2.1 Lévy Processes

In this section I closely follow Applebaum (2004) and Bertoin (1998). Let $X = (X_t, t \geq 0)$ be a scalar Lévy process defined on a probability space (Ω, \mathcal{F}, P) with given filtration $(\mathcal{F}_t)_{t \geq 0}$. From definition, a Lévy process X has independent and stationary increments, or more precisely, $X_s - X_t$ is independent of \mathcal{F}_t and has the same distribution as $X_{s-t} - X_0$ for all $0 \leq t \leq s$. It is also stochastically continuous. I restrict my analysis to the modification of X which exhibits càdlàg paths and hence its sample paths are right-continuous with left limits. By the Lévy-Ito decomposition every Lévy process can be decomposed as the sum of three independent processes: a linear drift, Brownian motion and a pure jump part. Accordingly, the log-characteristic function of a Lévy process is the sum of the log-characteristic functions of its Lévy components and is given by the Lévy-Khintchine formula. The characteristic function of the Lévy process is given by

$$\phi_{X_t}(u) = E[\exp(iuX_t)] = \exp(t\psi_x(u)), \quad t \geq 0, \quad (1)$$

where $u \in \mathbb{R}$ and $\psi_x(u)$ is called Lévy or the characteristic exponent of a given Lévy process. The Lévy-Khintchine formula determines the functional form of this exponent:

$$\psi_x(u) = ibu - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R} \setminus \{0\}} [\exp(iuz) - 1 - iuz1_{|z| < 1}]v(dz), \quad (2)$$

where $b \in \mathbb{R}$ controls the linear drift part, $\sigma \geq 0$ controls the Brownian part and v is a Lévy measure that characterizes the pure jump part of the Lévy process. The triplet (b, σ^2, v) completely characterizes the probabilistic behavior of the Lévy process. The Lévy measure v is a sigma-finite measure on $\mathbb{R} \setminus \{0\}$, not necessarily a finite measure, satisfying

$$\int_{\mathbb{R} \setminus \{0\}} \min(1, z^2)v(dz) < \infty. \quad (3)$$

The above condition implies finite quadratic variation of any Lévy process. We can extend the Lévy measure v to all \mathbb{R} , without loss of generality, assuming $v(\{0\}) = 0$. The Lévy measure has the interpretation that for any subset $E \subset \mathbb{R}$, $v(E)$ is the rate at which Lévy process takes jumps of size $x \in E$ and measures the numbers of jumps of size $x \in E$ in the unit time interval. The compound Poisson process is the only pure-jump ($\sigma = 0$) Lévy process, which satisfies $v(\mathbb{R}) < \infty$. For any other pure jump process $v([-\varepsilon, \varepsilon]) = \infty$ for any $\varepsilon > 0$ and hence in this case the Lévy process exhibits infinite number of small jumps in a finite time interval. For any Lévy process, however, the number of "large" jumps remains finite with $v(E) < \infty$ for any $E \subset \mathbb{R}$, $\bar{E} \cap \{0\} = \emptyset$. From now on we call Lévy process to have finite activity if and only if $v(\mathbb{R}) = \lambda < \infty$. The value λ is then called the Poisson intensity.

This class of processes is very general and contains Brownian motion and compound Poisson process as two special cases. Brownian motion is the only Lévy process with continuous sample paths and hence does not allow for discontinuous jumps. The compound Poisson jump process, however, represents special jump characteristics with its finite activity property. The sum of the Brownian part and the compound Poisson part, although a Lévy Process, does not allow for more

general jump structures and is one of the main critiques of asset returns models based on them. In this work I allow for more general properties of the jump structure by redefining the jump part of the underlying asset returns process and allowing for infinite activity. The pure jump Lévy process with infinite activity, however, can also be classified into two general sub-classes with respect to the total absolute variation of the process. The Lévy pure jump process is of finite total variation if the following condition is satisfied by its Lévy measure:

$$\int_{\mathbb{R} \setminus \{0\}} \min(1, |z|) v(dz) < \infty, \quad (4)$$

otherwise it is of infinite total variation.

Since infinite variation jumps resemble Brownian motion much closer than other types of jumps, I restrict my analysis to the Lévy α -stable pure jump process with index of stability $\alpha \in (1, 2)$, the Lévy process with infinite total variation¹. I also investigate another extreme case with finite activity Poisson type jumps, since its simplicity decreases an extent of possible overfitting problems.

2.2 Lévy α -stable Process

The building block of α -stable process is a stable distribution. Let $S(\alpha, \beta, \delta, \gamma)$ denote a stable distributed random variable with index of stability $\alpha \in (1, 2)$, skewness $\beta \in [-1, 1]$, scale parameter $\gamma > 0$, and location parameter $\delta \in \mathbb{R}$. In this paper I use the characteristic function specification as in Buckle (1995):

$$E(\exp(iuS)) = \begin{cases} \exp\{i\delta u - \gamma^\alpha |u|^\alpha \exp[-i\frac{\pi\beta}{2} \text{sgn}(u) \min(\alpha, 2 - \alpha)]\} & \text{for } \alpha \neq 1 \\ \exp\{i\delta u - \gamma |u| [1 + i\beta \frac{2}{\pi} \text{sgn}(u) \log(\gamma |u|)]\} & \text{for } \alpha = 1, \end{cases} \quad (5)$$

In this parametrization, parameter β controls the extent of skewness in the distribution with maximum positive skewness given by $\beta = -1$ and maximum negative skewness given by $\beta = 1$. The extent of skewness disappears as $\alpha \nearrow 2$, where the parameter α controls the fatness of tails, where at the limit $\alpha = 2$ we have the normal distribution. In general case, for $\alpha \in (1, 2)$, there is no closed form density function available. For a discussion of the parametrizations and properties of the stable distribution please refer to Nolan (2005). An efficient method to simulate stable random variables is presented by Chambers, Mallows and Stuck (1976). The parametrization (5), used in Buckle (1995), is applied in this paper unless otherwise stated.

The most widely used parametrization is given by Samorodnitsky and Taqqu (1994) and Zolotarev (1986) and denoted by $\bar{S}(\alpha, \bar{\beta}, \delta, \bar{\gamma})$, where:

$$\begin{aligned} \bar{\beta} &= \cot\left(\frac{\pi\alpha}{2}\right) \tan\left(\frac{\pi\beta}{2} \min(\alpha, 2 - \alpha)\right) \\ \bar{\gamma} &= \gamma \left[\cos\left(\frac{\pi\beta}{2} \min(\alpha, 2 - \alpha)\right)\right]^{1/\alpha}, \end{aligned} \quad (6)$$

and δ and α remain unchanged. It can be shown, that $\bar{S}(\alpha, \bar{\beta}, \delta, \bar{\gamma}) \stackrel{d}{=} S(\alpha, \beta, \delta, \gamma)$. The skewness parameter $\bar{\beta}$ has a different interpretation, where $\bar{\beta} > 0$ denotes positive skewness and $\bar{\beta} < 0$

¹Refer to Samorodnitsky and Taqqu (1994) and Janicki and Weron (1994) for the properties of stable random processes.

negative skewness. Note, that parameters $\bar{\gamma}$ and γ are proportional.

Since the stable distribution is infinitely divisible there exists a Lévy process $\{J_t^{SJ}\}$ with stable distributed increments - the Lévy α -stable process²:

$$J_t^{SJ} - J_s^{SJ} \stackrel{d}{=} S(\alpha, \beta, 0, \sigma(t-s)^{1/\alpha}) \text{ for } 0 \leq s \leq t.$$

I assume that there is no deterministic drift in this pure jump process with the restriction of $\delta = 0$. I will extensively apply the following scaling property of the stable distribution:

$$\sigma S(\alpha, \beta, \delta, \gamma) \stackrel{d}{=} S(\alpha, \beta, \delta, \sigma\gamma) \text{ for } \sigma > 0. \quad (7)$$

Since $\bar{\gamma}$ -scale parameter given by the translation formula (6) is proportional to γ , the scaling property holds for both parametrizations. Hence, without loss of generality, we can use both parametrizations to characterize the increments of Lévy α -stable process. Moreover, it can be shown, that Lévy α -stable process $\{J_t^{SJ}\}$ defined above is a pure jump Lévy process with infinite activity and infinite variation. For a more detailed exposition please refer to Samorodnitsky and Taqqu (1994) and Janicki and Weron (1994).

2.3 Dynamics of the Asset Returns Process

Let Y_t denote the logarithm of asset price or logarithm of the index level at time t and $Y_{t+1} - Y_t$ be the corresponding log-return. I consider several specifications that differ in the source of stochastic volatility in the returns process and the type of jump component in returns. In the following, $(B_t^{(1)}, B_t^{(2)})$ defines a two-dimensional standard Brownian motion on (Ω, \mathcal{F}, P) probability space defined above. Carr and Wu (2004) noted that stochastic volatility can be alternatively interpreted as the stochastic time change of the underlying processes. I define the following time-changed process $\{Y_t\}$, being a semimartingale:

$$\begin{aligned} dY_t &= \mu dt + dB_{T_t^B}^{(1)} + dJ_{T_t^J}^{(M)}, \\ dh_t &= \kappa_h(\theta_h - h_t)dt + \sigma_h(\rho dB_t^{(1)} + \sqrt{1-\rho^2}dB_t^{(2)}), \\ T_t^J &= \int_0^t \lambda(h_s)ds, \quad T_t^B = \int_0^t \xi(h_s)ds, \quad M \in \{PJ, SJ\}, \end{aligned} \quad (8)$$

where $\mu \in \mathbb{R}$ defines the drift part of the return process, $\kappa_h \in \mathbb{R}$ defines the speed of the mean reversion of the log-volatility h_t process towards its mean $\theta_h \in \mathbb{R}$, $\sigma_h > 0$ defines volatility of the volatility parameter. There are two types of jumps considered, where M denotes a model specification with values PJ and SJ respectively for Poisson and Lévy α -stable jumps. The stochastic volatility from diffusion is governed by the process $\xi(h_s) > 0$ and stochastic volatility from jumps by the process $\lambda(h_s) > 0$, both taking only positive values. I assume, that the functions $\lambda : \mathbb{R} \rightarrow \mathbb{R}^+$ and $\xi : \mathbb{R} \rightarrow \mathbb{R}^+$ are continuous. The parameter $\rho \in (-1, 1)$ controls the leverage effect. This can be seen by defining the process $B_t^{(3)}$:

$$dB_t^{(3)} = \rho dB_t^{(1)} + \sqrt{1-\rho^2}dB_t^{(2)}. \quad (9)$$

²For proof and further details please see Applebaum (2004).

It can be shown, that $B_t^{(3)}$ is a Brownian motion and $E_t(dB_t^{(1)}dB_t^{(3)}) = \rho dt$ and hence the parameter ρ can be interpreted as the instantaneous correlation between volatility and returns - the leverage effect. Since any pure jump component $J_t^{(M)}$ is independent from the continuous sample path Ornstein-Uhlenbeck (OU) process governing volatility, we cannot model leverage effect from jumps and I impose restriction $\rho = 0$ in specifications with stochastic volatility only from jumps. The log-volatility specification is borrowed from Jacquier, Polson and Rossi (1994) and leverage effect specification from Jacquier, Polson and Rossi (2004). The leverage effect has been extensively studied in recent research and was found to be an important characteristic of the asset return models both under statistical measure in Jacquier, Polson and Rossi (2004) and risk-neutral measure in Huang and Wu (2004).

- Models with Poisson jumps.

My model specifications with Poisson jumps draw from the work of Andersen, Benzoni and Lund (2002) and Eraker, Johannes and Polson (2003) among others. The compound Poisson jump process $J_t^{(PJ)}$ is characterized by its normally distributed jumps with mean $\mu_j \in \mathbb{R}$, variance σ_j^2 and unit jump intensity. The time changed process $J_{T_t^J}^{(PJ)}$ has an instantaneous Poisson arrival intensity $\lambda(h_s) > 0$ and the jump compensator $\lambda(h_t)\phi(x; \mu_j, \sigma_j^2)dxdt$, where $\phi(\cdot; \mu_j, \sigma_j^2)$ is a pdf of normal distribution with mean μ_j and variance σ_j^2 and hence $v^{(PJ)}(dx) = \phi(x; \mu_j, \sigma_j^2)dx$ is a Lévy measure of jumps of the compound Poisson process $J_t^{(PJ)}$.

- Models with Lévy α -stable jumps.

The idea of modelling asset returns with Lévy α -stable jumps is not new to the asset pricing literature. Carr and Wu (2003) and Huang and Wu (2004) applied models with both the diffusion and Lévy α -stable jumps to model asset returns under risk-neutral measure. I pursue similar specification with its application under statistical measure and I loosen up their assumption of maximum negative skewness. In the model above $J_t^{(SJ)}$ is a Lévy α -stable process with stable distributed increments:

$$J_t^{(SJ)} - J_s^{(SJ)} \stackrel{d}{=} S(\alpha, \beta, 0, (t-s)^{1/\alpha}), \text{ for } 0 \leq s < t,$$

with index of stability $\alpha \in (1, 2)$ and skewness parameter $\beta \in [-1, 1]$. The time-changed process $J_{T_t^J}^{(SJ)}$ has a jump compensator proportional to stochastic volatility from jumps and given by $\xi(h_t)v^{(SJ)}(dx)dt$, where $v^{(SJ)}(dx)$ denotes the Lévy measure of jumps of Lévy α -stable process $J_t^{(SJ)}$.

- Restrictions defining all model specifications.

The restrictions on the parameter ρ and the predictable functions $\lambda(\cdot)$, $\xi(\cdot)$ completely characterize all model specifications and are provided in Table I. I specify six model specifications, where the models (1), (2) and (3) have a Poisson jump component, and the models (4), (5) and (6) have a Lévy α -stable jump component. For each jump type I distinguish three sources of stochastic volatility: from diffusion, jumps and jointly: from the diffusion and jumps. For models with stochastic volatility only from the jump component, I consider specification without leverage

effect and I impose the restriction $\rho = 0$. In the other specifications I model the leverage effect and estimate $\rho \in (-1, 1)$. In order to model the source of stochastic volatility I have to define the functions $\lambda(h_s)$ and $\xi(h_s)$ governing the instantaneous "speed", or the time rate, of the business time respectively T_t^J and T_t^B . The affix PJ denotes Poisson type jumps and SJ denotes Lévy α -stable jumps.

In models with stochastic volatility only from diffusion (PJ, SJ), I specify $\lambda(\cdot)$ to be a positive constant, given by $\lambda^{PJ}(h_s) = \lambda_j > 0$ for the model with Poisson jumps (PJ) and $\lambda^{SJ}(h_s) = (\sigma_{SJ})^\alpha$ for the model with Lévy α -stable jumps (SJ) with $\sigma_{SJ} > 0$. I define the constant volatility from Lévy α -stable jumps to be a function of α , since the model simplifies significantly after discretization presented in Section 3.1. In order to specify stochastic volatility from diffusion in models PJ and SJ, I assume the log-volatility specification with $\xi(h_s) = \exp(h_s)$.³

In models with stochastic volatility from jumps (PJSV, SJSV), I specify $\lambda(h_s) = \exp(h_s)$ and $\xi(\cdot) = \sigma^2 > 0$ as a constant volatility from diffusion. In these models $\rho = 0$, since the pure jump part is independent from OU process governing the stochastic volatility and I do not model the leverage effect. Finally, in the class of models with joint stochastic volatility from both the diffusion and jumps (DiffPJSV, DiffSJSV), I specify $\lambda^{PJ}(h_s) = \exp(h_s)$, $\xi^{PJ}(h_s) = \sigma^2 \cdot \exp(h_s)$, $\lambda^{SJ}(h_s) = (\sigma_{SJ})^\alpha \cdot \exp(h_s)$, $\xi^{SJ}(h_s) = \exp(h_s)$ and $\sigma, \sigma_{SJ} > 0$. In these models the parameters $\sigma > 0$ and $\sigma_{SJ} > 0$ are identified, since stochastic volatility process $\{h_t\}$ drives both the diffusion and jumps, and hence drives the wedge between levels of log-volatilities for the diffusion and jump components. Without loss of generality I assume that σ drives this wedge via shift in the stochastic volatility from diffusion $\xi^{PJ}(\cdot)$ in the model with Poisson jumps and σ_{SJ} drives this wedge via stochastic volatility from jumps $\lambda^{SJ}(\cdot)$ in the model with Lévy α -stable jumps. This overcomes several estimation issues in models DiffPJSV and DiffSJSV.⁴

Summing up, I define three model specifications with Poisson jumps: model (1) PJ, model (2) PJSV and model (3) DiffPJSV. Accordingly, I have other three specifications with Lévy α -stable jumps: model (4) SJ, model (5) SJSV and model (6) DiffSJSV. The summary of all restrictions, defining each specification, is presented in Table I.

3 Estimation Method

3.1 Discretization scheme

In order to estimate the parameters of the continuous-time specifications I need to discretize the models. In the following I use first order Euler scheme⁵. $\varepsilon_t^{(1)}, \varepsilon_t^{(2)}, \varepsilon_t^{(J)}$ are independent *iid* $N(0, 1)$ distributed and all other random variables are also independent:

³For sake of notational simplicity I omit the superscripts PJ and SJ whenever specifications of volatility process is the same for both jump specifications.

⁴Note that both ways of introducing the joint stochastic volatilities are equivalent. This can be seen by reparametrization in the discretized versions of the models in Section 3.1.

⁵Refer to Kloeden and Platen (1992) for the details on the higher order approximation techniques. Jacod and Protter (1998) analyze the Euler scheme for SDEs with Lévy jumps.

- Models (1) - (3) with Poisson jumps:

$$\begin{aligned}
Y_{t+\delta} &= Y_t + \mu\delta + (\xi_t\delta)^{0.5}\varepsilon_{t+\delta}^{(1)} + q_{t+\delta}\varkappa_{t+\delta}, \\
h_{t+\delta} &= h_t + \kappa_h(\theta_h - h_t)\delta + \sigma_h\sqrt{\delta}(\rho\varepsilon_{t+\delta}^{(1)} + \sqrt{1-\rho^2}\varepsilon_{t+\delta}^{(2)}), \\
\varkappa_{t+\delta} &\sim iid N(\mu_j, \sigma_j^2), \\
q_{t+\delta}/\{h_t\} &\sim independent Bernoulli (\delta\lambda_t).
\end{aligned} \tag{10}$$

- Models (4) - (6) with Lévy α -stable jumps:

$$\begin{aligned}
Y_{t+\delta} &= Y_t + \mu\delta + (\xi_t\delta)^{0.5}\varepsilon_{t+\delta}^{(1)} + S_{t+\delta}(\alpha, \beta, 0, (\lambda_t\delta)^{\frac{1}{\alpha}}), \\
h_{t+\delta} &= h_t + \kappa_h(\theta_h - h_t)\delta + \sigma_h\sqrt{\delta}(\rho\varepsilon_{t+\delta}^{(1)} + \sqrt{1-\rho^2}\varepsilon_{t+\delta}^{(2)}).
\end{aligned} \tag{11}$$

In the above $t = 1, 2, \dots, T$, and $S_t(\alpha, \beta, 0, \gamma_{t-1})$, given $\alpha, \beta, \{\gamma_t\}$, is centered stable distributed with index of stability $\alpha \in (1, 2)$, skewness coefficient $\beta \in [-1, 1]$ and with respective scale parameters $\gamma_{t-1} > 0$ in the parametrization given by the characteristic function in (5). For notational simplicity I define $\lambda_t = \lambda(h_t)$ and $\xi_t = \xi(h_t)$. All other variables and parameters are defined in Section 2.3 with the respective constraints on the parameter ρ and the functions λ_t and ξ_t defining all model specifications.

The problem I face concerns a choice of $\delta > 0$ parameter, which governs the extent of the discretization bias. In this paper I fix $\delta = 1$ and use the data at daily frequency. As noted by Eraker, Johannes and Polson (2003) the discretization bias of daily data is not significant⁶.

Since my models are estimated at the daily frequency, in models (2) and (3) with Poisson jumps and stochastic volatility component from jumps, the volatility levels $\lambda_t = \exp(h_t)$ are close to zero. Hence, following Johannes and Polson (2003), I allow for maximum one jump per day. I consider the following approximation of the function governing stochastic volatility from Poisson jumps:

$$\lambda_t \equiv \exp(h_t) \approx (1 + \exp(-h_t))^{-1}, \quad h_t \ll 0. \tag{12}$$

The relative error of this approximation is given by $(1 + \exp(-h_t))^{-1}$ and is of negligible order at the daily frequency. Hence, λ_t takes a logistic form and is bounded from above by one. Since I allow for maximum one jump per unit of time, λ_t is an instantaneous probability of jump in a given time interval. This solves the problem of truncation of $\{h_t\}$ at zero to impose unit upper bound on the activity levels and guarantees continuity, which solves estimation problems for models with Poisson jumps. Similarly, in model (1) with Poisson jumps and constant volatility from jumps λ_j , I restrict the constant jump intensity $\lambda_j \in [0, 1]$ and allow for maximum one jump per day.

Since models with infinite activity jumps in returns have an infinite number of small jumps in a finite time, an identification problem arises if we are able to disentangle them from the continuous-path Brownian part. The recent work by Aït-Sahalia (2003) provides the positive theoretical answer for the simple model of asset returns with Cauchy jumps (stable jumps with

⁶Jones (1998) and Jones (2003) allow for an estimation with $\delta < 1$, where the data points at not observed frequencies are treated as latent variables. This approach is not feasible in the models with Lévy α -stable jumps due to the computational limitations.

$\alpha = 1$ and $\beta = 0$) and with constant volatility from diffusion. Finally, in Aït-Sahalia and Jacod (2008) a test is constructed to verify existence of jumps in the discretely observed continuous-time process. Since discretely sampled data allows to disentangle infinite activity jumps from diffusion, the test provides positive identification answer for models with infinite activity jumps and the diffusion.

