Journal of Econometrics 160 (2011) 12–21

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Econometric analysis of jump-driven stochastic volatility models

Viktor Todorov

Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208, United States

ARTICLE INFO

Article history:
Available online 6 March 2010

JEL classification:
G12
C51
C52

Keywords:
Lévy process
Method-of-moments
Power variation
Quadratic variation
Realized variance
Stochastic volatility

ABSTRACT

This paper introduces and studies the econometric properties of a general new class of models, which I refer to as jump-driven stochastic volatility models, in which the volatility is a moving average of past jumps. I focus attention on two particular semiparametric classes of jump-driven stochastic volatility models. In the first, the price has a continuous component with time-varying volatility and time-homogeneous jumps. The second jump-driven stochastic volatility model analyzed here has only jumps in the price, which have time-varying size. In the empirical application I model the memory of the stochastic variance with a CARMA(2,1) kernel and set the jumps in the variance to be proportional to the squared price jumps. The estimation, which is based on matching moments of certain realized power variation statistics calculated from high-frequency foreign exchange data, shows that the jump-driven stochastic volatility model containing continuous component in the price performs best. It outperforms a standard two-factor affine jump–diffusion model, but also the pure-jump jump-driven stochastic volatility model for the particular jump specification.

1. Introduction

Continuous-time stochastic volatility models have been used for a long time in the theoretical and empirical finance literature. The typical way of modeling the volatility process in these models is with (a superposition of) non-negative diffusion processes (e.g. square-root processes). However, empirical results in the studies of Eraker et al. (2003) and Pan (2002) among others suggest that including jumps in the stochastic volatility is necessary in order to account for sudden big changes in volatility. The main goal of this paper is to introduce and analyze some of the econometric properties of a general class of continuous-time models, referred to as jump-driven stochastic volatility (hereafter JDSV) models, and check empirically whether models in this class can provide good fit to high-frequency financial data. The distinctive characteristic of the JDSV models is that the stochastic variance is purely jump-driven. It is modeled as a general moving average of past positive jumps. This way of modeling the variance has several attractive features. First, it allows for flexible and at the same time parsimonious modeling of the persistence in the variance by an appropriate choice of the weights which past jumps receive in the current value of the variance. Second, since the variance is driven by jumps it has the natural ability to change quickly. Third, the linearity of the stochastic variance with respect to the jumps makes the JDSV models analytically tractable.

The paper analyzes two classes of JDSV models. In the first class, the price is comprised of a continuous component with stochastic volatility plus a time-homogeneous jump component. I refer to these models as jump–diffusion JDSV models. In the second class of models, referred to as pure-jump JDSV models, the price contains only jumps. The price jumps in this second class of models exhibit stochastic volatility. I derive moments of the return process in closed form for both classes of JDSV models.

In the empirical part of the paper I estimate semiparametric specifications of the JDSV models and compare their performance with a standard two-factor affine jump–diffusion model. The estimation of the models is based on matching moments of daily realized power variation statistics constructed from high-frequency FX data, and Monte Carlo study documents satisfactory performance of the estimator. In the estimated JDSV models the memory of the stochastic variance is modeled with a CARMA(2,1) kernel[1] and the jumps in the variance are set proportional to the squared price jumps without specifying the jump measure itself. The jump dependence in the estimated JDSV models is quite appealing — it bears analogy with GARCH models in discrete time, and at the same time it illustrates nicely the flexibility of the jump modeling in the JDSV models.

My empirical results show that the model providing best fit to the data is the jump–diffusion JDSV model for the particular

E-mail address: v-todorov@kellogg.northwestern.edu.

[1] CARMA stands for continuous-time autoregressive moving average. As later explained in the paper, the choice of CARMA(2,1) kernel in the current setting produces dynamics of the stochastic variance analogous to that of the traditional two-factor stochastic volatility models.