3.2 Markov chain Monte Carlo methods

In this section I briefly describe Markov chain Monte Carlo (MCMC) methods, with more detailed exposition in Chib and Greenberg (1996), Johannes and Polson (2003) and Jones (1998).

Let $Y = \{Y_t\}_{t=1}^T$ denote the observations, X are the unobserved (latent) state variables and θ are the parameters of the model. In the Bayesian inference we utilize the prior information on the parameters to derive the joint posterior distribution for both parameters and state variables. By the Bayes rule, we have:

$$p(\theta, X|Y) \propto p(Y|X, \theta) p(X|\theta) p(\theta), \quad (13)$$

where $p(Y|X, \theta)$ is the likelihood function of the model, $p(X|\theta)$ is the probability distribution of state variables conditional on the parameters and $p(\theta)$ is the prior probability distribution on the parameters of the model. Ideally we would like to know the analytical properties of the joint posterior distribution of X and θ , however, this is hardly feasible. The highly multidimensional joint posterior distribution is very often too complicated to work with and analytically intractable and hence even direct simulation from the joint posterior distribution is hard to perform. The remedy to this problem is to break the multidimensional distribution $p(\theta, X|Y)$ into its complete conditional distributions proposed by Clifford and Hammersley. They proved that the set of complete conditional distributions completely characterizes the joint distribution. In other words, knowledge of conditional distributions $p(\theta|X, Y)$ and $p(X|\theta, Y)$ determines the joint posterior distribution $p(\theta, X|Y)$. We can continue in this manner and characterize the joint posterior distribution $p(\theta, X|Y)$ by the set of complete one-dimensional conditional distributions, or group the variables in several blocks if we have knowledge on the respective higher dimensional conditional distributions. The MCMC algorithm can be defined as the way to construct a Markov chain, with invariant distribution as the desired target distribution, by consecutively drawing from the conditional posterior distributions. The simplest MCMC algorithm is a Gibbs sampler developed in Geman and Geman (1984). The proof of the Gibbs sampler, sufficient conditions and some applications can be found in Chib and Greenberg (1996).

The Gibbs sampler provides useful methods to draw samples from complicated and non-standard distributions. However, it assumes that we can sample directly from the set of all complete conditional distributions. If we face a problem of sampling from intractable distribution, we can replace the particular Gibbs sampler step by the Metropolis-Hastings (MH) step in Metropolis, Rosenbluth and Rosenbluth (1953). Further details about the MH algorithm can be found in Chib and Greenberg (1996).

In my work I am interested in obtaining random samples from the posterior distribution $p(\theta, X|Y)$. This allows for computation of several statistics including the sample means and higher moments from the desired marginal posterior distributions. The sample mean from the posterior distribution of the parameters is taken to be the population parameter estimate. Moreover, the

ergodic averaging theorem guarantees almost sure convergence to the true population moments (Johannes and Polson (2003)).

3.3 Bayesian Inference for Stable Distributions

In my application with latent Lévy α -stable jumps one of the sufficient conditions for the Gibbs sampler to converge needs to be carefully addressed. The constructed Markov chain should be constructed in a way, that guarantees strictly positive probability of visiting any subspace of the support of the target density. If Markov chain does not satisfy this condition, I call it a *separability problem*. Li, Wells and Yu (2008) do not correct for the separability problem in their MCMC algorithm derived for the latent stable jumps. This leaves their results questionable and demands alternative approach to the estimation of the latent stable distributed jumps.

The main problem in the application of the Bayesian MCMC methods for stable distributions is the nonexistence of its density function for index of stability $\alpha \in (1, 2)$. Buckle (1995) found a solution to this problem by introducing auxiliary variable, such that the joint density of the auxiliary variable and the stable distributed random variable exists. Let \tilde{S} and \tilde{Y} be the random variables with their joint density f , conditional on α, β, δ and γ : given by:

$$f : (-\infty, 0) \times (-\frac{1}{2}, l_{\alpha,\beta}) \cup (0, \infty) \times (l_{\alpha,\beta}, \frac{1}{2}) \longrightarrow (0, \infty)$$

$$f(\tilde{s}, \tilde{y}) = \frac{\alpha}{\gamma|\alpha - 1|} \exp \left\{ - \left| \frac{z}{t_{\alpha,\beta}(\tilde{y})} \right|^{\alpha/(\alpha-1)} \right\} \left| \frac{z}{t_{\alpha,\beta}(\tilde{y})} \right|^{\alpha/(\alpha-1)} \frac{1}{|z|}, \quad (14)$$

where

$$z = \frac{(\tilde{s} - \delta)}{\gamma},$$

$$t_{\alpha,\beta}(\tilde{y}) = \left(\frac{\sin[\pi\alpha\tilde{y} + \eta_{\alpha,\beta}]}{\cos \pi\tilde{y}} \right) \left(\frac{\cos \pi\tilde{y}}{\cos[\pi(\alpha - 1)\tilde{y} + \eta_{\alpha,\beta}]} \right)^{(\alpha-1)/\alpha}, \quad (15)$$

and $\alpha \in (1, 2)$, $\beta \in [-1, 1]$, $\delta \in (-\infty, +\infty)$, $\gamma > 0$, $\tilde{s} \in (-\infty, +\infty)$, $\tilde{y} \in (-\frac{1}{2}, \frac{1}{2})$, with $\eta_{\alpha,\beta} = \beta(2 - \alpha)\pi/2$ and $l_{\alpha,\beta} = -\eta_{\alpha,\beta}/\pi\alpha$. According to theorem 1. in Buckle (1995), f is a proper joint density of (\tilde{S}, \tilde{Y}) and the marginal distribution of \tilde{S} is $S(\alpha, \beta, \delta, \gamma)$. It is important to note that the domain of the density function is $(\tilde{s}, \tilde{y}) \in (-\infty, 0) \times (-\frac{1}{2}, l_{\alpha,\beta}) \cup (0, \infty) \times (l_{\alpha,\beta}, \frac{1}{2})$. Hence we have the following dependence between \tilde{S} and \tilde{Y} random variables:

$$\tilde{S} > 0 \iff \tilde{Y} \in (l_{\alpha,\beta}, \frac{1}{2}) \quad (16)$$

and

$$\tilde{S} < 0 \iff \tilde{Y} \in (-\frac{1}{2}, l_{\alpha,\beta}). \quad (17)$$

In my application I have to draw S_t for all t conditional on all other state variables and parameters as in the Gibbs sampler. Since one of the conditioning state variables is the auxiliary variable \tilde{Y}_t , it uniquely determines the sign of the draw $S_t^{(i)}$ at the $i - th$ step of the Gibbs sampler.

This violates one of the main assumptions of the MCMC method since the state space cannot be separated into two subspaces according to the sign of the starting value of $S_t^{(1)}$ - the sign that it would never leave. To illustrate the problem, let the starting values in the Gibbs sampler specify $S_t^{(1)} > 0$, $\tilde{Y}_t^{(1)} \in (l_{\alpha^{(1)},\beta^{(1)}}, \frac{1}{2})$ for some t , and all other parameters, including $\alpha^{(1)}$, $\beta^{(1)}$, $\delta^{(1)}$, $\gamma^{(1)}$ (consistent with chosen values $S_t^{(1)}$ and $\tilde{Y}_t^{(1)}$, in the support of the joint distribution and with other state variables). Suppose, without loss of generality, that we have to first update the jump size S_t in the algorithm. Since $\tilde{Y}_t^{(1)} \in (l_{\alpha^{(1)},\beta^{(1)}}, \frac{1}{2})$ we have to draw $S_t^{(2)} > 0$. In the next step the draw of all other jump specific parameters $\alpha^{(2)}$, $\beta^{(2)}$, $\delta^{(2)}$, $\gamma^{(2)}$ have to be consistent with $S_t^{(2)} > 0$ and $\tilde{Y}_t^{(1)} \in (l_{\alpha^{(2)},\beta^{(2)}}, \frac{1}{2})$. At the end we have to update the auxiliary variable $\tilde{Y}_t^{(2)}$ in support of the joint distribution, hence $\tilde{Y}_t^{(2)} \in (l_{\alpha^{(2)},\beta^{(2)}}, \frac{1}{2})$. Continuing in this manner we construct an MCMC chain that never visits negative values of jump sizes at time t . The algorithm has to draw $S_t^{(i)}$ for all iterations i with the same sign as the starting value $S_t^{(1)}$. However, if we do not treat the jump variables as latent and we observe the jump sizes S_t for all t as in the Buckle (1995), there is no update step of the jump sizes and there are no MCMC separability issues.

In this paper I offer a solution to this problem by construction of the mixture distribution of two, truncated at zero, stable distributions. Lets define the following probability:

$$p_{\alpha,\beta,\delta} = P(\tilde{S} > 0, \tilde{Y} \in (l_{\alpha,\beta}, \frac{1}{2})) = P(\tilde{S} > 0), \quad (18)$$

which for $\delta = 0$ can be found analytically to be:

$$p_{\alpha,\beta} = p_{\alpha,\beta,\delta=0} = 0.5 + \frac{\arctan[\hat{\beta} \tan(\pi\alpha/2)]}{\pi\alpha}, \quad (19)$$

where $\hat{\beta}$ is given by the translation of parameters formula in eq. (6) $\hat{\beta} = \cot(\frac{\pi\alpha}{2}) \tan(\frac{\pi\beta}{2} \min(\alpha, 2 - \alpha))$ to the Buckle (1995) characteristic function specification. This formula is based on the value of the stable distribution function at zero in Nolan (2005). In the next step I have to define the distribution of truncated stable variables \tilde{S}^+ , \tilde{S}^- and their respective auxiliary variables \tilde{Y}^+ , \tilde{Y}^- . Let $(\tilde{S}^+, \tilde{Y}^+)$ and $(\tilde{S}^-, \tilde{Y}^-)$ have joint distributions defined respectively by the following density functions:

$$f^+(\tilde{s}^+, \tilde{y}^+) = \begin{cases} \frac{1}{p_{\alpha,\beta}} f(\tilde{s}^+, \tilde{y}^+; \delta = 0) & \text{if } \tilde{s}^+ > 0 \text{ and } \tilde{y}^+ \in (l_{\alpha,\beta}, \frac{1}{2}) \\ 0 & \text{otherwise} \end{cases}, \quad (20)$$

$$f^-(\tilde{s}^-, \tilde{y}^-) = \begin{cases} \frac{1}{1-p_{\alpha,\beta}} f(\tilde{s}^-, \tilde{y}^-; \delta = 0) & \text{if } \tilde{s}^- < 0 \text{ and } \tilde{y}^- \in (-\frac{1}{2}, l_{\alpha,\beta}) \\ 0 & \text{otherwise} \end{cases}, \quad (21)$$

where the density f is defined in eq. (14). Moreover, let \tilde{U} be a Bernoulli distributed (conditional on α and β) random variable with probability of success $p_{\alpha,\beta}$. Let \tilde{U} , $(\tilde{S}^+, \tilde{Y}^+)$, $(\tilde{S}^-, \tilde{Y}^-)$ be independent, then it is straightforward to show, that:

$$\tilde{S} \stackrel{d}{=} \tilde{U} \cdot \tilde{S}^+ + (1 - \tilde{U}) \cdot \tilde{S}^- \sim S(\alpha, \beta, \delta = 0, \gamma). \quad (22)$$

This specification complicates the MCMC algorithm by introducing mixing variable \tilde{U} and auxiliary variables \tilde{Y}^+ and \tilde{Y}^- for respectively positive jumps \tilde{S}^+ and negative jumps \tilde{S}^- . All of

them are updated in the MCMC algorithm. However, the above specification solves the problem of separability of the resulting Markov chain state-space in the models with latent stable jumps.

3.4 Estimation Algorithm

In this section I briefly describe the set of complete conditional distributions to be used in the MCMC algorithm. My algorithm allows for the most general stochastic volatility specifications, both from the diffusion and infinite activity jumps, the new feature in the asset pricing literature under statistical measure. Since models with Poisson jumps in returns have already been studied in the literature, I postpone their model specific derivations to the appendix.

In the following I concentrate attention on the jump specific parameters and state variables in the model with Lévy α -stable jumps, joint stochastic volatility and leverage effect - DiffSJSV specification. Other model specifications can be approached in a similar way with specific constraints on the parameters ρ and the functions governing stochastic volatility ξ_t and λ_t described in Section 2.3 and Table I.⁷

In the sequel I assume the number of daily observations T and discretization parameter $\delta = 1$. I present the detailed discussion of updating pure jump sizes $\{\tilde{S}_t^+\}_{t=2}^T$, $\{\tilde{S}_t^-\}_{t=2}^T$, their respective auxiliary variables $\{\tilde{Y}_t^+\}_{t=2}^T$, $\{\tilde{Y}_t^-\}_{t=2}^T$, the mixing variables $\{\tilde{U}_t\}_{t=2}^T$ and the jump specific parameters α, β .

Let $\Xi^{-\Upsilon} = \Xi \setminus \{\Upsilon\}$ for sets Ξ and Υ , $S_t = \tilde{U}_t \tilde{S}_t^+ + (1 - \tilde{U}_t) \tilde{S}_t^-$ with truncation at zero defined above for S_t with Lévy α -stable distribution, and $\varepsilon_t^{(3)} = \rho \varepsilon_t^{(1)} + \sqrt{1 - \rho^2} \varepsilon_t^{(2)}$.

Note that in model DiffSJSV, given $\alpha, \beta, \sigma_{SJ}, \{h_t\}$ we have $S_t \sim \text{independent } S(\alpha, \beta, 0, (\lambda_{t-1} \delta)^{\frac{1}{\alpha}})$. Hence, \tilde{S}_t^+ and \tilde{S}_t^- are, given $\alpha, \beta, \sigma_{SJ}, \{h_t\}$, the respective jointly independent, truncated (at zero) parts of S_t . Moreover, the realization of $\{h_t\}$, having its impact only on the scale parameter, does not affect the distribution governing the mixing variables \tilde{U}_t which are still Bernoulli with parameter $p_{\alpha, \beta}$ in eq. (19).

Let $\theta = (\mu, \sigma_{SJ}, \kappa_h, \theta_h, \sigma_h, \alpha, \beta, \rho)$, $Y = (Y_1, Y_2, \dots, Y_T) = \{Y_t\}_{t=1}^T$ and

$$X = (\{h_t\}_{t=1}^T, \{\tilde{U}_t\}_{t=2}^T, \{\tilde{S}_t^+\}_{t=2}^T, \{\tilde{S}_t^-\}_{t=2}^T, \{\tilde{Y}_t^+\}_{t=2}^T, \{\tilde{Y}_t^-\}_{t=2}^T) \quad (23)$$

be respectively the vector of parameters, the observed log-asset prices and the vector of state variables.

3.4.1 Updating auxiliary variables $\tilde{Y}_t^+, \tilde{Y}_t^-$

Define the following change of variables $v_t^\pm = t_{\alpha, \beta}(\tilde{Y}_t^\pm)$. As proved in Buckle (1995), function $t_{\alpha, \beta} : (-\frac{1}{2}, \frac{1}{2}) \mapsto \mathbb{R}$ in eq. (15) is increasing for given parameters $\alpha \in (1, 2)$ and $\beta \in [-1, 1]$. Moreover, for $\beta \in (-1, 1)$, $t_{\alpha, \beta} \nearrow \infty$ as $\tilde{Y}_t^+ \nearrow \frac{1}{2}$ and $t_{\alpha, \beta} \searrow -\infty$ as $\tilde{Y}_t^- \searrow -\frac{1}{2}$.

From (14) the conditional posterior for \tilde{Y}_t^\pm , $t = 2, \dots, T$, is given by:

$$p(\tilde{Y}_t^\pm | \theta, X^{-\{\tilde{Y}_t^\pm\}}, Y) \propto \exp \left\{ - \left| \frac{\tilde{S}_t^\pm}{(\lambda_{t-1} \delta)^{\frac{1}{\alpha}} t_{\alpha, \beta}(\tilde{Y}_t^\pm)} \right|^{\frac{\alpha}{\alpha-1}} + 1 \right\} \left| \frac{\tilde{S}_t^\pm}{(\lambda_{t-1} \delta)^{\frac{1}{\alpha}} t_{\alpha, \beta}(\tilde{Y}_t^\pm)} \right|^{\frac{\alpha}{\alpha-1}} \quad (24)$$

⁷Note that the model with joint stochastic volatility cannot be treated as simple generalization of specifications with stochastic volatility only from jumps or diffusion.

The support of the above distribution is given by $\tilde{S}_t^+ > 0, \tilde{Y}_t^+ \in (l_{\alpha,\beta}, \frac{1}{2})$ for the positive part and $\tilde{S}_t^- < 0, \tilde{Y}_t^- \in (-\frac{1}{2}, l_{\alpha,\beta})$ for the negative part. Following Buckle (1995), for $\alpha \in (1, 2)$ and $\beta \in (-1, 1)$ we have

$$t_{\alpha,\beta}(\pm\frac{1}{2}) = \pm\infty \text{ and } t_{\alpha,\beta}(l_{\alpha,\beta}) = 0 \quad (25)$$

and hence

$$p(\pm\frac{1}{2}|\theta, X^{-(\tilde{Y}_t^\pm)}, Y) = p(l_{\alpha,\beta}|\theta, X^{-(\tilde{Y}_t^\pm)}, Y) = 0. \quad (26)$$

Moreover, since $t_{\alpha,\beta}(\cdot)$ is monotonic, it does not contribute to the maximum of (24), so density in eq. (24) is unimodal with its mode at $\tilde{Y}_t^\pm = t_{\alpha,\beta}^{-1}\left(\frac{\tilde{S}_t^\pm}{(\lambda_{t-1}\delta)^{\frac{1}{\alpha}}}\right)$ and value at maximum $p(t_{\alpha,\beta}^{-1}\left(\frac{\tilde{S}_t^\pm}{(\lambda_{t-1}\delta)^{\frac{1}{\alpha}}}\right)|\theta, X^{-(\tilde{Y}_t^\pm)}, Y) = 1$. This makes the rejection sampling algorithm a suitable method to draw from the distribution in (24). For details on the rejection sampling please see Devroye (1986). In this paper, however, I use adaptive rejection sampling algorithm (ARS) of Gilks and Wild (1992), which utilizes rejected draws⁸. This makes the sampling much more efficient especially in the case where the distribution is highly spiked around the mode. Since $v_t^\pm = t_{\alpha,\beta}(\tilde{Y}_t^\pm)$ is a bijection, we obtain the draws on v_t^+ and v_t^- . In the derivation of other conditional distributions, we condition on $\{v_t^+\}_{t=2}^T$ and $\{v_t^-\}_{t=2}^T$, which solves several multimodality problems as described below⁹. In the following I redefine the vector of state variables by replacing \tilde{Y}_t^+ by v_t^+ and \tilde{Y}_t^- by v_t^- for all $t = 2, \dots, T$:

$$X = (\{h_t\}_{t=1}^T, \{\tilde{U}_t\}_{t=2}^T, \{\tilde{S}_t^+\}_{t=2}^T, \{\tilde{S}_t^-\}_{t=2}^T, \{v_t^+\}_{t=2}^T, \{v_t^-\}_{t=2}^T). \quad (27)$$

Note that in the above $\beta \in (-1, 1)$ and some of the properties in eq. (25) do not hold for $\beta = \pm 1$. Since we are interested in the negative skewness, we have the following proposition for the maximum negative skewness $\beta = 1$:

Proposition 1 For $\alpha \in (1, 2)$ and $\beta = 1$, we have $t_{\alpha,\beta}(\tilde{Y}_t^+) \nearrow \alpha \left(\frac{1}{\alpha-1}\right)^{(\alpha-1)/\alpha} < \infty$ as $\tilde{Y}_t^+ \nearrow \frac{1}{2}$.

Proof. Apply L'Hospital's rule for $\frac{\cos \pi \tilde{y}}{\cos[\pi(\alpha-1)\tilde{y} + \eta_{\alpha,\beta}]}$ for $\beta = 1$. Then substitute the limit to the formula for $t_{\alpha,\beta}$ in eq. (15). The result follows immediately. ■

This result shows, that the update procedure described above cannot be directly applied for $\beta = 1$. Since Li, Wells and Yu (2008) applied a similar update procedure for the model with stochastic volatility from diffusion and the maximum negative skewness $\beta = 1$, this leaves their update method incorrect.¹⁰ The first source of their misspecification is the separability problem of the MCMC and the second is their application of the Buckle (1995) updating method for $\beta = 1$. My algorithm corrects for both of these problems in the models with Lévy α -stable jumps by construction of the MCMC free of the separability issue and by estimation of $\beta \in (-1, 1)$.

⁸I do not construct the envelope function as in Gilks in Wild (1992) but follow closely Buckle (1995).

⁹The draw of v_t^+ and v_t^- contains information on the conditioning α and β parameters and hence it changes the property of updating procedure of parameter α and β if we condition on v_t^+ and v_t^- and not on \tilde{Y}_t^+ and \tilde{Y}_t^- .

¹⁰Li, Wells and Yu (2008) do not estimate the parameter β and fix it at $\beta = 1$ maximum negative skewness.

3.4.2 Updating jump size variables \tilde{S}_t^+ , \tilde{S}_t^-

By application of the Bayes rule the conditional posterior distribution for \tilde{S}_t^+ and \tilde{S}_t^- is given by:

$$p(\tilde{S}_t^\pm | \theta, X^{-(\tilde{S}_t^\pm)}, Y) \propto p(Y_t | Y_{t-1}, \theta, X, Y^{-(Y_t, Y_{t-1})}) p(\tilde{S}_t^\pm | v_t^\pm, \alpha, \beta, \sigma_{SJ}, h_{t-1}) \quad (28)$$

$$\propto \exp \left\{ -\frac{1}{2} \frac{(Y_t - Y_{t-1} - \mu\delta - (\xi_{t-1}\delta)^{\frac{1}{2}} \rho \varepsilon_t^{(3)} - S_t)^2}{\xi_{t-1}\delta(1 - \rho^2)} \right\} \cdot \exp \left\{ -\left| \frac{\tilde{S}_t^\pm}{(\lambda_{t-1}\delta)^{\frac{1}{\alpha}} v_t^\pm} \right|^{\frac{\alpha}{\alpha-1}} \right\} \left| \tilde{S}_t^\pm \right|^{\frac{1}{\alpha-1}}.$$

By definition $S_t = \tilde{U}_t \tilde{S}_t^+ + (1 - \tilde{U}_t) \tilde{S}_t^-$, so in the above density function the first exponential part is a function of \tilde{S}_t^+ or \tilde{S}_t^- only if $\tilde{U}_t = 1$ or respectively $\tilde{U}_t = 0$. This property is intuitive, since there is no information contained in the sample about positive (negative) jump if there is a negative (positive) jump in the returns at time t . The support of the density is $\tilde{S}_t^+ > 0$ and $\tilde{S}_t^- < 0$ respectively. Lets define the following bijection:

$$x_t^\pm = \frac{\tilde{S}_t^\pm}{\left| (\lambda_{t-1}\delta)^{\frac{1}{\alpha}} v_t^\pm \right|}. \quad (29)$$

Using the change of variable formula, the density of x_t^+ and x_t^- is unimodal and log-concave. Moreover, this property is not affected if the sample contains information about the jump or not. Hence, I apply ARS algorithm by Gilks and Wild (1992). I significantly improve the ARS algorithm by supplying the unique maximum of the density for x_t^+ and x_t^- . When the data contains no information about jump, it can be computed analytically by a simple differentiation. In the other case I found the Newton's method to be efficient in computation. After the draw of x_t^\pm we obtain the draw of \tilde{S}_t^\pm by inverting the function in eq. (29).

3.4.3 Updating index of stability parameter α

The next problem is the choice of bounds for the parameter $\alpha \in (1, 2)$. This is a delicate matter since as $\alpha \searrow 1$ the power coefficient $\frac{\alpha}{\alpha-1}$ in eq. (14) approaches infinity. Moreover, as $\alpha \nearrow 2$ we approach normal distribution and lose identification. Taking the above into account, I assume the uniform prior distribution on $\alpha \in [1.05, 1.99]$ to avoid overflow computation problems. This not a restrictive assumption, since bounds are barely (or not at all) hit by the sampler.

As noted by Buckle (1995) updating the index of stability α is the most difficult part in the Bayesian inference of stable jumps. I modify his approach to accommodate for the mixture of truncated stable distributions. He solved the problem of multimodality of complete conditional distribution by the above described change of variables from the auxiliary variables \tilde{Y}_t^\pm to v_t^\pm using the transformation $v_t^\pm = t_{\alpha, \beta}(\tilde{Y}_t^\pm)$. If we condition not on $\left\{ \tilde{Y}_t^+ \right\}_{t=2}^T$ and $\left\{ \tilde{Y}_t^- \right\}_{t=2}^T$ but instead on $\left\{ v_t^+ \right\}_{t=2}^T$ and $\left\{ v_t^- \right\}_{t=2}^T$, the complete conditional distribution of α is given by:

$$p(\alpha | \theta^{-(\alpha)}, X, Y) \propto p(\left\{ \tilde{U}_t \right\}_{t=2}^T | \alpha, \beta) p(\alpha | \left\{ \tilde{S}_t^+, v_t^+ \right\}_{t=2}^T, \left\{ \tilde{S}_t^-, v_t^- \right\}_{t=2}^T, \beta, \sigma_{SJ}, \{h_t\}) \propto$$

$$p(\left\{ \tilde{U}_t \right\}_{t=2}^T | \alpha, \beta) p(\left\{ \tilde{S}_t^+, v_t^+ \right\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) p(\left\{ \tilde{S}_t^-, v_t^- \right\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) p(\alpha),$$

where $p(\alpha)$ is the prior distribution on α , $p(\alpha) \sim U(1.05, 1.99)$, independent of other parameters'

priors, and:

$$p(\{\tilde{U}_t\}_{t=2}^T | \alpha, \beta) = (p_{\alpha, \beta})^{\sum_{t=2}^T \tilde{U}_t} (1 - p_{\alpha, \beta})^{\sum_{t=2}^T (1 - \tilde{U}_t)},$$

$$p(\{\tilde{S}_t^+, v_t^+\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) \propto \left(\frac{1}{p_{\alpha, \beta}}\right)^{T-1} \left(\frac{\alpha}{\alpha - 1}\right)^{T-1} \cdot \exp \left\{ - \sum_{t=2}^T \left| \frac{\tilde{S}_t^+}{(\lambda_{t-1} \delta)^{\frac{1}{\alpha}} v_t^+} \right|^{\frac{\alpha}{\alpha-1}} \right\} \prod_{t=2}^T \left(\left| \frac{\tilde{S}_t^+}{(\lambda_{t-1} \delta)^{\frac{1}{\alpha}} v_t^+} \right|^{\frac{\alpha}{\alpha-1}} \left| \frac{\partial t_{\alpha, \beta}(\tilde{Y}_t^+)}{\partial \tilde{Y}_t^+} \right|_{t_{\alpha, \beta}(\tilde{Y}_t^+) = v_t^+}^{-1} \right),$$

$$p(\{\tilde{S}_t^-, v_t^-\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) \propto \left(\frac{1}{1 - p_{\alpha, \beta}}\right)^{T-1} \left(\frac{\alpha}{\alpha - 1}\right)^{T-1} \cdot \exp \left\{ - \sum_{t=2}^T \left| \frac{\tilde{S}_t^-}{(\lambda_{t-1} \delta)^{\frac{1}{\alpha}} v_t^-} \right|^{\frac{\alpha}{\alpha-1}} \right\} \prod_{t=2}^T \left(\left| \frac{\tilde{S}_t^-}{(\lambda_{t-1} \delta)^{\frac{1}{\alpha}} v_t^-} \right|^{\frac{\alpha}{\alpha-1}} \left| \frac{\partial t_{\alpha, \beta}(\tilde{Y}_t^-)}{\partial \tilde{Y}_t^-} \right|_{t_{\alpha, \beta}(\tilde{Y}_t^-) = v_t^-}^{-1} \right).$$

Note that the state vector X is already redefined in eq. (27) and contains information on $\{v_t^+\}_{t=2}^T$ and $\{v_t^-\}_{t=2}^T$. In order to compute the value of the above conditional distribution we need to inverse the function $t_{\alpha, \beta}$ for given v_t^+ and v_t^- and find the respective values of \tilde{Y}_t^+ and \tilde{Y}_t^- . This can be done efficiently using Newton's method as suggested in Buckle (1995). Since the complete conditional distribution of α is of unknown form, I rely on the MH step to sample from this distribution. The random walk MH step with normal proposal distribution has been found to be efficient.

3.4.4 Updating skewness parameter β

Since I want to model the negative skewness of asset returns, I consider the restriction $\beta > 0$. In order to control the degree of skewness, I relax the maximum negative skewness ($\beta = 1$) assumption of Carr and Wu (2003). Their assumption is needed to price derivative securities but is not required under statistical measure.

In my setting I have to restrict $\beta \neq 1$, since according to proposition (1), one cannot guarantee unimodality of the distribution in eq. (24) for \tilde{Y}_t^+ . The choice of the uniform, independent prior distribution $p(\beta) \sim U(0.01; 0.99)$ addresses these issues and avoids overflow computation problems.

Updating skewness parameter β is similar to updating α :

$$p(\beta | \theta^{-(\beta)}, X, Y) \propto p(\{\tilde{U}_t\}_{t=2}^T | \alpha, \beta) p(\beta | \{\tilde{S}_t^+, v_t^+\}_{t=2}^T, \{\tilde{S}_t^-, v_t^-\}_{t=2}^T, \alpha, \sigma_{SJ}, \{h_t\}) \propto p(\{\tilde{U}_t\}_{t=2}^T | \alpha, \beta) p(\{\tilde{S}_t^+, v_t^+\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) p(\{\tilde{S}_t^-, v_t^-\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) p(\beta),$$

where all components are derived in the updating procedure for α but should be treated as the functions of β . By conditioning on $\{v_t^+\}_{t=2}^T$ and $\{v_t^-\}_{t=2}^T$ I achieve unimodality of complete conditional posterior of β as in Buckle (1995). The random walk MH step with normal proposal

distribution is found to be efficient in this case.

3.4.5 Updating mixing variables \tilde{U}_t

Since $p(\tilde{U}_t|\alpha, \beta)$ is a Bernoulli distribution with probability of success $p_{\alpha,\beta}$, the complete conditional posterior is also Bernoulli and is given by:

$$p(\tilde{U}_t = 1|\theta, X^{-(\tilde{U}_t)}, Y) \propto p(Y_t|\theta, X^{-\tilde{U}_t}, \tilde{U}_t = 1, Y_{t-1})p(\tilde{U}_t = 1|\theta) \propto \exp\left\{-\frac{1}{2}\frac{(Y_t - Y_{t-1} - \mu\delta - \tilde{S}_t^+ - (\xi_{t-1}\delta)^{\frac{1}{2}}\rho\varepsilon_t^{(3)})^2}{\delta\xi_{t-1}(1-\rho^2)}\right\}p_{\alpha,\beta}, \quad (30)$$

and:

$$p(\tilde{U}_t = 0|\theta, X^{-(\tilde{U}_t)}, Y) \propto p(Y_t|\theta, X^{-(\tilde{U}_t)}, \tilde{U}_t = 0, Y_{t-1})p(\tilde{U}_t = 0|\theta) \propto \exp\left\{-\frac{1}{2}\frac{(Y_t - Y_{t-1} - \mu\delta - \tilde{S}_t^- - (\xi_{t-1}\delta)^{\frac{1}{2}}\rho\varepsilon_t^{(3)})^2}{\delta\xi_{t-1}(1-\rho^2)}\right\}(1-p_{\alpha,\beta}). \quad (31)$$

We can directly calculate the above probabilities and normalize their sum to unity. The draw from this distribution is then straightforward.

3.5 Auxiliary Particle Filter

In order to perform density forecast analysis, I fix vector of parameters θ for each model at the respective posterior mean and calculate the following probabilities:

$$\Pr(R_{t+L} < r_{t+L}|\theta, R^t, m), \quad (32)$$

$$L = 1, t \in \mathbb{N}_L, R^t = \{R_s\}_{s=1,\dots,t}$$

where θ denotes parameter vector for model m specification, $m \in \{1, \dots, 10\}$, $R_{t+L} = Y_{t+L} - Y_{t+L-1}$ is a daily log-return on the asset at time $t + L$ with its law determined by the model specification and r_{t+L} is observed value of this log-return at time $t + L$. In the above \mathbb{N}_L denotes the subset of natural numbers less than $T - L$ and divisible by L . This guarantees that I analyze only "non-overlapping" periods and can be further extended for other forecasting horizons. In my paper I focus attention on the one-day horizon forecasts ($L = 1$).

Note that I condition on the estimate of parameter value θ and do not integrate it out. Hence, I do not take into account the parameter estimation uncertainty. Since I have a relatively long sample size, the parameters are estimated with high precision¹¹. The effect of parameter estimation uncertainty is beyond the scope of this paper. In the notation below, I omit the explicit dependence on the model m specification, since it suffices to induce it from θ vector of parameter estimates.

In this work I consider one-day ahead ($L = 1$) time horizon for density forecast analysis, which makes it possible to assess a model ability to forecast one-day ahead daily log-return distribution. Note that (32) can be calculated not only for the in-sample period but also for the out-of sample period, whenever we have data available. We can study quantile forecast (VaR) performance of

¹¹The only parameters that are not estimated with high precision are those governing skewness of returns μ_j and β respectively in Poisson and Lévy α -stable jumps.

the model by comparison of given significance levels and unconditional covering frequencies of each model implied by the probabilities in (32). Moreover, if the model is correctly specified, binary variables indicating if the data points are contained in the VaR interval, should be independently distributed. Hence, there should be no "clustering" in time of their respective realizations.

We can estimate values in eq. (32) by:

$$z_t^{(L)} = \widehat{\text{Pr}}(R_{t+L} < r_{t+L} | \theta, R^t) = \frac{1}{K} \sum_{k=1}^K \text{Pr}(R_{t+L} < r_{t+L} | \theta, R^t, J_{t+L}^{(k)}, h_{t+L}^{(k)}, h_{t+L-1}^{(k)}), \quad (33)$$

where $(J_{t+L}^{(k)}, h_{t+L}^{(k)}, h_{t+L-1}^{(k)}) \sim iid p(J_{t+L}, h_{t+L}, h_{t+L-1} | \theta, R^t)$.

The draws from this distribution can be performed by utilizing the following condition:

$$p(J_{t+L}, h_{t+L}, h_{t+L-1}, \dots, h_t | \theta, R^t) = p(J_{t+L} | h_{t+L-1}, \theta) \cdot p(h_{t+L} | h_{t+L-1}, \dots, h_t, \theta) \cdot p(h_t | \theta, R^t) \quad (34)$$

By discarding draws for variables that do not directly enter in equation (33), we have draws from the desired $p(J_{t+L}, h_{t+L}, h_{t+L-1} | \theta, R^t)$ distribution. It is important to note that the above holds for all models with stochastic volatility from diffusion, jumps (or both) with jump sizes:

$$J_t(m) = \begin{cases} q_t \varkappa_t : \text{models with Poisson jumps} \\ S_t(\alpha, \beta, 0, (\lambda_{t-1} \delta)^{1/\alpha}) : \text{models with Lévy } \alpha\text{-stable jumps} \end{cases} \quad (35)$$

The work of Christoffersen (1998) on the evaluation of the interval forecasts and its further extension by Diebold, Gunther and Tay (1998) to the context of the density forecast allow us to draw conclusions based on the following criterion. A given model is correctly specified if $z_t^{(L)}$ (for $t \in \mathbb{N}_L$) is *iid* $U(0, 1)$ distributed. By transformation using the inverse cdf of the standard normal distribution, I define:

$$\widehat{z}_t^{(L)} = \Phi^{-1}(z_t^{(L)}). \quad (36)$$

The distribution of $z_t^{(L)}$ implies, that transformed variables $\widehat{z}_t^{(L)}$ should be *iid* $N(0, 1)$ distributed. This fact is later used for model evaluation in view of the density forecast and quantile forecast (VaR) performance.

In order to sample from distributions in eq. (34) we have to sample from filtering density $p(h_t | \theta, R^t)$ and then, conditional on this draw, sample from all predicting densities $p(h_{t+l} | h_{t+l-1}, \theta)$ and $p(J_{t+L} | h_{t+L-1}, \theta)$. Sampling from these densities is rather straightforward. The most difficult problem involves approximation of the filtering density $p(h_t | \theta, R^t)$ by the auxiliary particle filter, as developed in Pitt and Shephard (1999). Chib, Nardari and Shephard (2002) extend basic auxiliary particle filter of Pitt and Shephard (1999) for Poisson type jumps in returns but do not include leverage effect. Johannes, Polson and Stroud (2008) offer further refinements to the auxiliary particle filter algorithm for models with jumps and stochastic volatility. However, their algorithm cannot be applied to the specifications with Lévy α -stable jumps.¹² Durham (2006) extends the

¹²The sampling step in their algorithm requires the closed form of stable density, which is unavailable. Moreover, the draw should be performed from the density obtained as a multiplication of the normal and stable kernels so it is not standard. All of the above renders the Johannes, Polson and Stroud (2008) method inapplicable in our

basic particle filter for models with leverage effect but does not include jumps in returns, moreover, he works with particle filter and does not apply auxiliary particle filter involving index parameters draws¹³. In this paper I present auxiliary particle filter for jump-diffusion models with the leverage effect. Moreover, I allow for different sources of stochastic volatility and most importantly for Lévy α -stable jumps in returns.

Lets first notice, that:

$$p(h_{t+1}, J_{t+1}(m)|\theta, R^{t+1}) = \int \frac{p(h_{t+1}, h_t, J_{t+1}|R^{t+1}, \theta)}{p(h_t|\theta, R^t)} dp(h_t|\theta, R^t) \quad (37)$$

and

$$\begin{aligned} p(h_{t+1}, h_t, J_{t+1}|R^{t+1}, \theta) &= \frac{p(h_{t+1}, h_t, J_{t+1}, R_{t+1}|R^t, \theta)}{p(R_{t+1}|R^t)} = \\ &= \frac{p(R_{t+1}|h_{t+1}, h_t, J_{t+1}, R^t, \theta)p(h_{t+1}|h_t, \theta)p(J_{t+1}|h_t, \theta)p(h_t|\theta, R^t)}{p(R_{t+1}|R^t)}. \end{aligned} \quad (38)$$

Substituting (38) into (37) we have:

$$\begin{aligned} p(h_{t+1}, J_{t+1}|\theta, R^{t+1}) &\propto \\ &\int p(R_{t+1}|h_{t+1}, h_t, J_{t+1}, R^t, \theta)p(h_{t+1}|h_t, \theta)p(J_{t+1}|h_t, \theta)dp(h_t|\theta, R^t) \end{aligned} \quad (39)$$

Auxiliary particle filter is a recursive algorithm to approximate filtering densities $p(h_t|\theta, R^t)$ for $t = 0, \dots, T$ by a finite number K of "particles" for each t . These particles define discrete probability distribution filter $\bar{p}(h_t|\theta, R^t)$. I denote particles for filter at time t as $\bar{h}_t^{(k)}$, where $k = 1, 2, \dots, K$. Given K particles defining discrete probability distribution filter at time t , we obtain approximation $\bar{p}(h_{t+1}|\theta, R^{t+1})$ for $t + 1$ defined by its respective K particles using relation in (39):

1. draw $N \geq K$ indexes k_1, k_2, \dots, k_N from the discrete probability distribution $g(k|R^{t+1})$ with support of $k = 1, 2, \dots, K$. The choice of $g(k|R^{t+1})$ should reflect an information content of the future return R_{t+1} on the choice of index mixture k and hence the particle $\bar{h}_t^{(k)}$. Note that k represents an index on mixture in (39) as in Pitt and Shephard (1999). I specify the weights to be proportional to:

$$g(k|R^{t+1}) \propto \phi(R_{t+1}; \mu_{(m)}, \sigma_{(m)}^2) \quad (40)$$

where ϕ denotes normal pdf calculated at $R_{t+1} = r_{t+1}$ with mean $\mu_{(m)}$ and variance $\sigma_{(m)}^2$ given by:

$$\mu_{(m)} = \begin{cases} \mu + \lambda_t \mu_j & \text{for models with Poisson jumps} \\ \mu & \text{for models with Lévy } \alpha\text{-stable jumps} \end{cases}$$

setting.

¹³For discussion on improvements of auxiliary particle filter over particle filter please refer to Pitt and Shephard (1999).

$$\sigma_{(m)}^2 = \begin{cases} (1 - \rho^2)\xi_t + \lambda_t(\sigma_j^2 + \mu_j^2) - (\lambda_t\mu_j)^2 & \text{for models with Poisson jumps} \\ (1 - \rho^2)\xi_t + \lambda_t^{\frac{2}{\alpha}} & \text{for models with Lévy } \alpha\text{-stable jumps} \end{cases}$$

and where we substitute all constraints as in Table I. My choice of $\mu_{(m)}$ and $\sigma_{(m)}^2$ for models with Poisson jumps coincides respectively with $E_t(R_{t+1}|\theta, \varepsilon_{t+1}^{(3)} = 0)$ and $var_t(R_{t+1}|\theta, \varepsilon_{t+1}^{(3)} = 0)$. In case of models with Lévy α -stable jumps I pursue the same strategy as for Poisson models, but in the first step I approximate the distribution $S_{t+1}(\alpha, \beta, 0, (\lambda_t)^{\frac{1}{\alpha}})$, given λ_t , by $N(0, (\lambda_t)^{\frac{2}{\alpha}})$. Hence, I simply replace α -stable distribution by normal distribution with the same scale parameter. Finally, I denote the index draws as $\widehat{k}_1, \widehat{k}_2, \dots, \widehat{k}_N$ and record the respective values of $h_t^{(\widehat{k}_i)}$ for $i = 1, \dots, N$, where $h_t^{(\widehat{k}_i)} \equiv \bar{h}_t^{(\widehat{k}_i)}$. Note that the above specification of $g(k|R^{t+1})$ allows to draw the indexes k on particles $\bar{h}_t^{(k)}$ and this draw is consistent with the scale and mean implied by the next-period return R_{t+1} . This makes my algorithm efficient and easy to implement for all considered models.

2. Draw proposal particles $h_{t+1}^{(n)}$, $n = 1, 2, \dots, N$, given mixture index and particle from the preceding filter using the respective prediction density:

$$h_{t+1}^{(n)} \sim p(h_{t+1}|h_t^{(\widehat{k}_n)}, \theta)$$

3. Draw jump increments $J_{t+1}^{(n)}$, $n = 1, 2, \dots, N$, from $p(J_{t+1}|h_t^{(\widehat{k}_n)}, \theta)$.
4. Reweight the draws $(h_{t+1}^{(n)}, J_{t+1}^{(n)})$, $n = 1, 2, \dots, N$, by drawing K times (with replacement) from the discrete probability distribution with weights proportional to:

$$\varpi_n = \frac{p(r_{t+1}|h_{t+1}^{(n)}, h_t^{(\widehat{k}_n)}, J_{t+1}^{(n)}, R^t, \theta)}{\phi(R_{t+1}; \mu_{(m)}, \sigma_{(m)}^2)} \quad (41)$$

for $n = 1, 2, \dots, N$. We finally get K draws defining discrete filter distribution $\bar{p}(h_{t+1}|\theta, R^{t+1})$ by discarding draws on J_{t+1} . Denote the new particles as $\bar{h}_{t+1}^{(k)}$, where $k = 1, 2, \dots, K$.