0304-4076/$ – see front matter
doi:10.1016/j.jeconom.2010.03.009

jump specification. This model is found to outperform the standard two-factor affine jump–diffusion model. On the other hand the estimation results show that the pure-jump JDSV model in which variance jumps are proportional to the squared price jumps does not fit the high-frequency data. Intuitively, the two-factor affine jump–diffusion model cannot generate enough volatility in the stochastic variance. On the other hand, the pure-jump JDSV model generates too much volatility of the stochastic variance to be consistent with the value of the fourth power variation observed in the data. Finally, the estimation results provide empirical support for the CARMA Lévy-driven modeling of the volatility, which was recently analyzed theoretically in Brockwell (2001).

I proceed with a short comparison of the current paper with the related literature. Barndorff-Nielsen and Shephard (2001) were the first to propose a model in which the stochastic variance is purely jump-driven. Their Non-Gaussian OU-SV model is nested in the jump–diffusion JDSV class defined here. In the Non-Gaussian OU-SV model the memory function is exponential and the jumps in the price are proportional to the jumps in the variance. Brockwell (2001) extended the Non-Gaussian OU-SV model by considering CARMA memory functions.[2] The jump-driven JDSV class that I propose here improves/generalizes these models in the following directions: (i) the memory function is allowed to be completely general (subject to integrability conditions), (ii) the dependence between the jumps in the price and the variance encompasses all possible cases (and not just that of a perfect linear dependence that was previously considered), (iii) the activity of the price jumps is unrestricted (it can be even of infinite variation).

While the first class of JDSV models generalizes a particular model that was already available in the literature, the second class, i.e. the pure-jump JDSV class, has no known precedents. The closest model is the pure-jump COGARCH of Klüppelberg et al. (2004) — it is comparable with a pure-jump JDSV model in which the jumps in the variance are proportional to the squared price jumps. Common for both models is that the time-variation in the price jumps is introduced through time-variation in the jump size.[3] The difference is that in the COGARCH model the stochastic variance is driven by the price jumps which are time-varying, while in the jump-driven JDSV models the driving jumps in the variance specification are time-homogeneous.

Turning to the estimation, the current work is naturally related to previous studies which derive in closed form moments of returns associated with continuous-time stochastic volatility models. Das and Sundaram (1999) derive unconditional moments of the Heston's model. Pan (2002) derives joint moments of returns and spot volatility in one-factor affine jump–diffusion model. She uses these moments for joint inference in an implied-state GMM estimation based on spot and option data. Meddahi (2002) derives both conditional and unconditional moments of general eigenfunction stochastic volatility models, which nest most of the diffusive-volatility continuous-time models used in finance. All these studies consider estimation at low frequency.

In contrast, in this paper the estimation is based on aggregating high-frequency data in daily statistics and matching moments of these statistics. This estimation method has been analyzed theoretically in a time-homogeneous context in Ait-Sahalia (2004). Empirical applications so far have restricted attention to matching moments only of the realized variance. Bollerslev and Zhou (2002) treat the realized variance as the unobservable integrated variance in a GMM estimation of affine jump–diffusion models.[4] Barndorff-Nielsen and Shephard (2002, 2006) have used the realized variance to estimate via QMLE part of the parameters of Non-Gaussian OU-SV models and time-changed Lévy models.[5] But because Barndorff-Nielsen and Shephard (2002, 2006) use only the realized variance, they are unable to separate the continuous and discontinuous price components. In this paper I use not only the realized variance in the estimation, but also the realized fourth power variation. That enables me to separate the two price components. As the results show including the forth power variation in the estimation penalizes strongly for the omission of price jumps from the estimated model.

The rest of the paper is organized as follows. In Section 2 I introduce the JDSV models in their general form and in Sections 2.1 and 2.2 I analyze two classes of these models — the jump–diffusion and the pure-jump JDSV models respectively. Section 3 contains the empirical part of the paper. Section 3.1 gives details on the method of estimation that is used. In Sections 3.2 and 3.3 I specify the memory function and the jump dependence in the estimated JDSV models. Section 3.4 specifies the two-factor affine jump–diffusion model that is used to compare with the JDSV models. Section 3.5 contains a Monte Carlo study to assess the performance of the estimator. Finally, in Section 3.6 I discuss the estimation results. Section 4 concludes the paper and points out directions for future research. The proofs of all results in the paper are contained in Appendix available upon request.