5. Go to point 1. and increment t .

By comparing weights in (41) to the first integrand component in (39), the validity of the whole algorithm is based on the importance sampling principle. In my applications I take $N = 10,000,000$ and $K = 1,000,000$. I do not draw from the discrete auxiliary particle filter distribution $\bar{p}(h_t|\theta, R^t)$ but directly utilize all derived particles from the filter. The above choice of N and K is sufficient to induce a low variability of statistics calculated using derived particles among different starting seeds of random number generator.

4 Empirical Application

4.1 The Data

In this paper the data on the S&P 500 index extends from 01/02/1981 to 12/31/2007 and comprises of 6813 daily observations available from CRSP database. The S&P 500 index levels are reported

at the closing times in each business day. All six model specifications are estimated using this data set. It allows for modelling the market crash of 1987 and the "dot.com" corrections from 1999-2001.

In Table II and Figure 1, I present respectively the descriptive statistics of daily log-returns on the S&P 500 index and graphs of S&P 500 index log-level and S&P 500 index log-returns. The data on the S&P log-returns indicate that there exists significant negative skewness of -1.7465 and kurtosis of 42.79 . In Figure 1, I also report the skewness and kurtosis as a term structure of S&P 500 returns. The term structure of skewness and kurtosis determine volatility smiles for options across all maturities.¹⁴ Carr and Wu (2003) find that the volatility smiles do not flatten completely as maturity increases and propose the log-stable model of asset returns, where asset returns have infinite variance and higher moments, and the CLT does not work. This also motivates my specification with Lévy α -stable jumps¹⁵.

4.2 Estimation Procedure

Since MCMC algorithms require a choice of starting values for all parameters and latent variables, I first list them for all estimated models. The parameter estimates were found not to be affected by different choice of starting values for the MCMC algorithms. I take the posterior mean for each model to be an estimate of the respective parameters and reported in Table III for all model specifications. In Figures 2 and 3, I present respectively the smoothed estimates of jump sizes in eq. (35) and stochastic volatility $\{h_t\}$.

For models with Lévy α -stable jumps in returns, the starting values for mean/drift parameters μ , θ_h are zero, for the scale parameters σ_h , σ_{SJ} , σ are one, for the correlation parameter ρ is zero, for κ_h is one. For jump specific parameters I specify $\alpha^{(1)} = 1.5$ and $\beta^{(1)} = 0.5$. The choice of starting values for the latent variables involves the choice of $\tilde{U}_t^{(1)} = 1$ (only positive jumps), $h_t^{(1)} = 0$, and finally $\tilde{S}_t^{+(1)} = 0.01$ and $\tilde{S}_t^{-(1)} = -0.01$ for all t . Since I update stable jumps auxiliary variables \tilde{Y}_t^+ and \tilde{Y}_t^- at the beginning of the MCMC algorithm, I do not need to specify their starting values. In all models the choice of starting values for the MCMC does not affect the estimation results. In models SJ and SJSV with stochastic volatility from either diffusion or jumps but not both, I draw 400,000 realizations from the MCMC chains, where the first 200,000 draws are treated as the burn-in period and the last 200,000 as draws from the stationary distribution. In the model with joint stochastic volatility DiffSJSV, I choose the same starting values but draw 700,000 realizations and I double the size of the burn-in period to the first 400,000 draws compared to other models with Lévy α -stable jumps. I run simulation in model PJ for 500,000 draws, in model PJSV for 1,000,000 draws and in model DiffPJSV for 1,500,000 draws where I treat the first 300,000, 800,000 and 1,000,000 draws as the burn-in period, respectively.¹⁶ For models with Poisson jumps the drift and log-volatility related parameters and latent variables are given the same starting values as for models with infinite activity jumps.

¹⁴The volatility smiles depend on both the distribution of asset returns under the statistical measure and the risk premia.

¹⁵With an exception of maximum negative skewness ($\beta = 1$), models with Lévy α -stable jumps require further refinements in order to be applied in the option pricing, e.g. tempering of Lévy measure.

¹⁶MCMC converges in smaller number of realizations for models with Poisson jumps. Since the draws for these models are not computationally demanding, we decide to run longer draws.

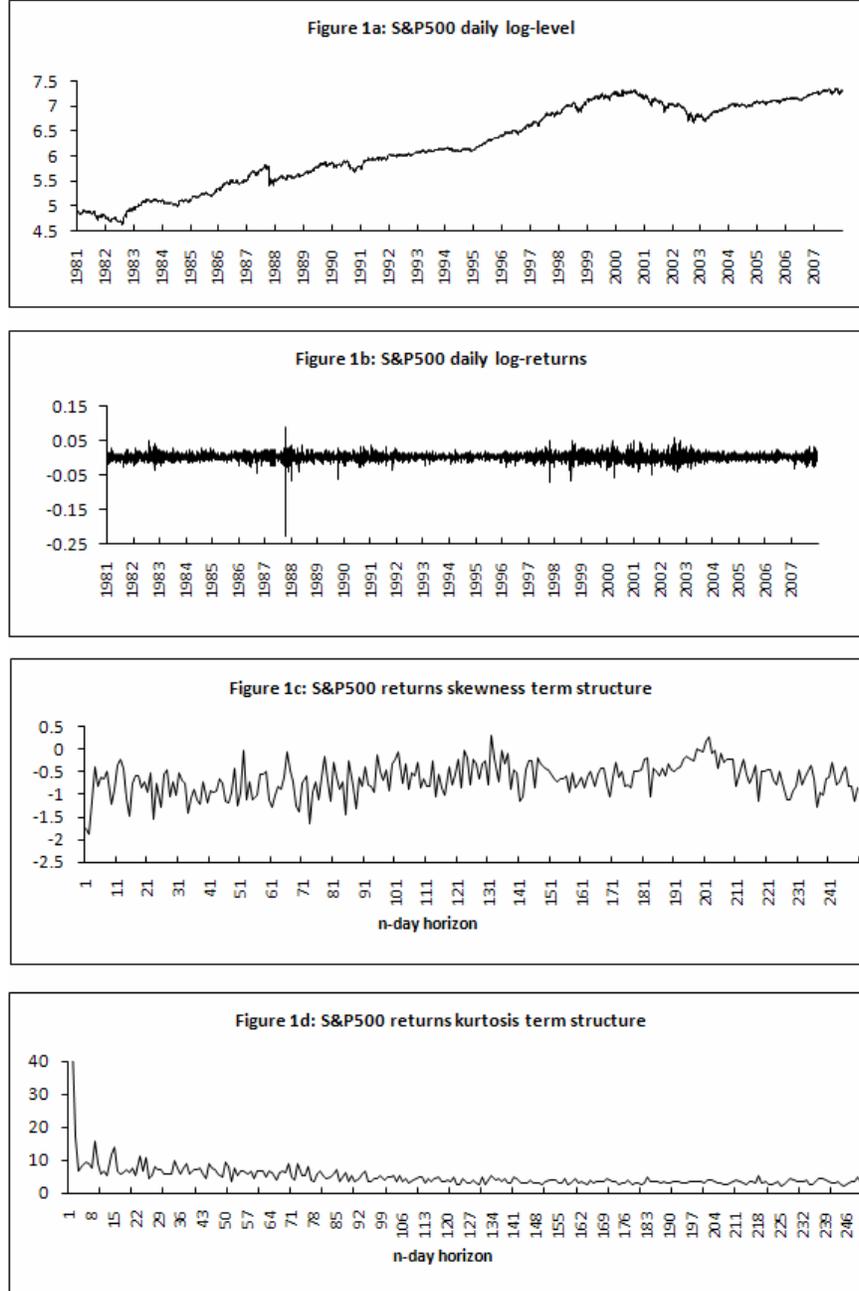


Figure 1. (a) S&P 500 index daily log-levels, 01/02/1981-12/31/2007; (b) S&P 500 index daily log-returns, 01/02/1981-12/31/2007; (c) Skewness term structure of S&P 500 daily log-returns, 01/02/1981-12/31/2007; (d) Kurtosis term structure of S&P 500 daily log-returns, 01/02/1981-12/31/2007

Poisson jump specific parameters are given starting values $\mu_j^{(1)} = 0$, $\sigma_j^{(1)} = 1$ and $\lambda_j^{(1)} = 0.5$ and for latent variables I assume no jumps $q_t^{(1)} = \varkappa_t^{(1)} = 0$ for all t . Moreover, in the model with joint stochastic volatility I choose $\sigma^{(1)} = 1$.

In the following I implement the model selection criteria developed in Jones (2003). Recall that in all model specifications $\varepsilon_t^{(1)}$ and $\varepsilon_t^{(2)}$ are assumed to be jointly independent and *iid* $N(0, 1)$. In the following lets call $\varepsilon_t^{(1)}$ the residuals from returns equation and $\varepsilon_t^{(3)} = \rho\varepsilon_t^{(1)} + \sqrt{1 - \rho^2}\varepsilon_t^{(2)}$ the

residuals from the log-volatility equation. We may view those residuals as latent variables. Hence, we can construct posterior distributions for functions of these latent variables by evaluating these functions at each step of the MCMC algorithm. Since model residuals are *iid* $N(0, 1)$ I calculate mean, standard deviation, skewness, kurtosis and first-order autocorrelation. Then I calculate the median and 95% confidence intervals for those statistics reported in Table IV. A correct model specification implies that mean is zero, standard deviation one, skewness zero, kurtosis three and autocorrelation zero.

4.3 Estimation Results for S&P 500 Data

Since I have the same log-volatility specification as in Jacquier, Polson and Rossi (2004), their results shed light on the importance of the leverage effect. In their work, the leverage effect is found to correct for a possible misspecification resulting in the biased estimates of the volatility states and the parameters of the log-volatility process. Hence, in my work I consider specifications with leverage effect and focus attention on the source of stochastic volatility and the jump structure. In models with either joint stochastic volatility or stochastic volatility from diffusion I find the leverage effect to be statistically significant with the estimates of -0.5891 , -0.5880 , -0.7496 , -0.6428 respectively for models PJ, DiffPJSV, SJ and DiffSJSV. Since the only models that do not allow for the leverage effect are the models with stochastic volatility from jumps, I restrict $\rho = 0$ for these specifications.¹⁷ In my analysis the estimation of all six model specifications allows us to draw conclusions about the marginal importance of the different jump structures and the source of stochastic volatility.

In all models the parameters are precisely estimated with an exception of the parameters governing skewness of returns μ_j and β , for models with Poisson and Lévy α -stable jumps respectively. In model PJ the parameter λ_j is estimated at the level 0.0022, which gives approximately one jump per two calendar years. Similarly, in models PJSV and DiffPJSV the activity rate of the Poisson jumps, governed by $\lambda_t = \exp(h_t)$ process, also indicates a similar average jump intensity. The small number of realized Poisson jumps limits the ability to precisely estimate the mean of the jump sizes μ_j and results in the relative estimation errors of 65.1%, 86.4%, 63.9% for models PJ, PJSV and DiffPJSV, respectively. In models with Lévy α -stable jumps, the lack of precision in the estimation of β is also a consequence of limited information in the sample about the tails of the returns distribution and implies that the relative estimation errors for parameter β are 42%, 65.7% and 16.7% for specifications SJ, SJSV and DiffSJSV, respectively. In model DiffSJSV the parameter σ_{SJ} controls for the relative importance of diffusion and Lévy α -stable jumps to the total volatility of returns. The relative error of estimation of 20% suggests that there is enough information in the sample to disentangle diffusion from infinite activity Lévy α -stable jumps. In Figures 2 and 3, I present smoothed jump sizes and log-volatility estimates respectively, where the former are defined in (35). The models with stochastic volatility arising only from diffusion violate *iid* property of jumps, since in models PJ and SJ I visually find an evidence of jump clustering. In specifications with joint stochastic volatility DiffPJSV (DiffSJSV) a jump clustering is a built-in characteristic of the model but clustering is allowed to arise only from the stochastic volatility.

I test for independence of the jump increments by using the standard Ljung-Box test.¹⁸

¹⁷Refer to section 2.3 for further details on the models with stochastic volatility from jumps.

¹⁸Note that stable increments do not have finite second moments, however, the sample autocorrelations can be always computed. More formally, the Ljung-Box test can be performed by truncating the increments at some given

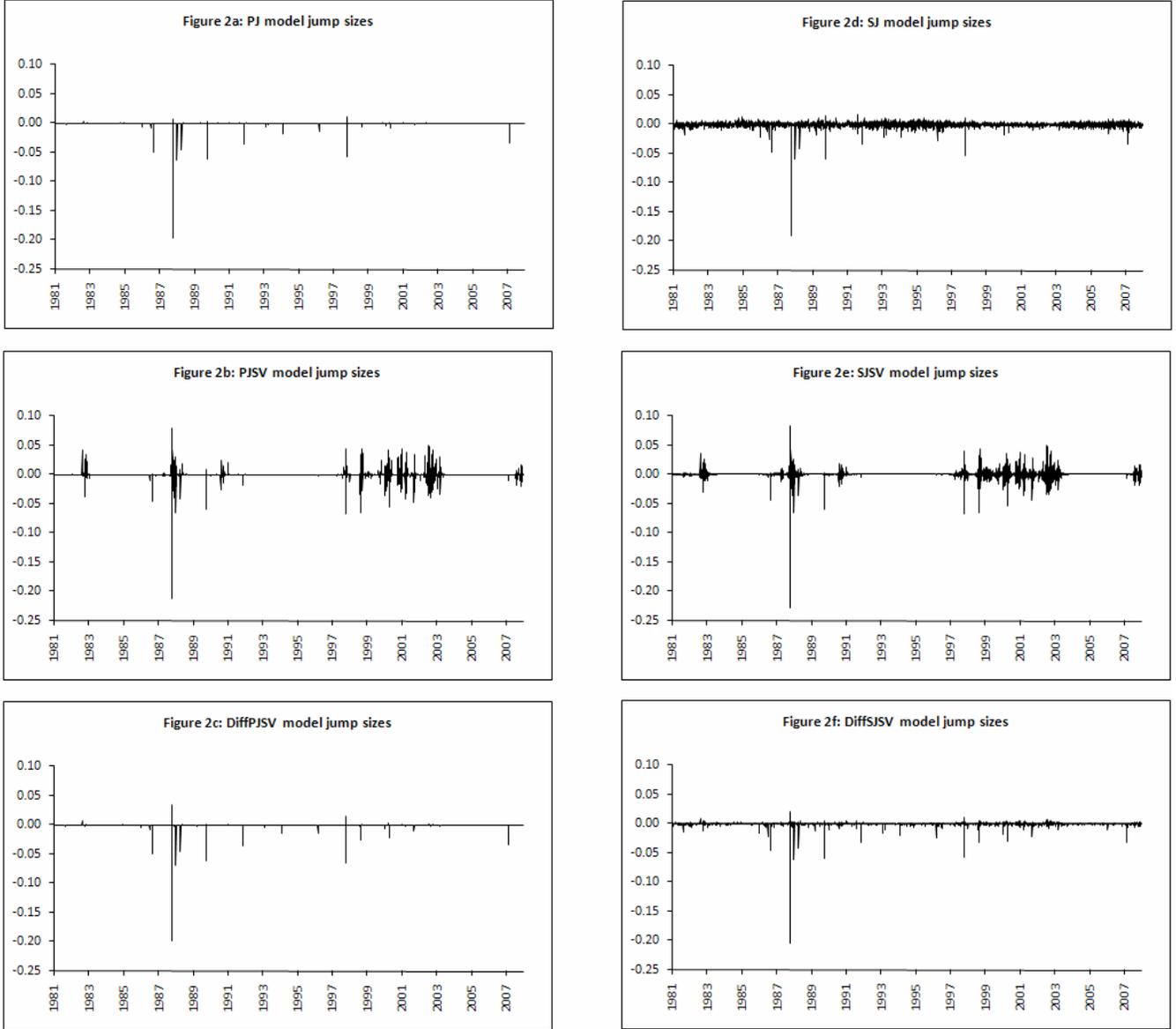


Figure 2. (a, b, c) smoothed jump size estimates for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV; (d, e, f) smoothed jump size estimates for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV. The jump sizes are defined for all models in eq. (35).

Since Poisson jumps are rare events I concentrate attention only on the Lévy α -stable jumps specifications SJ and DiffSJSV. In Figure 4, I illustrate smoothed jump increments for models SJ and DiffSJSV, where the latter are corrected for the varying intensity, or in other words, are scaled by the stochastic volatility. I present smoothed estimates of $S_{t+\delta}(\alpha, \beta, 0, \sigma_{SJ} \cdot \delta^{\frac{1}{\alpha}})$ for model DiffSJSV and call it in the sequel as *descaled jump increments*. The descaled jump increments are *iid* by construction. By using the scalability property for stable distribution and applying it

threshold. I assume that the threshold has never been met in the sample.

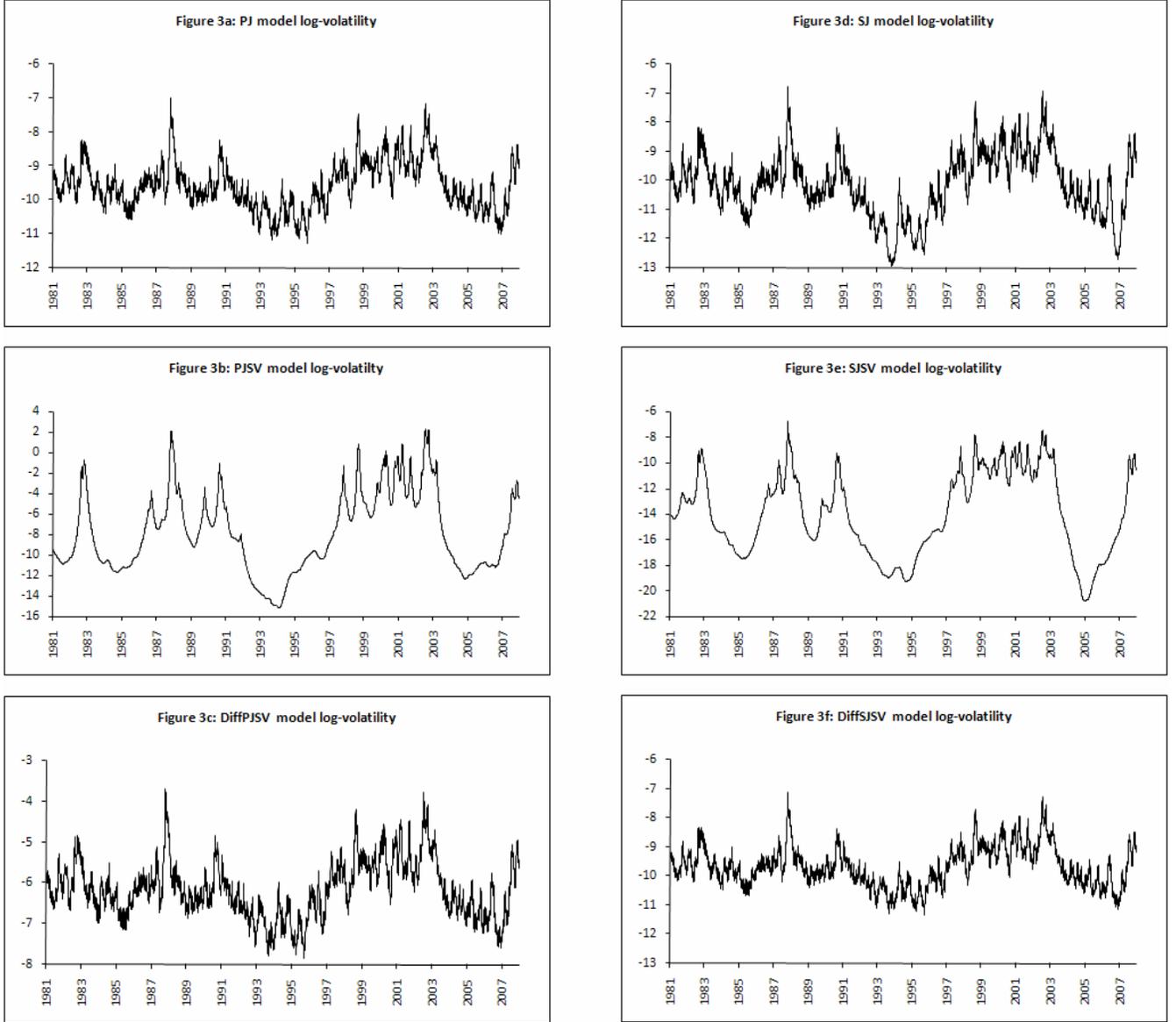


Figure 3. (a, b, c) smoothed log-volatility estimates $\{h_t\}$ for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV; (d, e, f) smoothed log-volatility estimates $\{h_t\}$ for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV

to the discretized version of the model in eq. (11) we have

$$S_{t+\delta}(\alpha, \beta, 0, \sigma_{SJ} \cdot \delta^{\frac{1}{\alpha}}) \stackrel{d}{=} \frac{S_{t+\delta}(\alpha, \beta, 0, (\lambda_t \delta)^{\frac{1}{\alpha}})}{\exp(h_t/\alpha)} \sim iid S(\alpha, \beta, 0, \sigma_{SJ} \cdot \delta^{\frac{1}{\alpha}}) \quad (42)$$

where the descaled jumps are the jump sizes divided by the instantaneous volatility. In model SJ I do not have to follow this procedure, since λ_t is constant and hence for this model I present the jump size estimates as in Figure 2 and eq. (35). I find that model DiffSJSV produces in general higher p -values at lags 1-500 than model SJ.¹⁹ Hence, there is less degree of dependence between

¹⁹Since I use Ljung-Box test, each time I include all lags up to, and including, the specified level. Under the null hypothesis the sample is random and under the alternative there is dependence.

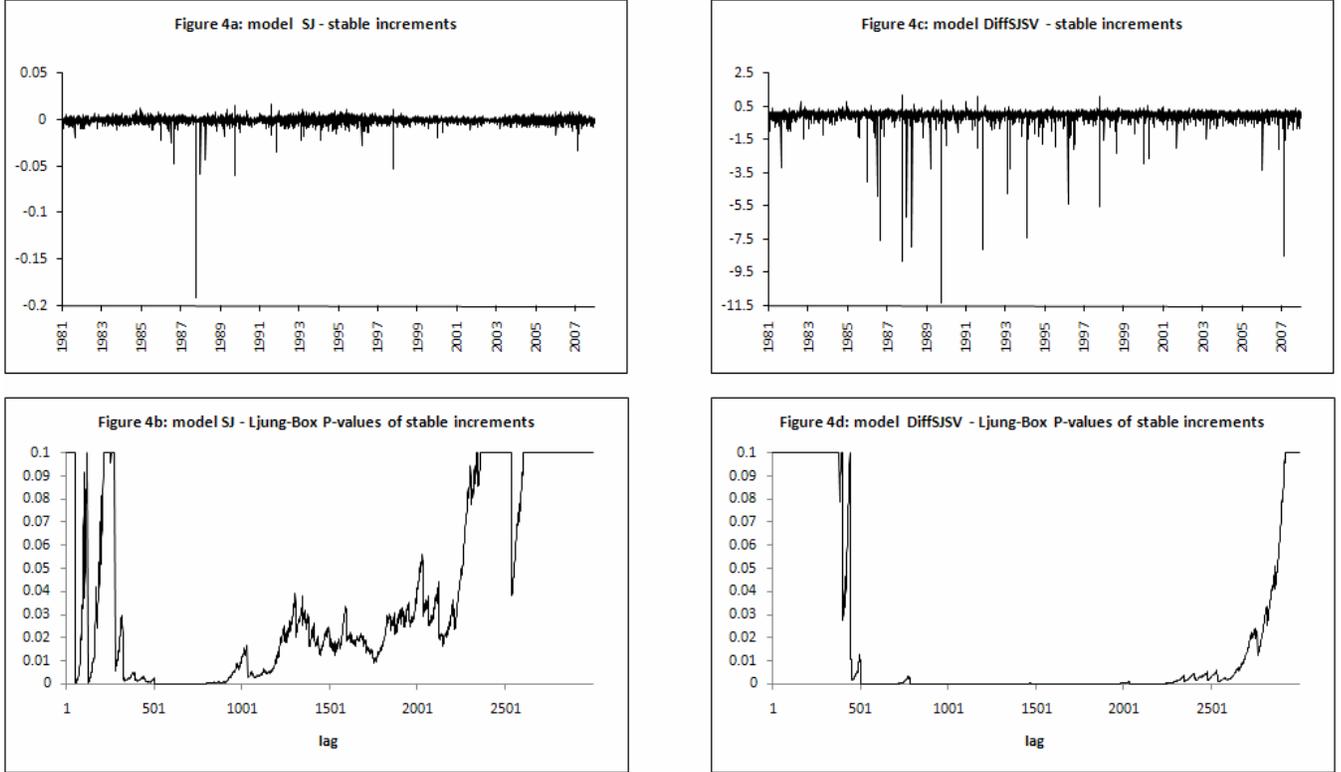


Figure 4. (a, c) smoothed jump increments for model SJ and descaled jump increments for model DiffSJSV (scaled back by realizations of stochastic volatility defined in eq. (42)); (b, d) Ljung-Box test p -values of smoothed jump increments and descaled jump increments are calculated for different maximum number of lags (1-3000) for models SJ and DiffSJSV respectively. The values at lags (1-3000) at the level 0.1 correspond to p -values at, or exceeding 0.1. The higher maximum lags (3001-6811) are omitted, since they are found to have p -values higher than 0.1

the neighboring jumps with distance of up to 2 years in model DiffSJSV. At horizons ranging from 500-2500 both models perform poorly, although model SJ performs somewhat better. However, none of these models reach the p -value of 0.05 at lags 500-2000. Since in the following sections I am mostly interested in the short-horizon density forecasts, model DiffSJSV having superior fit at shorter lags is better suited for this task.

4.3.1 The Source of Stochastic Volatility

Focusing attention on models with stochastic volatility from jumps PJSV and SJSV, we can evidently eliminate them as they are outperformed by other models with the same jump structure. In both models the parameters governing skewness of returns μ_j and β are estimated with the lowest precision among all specifications. Moreover, the speed of mean reversion parameters κ_h are much closer to the non-stationarity level and have the highest relative estimation errors among all stochastic volatility specifications of 60% and 55.5% for models PJSV and SJSV, respectively. There are also significantly higher relative errors of estimation of parameters θ_h of respectively 38.3% and 32.7%.

In terms of goodness of fit analysis presented in Table IV, models PJSV and SJSV perform much worse than their counterparts with the same jump specification. Although they perform

relatively well with respect to the skewness of returns, they cannot represent leptokurtic property of returns. This is documented by too small standard deviation of residuals of 0.9084 for model PJSV and 0.8965 for models SJSV, as well as by too high kurtosis of residuals, respectively 3.7349 and 3.7059. Note that even much richer specification of infinite activity Lévy α -stable jumps do not alleviate these problems since model SJSV do not fit the data better than the simple Poisson jump model PJSV.

I find almost perfect fit with respect to the log-volatility equation for all model specifications irrespective of the source of stochastic volatility and jump structure.

In models with stochastic volatility from diffusion (PJ) and joint stochastic volatility (Diff-PJSV) with Poisson jumps I do not find significant differences with respect to the precision of parameter estimates and goodness of fit, that can in a decisive way point out the best stochastic volatility specification. However, in models SJ and DiffSJSV with Lévy α -stable jumps the differences in goodness of fit can be found in the degree of kurtosis 3.059 and 3.1397, respectively, in Table IV. However, the latter still dominates all other models including all specifications with Poisson jumps. On the other hand, I find in the previous section that model SJ is dominated by model DiffSJSV with joint stochastic volatility when satisfying the independence assumption of descaled jump increments at shorter autocorrelation horizons of up to 2 years.

Summing up, I reject models with stochastic volatility from jumps and find that diffusion is an important feature, since it has to be a source of stochastic volatility. I postpone the final choice between models with stochastic volatility from diffusion and joint stochastic volatility to density forecast and VaR analysis in Section 4.4.

4.3.2 Modelling Jumps in Returns

In this section I provide an evidence in favor of models with infinite activity Lévy α -stable jumps. I restrict my analysis to models with either joint stochastic volatility or stochastic volatility from diffusion, since they dominate the models with stochastic volatility from jumps with respect to the estimation precision and goodness of fit. Since all considered models are estimated with high degree of precision with the exception of parameters governing skewness of returns, I concentrate attention on the goodness of fit analysis presented in Table IV.

I find that Poisson jumps are suited to fit only big jumps, which agrees with findings in Li, Wells and Yu (2008). My estimates of jump intensity λ_j for models with stochastic volatility from diffusion and $\exp(\hat{\theta}_h)$ for joint stochastic volatility imply only about one jump per two years. Hence "small", frequent, and more subtle jumps are simply not represented by the models with Poisson jumps in returns, even if we include joint stochastic volatility. The above can be seen by a comparison of smoothed jump size estimates for Poisson models with the respective estimates for Lévy α -stable models in Figure 2. As expected the skewness, affected by large jumps in the very left tail of the return distribution, is much better represented than kurtosis of returns in models with Poisson jumps. This is documented by the skewness of residuals of -0.0484 and -0.0470 , and kurtosis of residuals of 3.2578 and 3.2481 respectively for models PJ and DiffPJSV. Models SJ and DiffSJSV, with Lévy α -stable jumps in returns and infinite number of "small" jumps in the finite time interval, have a very good fit both with respect to skewness and kurtosis of returns and dominate other model specifications.

4.4 Density Forecast and VaR analysis

In this section I apply auxiliary particle filter described in Section 3.5 to evaluate one-day horizon forecast and quantile forecast (VaR) performance of all models. In Table V, I present descriptive statistics of $\{\widehat{z}_t^{(L)}\}$ distribution ($L = 1$), with $\widehat{z}_t^{(L)}$ defined in eq. (36) calculated for different model specifications with their respective parameters fixed at the MCMC estimates as in Table III for S&P500 data on the period 01/02/1981-12/31/2007. S&P500 daily log-returns, used for $\widehat{z}_t^{(L)}$ calculation, are derived from S&P500 index level for all available observations. A correct specification implies that $\{\widehat{z}_t^{(L)}\}$ is *iid* $N(0, 1)$ distributed, hence the mean is zero, standard deviation one, skewness zero, kurtosis three. Note that excess kurtosis and negative skewness of generalized residuals $\widehat{z}_t^{(1)}$ implies respectively too small kurtosis and not enough negative skewness in the model implied forecasting distribution only if the scale of the generalized residuals is correctly represented and close to one.²⁰ Moreover, there should be no autocorrelation in the levels and the squares of generalized residuals.

In Figure 5, I present quantile-quantile plots (qq-plots) of generalized residuals, that show all deviations from assumption of normality. In Table V, I present the calculated statistics (and the p -values) of Jarque-Bera test for normality. In Figures 6 and 7, I include autocorrelation functions for levels and squares of generalized residuals respectively. This allows us to draw conclusions on whether the independence assumption is satisfied. Serial correlation in the squares of generalized residuals is an indication of the lack of ability of the model to represent the volatility of returns. In Figures 8 and 9, I present p -values of the Ljung-Box test for dependence calculated at different maximum number of lags in the autocorrelation expansion for levels and squares of generalized residuals. In Table VI, I also present one-day horizon VaR performance. I calculate values of $\{z_t^{(L)}\}$ defined in equation (33), given estimated model parameters, and then compute empirical coverage frequencies for significance levels of 1%, 5% and 10%. Note that the density forecast analysis deals with the whole shape of the predictive density, while VaR analysis refers only to its very left tail.

The density forecast analysis in general, and VaR analysis in specific, stress the importance of correct specification of stochastic volatility. If model misspecifies stochastic volatility, it also performs poorly in density forecast and VaR analysis. To illustrate this, note that for goodness of fit analysis we utilize all available information in the sample, conditioning on all observed asset returns, while in the forecast and VaR analysis we only condition on the filtered volatility states and have available only current and past values of returns determining latent volatility. Hence, the behavior of stochastic volatility, as a state variable, and an ability to filter its values, is of fundamental importance in the density forecast. In this light a correct specification of the source of stochastic volatility determines a forecasting ability of the model.

4.4.1 Models with Poisson Jumps

I reject model PJSV with stochastic volatility from jumps, since it is outperformed by other specifications and has poor performance with respect to both skewness -0.2335 and kurtosis 4.9226 in the density forecast analysis in Table V. Most importantly, the scale of the forecast is incorrect at 0.8834 . The above results in the rejection of normality by the Jarque-Bera test. The

²⁰In the sequel I use Durham (2006) interpretation of $\{\widehat{z}_t^{(L)}\}$ realizations as "generalized residuals".

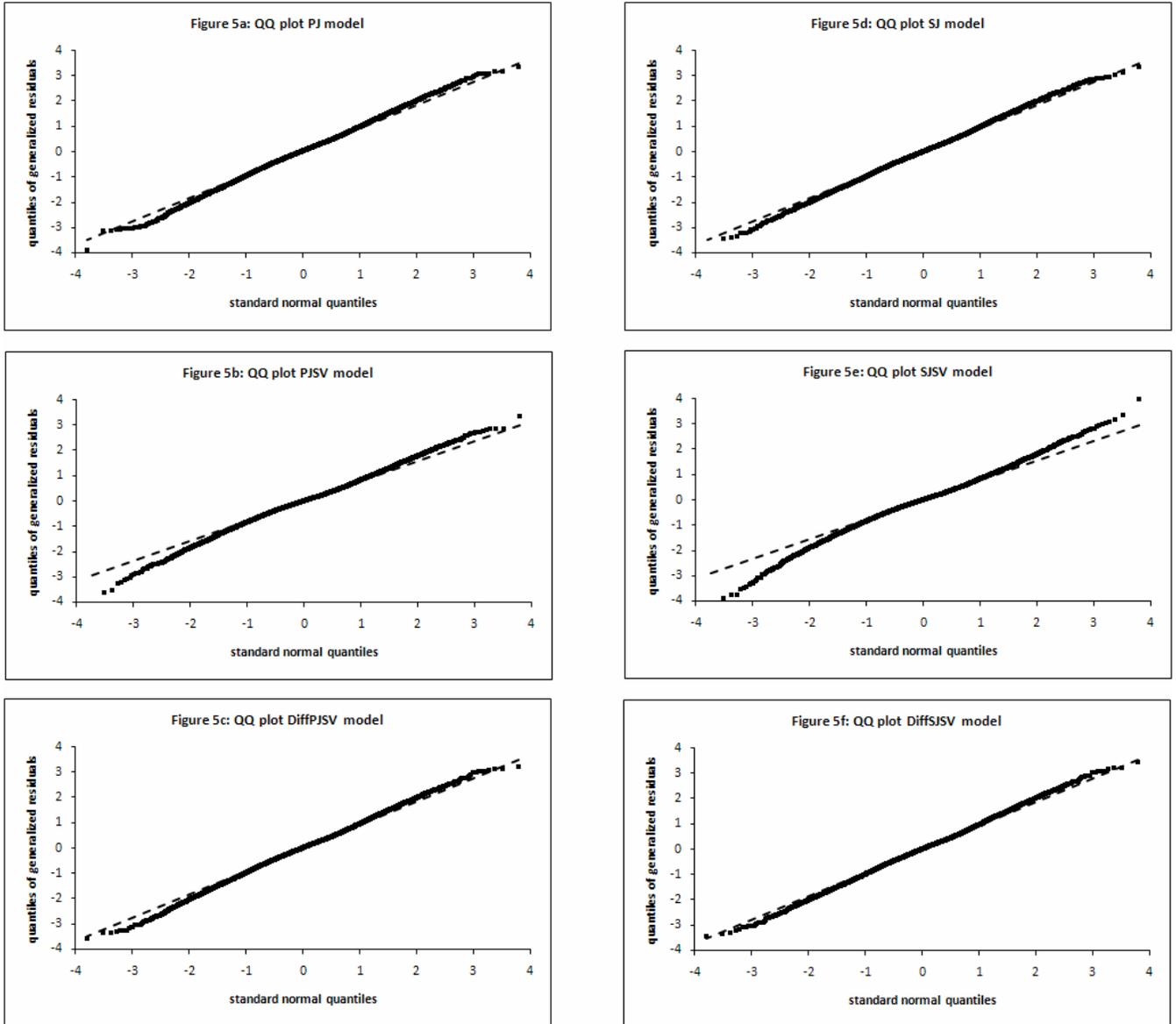


Figure 5. (a, b, c) quantile-quantile plots of the generalized residuals $\hat{z}_t^{(1)}$ for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV; (d, e, f) quantile-quantile plots of the generalized residuals $\hat{z}_t^{(1)}$ for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV

qq-plot in Figure 5b illustrates the problem in the tails of the forecasting distribution²¹. Ljung-Box test statistics for the levels of $\{\hat{z}_t^{(1)}\}$ in Figure 8b do not differ from the other model specifications and accept independence in the levels at 1% significance level if we include small number of lags in the test of up to 2 years apart. However, in the test for squared residuals the model completely falls behind other specifications with the p -values of the Ljung-Box test close to 0 in Figure 9b. This is an evidence of model PJSV's inability to represent not only the distribution of one-day ahead forecasted returns but also the dynamics of volatility. In Figures 6b and 7b an inspection of autocorrelation functions of levels and squares of generalized residuals visualizes the problem.

²¹We cannot simply conclude that the model imposes to little skewness and kurtosis, since the scale of the generalized residuals is also incorrect and less than 1. VaR analysis confirms this finding.

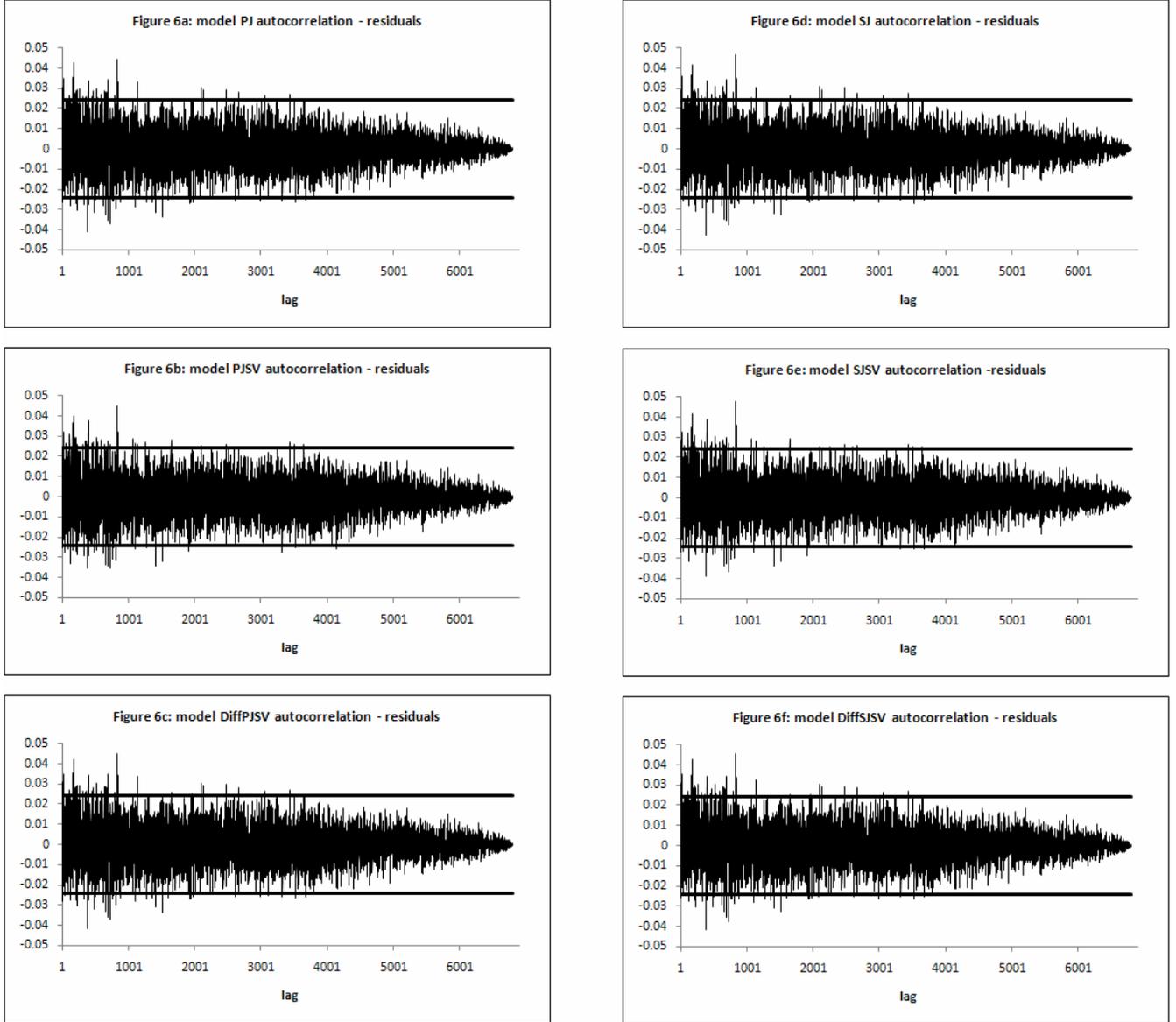


Figure 6. (a, b, c) autocorrelation function of the generalized residuals $\hat{z}_t^{(1)}$ for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV; (d, e, f) autocorrelation function of the generalized residuals $\hat{z}_t^{(1)}$ for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV. 95% confidence intervals depicted as horizontal lines. One of the possible explanations is an incorrect source of stochastic volatility.