2. Jump-Driven stochastic volatility model

The JDSV model for the logarithmic asset price $p(t)$ is specified with the following two equations

$$\begin{aligned} p(t) &= p(0) + \alpha t + \int_0^t \sigma_1(s-)\mathrm{d}W(s) \\ &\quad + \int_0^t \int_{\mathbb{R}_0^n} \sigma_2(s-)g(\mathbf{x})\tilde{\mu}(\mathrm{d}s, \mathrm{d}\mathbf{x}), \end{aligned} \tag{1}$$

$$\sigma_i^2(t) = \bar{\sigma}_{i0} + \int_{-\infty}^t \int_{\mathbb{R}_0^n} f_i(t-s)k_i(\mathbf{x})\mu(\mathrm{d}s, \mathrm{d}\mathbf{x}) \quad \text{for } i = 1, 2, \tag{2}$$

where $W(t)$ is a standard Brownian motion; μ denotes homogeneous Poisson random measure on $\mathbb{R} \times \mathbb{R}_0^n$ with compensator $\nu(\mathrm{d}s, \mathrm{d}\mathbf{x}) = \mathrm{d}sG(\mathrm{d}\mathbf{x})$ for some positive σ-finite measure $G(\cdot)$ and $\tilde{\mu}$ is the compensated version of μ, i.e. $\tilde{\mu} = \mu - \nu$; $g : \mathbb{R}_0^n \to \mathbb{R}_0$; $k_i : \mathbb{R}_0^n \to \mathbb{R}^+$, $f_i : \mathbb{R} \to \mathbb{R}^+$ and $\bar{\sigma}_{i0} \geq 0$ for $i = 1, 2$. Since $f_i(\cdot)$ and $k_i(\cdot)$ take only non-negative values the integral in (2) is non-negative. The assumption for a constant drift term in Eq. (1) is for simplicity.

The main distinctive feature of this class of models is that the state variables (when time-varying) are modeled as moving averages of past jumps. $\sigma_i^2(t)$ can be written equivalently as (with the normalization $f_i(0) = 1$)

$$\sigma_i^2(t) = \bar{\sigma}_{i0} + \sum_{s \leq t} f_i(t-s)\Delta\sigma_i^2(s).$$

From this representation it is clear that the function $f_i(\cdot)$ determines the weight which past jumps have in the current value

[2] To generate richer autocorrelation structure of the stochastic variance, Barndorff-Nielsen and Shephard (2001) propose also superpositions of Lévy-driven OU processes for modeling the stochastic variance.

[3] An alternative way of introducing stochastic volatility in pure-jump models is to introduce time-variation in the compensator of the price jumps. Examples of such models in the literature are the time-changed Lévy processes studied in Carr et al. (2003); Carr and Wu (2004) and Barndorff-Nielsen and Shephard (2006).

[4] A Similar approach is adopted by Garcia et al. (2006) in the estimation of objective and risk-neutral distributions.

[5] Roberts et al. (2004) and Griffin and Steel (2006) use MCMC techniques to estimate, based on daily data, parametric specifications of the Non-Gaussian OU-SV model.

of the state variable $\sigma_i^2(t)$. Therefore the function $f_i(\cdot)$ is referred to as the memory function or kernel. Thus, the memory functions $f_1(\cdot)$ and $f_2(\cdot)$ determine the pattern of the persistence in the stochastic variance. The specification of $\sigma_i^2(t)$ in (2) is to be contrasted with the traditional way of modeling the stochastic variance, which is done by a superposition of independent variance factors.[6] The advantage of modeling $\sigma_i^2(t)$ with a moving average of jumps is that it allows for a succinct way of generating flexible dependence in the stochastic variance without the need of introducing many factors which in general are hard to identify and price. Drawing on an analogy with time series modeling in discrete time, a natural choice for $f_i(\cdot)$ is a CARMA kernel studied in Brockwell (2001). Intuitively, with a CARMA kernel for $f_i(\cdot)$, $\sigma_i^2(t)$ is a continuous-time analogue of the discrete-time ARMA process with non-Gaussian innovations. However, other kernels which satisfy certain integrability conditions (given below) could also be used.