Models PJ and DiffPJSV with diffusion included as a source of stochastic volatility dominate model PJSV. An inspection of the qq-plots in Figure 5 and descriptive statistics of generalized residuals in Table V reveal the clear advantage of models PJ and DiffPJSV over model PJSV in representing forecasting distribution. However, there is no significant difference in performance between models PJ and DiffPJSV. In the density forecast analysis presented in Table V both models perform on par and dominate model PJSV. Although they represent better forecasting distribution compared to model PJSV, they still fall short in this respect with Jarque-Bera p -values of 0.0061 and 0.0027 respectively. Hence, it still does not suffice to accept the null hypothesis of normally distributed generalized residuals even at the 1% level. In order to identify a source of the problem I inspect descriptive statistics of generalized residuals and find that both models

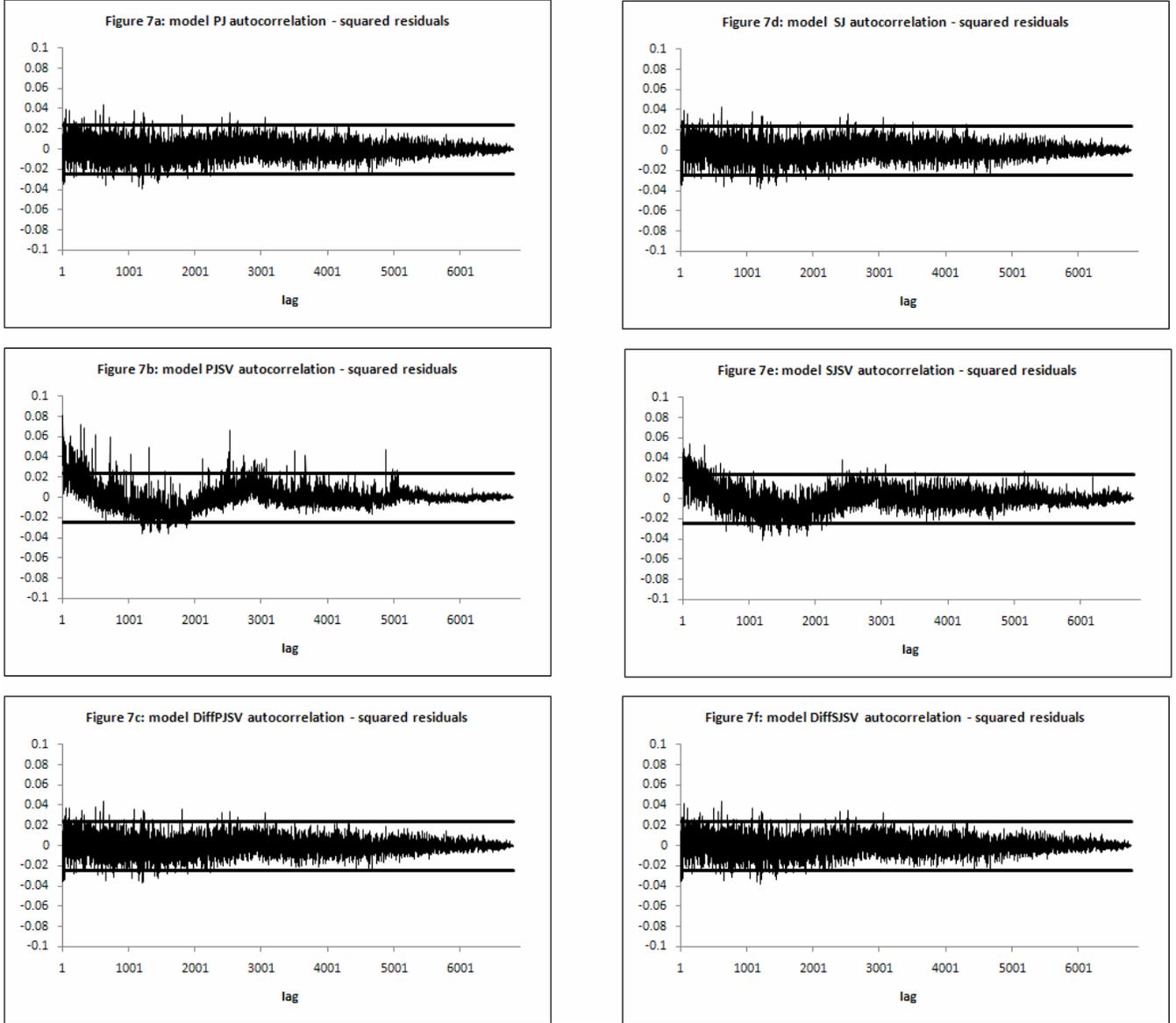


Figure 7. (a, b, c) autocorrelation function of the squared generalized residuals $\left(\hat{z}_t^{(1)}\right)^2$ for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV; (d, e, f) autocorrelation function of the squared generalized residuals $\left(\hat{z}_t^{(1)}\right)^2$ for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV. 95% confidence intervals depicted as horizontal lines.

fail with respect to kurtosis of the forecasting distribution but represent skewness slightly better. The high kurtosis values of 3.1517 and 3.1642, for models PJ and DiffPJSV respectively, are the main driving factor of high Jarque-Bera test statistics found in both models. This result stays in line with my previous findings from goodness of fit analysis in Section 4.3.2, where the models with Poisson jumps PJ and DiffPJSV represent skewness better than kurtosis of returns. Since the scale of the forecasting distribution is well represented we can draw conclusion that there is too small kurtosis in the PJ and DiffPJSV model implied forecasting distribution which is later verified in the VaR analysis. In the Ljung-Box test for dependence in the squares of generalized residuals presented in Figure 9 the test p -values are higher than 1% with an exception for the maximum lags

of approximately 25 – 400 and 1200 – 1800 business days. This constitutes a colossal improvement compared to model PJSV in representing the dynamics of volatility. The same conclusions can be drawn from an inspection of autocorrelation functions for the squares of generalized residuals in Figure 7. Most importantly, I do not find significant differences between models PJ and DiffPJSV.

Finally, model PJSV is outperformed by models PJ and DiffPJSV in the VaR analysis presented in Table VI. Model PJSV overestimates VaR values in the estimation sample and therefore underestimates the empirical coverage frequencies by inducing too high skewness and kurtosis in the forecasting distribution. Both models PJ and DiffPJSV perform on par in the VaR analysis underestimating VaR values at 1% with coverage frequencies of approximately 1.2%. Both models produce very good results at the 5% and 10% levels with a general tendency to perform better at higher significance levels.

Since models PJ and DiffPJSV include diffusion component as a source of stochastic volatility, it is the diffusion that contains the most information about latent stochastic volatility. I conclude that diffusion is the primary source of stochastic volatility in models with Poisson jumps which is intuitive, since Poisson jumps are rare. As discussed in the previous section, the correct specification of stochastic volatility is of major importance. It affects how the model performs in the density forecast and VaR analysis.

Summing up, model PJSV with stochastic volatility only from jumps is rejected not only with respect to the goodness of fit but also with respect to the density forecast and VaR performance, since it is outperformed by other specifications with Poisson jumps. Models PJ and DiffPJSV both perform on par and hence the benefits of additional source of stochastic volatility from the Poisson jumps in model DiffPJSV are rather minor, if any. The diffusion component serves as the primary source of stochastic volatility in the models with Poisson jumps.

4.4.2 Models with Lévy α -stable Jumps

In forecast analysis the main objective is to correctly represent the forecasting distribution, where filtered latent volatility states play the first role. Since Lévy α -stable jumps have infinite activity property, they are able to represent not only big and rare Poisson type jumps, but also more frequent and subtle jumps. Hence, when the Lévy α -stable jumps component is included as a source of stochastic volatility, it should provide additional information about latent stochastic volatility and therefore improve the density forecast performance. I also analyze an extreme case where stochastic volatility arises only from the pure jump Lévy α -stable process. It allows to verify if diffusion still plays the fundamental role as a source of stochastic volatility in the models with infinite activity, infinite variation jumps.

I find that we cannot simply exclude the diffusion as a component driving stochastic volatility even with Lévy α -stable jumps in returns. This is illustrated by the poor performance of model SJSV with respect to the density forecast and VaR analysis presented in Tables V and VI respectively. In the density forecast model SJSV shares similarities with model PJSV. It fails to represent scale, skewness and kurtosis of the forecasting distribution with the respective statistics of 0.891, -0.1221 and 3.7942 calculated for generalized residuals. This implies the high value of Jarque-Bera test statistic of 195.96 and rejection of normality, although this statistic is significantly improved compared to model PJSV. In Figure 5e, I present qq-plot that visualizes the failure of the SJSV specification to represent the forecasting distribution. Since the scale of the forecasting distribution is misspecified as in model PJSV, I have to postpone a conclusion about skewness

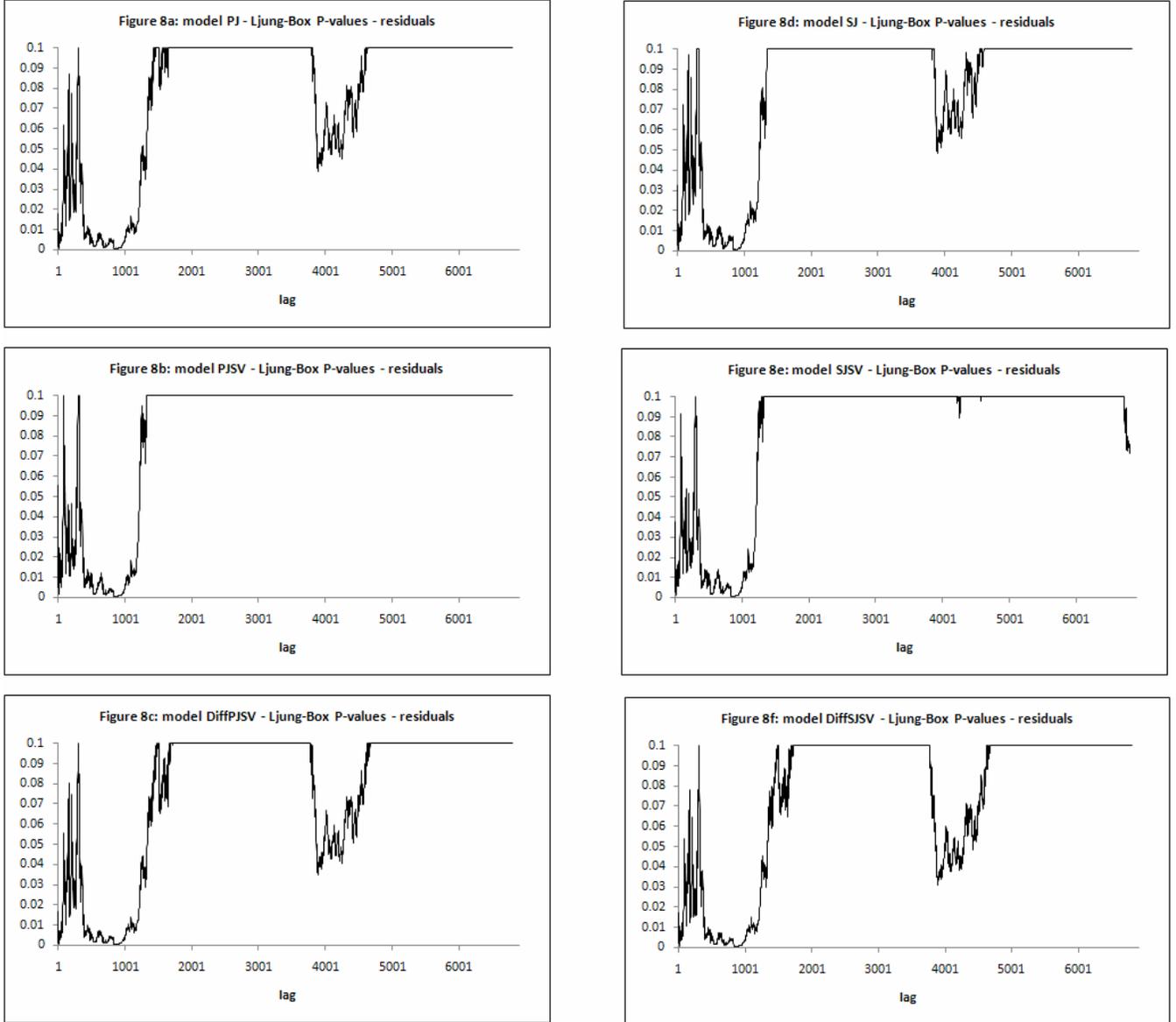


Figure 8. (a, b, c) Ljung-Box test p -values of the generalized residuals $\hat{z}_t^{(1)}$ for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV, calculated as a function of different maximum number of lags; (d, e, f) Ljung-Box test p -values of the generalized residuals $\hat{z}_t^{(1)}$ for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV, calculated as a function of different maximum number of lags. The values at the level 0.1 correspond to p -values at, or exceeding 0.1.

and kurtosis of the implied forecasting distribution to the VaR analysis. In the test for dependence in the squared generalized residuals in Figure (9) model SJSV completely falls behind other specifications with its Ljung-Box test p -values close to zero at all maximum lags considered (2 – 6811) with an exception of only one lag (1), where it equals 0.15. Hence, model SJSV can neither represent the distribution of one-day ahead forecasted returns nor the dynamics of volatility. However, both models PJSV and SJSV produce similar results to other specifications in the Ljung-Box test for dependence in levels of the generalized residuals in Figure (8). In Figures 6 and 7 an inspection of autocorrelation functions of levels and squares of generalized residuals visualize these findings.

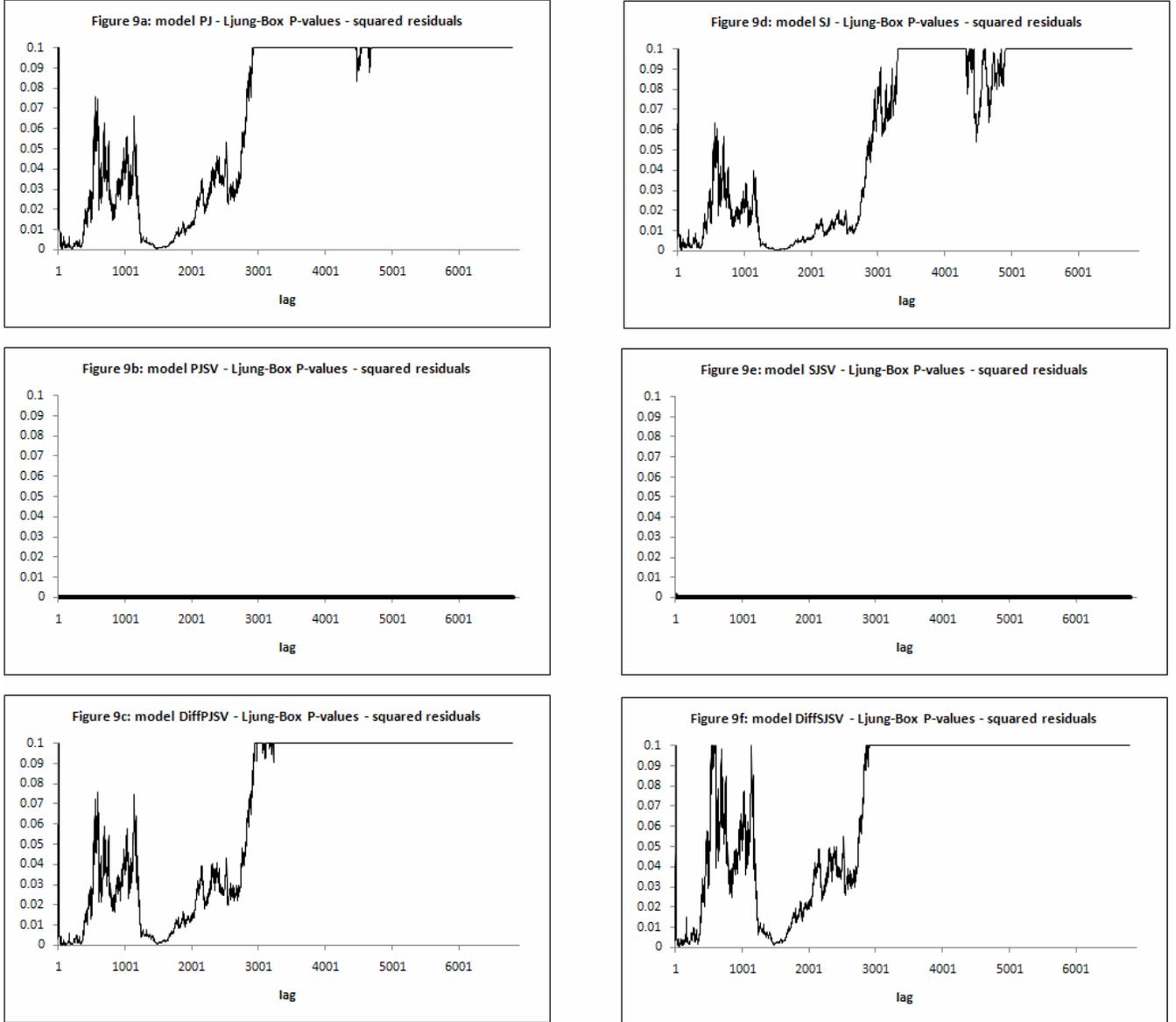


Figure 9. (a, b, c) Ljung-Box test p -values of the squared generalized residuals $\left(\hat{z}_t^{(1)}\right)^2$ for models with Poisson jumps, respectively PJ, PJSV, DiffPJSV, calculated as a function of different maximum number of lags; (d, e, f) Ljung-Box test p -values of the squared generalized residuals $\left(\hat{z}_t^{(1)}\right)^2$ for models with Lévy α -stable jumps, respectively SJ, SJSV, DiffSJSV, calculated at different maximum number of lags. The values at the level 0.1 correspond to p -values at, or exceeding 0.1.

Finally, I concentrate on models with diffusion as a source of stochastic volatility. In model DiffSJSV the joint stochastic volatility enables us to extract information about latent volatility from both diffusion and infinite activity jumps. The above produces the best performance in the density forecast analysis across all model specifications. This model performs the best in terms of representing skewness and kurtosis of predictive distribution with skewness of -0.0014 and kurtosis of 3.0981 compared to -0.0266 and 3.1042 respectively for model SJ. The Jarque-Bera p -value of 0.255 confirms superiority of model DiffSJSV in representing the forecasting distribution, although

model SJ with the p -value of 0.1432 also performs well. Both mean and standard deviation of generalized residuals are close to the theoretical values in models SJ and DiffSJSV. An inspection of qq-plot in Figure 5 verifies normality assumption of generalized residuals for both considered models. Model DiffSJSV's advantage over other models, including the SJ specification, shows up in the very left tail of the generalized residuals. As I find above, the autocorrelation functions in Figure 6 for levels of generalized residuals are not significantly affected by the model specification and all models perform similarly in the Ljung-Box test for levels of generalized residuals in Figure 8. On the other hand, I find significant differences between models SJ and DiffSJSV in the analysis of the squared generalized residuals and hence the model implied dynamics of volatility. Even though the autocorrelation functions in Figure 7 do not provide any significant proof of this, in Figure 9 I find that model DiffSJSV performs superior to any other model and has the highest Ljung-Box test p -values. Most importantly, comparing directly models SJ and DiffSJSV in Figures 9d,f I find that model DiffSJSV outperforms model SJ especially at maximum lags in the Ljung-Box test of up to about 5 years. This implies that model DiffSJSV is superior in representing the dynamics of volatility, which shows up in its density forecast performance.

In the VaR analysis in Table VI models SJ and DiffSJSV produce very good results among all specifications. Model DiffSJSV dominates all other specifications in the VaR analysis in the very left tail at 1% and 5% levels with empirical coverage frequencies of 0.98% and 5.06%. At 10% level it produces second best result with the coverage frequency of 9.97% comparing to 10% in model PJ. I also find that model SJ performs very well in the VaR analysis but is dominated by model DiffSJSV with a joint stochastic volatility specification.

The biggest problem in the VaR analysis arises in model SJSV at the 5% and 10% levels and this result is similar to model PJSV. The model improves at the 1% level. Model SJSV, although dominated in the VaR analysis by other models with Lévy α -stable jumps, performs better than model PJSV.

Summing up, the jump component cannot serve as the only source of stochastic volatility even in the models with Lévy α -stable jumps. In case of models SJ and DiffSJSV with diffusion as a source of stochastic volatility, I find evidence in favor of model DiffSJSV with stochastic volatility arising from both diffusion and jump components. This stays in contrast to the results for the Poisson jumps, where the marginal importance of the stochastic volatility from the jump component does not have a first order importance in the density forecast analysis.

4.4.3 Poisson or Lévy α -stable Jumps?

I focus attention on the choice of the jumps specification: compound Poisson or Lévy α -stable jumps. The model with Lévy α -stable jumps and joint stochastic volatility DiffSJSV outperforms all other models with respect to the density forecast and VaR analysis, which stems from the fact that it offers one important advantage in the modelling of asset returns. Since joint stochastic volatility makes possible to extract information about latent volatility from both diffusion and jumps and since jumps have infinite activity, they are more informative about the latent volatility. The infinite activity guarantees that it can represent small and frequent jumps as oppose to the Poisson jumps model, which is found to fit only big and rare jumps in returns. Since Poisson jumps occur very rarely, they have only limited information about latent stochastic volatility and even if this information is extracted in the form of stochastic volatility arising from both the diffusion and the jumps in model DiffPJSV, its forecasting performance is dominated by model DiffSJSV.

5 Conclusions

In this paper I address the choice of jump structure and source of stochastic volatility in the continuous-time jump diffusion models of asset returns. I consider two types of jump structures - compound Poisson and infinite activity Lévy α -stable jumps. The source of stochastic volatility comes from the diffusion component, the pure jump component or both. I use data on daily S&P500 index returns since it is a broad indicator of the equity markets. I perform estimation under statistical measure - this allows us to not only answer how the models fit the data but also how they perform in the density forecast and VaR analysis. The large family of models considered lets us marginalize the effects of different jump structures and source of stochastic volatility with respect to goodness of fit and density forecast performance.

I face the problem of parameters estimation in the presence of latent stochastic volatility and latent jump sizes. I perform estimation using Bayesian methods and propose a new algorithm for models with Lévy α -stable jumps. My method solves the problem of MCMC state-space separability and thus allows for the estimation of a broad class of continuous-time jump diffusion models with Lévy α -stable jumps and various sources of stochastic volatility.

Lévy α -stable jumps dominate compound Poisson jumps specifications with respect to goodness of fit analysis, since the latter are only suited to fit big and rare jumps. Moreover, models with Lévy α -stable jumps can adequately represent kurtosis of the underlying data and skewness of the returns distribution but only if diffusion is included as a source of stochastic volatility. It is important to note that models with stochastic volatility arising only from the pure jump component fail to fit the returns and this feature is irrespective of the jump structure specification. On the other hand, one cannot in a decisive way point out if there is a need for jump component as the second source of stochastic volatility by restricting analysis only to goodness of fit. This conclusion holds for the models with all considered jump structures including infinite activity Lévy α -stable jumps.

The density forecast and VaR analysis shed new light on the application of continuous-time jump diffusion models of asset returns. I find that correct specification of the source of stochastic volatility is of fundamental importance in the density forecast and VaR analysis. The performance of the compound Poisson jump models do not significantly change with the addition of the jump component to the diffusion as the source of stochastic volatility. On the contrary, the models with Lévy α -stable jumps improve in the density forecast and VaR performance with the inclusion of both sources of stochastic volatility, thus dominating all other model specifications. The joint stochastic volatility enables us to extract information about latent volatility from both diffusion and jumps, where the jumps are more informative with their infinite activity property. However, we cannot go further and exclude the diffusion from the source of stochastic volatility. This conclusion does not depend on the jump structure and agrees with the goodness of fit analysis.

A line for future research is to study the density forecast and VaR performance using data from the underlying and option prices. Since option prices contain information about latent volatility, it is important to investigate their potential explanatory power. Moreover, in this paper I analyze diffusion as the only source of leverage effect and further research could involve removing this restriction. This can help to answer a question of whether diffusion is still an important feature of the model when jump component is a source of leverage effect.

6 Appendix: Complete Conditional Posteriors

Since models DiffPJSV and DiffSJSV contain both: stochastic volatility from diffusion and jumps, we can derive complete conditional distributions for them and then apply the constraints in order to employ the derived distributions for other model specifications. However, note that they are not simple generalizations of the other specifications.

6.1 Models DiffPJSV and DiffSJSV - complete conditional posteriors for non-jump specific parameters and latent variables

Let use the notation from Sections 3.1 and 3.4. Let denote the jump specific parameters and latent variables depending on the model specification with Poisson or Lévy α -stable jumps by the index $m \in \{PJ, SJ\}$. In the following $\theta^{NJ}(PJ) = (\mu, \kappa_h, \theta_h, \sigma_h, \rho, \sigma)$, $\theta^{NJ}(SJ) = (\mu, \kappa_h, \theta_h, \sigma_h, \rho, \sigma_{SJ})$ are the vectors of non-jump specific parameters, $X^{NJ} = (\{h_t\}_{t=1}^T)$ latent log-volatility states and $J = (J_2, \dots, J_T) = \{J_t\}_{t=2}^T$ the jump sizes defined in (35). Moreover, let define the jump specific parameters and latent variables:

$$\begin{aligned}\theta^J(PJ) &= (\mu_j, \sigma_j), \\ \theta^J(SJ) &= (\alpha, \beta), \\ X^J(PJ) &= (\{q_t\}_{t=2}^T, \{\varkappa_t\}_{t=2}^T), \\ X^J(SJ) &= (\{\tilde{U}_t\}_{t=2}^T, \{\tilde{S}_t^+\}_{t=2}^T, \{\tilde{S}_t^-\}_{t=2}^T, \{v_t^+\}_{t=2}^T, \{v_t^-\}_{t=2}^T),\end{aligned}$$

and define $X \equiv X^{NJ} \cup X^J$ and $\theta = \theta^{NJ} \cup \theta^J$ for vector of latent variables and vector of parameters respectively.

1. updating μ

I choose the following prior distribution on μ : $\mu \sim N(a, A)$. I set $a = 0$ and $A = 10$, which is a relatively flat prior for the mean of asset returns. The conditional posterior distribution is conjugate to prior and given by:

$$p(\mu | \sigma_y, \kappa_h, \theta_h, \sigma_h, \rho, X^{NJ}, Y, J) \sim N(a^*, A^*),$$

where $A^* = (A^{-1} + \frac{\delta}{(1-\rho^2)} \sum_{t=1}^{T-1} \frac{1}{\xi_t})^{-1}$ and $a^* = A^* \cdot (A^{-1}a + \frac{1}{1-\rho^2} \sum_{t=1}^{T-1} \left[\frac{(Y_{t+1} - Y_t - J_{t+1})}{\xi_t} - \frac{\rho \varepsilon_{t+1}^{(3)} \delta^{0.5}}{\xi_t^{0.5}} \right])$.

2. updating κ_h

The prior on κ_h is $\kappa_h \sim$ truncated $N(b, B)$ with the support $\kappa_h \in (0, \frac{2}{\delta})$ and $b = \frac{1}{\delta}$, $B = \frac{6}{\delta}$, which is also a relatively flat prior that imposes stationarity on the log-volatility process h_t . Hence, the conditional posterior is also truncated and conjugate to prior:

$$p(\kappa_h | \mu, \sigma_y, \theta_h, \sigma_h, \rho, X^{NJ}, Y, J) \sim N(b^*, B^*),$$

where $\kappa_h \in (0, \frac{2}{\delta})$, $B^* = (B^{-1} + \frac{1}{1-\rho^2} \sum_{t=1}^{T-1} \frac{(\theta_h - h_t)^2 \delta}{\sigma_h^2})^{-1}$ and $b^* = B^* \cdot (B^{-1}b + \frac{1}{\sigma_h^2(1-\rho^2)} \sum_{t=1}^{T-1} \left[(h_{t+1} - h_t)(\theta_h - h_t) - \rho \sigma_h \delta^{0.5} \varepsilon_{t+1}^{(1)} (\theta_h - h_t) \right])$.

3. updating θ_h

The prior on θ_h is $\theta_h \sim N(c, C)$, with $c = 0$ and $C = 10$. The conditional posterior is conjugate to prior:

$$p(\theta_h | \mu, \kappa_h, \sigma_h, \rho, X^{NJ}, Y, J) \sim N(c^*, C^*),$$

where $C^* = (C^{-1} + \frac{(T-1)\kappa_h^2\delta}{\sigma_h^2(1-\rho^2)})^{-1}$, $c^* = C^* \cdot (C^{-1}c + \frac{\kappa_h}{\sigma_h^2(1-\rho^2)} \sum_{t=1}^{T-1} [h_{t+1} - h_t + \kappa_h h_t \delta - \sigma_h \delta^{1/2} \rho \varepsilon_{t+1}^{(1)}])$.

4. updating σ_h and ρ

I update σ_h and ρ as a block following Jacquier, Polson and Rossi (2004) (JPR). Let define the following bijective correspondence:

$$\phi_h = \sigma_h \rho \text{ and } \omega_h = \sigma_h^2(1 - \rho^2). \quad (43)$$

I choose the joint prior distribution on the transformed parameters ϕ_h and ω_h specified by $\omega_h \sim IG(d, D)$ and $\phi_h | \omega_h \sim N(0, \frac{1}{2}\omega_h)$ as in JPR. In my application I choose $d = 3$ and $D = \frac{1}{20}$ for an uninformative prior. The conditional posterior of ω_h is conjugate to prior²² and given by:

$$p(\omega_h | \mu, \kappa_h, \theta_h, X^{NJ}, Y, J) \sim IG(d^*, D^*),$$

where $d^* = d + \frac{T-1}{2}$, $D^* = D + \frac{1}{2}(err' * err) + b'_{par} \cdot (\sum_{t=1}^{T-1} (\varepsilon_{t+1}^{(1)})^2 + 2)^{-1} \cdot (\sum_{t=1}^{T-1} (\varepsilon_{t+1}^{(1)})^2) \cdot b_{par}$ and err is the vector of regression residuals and b_{par} is the OLS estimator of ϕ_h in the following regression model for $t = 1, 2, \dots, T-1$:

$$\frac{h_{t+1} - h_t}{\sqrt{\delta}} - \kappa_h(\theta_h - h_t)\sqrt{\delta} = \phi_h \varepsilon_{t+1}^{(1)} + \sqrt{\omega_h} \eta_{t+1}, \quad \eta_{t+1} \sim N(0, 1)$$

The conditional posterior of ϕ_h is also conjugate to prior distribution and is given by:

$$p(\phi_h | \omega_h, \mu, \kappa_h, \theta_h, X^{NJ}, Y, J) \sim N(m^*, M^*),$$

where $M^* = \omega_h \cdot (2 + \sum_{t=1}^{T-1} (\varepsilon_{t+1}^{(1)})^2)^{-1}$, $m^* = M^* \cdot \sum_{t=1}^{T-1} (\varepsilon_{t+1}^{(1)} [\frac{h_{t+1} - h_t}{\sqrt{\delta}} - \kappa_h(\theta_h - h_t)\sqrt{\delta}])$. After the draw of (ϕ_h, ω_h) we find (σ_h, ρ) by an inverse of the correspondence in (43) given by $\sigma_h^2 = \omega_h + \phi_h^2$, $\rho = \frac{\phi_h}{\sigma_h}$.

5. updating volatility states h_t for $1 < t < T$

By application of the Bayes rule, we have:

$$p(h_t | \theta, X^{-(h_t)}, Y) \propto p(Y_{t+1} | \theta, X^{NJ}, Y_t, J) p(Y_t | \theta, X^{NJ}, Y_{t-1}, J) p(h_{t+1} | h_t, \theta) p(h_t | h_{t-1}, \theta) \tilde{p}_t(m),$$

where

$$p(Y_{\tau+1} | \theta^{NJ}, X^{NJ}, Y_{\tau}, J) \propto \frac{1}{\xi_{\tau}^{0.5}} \exp \left\{ -\frac{1}{2} \frac{[Y_{\tau+1} - Y_{\tau} - \mu\delta - J_{\tau+1} - (\xi_{\tau}\delta)^{\frac{1}{2}} \rho \varepsilon_{\tau+1}^{(3)}]^2}{\delta(1 - \rho^2)\xi_{\tau}} \right\}, \quad (44)$$

²²Note, that we do not condition on ϕ_h .

$$p(h_{t+1}|h_t, \theta^{NJ})p(h_t|h_{t-1}, \theta^{NJ}) \sim N\left(\frac{(1 - \kappa_h \delta)(h_{t+1} + h_{t-1}) + \kappa_h^2 \delta^2 \theta_h}{(1 - \kappa_h \delta)^2 + 1}, \frac{\sigma_h^2 \delta}{(1 - \kappa_h \delta)^2 + 1}\right),$$

where $\tau = t - 1, t$ and the last term $\tilde{p}_t(m)$ depends on the model specification:

$$\tilde{p}_t(PJ) = p(q_{t+1}|h_t) \text{ and } \tilde{p}_t(SJ) = p((\tilde{S}_{t+1}^+, v_{t+1}^+)|\alpha, \beta, h_t, \sigma_{SJ})p((\tilde{S}_{t+1}^-, v_{t+1}^-)|\alpha, \beta, h_t, \sigma_{SJ}). \quad (45)$$

I apply MH algorithm, with a proposal density given by $p(h_{t+1}|h_t, \theta^{NJ})p(h_t|h_{t-1}, \theta^{NJ})$. I found it to be very efficient with high acceptance rates for most data sets.

6. updating volatility state h_t for $t = 1$

By application of the Bayes rule, we have:

$$p(h_1|\theta^{NJ}, \theta^J, X^J, X^{NJ-(h_1)}, Y) \propto p(Y_2|\theta^{NJ}, X^{NJ}, Y^{-(Y_2)}, J)p(h_1|h_2, \theta^{NJ})\hat{p}_1(m),$$

where the first component is given in (44) with $\tau = 1$, the second component is given by the symmetry formula for autoregressive models (AR):

$$p(h_1|h_2, \theta^{NJ}) \sim N(\theta_h \kappa_h \delta + (1 - \kappa_h \delta)h_2, \sigma_h^2 \delta),$$

and the last component \hat{p}_1 is given in (45). I use the MH step to sample from this distribution with proposal density given by $p(h_1|h_2, \theta^{NJ})$.

7. updating volatility state h_t for $t = T$

By application of the Bayes rule, we have:

$$p(h_T|\theta^{NJ}, X^{NJ-(h_T)}, Y, J) \propto p(Y_T|\theta^{NJ}, X^{NJ}, Y^{-(Y_T)}, J)p(h_T|h_{T-1}, \theta^{NJ}),$$

and after simplifying:

$$p(h_T|\theta^{NJ}, X^{NJ-(h_T)}, Y, J) \sim N\left(\xi - \frac{\rho\gamma\sigma_h}{\xi_{t-1}^{0.5}}, \sigma_h^2 \delta(1 - \rho^2)\right),$$

where $\xi = h_{T-1} + \kappa_h(\theta_h - h_{T-1})\delta$, $\gamma = -Y_T + Y_{T-1} + \mu\delta + J_T$.

8. updating parameter σ (models PJSV, SJSV and DiffPJSV only)

Lets define the function $g^{(m)}(h_t) = 1$ for models $m \in \{PJSV, SJSV\}$ and $g^{(m)}(h_t) = \exp(h_t)$ for model $m = DiffPJSV$. Update of parameter $\sigma > 0$ is equivalent to update in the following regression model:

$$\frac{Y_{t+1} - Y_t - \mu\delta - J_{t+1}}{(\delta g^{(m)}(h_t))^{0.5} \sqrt{1 - \rho^2}} = \sigma \cdot \left(\frac{\rho \varepsilon_{t+1}^{(3)}}{\sqrt{1 - \rho^2}} \right) + \sigma \cdot \eta_{t+1},$$

where $\eta_t \sim iid N(0, 1)$. Lets define $\Phi_t = \frac{Y_{t+1} - Y_t - \mu\delta - J_{t+1}}{(\delta g^{(m)}(h_t))^{0.5} \sqrt{1 - \rho^2}}$ and $\Psi_t = \frac{\rho \varepsilon_{t+1}^{(3)}}{\sqrt{1 - \rho^2}}$. We have the following

complete conditional posterior for σ^2 :

$$p(\sigma^2 | \theta^{NJ-(\sigma)}, X^{NJ}, J, Y) \propto p(\{R_t\} | \theta^{NJ}, X^{NJ}, J, \sigma) p(\sigma^2),$$

where $R_t = Y_t - Y_{t-1}$. The first component above is given by:

$$p(\{R_t\} | \theta^{NJ}, X^{NJ}, J, \sigma) \propto \frac{1}{\sigma^{T-1}} \exp\left(-\frac{\sum_{t=1}^{T-1} (\Phi_t)^2}{2\sigma^2} + \frac{\sum_{t=1}^{T-1} (\Phi_t \Psi_t)}{\sigma}\right).$$

I assume $p(\sigma^2) \sim IG(e, E)$ prior, with $e = 3$, $E = \frac{1}{20}$. In models PJSV and SJSV we have $\rho = 0$ and hence this posterior is conjugate to prior and given by $IG(e^*, E^*)$, where $e^* = e + \frac{T-1}{2}$ and $E^* = E + 0.5 \sum_{t=1}^{T-1} (\Phi_t)^2$. In the model DiffPJSV I use the normal random walk MH algorithm to draw from this conditional posterior.

9. updating σ_{SJ} (models SJ and DiffSJSV)

In the model SJ let assume an inverse gamma prior on σ_{SJ} , $p(\sigma_{SJ}) \sim IG(dd, DD)$, where $dd = 3$ and $DD = \frac{1}{20}$. In the model DiffSJSV I assume completely flat prior $p(\sigma_{SJ}) \sim U(0, 10)$. Let define the function $g^{(m)}(h_t) = 1$ for model $m = SJ$ and $g^{(m)}(h_t) = \exp(h_t)$ for model $m = DiffSJSV$. We have the following complete conditional posterior for σ_{SJ} :

$$p(\sigma_{SJ} | \theta^{-(\sigma_{SJ})}, X, Y) \propto p(\{\tilde{S}_t^+, v_t^+\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) p(\{\tilde{S}_t^-, v_t^-\}_{t=2}^T | \alpha, \beta, \sigma_{SJ}, \{h_t\}) p(\sigma_{SJ}) \propto \exp\left\{-\left(\frac{1}{\sigma_{SJ}}\right)^{\frac{\alpha}{\alpha-1}} \left[\sum_{t=2}^T \left(\left|\frac{\tilde{S}_t^-}{(g^{(m)}(h_{t-1})\delta)^{\frac{1}{\alpha}} v_t^-}\right|^{\frac{\alpha}{\alpha-1}} + \left|\frac{\tilde{S}_t^+}{(g^{(m)}(h_{t-1})\delta)^{\frac{1}{\alpha}} v_t^+}\right|^{\frac{\alpha}{\alpha-1}}\right)\right]\right\} \left(\frac{1}{\sigma_{SJ}^{\frac{\alpha}{\alpha-1}}}\right)^{2(T-1)} \cdot p(\sigma_{SJ}).$$

Let define the following proposal density in the MH step for updating σ_{SJ} :

$$q(\sigma_{SJ}) \propto \exp\left\{-\left(\frac{1}{\sigma_{SJ}}\right)^{\frac{\alpha}{\alpha-1}} \left[\sum_{t=2}^T \left(\left|\frac{\tilde{S}_t^-}{(g^{(m)}(h_{t-1})\delta)^{\frac{1}{\alpha}} v_t^-}\right|^{\frac{\alpha}{\alpha-1}} + \left|\frac{\tilde{S}_t^+}{(g^{(m)}(h_{t-1})\delta)^{\frac{1}{\alpha}} v_t^+}\right|^{\frac{\alpha}{\alpha-1}}\right)\right]\right\} \left(\frac{1}{\sigma_{SJ}^{\frac{\alpha}{\alpha-1}}}\right)^{2(T-1)}.$$

We can directly draw from this distribution by the change of variable. Let $\hat{\sigma}_{SJ} = \sigma_{SJ}^{\frac{\alpha}{\alpha-1}}$, then by the change of variable formula:

$$\hat{\sigma}_{SJ} \sim IG\left(\left(2(T-1) + \frac{1}{\alpha} - 1\right), \sum_{t=2}^T \left(\left|\frac{\tilde{S}_t^-}{(g^{(m)}(h_{t-1})\delta)^{\frac{1}{\alpha}} v_t^-}\right|^{\frac{\alpha}{\alpha-1}} + \left|\frac{\tilde{S}_t^+}{(g^{(m)}(h_{t-1})\delta)^{\frac{1}{\alpha}} v_t^+}\right|^{\frac{\alpha}{\alpha-1}}\right)\right),$$

After drawing from this distribution, we have to solve for $\sigma_{SJ} = \hat{\sigma}_{SJ}^{\frac{\alpha-1}{\alpha}}$. In the model SJ I use MH algorithm with the above choice of proposal density, which is very efficient with acceptance rate of above 98%. This also reflects that my prior in the model SJ is uninformative, since for completely flat prior on σ_{SJ} the acceptance rate should be 100%. In the model DiffSJSV I directly draw from the conditional posterior, since I impose the uniform prior $U(0, 10)$ and there is no need to reweight the draws.

6.2 Models with Poisson Jumps: complete conditional posteriors for jump specific parameters and latent variables

1. updating latent jump times $\{q_t\}_{t=2}^T$

Since $p(q_{t+1}|\lambda(h_t))$ is *Bernoulli*($\delta\lambda_t$), the conditional posterior is also *Bernoulli*:

$$p(q_{t+1} = j|\theta, X^{-(q_{t+1})}, Y) \propto p(Y_{t+1}|Y_t, X^{-(q_{t+1})}, \theta, q_{t+1} = j)p(q_{t+1} = j|\lambda_t)$$

where $j = 0, 1$ and the sum of probabilities across j is normalized to one. In the above the first component is given by:

$$p(Y_{t+1}|Y_t, \theta, X) \propto \exp\left(-\frac{1}{2} \frac{[Y_{t+1} - Y_t - \mu\delta - q_{t+1}\varkappa_{t+1} - (\xi_t\delta)^{\frac{1}{2}}\rho\varepsilon_{t+1}^{(3)}]^2}{\delta(1-\rho^2)\xi_t}\right).$$

2. updating latent jump sizes $\{\varkappa_t\}_{t=2}^T$

By application of the Bayes rule, we have:

$$p(\varkappa_{t+1}|\theta, X^{-(\varkappa_{t+1})}, Y) \propto p(Y_{t+1}|Y_t, X^{-(\varkappa_{t+1})}, \theta, \varkappa_{t+1})p(\varkappa_{t+1}|\mu_j, \sigma_j)$$

If q_{t+1} takes the value 0, then the first component is not a function of \varkappa_{t+1} and we are left with the prior distribution $p(\varkappa_{t+1}|\mu_j, \sigma_j) \sim N(\mu_j, \sigma_j^2)$. If q_{t+1} takes the value 1, then both components are the function of \varkappa_{t+1} and we have:

$$p(\varkappa_{t+1}|\theta, X^{-(\varkappa_{t+1})}, Y) \sim N\left(-\frac{1}{2} \frac{B}{A}, \frac{(1-\rho^2)\delta\xi_t\sigma_j^2}{A}\right),$$

where $A = \sigma_j^2 + (1-\rho^2)\delta\xi_t$, and $B = -2\sigma_j^2(Y_{t+1} - Y_t - \mu\delta - (\xi_t\delta)^{\frac{1}{2}}\rho\varepsilon_{t+1}^{(3)}) - 2\mu_j(1-\rho^2)\delta\xi_t$.

3. updating mean jump size parameter μ_j

I take the prior $p(\mu_j) \sim N(a, A)$ with $a = 0$, $A = 10$. By application of the Bayes rule, we have:

$$p(\mu_j|\theta^{-(\mu_j)}, X, Y) \propto p(\{\varkappa_{t+1}\}|\mu_j, \sigma_j^2)p(\mu_j) \sim N(a^*, A^*),$$

where $A^* = (A^{-1} + (T-1)\sigma_j^{-2})^{-1}$ and $a^* = A^*(A^{-1}a + \sigma_j^{-2} \sum_{t=2}^T \varkappa_t)$.