In addition to the convenient way of generating persistence in the volatility, the modeling of $\sigma_i^2(t)$ as a moving average of positive jumps has the natural ability of generating sudden big moves in the variance. As already mentioned in the introduction, this is found empirically to be an important characteristic of the variance. Thus, these two important features of the variance, persistence and ability to change quickly, can be easily captured by the JDSV models. Importantly, this is achieved without nonlinear transformations of underlying stochastic processes (like in the standard log-normal stochastic volatility model for example). This makes possible deriving moving average type representation for the integrated state variables (given in Eq. (4) below), which is a key for many of the subsequent results.

The modeling of the jumps in the JDSV models provides a very convenient framework for describing the dependence between the jumps in the price and in the state variables. All jumps in the JDSV models are written as different functions of the jumps associated with a common Poisson measure, which is of arbitrary big dimension. Thus, for example, independence between the jumps in the price and in the state variables can be generated by specifying $\mathbf{x}$ to be multidimensional with independent marginals and letting the jump functions $k_i(\mathbf{x})$ and $g(\mathbf{x})$ depend on different elements of the vector $\mathbf{x}$. On the other extreme, perfect linear dependence between the jumps can be generated by making $k_i(\mathbf{x}) \propto g(\mathbf{x})$. The empirical section further illustrates the convenience of this way of modeling the jumps. It should be noted that the dependence between the price jumps and the jumps in the state variables is the only way through which the so-called "leverage effect" can be generated in the JDSV models. Therefore, in these models modeling of the jump dependence is at the same time modeling of the "leverage effect".

I continue with fixing some notation associated with the JDSV model, which will be used throughout. I denote the continuously-compounded return over the period $(t, t+a]$ with $r_a(t) = p(t+a) - p(t)$. The quadratic variation of the price process over the period $(t, t+a]$ is given by

$$\begin{aligned}
[p,p]_{(t,t+a]} &= \int_t^{t+a} \sigma_1^2(s)\mathrm{d}s + \int_t^{t+a}\int_{R_0^n} \sigma_2^2(s) g^2(\mathbf{x})\mu(\mathrm{d}s,\mathrm{d}\mathbf{x}) \\
&= \int_t^{t+a} \sigma_1^2(s)ds + \int_t^{t+a} \sigma_2^2(s)\mathrm{d}s \int_{R_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) \\
&\quad + \int_t^{t+a}\int_{R_0^n} \sigma_2^2(s) g^2(\mathbf{x})\tilde{\mu}(\mathrm{d}s,\mathrm{d}\mathbf{x}). \qquad (3)
\end{aligned}$$

An easy application of Fubini's theorem gives the following representation for the integrated state variables over the period $(t, t+a]$[7]

$$\begin{aligned}
IV_a^i(t) &:= a\bar{\sigma}_{i0} + \int_t^{t+a} \sigma_i^2(s)\mathrm{d}s = a\bar{\sigma}_{i0} \\
&\quad + \int_{-\infty}^{t+a}\int_{\mathbb{R}_0^n} H_i^a(t,s)k_i(\mathbf{x})\mu(\mathrm{d}s,\mathrm{d}\mathbf{x}), \qquad (4)
\end{aligned}$$

where

$$H_i^a(t,s) = \begin{cases} \int_t^{t+a} f_i(z-s)\mathrm{d}z & \text{if } s < t \\ \int_s^{t+a} f_i(z-s)\mathrm{d}z & \text{if } t \le s < t+a. \end{cases} \qquad (5)$$

Eq. (4) shows that, like $\sigma_i^2(t)$, the integrated quantity $IV_a^i(t)$ is a weighted sum of past jumps. The only difference between $\sigma_i^2(t)$ and $IV_a^i(t)$ is in the weights which past jumps receive.