4. updating variance jump size parameter σ_j^2

I take the inverse gamma prior $p(\sigma_j^2) \sim IG(b, B)$ with $b = 3$ and $B = \frac{1}{20}$. By application of the Bayes rule, we have:

$$p(\sigma_j^2|\theta^{-(\sigma_j^2)}, X, Y) \propto p(\{\varkappa_{t+1}\}|\mu_j, \sigma_j^2)p(\sigma_j^2) \sim IG(b^*, B^*),$$

and the posterior is conjugate with $b^* = (T-1)/2 + b$ and $B^* = B + 0.5 \sum_{t=2}^T (\varkappa_t - \mu_j)^2$.

5. updating constant jump intensity λ_j in model PJ

I specify the beta prior distribution on $\lambda_j \sim \text{Beta}(0.5, 0.5)$. Then the conditional posterior is conjugate to prior:

$$p(\lambda_j | \theta^{-(\lambda_j)}, X, Y) \propto p(\{q_t\}_{t=2}^T | \lambda_j) p(\lambda_j) \sim \text{Beta}\left(\sum_{t=2}^T q_t + 0.5, \sum_{t=2}^T (1 - q_t) + 0.5\right).$$

6.3 Models with Lévy α -stable Jumps: complete conditional posteriors for jump specific parameters and latent variables

1. updating volatility states $\{\frac{h_t}{\alpha}\}$ in model SJSV

Since in this model $\{h_t\}$ process drives only jumps, I redefine SJSV model to simplify its update procedure by using a change of variable $\tilde{h}_t \equiv h_t/\alpha$ and then directly update $\{\tilde{h}_t\}$. I also redefine the set X^{NJ} to replace $\{h_t\}$ by $\{\tilde{h}_t\}$.²³ The OU process $\{\tilde{h}_t\}$ is given by:

$$\tilde{h}_{t+\delta} = \tilde{h}_t + \tilde{\kappa}_h(\tilde{\theta}_h - \tilde{h}_t)\delta + \tilde{\sigma}_h\sqrt{\delta}\varepsilon_{t+\delta}^{(2)},$$

where the parameters satisfy $\tilde{\kappa}_h \equiv \kappa_h$, $\tilde{\theta}_h \equiv \frac{\theta_h}{\alpha}$, $\tilde{\sigma}_h \equiv \frac{\sigma_h}{\alpha}$ and replace respectively κ_h , θ_h and σ_h in the vector θ^{NJ} . I later derive the estimates for κ_h , θ_h and σ_h by calculating their respective values at the end of each MCMC step. By ergodic theorem, their average converges to the posterior mean.

The complete conditional posterior for $\{\tilde{h}_t\}$, $1 < t < T$, is given by:

$$p(\tilde{h}_t | \theta, X^{-(\tilde{h}_t)}, Y) \propto p(Y_{t+1} | Y^{-(Y_{t+1})}, \theta, X) p(\tilde{h}_t | \tilde{h}_{t+1}, \tilde{h}_{t-1}, \tilde{\theta}_h, \tilde{\kappa}_h, \tilde{\sigma}_h),$$

where

$$p(\tilde{h}_t | \tilde{h}_{t+1}, \tilde{h}_{t-1}, \tilde{\theta}_h, \tilde{\kappa}_h, \tilde{\sigma}_h) \sim N\left(\frac{(1 - \tilde{\kappa}_h\delta)(\tilde{h}_{t+1} + \tilde{h}_{t-1}) + (\tilde{\kappa}_h\delta)^2\tilde{\theta}_h}{(1 - \tilde{\kappa}_h\delta)^2 + 1}, \frac{\tilde{\sigma}_h^2\delta}{(1 - \tilde{\kappa}_h\delta)^2 + 1}\right),$$

and

$$p(Y_{\tau+1} | Y^{-(Y_{\tau+1})}, \theta, X) \propto \exp\left(-\frac{1}{2} \frac{(Y_{\tau+1} - Y_\tau - \mu\delta - \exp(\tilde{h}_\tau) \left(\frac{S_{\tau+1}}{\exp(\tilde{h}_\tau)}\right))^2}{\sigma^2\delta}\right) \quad (46)$$

for $\tau = t$. Note, that the conditional distribution of $\frac{S_{\tau+1}}{\exp(\tilde{h}_\tau)}$ has a constant scale which does not depend on \tilde{h}_τ . Hence, I also directly update positive and negative parts for $S'_{\tau+1} \equiv \frac{S_{\tau+1}}{\exp(\tilde{h}_\tau)}$ and replace the respective positive and negative parts of $S_{\tau+1}$ in the set X^J . Moreover, for $t = 1$ we have:

$$p(\tilde{h}_1 | \theta, X^{-(\tilde{h}_1)}, Y) \propto p(Y_2 | Y^{-(Y_2)}, \theta, X) p(\tilde{h}_1 | \tilde{h}_2, \tilde{\theta}_h, \tilde{\kappa}_h, \tilde{\sigma}_h),$$

²³This simplifies the update of α presented in Section 3.4.3, where we condition on $\{\tilde{h}_t\}$ and not on $\{h_t\}$.

where the second component is given by the symmetry formula for autoregressive models (AR):

$$p(\tilde{h}_1|\tilde{h}_2, \tilde{\theta}_h, \tilde{\kappa}_h, \tilde{\sigma}_h) \sim N(\tilde{\theta}_h\tilde{\kappa}_h\delta + (1 - \tilde{\kappa}_h\delta)\tilde{h}_2, \tilde{\sigma}_h^2\delta),$$

and the first by (46) calculated at $\tau = 1$. After a careful study of the conditional posterior for \tilde{h}_t for $t = 1, 2, \dots, (T-1)$, I find that it is either bimodal or unimodal and in general in the unimodal case it is not log-concave. However, I can numerically compute all local maxima using Newton's method without much computational burden and then apply rejection algorithm utilizing the rejected points. Finally, the complete conditional posterior for \tilde{h}_T is $N(\tilde{\kappa}_h\delta\tilde{\theta}_h + (1 - \tilde{\kappa}_h\delta)\tilde{h}_{T-1}, \tilde{\sigma}_h^2\delta)$.

References

- [1] Aït-Sahalia, Y., 2003, Disentangling volatility from jumps, Working Paper
- [2] Aït-Sahalia, Y. and J. Jacod, 2008, Testing for jumps in a discretely observed process, *Annals of Statistics*, forthcoming
- [3] Andersen, T.G., L. Benzoni and J. Lund, 2002, An empirical investigation of continuous-time equity return models, *Journal of Finance*, vol. 57 (3, June), 1239-1284.
- [4] Applebaum, D., 2004, *Lévy processes and stochastic calculus* (Cambridge University Press).
- [5] Bertoin, J., 1998, *Lévy processes* (Cambridge University Press, Cambridge)
- [6] Black, F. and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy*, 81, 637-659.
- [7] Buckle, D.J., 1995, Bayesian inference for stable distributions, *Journal of the American Statistical Association*, 90, 605-613.
- [8] Carr, P., H. Geman, D.B. Madan and M. Yor, 2002, The Fine structure of asset returns: An empirical investigation, *Journal of Business*, vol. 75, no.2, 305-332.
- [9] Carr, P. and L. Wu, 2003, The finite moment log stable process and option pricing, *The Journal of Finance*, vol. LVIII, no.2, 753-778.
- [10] Carr, P. and L.Wu, 2004, Time-changed Lévy processes and option pricing, *Journal of Financial Economics*, 71, 113-141.
- [11] Chambers, J.M., C.L. Mallows and B.W. Stuck, 1976, A method for simulating stable random variables, *Journal of the American Statistical Association*, vol. 71, no. 354, 340-344.
- [12] Chernov, M. and E. Ghysels, 2000, A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of option valuation, *Journal of Financial Economics*, 56, 407-458.
- [13] Chib, S. and E. Greenberg, 1996, Markov Chain Monte Carlo simulation methods in econometrics, *Econometric Theory*, 12, 409-431.
- [14] Chib, S., F. Nardari and N. Shephard, 2002, Markov chain Monte Carlo methods for stochastic volatility models, *Journal of Econometrics*, 108, 281-316.
- [15] Christoffersen, P., 1998, Evaluating interval forecasts, *International Economic Review*, vol. 39, no.4, 841-862.
- [16] Cox, J.C., J.E. Ingersoll, Jr., and S.A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385-407.
- [17] Das, S.R., R.K. Sundaram, 1999, Of smiles and smirks: a term structure perspective, *Journal of Financial and Quantitative Analysis*, vol. 34, no. 2, 211-239.
- [18] Devroye, L., 1986, *Nonuniform random variate generation* (Springer Verlag, New York)
- [19] Diebold, F., T. Gunther and A. Tay, 1998, Evaluating density forecasts with applications to financial risk management, *International Economic Review*, vol. 39, 863-883.
- [20] Durham, G., 2006, Monte Carlo methods for estimating, smoothing and filtering one- and two-factor stochastic volatility models, *Journal of Econometrics*, vol. 133, 273-305.

- [21] Eraker, B., 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices, *The Journal of Finance*, vol. LIX, No. 3, 1367-1403
- [22] Eraker, B., M. Johannes and N. Polson, 2003, The impact of jumps in volatility and returns, *The Journal of Finance*, vol. LVIII, no.3, 1269-1300.
- [23] Gallant, A.R. and G. Tauchen, 1996, Which moments to match?, *Econometric Theory*, 12, 657-681.
- [24] Geman, S. and D. Geman, 1984, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 609-628.
- [25] Gilks, W.R. and P. Wild, 1992, Adaptive rejection sampling for Gibbs sampling, *Applied Statistics*, vol. 41, no. 2, 337-348.
- [26] Heston, S., 1993, A closed - form solution for options with stochastic volatility with applications to bonds and currency options, *Review of Financial Studies*, 6, 327-344.
- [27] Huang, J.Z. and L. Wu, 2004, Specification analysis of option pricing models based on time-changed Lévy processes, *Journal of Finance*, vol. LIX, No.3, 1405-1439.
- [28] Jacod, J. and P. Protter, 1998, Asymptotic error distributions for the Euler method for stochastic differential equations, *The Annals of Probability*, vol. 26, no. 1, 267-307.
- [29] Jacquier, E., N.G. Polson and P.E. Rossi, 1994, Bayesian analysis of stochastic volatility models, *Journal of Business & Economic Statistics*, vol. 12, no. 4, 371-389.
- [30] Jacquier, E., N.G. Polson and P.E. Rossi, 2004, Bayesian analysis of stochastic volatility models with fat-tails and correlated errors, *Journal of Econometrics*, 122, 185-212.
- [31] Janicki, A. and A. Weron, 1994, *Simulation and chaotic behavior of α -stable stochastic processes* (Marcel Dekker, New York)
- [32] Johannes, M. and N. Polson, 2003, MCMC methods for Financial Econometrics, 2003, Working Paper
- [33] Johannes, M., N. Polson and J. Stroud, 2008, Optimal filtering of jump-diffusions: extracting latent states from asset prices, Working Paper
- [34] Jones, C.S., 1998, Bayesian estimation of continuous - time Finance models, Working Paper
- [35] Jones, C.S., 2003, The dynamics of stochastic volatility: evidence from underlying and options market, *Journal of Econometrics* 116, 181-224.
- [36] Kloeden, P.E. and E. Platen, 1992, *Numerical simulation of stochastic differential equations* (Springer-Verlag, New York, NY).
- [37] Li, H., M.T. Wells and C.L. Yu, 2008, A Bayesian analysis of return dynamics with Lévy jumps, *The Review of Financial Studies*, vol. 21, issue 5, p. 2345-2378
- [38] Madan, D.B., P. Carr and E.C. Chang, 1998, The variance gamma process and option pricing, *European Finance Review*, 2, 79-105.
- [39] Merton, R.C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics*, vol. 3, No. 1-2, pp. 125-144

- [40] Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, 1953, Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, 21, 1087-1092.
- [41] Nolan, J.P., 2005, *Stable Distributions, models for heavy tailed data*, American University Manuscript
- [42] Pitt, M. and N. Shephard, 1999, Filtering via simulation: auxiliary particle filter, *Journal of the American Statistical Association*, 94, 590-599.
- [43] Samorodnitsky, G. and M.S. Taqqu, 1994, *Stable non-Gaussian random processes: stochastic models with infinite variance* (Chapman & Hall, New York)
- [44] Zolotarev, V.M., 1986, *One - dimensional stable distributions*, Amer. Math. Soc. Transl. of Math. Monographs, vol. 65, American Mathematical Society, Providence, RI (translation of the original 1983 Russian)

Table I**The List of Restrictions Defining All Model Specifications**

The table reports the list of restrictions for models presented in SDE (8).

	ρ	$\xi(x)$	$\lambda(x)$
<i>Stoch. Volatility from Diffusion</i>			
model (1) PJ	$-1 < \rho < 1$	$\exp(x)$	λ_j
model (4) SJ	$-1 < \rho < 1$	$\exp(x)$	$(\sigma_{SJ})^\alpha$
<i>Stoch. Volatility from Jumps</i>			
model (2) PJSV	0	$\sigma^2 > 0$	$\exp(x)$
model (5) SJSV	0	$\sigma^2 > 0$	$\exp(x)$
<i>Joint Stochastic Volatility</i>			
model (3) DiffPJSV	$-1 < \rho < 1$	$\sigma^2 \exp(x)$	$\exp(x)$
model (6) DiffSJSV	$-1 < \rho < 1$	$\exp(x)$	$(\sigma_{SJ})^\alpha \exp(x)$

Table II**Descriptive statistics of S&P500 daily log-returns**

The table reports the mean, standard deviation, skewness and kurtosis of S&P500 daily log-returns (x100).

S&P 500 daily log-returns (x100)	Date	Mean	St. Deviation	Skewness	Kurtosis
estimation sample T=6812	01/02/1981 - 12/31/2007	3.49%	1.03	-1.7465	42.79

Table III

Parameter Estimates from S&P500 Returns Data

The table reports the descriptive statistics of posterior distributions for S&P500 return models with S&P500 index data on $T = 6813$ observations, 01/02/1981-12/31/2007. Parameter estimates and standard errors (in parentheses) are the mean and standard deviation of the respective posterior distributions.

Poisson Jumps, Stochastic Volatility from Diffusion								
	μ	κ_h	θ_h	σ_h	λ_j	μ_j	σ_j	ρ
PJ	3.678E-04 (9.32E-05)	0.0143 (0.0027)	-9.5555 (0.1158)	0.133 (0.0102)	0.0022 (8.16E-04)	-0.0436 (0.0284)	0.0886 (0.0181)	-0.5891 (0.0411)
Poisson Jumps, Stochastic Volatility from Jumps								
	μ	κ_h	θ_h	σ_h	σ	μ_j	σ_j	-
PJSV	4.630E-04 (1.18E-04)	0.0045 (0.0027)	-5.9237 (2.2717)	0.4827 (0.0995)	0.0091 (1.05E-04)	-0.0022 (0.0019)	0.0321 (0.0019)	- - -
Poisson Jumps, Joint Stochastic Volatility								
	μ	κ_h	θ_h	σ_h	σ	μ_j	σ_j	ρ
DiffPJSV	3.696E-04 (9.32E-05)	0.0141 (0.0028)	-6.1545 (0.3666)	0.1309 (0.0109)	0.1849 (0.0313)	-0.0382 (0.0244)	0.0845 (0.0154)	-0.588 (0.0411)
Stable Jumps, Stochastic Volatility from Diffusion								
	μ	κ_h	θ_h	σ_h	σ_{SJ}	α	β	ρ
SJ	2.824E-04 (1.01E-04)	0.0107 (0.0022)	-10.0638 (0.2144)	0.179 (0.0133)	0.003 (1.68E-04)	1.8528 (0.0319)	0.4869 (0.2045)	-0.7496 (0.0466)
Stable Jumps, Stochastic Volatility from Jumps								
	μ	κ_h	θ_h	σ_h	σ	α	β	-
SJSV	4.606E-04 (1.15E-04)	0.0027 (0.0015)	-11.7911 (3.8597)	0.2874 (0.0328)	0.0086 (9.57E-05)	1.8974 (0.0394)	0.3882 (0.2549)	- - -
Stable Jumps, Joint Stochastic Volatility								
	μ	κ_h	θ_h	σ_h	σ_{SJ}	α	β	ρ
DiffSJSV	2.547E-04 (9.823E-05)	0.0145 (0.0027)	-9.6756 (0.1179)	0.1307 (0.0101)	0.5463 (0.1094)	1.6212 (0.0815)	0.8256 (0.1379)	-0.6428 (0.0443)

Table IV

In-Sample Fit of Returns and Volatility

The table reports posterior medians and 95% confidence intervals (in parentheses) of mean, std. deviation, skewness, kurtosis and daily autocorrelation of model residuals calculated at each step of MCMC algorithms for different model specifications. A correct specification implies, that mean is zero, standard deviation one, skewness zero, kurtosis three and daily autocorrelation zero.

	Mean	Std. Deviation	Skewness	Kurtosis	Daily autocorr
Return Equation					
PJ	0.0035 (-0.0190;0.0260)	1.0002 (0.9826;1.0164)	-0.0484 (-0.0869;-0.0081)	3.2578 (3.1666;3.3654)	0.0226 (0.0166;0.0284)
PJSV	3.97E-05 (-0.0237;0.0237)	0.9084 (0.8929;0.9241)	-0.0382 (-0.0831;0.0076)	3.7349 (3.6179;3.8528)	0.0166 (0.0046;0.0286)
DiffPJSV	0.0035 (-0.0191;0.0261)	1.0002 (0.9834;1.0169)	-0.047 (-0.0860;-0.0082)	3.2481 (3.1528;3.3590)	0.0232 (0.0175;0.0291)
SJ	0.005 (-0.0172;0.0271)	0.9995 (0.9829;1.0168)	-0.0074 (-0.0630;0.0482)	3.059 (2.9532;3.1779)	0.0101 (-0.0080;0.0279)
SJSV	0.0001 (-0.0236;0.0238)	0.8965 (0.8812;0.9119)	-0.0301 (-0.0825;0.0227)	3.7059 (3.5671;3.8480)	0.0209 (0.0061;0.0356)
DiffSJSV	0.0024 (-0.0203;0.0251)	0.9999 (0.9835;1.0165)	0.0076 (-0.0416;0.0561)	3.1397 (3.0444;3.2429)	0.0198 (0.0078;0.0307)
Log-Volatility Equation					
PJ	-0.003 (-0.0263;0.0203)	0.9998 (0.9832;1.0166)	0.0064 (-0.0521;0.0651)	3.0302 (2.9168;3.1570)	0.0077 (-0.0145;0.0295)
PJSV	-1.07E-02 (-0.0319;0.0116)	1.0003 (0.9835;1.0171)	0.0003 (-0.0578;0.0585)	2.9974 (2.8878;3.1202)	0.0005 (-0.0233;0.0242)
DiffPJSV	-0.0027 (-0.0260;0.0207)	0.9999 (0.9832;1.0167)	0.0057 (-0.0525;0.0642)	3.0284 (2.9148;3.1566)	0.0078 (-0.0144;0.0298)
SJ	-0.0049 (-0.0277;0.0180)	0.9998 (0.9830;1.0168)	0.001 (-0.0568;0.0592)	3.0186 (2.9072;3.1428)	0.0052 (-0.0168;0.0271)
SJSV	-0.0063 (-0.0289;0.0173)	1 (0.9833;1.0169)	0.0011 (-0.0571;0.0592)	2.998 (2.8886;3.1213)	0.0001 (-0.0237;0.0239)
DiffSJSV	-0.0023 (-0.0257;0.0210)	0.9998 (0.9831;1.0166)	-0.0045 (-0.0624;0.0539)	3.0269 (2.9150;3.1522)	0.0067 (-0.0151;0.0285)

Table V**Density Forecast Analysis**

Descriptive statistics of $\{\hat{z}_t^{(1)}\}$ distribution with $\hat{z}_t^{(1)}$ defined in eq. (36) and calculated for different model specifications with their parameters fixed at the MCMC estimates for S&P500 data from 01/02/1981 to 12/31/2007. Data on S&P500 daily returns, used for $\hat{z}_t^{(1)}$ calculation, are derived from S&P500 index levels for the full available sample from 01/02/1981 to 12/31/2007. A correct specification implies that mean is zero, standard deviation one, skewness zero and kurtosis three. Jarque-Bera test statistics for normality are presented for the $\hat{z}_t^{(1)}$ levels with the respective p - values (in parantheses).

	Mean	Std. Deviation	Skewness	Kurtosis	Jarque-Bera
One-Day Ahead Forecast, $L=1$					
PJ	0.0022	0.9937	-0.0569	3.1517	10.2113 (0.0061)
PJSV	-0.0054	0.8834	-0.2335	4.9226	1.11E+03 (0)
DiffPJSV	0.0022	0.9922	-0.0608	3.1642	11.8461 (0.0027)
SJ	0.0033	0.9829	-0.0266	3.1042	3.8866 (0.1432)
SJSV	-0.0133	0.8910	-0.1221	3.7942	195.9642 (0)
DiffSJSV	0.0024	0.9931	-0.0014	3.0981	2.7327 (0.2550)

Table VI**Value-at-Risk Analysis**

One-day horizon VaR performance is presented for all model specifications. For each model the parameters are fixed at the MCMC estimates for S&P500 data from 01/02/1981 to 12/31/2007. I report (left sided) significance levels of 1%, 5% and 10% and then compute empirical coverage frequencies. I derive empirical coverage frequencies using $z_t^{(1)}$ values defined in eq. (33).

One-Day Ahead VaR Analysis, $L=1$			
01/02/1981-12/31/2007 T=6812			
	1%	5%	10%
PJ	0.0120	0.0511	0.1000
PJSV	0.0081	0.0367	0.0730
DiffPJSV	0.0125	0.0509	0.0997
SJ	0.0103	0.0484	0.0979
SJSV	0.0094	0.0388	0.0768
DiffSJSV	0.0098	0.0506	0.0997