In the next two subsections I study two classes of the JDSV models in which either $\sigma_1^2(t)$ or $\sigma_2^2(t)$ is equal to a constant. Therefore, to simplify notation, I remove the index i of the non-constant state variable $\sigma_i^2(t)$ in these models[8] and refer to $\sigma^2(t)$ and $\int_t^{t+a}\sigma^2(s)ds$ as stochastic variance and integrated variance respectively.

2.1. Jump–diffusion JDSV model

The jump–diffusion JDSV model is a special case of the JDSV model in which the price jumps do not exhibit time variation, i.e. $\sigma_2^2(t) = 1$[9] and thus its dynamics is given by

$$p(t) = p(0) + \alpha t + \int_0^t \sigma(s-)\mathrm{d}W(s) + \int_0^t\int_{\mathbb{R}_0^n} g(\mathbf{x})\tilde{\mu}(\mathrm{d}s,\mathrm{d}\mathbf{x})$$

$$\text{with } \sigma^2(t) = \int_{-\infty}^t\int_{\mathbb{R}_0^n} f(t-s)k(\mathbf{x})\mu(\mathrm{d}s,\mathrm{d}\mathbf{x}).$$

This way of modeling the price is "traditional" in terms of the role which price jumps play. The continuous component of the price is expected to generate most of the small moves of the price, while the jump component of the price should account for the sudden relatively bigger changes in the price. However, note that unlike most jump–diffusion models in the financial literature here the price jumps are left unrestricted with respect to their activity — they can be even of infinite variation. This is due to the modeling of the price jumps as integrals with respect to the compensated jump measure, $\tilde{\mu}$, which is adopted in this paper.

For the existence of moments of the return process we need integrability conditions, which are given in assumptions **H1**–**H4** below[10]

H1. $\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$,

H2. $\int_0^\infty f(s)\mathrm{d}s < \infty$ and $\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$,

H3. $\int_0^\infty f^2(s)\mathrm{d}s < \infty$ and $\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$,

H4. $\int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$.

In the next theorem I derive moments of the return process which will be used in the estimation.

Theorem 1 (*Moments of The Jump–Diffusion JDSV Model*)**.** *For the jump–diffusion JDSV model assume that conditions* **H1**–**H4** *are satisfied. Then we have*

---

6 Superposition of processes can be nested in the JDSV models by making the functions $f_i(\cdot)$ and $k_i(\cdot)$ conformable vectors.

7 Note that $\sigma_i^2(t)$ is defined as a pathwise integral, which is almost surely finite.

8 This applies also for all quantities associated with the process $\sigma_i^2(t)$ like $f_i(\cdot)$ and $k_i(\cdot)$.

9 And, as mentioned above, to ease notation I set $\sigma_1^2(t) = \sigma^2(t)$ and drop the index 1 from all other quantities associated with the process $\sigma_1^2(t)$.

10 Of course, these conditions imply that the integrals used in defining $p(t)$ and $\sigma^2(t)$ are square-integrable.

$$Var(r^a(t)) = a\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x}) + a\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x}), \quad (6)$$

$$\mathbb{E}(r^a(t) - \mathbb{E}(r^a(t)))^3 = 3\int_0^a H^a(0,u)\mathrm{d}u\int_{\mathbb{R}_0^n} k(\mathbf{x})g(\mathbf{x})G(\mathrm{d}\mathbf{x}) + a\int_{\mathbb{R}_0^n} g^3(\mathbf{x})G(\mathrm{d}\mathbf{x}), \quad (7)$$

$$\begin{aligned}\mathbb{E}(r^a(t) - \mathbb{E}(r^a(t)))^4 &= 3\int_{-\infty}^a (H^a(0,u))^2\mathrm{d}u\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ 3\left(a\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\right)^2 + a\int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ 6a^2\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ 6\int_0^a H^a(0,u)\mathrm{d}u\int_{\mathbb{R}_0^n} g^2(\mathbf{x})k(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ 3a^2\left(\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\right)^2 \end{aligned} \quad (8)$$

and for $h = a, 2a, 3a \dots$ we also have

$$\begin{aligned}&Cov((r_a(0) - \mathbb{E}(r_a(0)))^2, (r_a(h) - \mathbb{E}(r_a(h)))^2)\\ &= \int_0^a H^a(h,u)\mathrm{d}u\int_{\mathbb{R}_0^n} g^2(\mathbf{x})k(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ \int_{-\infty}^a H^a(h,u)H^a(0,u)\mathrm{d}u\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x}).\end{aligned} \quad (9)$$

Eq. (7) reveals the source of the skewness in the returns in this model. The first term in (7) is from the presence of leverage effect, which in this model reduces to negative linear relationship between the jumps in the price and the jumps in the stochastic variance $\sigma^2(t)$. In addition, skewness in the returns could be generated through skewness in the jump component of the price, which is the second component of Eq. (7). Turning to the covariance between the demeaned squared returns, we see from Eq. (9) that it consists of two terms. The first term is due to the link between jumps in the price and those in the variance. If the jumps in the price and the variance are independent then this term will disappear. The second term in Eq. (9) is due to the time-variation in $\sigma^2(t)$.

2.2. Pure-jump JDSV model

The pure-jump JDSV model is a special case of the JDSV model in which the price does not contain continuous component, i.e. in which $\sigma_1^2(t) = 0$[11] and therefore its dynamics is

$$p(t) = p(0) + \alpha t + \int_0^t\int_{\mathbb{R}_0^n} \sigma(s-)g(\mathbf{x})\tilde{\mu}(\mathrm{d}s, \mathrm{d}\mathbf{x})$$

$$\text{with } \sigma^2(t) = \int_{-\infty}^t\int_{\mathbb{R}_0^n} f(t-s)k(\mathbf{x})\mu(\mathrm{d}s, \mathrm{d}\mathbf{x}).$$

In contrast to the jump–diffusion JDSV model, in this model the jumps should account for big as well as very small moves in the price. Therefore, it is expected that the price jumps will exhibit paths of infinite variation, so that enough small moves in the price can be generated. The modeling of the asset price as solely driven by jumps was recently used in finance by Carr et al. (2003) and Barndorff-Nielsen and Shephard (2006). In these papers the stochastic volatility is generated by time-changing time-homogeneous (i.e. Lévy) jumps. In contrast, in the pure-jump JDSV model stochastic volatility is generated by introducing time-variation in the size of the price jumps.

Conditions similar to **H1**–**H4** are assumed to hold here as well

S1. $\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$,

S2. $\int_0^\infty f(s)\mathrm{d}s < \infty$ and $\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$,

S3. $\int_0^\infty f^2(s)\mathrm{d}s < \infty$ and $\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$,

S4. $\int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x}) < \infty$.

As in the jump–diffusion JDSV model, these conditions guarantee that the returns and $\sigma^2(t)$ are weakly stationary. Condition **S4** is needed for deriving the fourth moment of the returns as well as for the covariance of the squared returns and it guarantees the finiteness of these moments. In addition, for deriving some of the moments of the return process we need the following additional condition

S5. $\int_{\mathbb{R}_0^n} g(\mathbf{x})k(\mathbf{x})G(\mathrm{d}\mathbf{x}) = 0$ and $\int_{\mathbb{R}_0^n} g^3(\mathbf{x})G(\mathrm{d}\mathbf{x}) = 0$.

Assumptions **S1**–**S4** are not restrictive in the sense that they are the minimal assumptions needed to estimate the model by a method-of-moments type estimator. Condition **S5** on the other hand rules out linear dependence between the jumps and the variance and it is used for deriving closed form analytical expression for the fourth moment of the returns and covariance of the squared returns. Therefore, condition **S5** is restrictive as it rules out leverage effect in the model. However, note that condition **S5** does not preclude dependence between the jumps in $p(t)$ and $\sigma^2(t)$. It only rules out linear dependence.

I finish this section with a theorem which gives moments of the return process for the pure-jump JDSV model to be used in the empirical application in Section 3.

Theorem 2 (*Moments of the Pure-Jump JDSV Model*)**.** *For the pure-jump JDSV model assume conditions* **S1**–**S5** *hold. Then we have*

$$Var(r_a(t)) = a\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x}), \quad (10)$$

$$\begin{aligned}&E(r^a(t) - \mathbb{E}(r_a(t)))^4\\ &= a\int_0^\infty f^2(s)\mathrm{d}s\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ a\left(\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\right)^2\int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ 6\int_0^a\int_0^s f(s-u)\mathrm{d}u\mathrm{d}s\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &\times\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\int_{\mathbb{R}_0^n} g^2(\mathbf{x})k(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &+ 3a^2\left(\int_0^\infty f(s)\mathrm{d}s\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\right)^2\\ &+ 6\int_0^a\int_{-\infty}^u H^u(0,s)f(u-s)\mathrm{d}s\mathrm{d}u\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\\ &\times\left(\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\right)^2,\end{aligned} \quad (11)$$

and for $h = a, 2a, \dots$

$$Cov(r_a^2(t), r_a^2(t+h)) = \int_{-\infty}^a H^a(h,u)H^a(0,u)\mathrm{d}u\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$$

[11] Again, to simplify notation, I set $\sigma_2^2(t) = \sigma^2(t)$ and drop the index 2 from all quantities associated with the process $\sigma_2^2(t)$.

$$\times \left(\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})\right)^2 + \int_0^\infty f(s)\mathrm{d}s \int_0^a H^a(h,u)\mathrm{d}u$$
$$\times \int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x}) \int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) \int_{\mathbb{R}_0^n} g^2(\mathbf{x})k(\mathbf{x})G(\mathrm{d}\mathbf{x}). \quad (12)$$

3. Empirical application

The goal of this section is to estimate models in the two classes of JDSV models that were analyzed in the previous section and to compare their performance with the widely used in the finance literature affine jump–diffusion models. I start this section with a description of the method of estimation.

3.1. Realized power variation and estimation

The estimation in this paper is based on matching moments of daily returns and realized (daily) power variation. The realized p-power variation over a day t computed from high-frequency observations with length δ is defined as

$$RPV_\delta^p(t) = \sum_{i=1}^{M} |r_\delta(t+(i-1)\delta)|^p, \quad p>0, M=\lfloor 1/\delta \rfloor. \quad (13)$$

In this paper the realized power variation statistics that are used are the Realized Variance ($RV_\delta(t) := \sum_{i=1}^M |r_\delta(t+(i-1)\delta)|^2$, hereafter RV) and the Realized Fourth Power Variation ($FV_\delta(t) := \sum_{i=1}^M |r_\delta(t+(i-1)\delta)|^4$, hereafter FV). As $\delta \downarrow 0$, $RV_\delta(t)$ converges in probability to the quadratic variation over day t, while $FV_\delta(t)$ converges to the sum of the price jumps raised to the power four (over day t). Therefore, using RV and FV in the estimation we can identify the parameters controlling the variance of the continuous and discontinuous components of the price. Of course, to improve efficiency in the estimation potentially we need to consider also other realized power variation statistics. The choice of RV and FV for the estimation here is driven by the fact that these statistics can be used to disentangle continuous and discontinuous components of the price and importantly moments of RV and FV can be computed in closed form for the JDSV models which makes the estimation easy to apply. Moreover, the method of estimation used here has the advantage that it can provide an answer whether a whole class of models is appropriate in modeling asset prices without being fully parametric.

The particular moments used in the estimation of all models in this paper are the following: mean, variance and autocorrelation of RV; mean of FV and mean of fourth moment of daily returns. These moments can be computed in closed-form for the JDSV models introduced in Sections 2.1 and 2.2 using Theorems 1 and 2 respectively. For the autocorrelation of RV I use lags one, four, seven, ten as well as the sum of the autocorrelations from lag twenty till lag forty. Thus, altogether I end up with nine conditions. Adding more conditions could increase the asymptotic efficiency of the estimator, but in a finite sample this is also associated with a lower precision in estimating the optimal weighting matrix (see the Monte Carlo evidence in Andersen and Sørensen (1996)).

I proceed with further details about the estimation. In the estimation of all models I set the drift term to zero and use demeaned stock market return data. For the optimal weighting matrix I use a HAC estimator of the covariance matrix with a Bartlett kernel and a lag-length of eighty. The estimation is performed using the MCMC approach of Chernozhukov and Hong (2003) of treating the Laplace transform of the objective function as an unnormalized likelihood function and applying MCMC to the pseudo posterior. Then the parameter estimates are the resulting mode of the pseudo posterior. In the estimation I impose all restrictions on the parameters that guarantee non-negativity and stationarity of the state variables in the estimated models.[12]

To estimate the JDSV models with the method-of-moments type estimator proposed here we need to impose more structure on these models. In particular, we need to specify the memory function $f(\cdot)$ and in addition we need to model the jumps in the price and the variance. This is done in the next two subsections.

3.2. Modeling the memory of the stochastic variance

For the memory function $f(\cdot)$ in the estimation I use CARMA(2,1) kernel with two distinct real autoregressive roots. The CARMA(2,1) kernel generates the same autocorrelation in $\sigma^2(t)$ as the two-factor stochastic volatility models, which in turn are found to be successful in fitting financial asset prices. Based on previous estimation of the two-factor models, one of the autoregressive roots is expected to be slowly mean reverting, corresponding to a persistent factor in the variance. The second autoregressive root is expected to be fast mean reverting which corresponds to a less persistent factor in the variance.

The (normalized) CARMA(2,1) kernel with two distinct negative autoregressive roots is given by (see Brockwell, 2001 for details)

$$f(u) = \frac{b_0+\rho_1}{\rho_1-\rho_2}e^{\rho_1 u} + \frac{b_0+\rho_2}{\rho_2-\rho_1}e^{\rho_2 u}, \quad u \geq 0. \quad (14)$$

For this choice of $f(\cdot)$ the kernel of the integrated variance $H^a(t,s)$ in (5) becomes

$$H^a(t,s) = \begin{cases} \dfrac{b_0+\rho_1}{\rho_1-\rho_2}\dfrac{e^{\rho_1 a}-1}{\rho_1}e^{\rho_1(t-s)} \\ \quad + \dfrac{b_0+\rho_2}{\rho_2-\rho_1}\dfrac{e^{\rho_2 a}-1}{\rho_2}e^{\rho_2(t-s)} & \text{if } s<t \\ \dfrac{b_0+\rho_1}{\rho_1-\rho_2}\dfrac{e^{\rho_1(t-s+a)}-1}{\rho_1} \\ \quad + \dfrac{b_0+\rho_2}{\rho_2-\rho_1}\dfrac{e^{\rho_2(t-s+a)}-1}{\rho_2} & \text{if } t \leq s < t+a. \end{cases} \quad (15)$$

The necessary and sufficient condition for the CARMA(2,1) kernel to be non-negative is $b_0 \geq -\max\{\rho_1, \rho_2\} > 0$ (see Todorov and Tauchen, 2006). The CARMA(2,1) kernel in Eq. (14) reduces to a CARMA(1,0) kernel when $b_0 = -\min\{\rho_1, \rho_2\}$. Therefore, the results for the CARMA(2,1) kernel could be specialized for the CARMA(1,0) case. In the empirical part I estimate both JDSV models with CARMA(2,1) and CARMA(1,0) kernel. Note that the CARMA(1,0) choice for the kernel $f(\cdot)$ corresponds to the case where the stochastic variance follows a Lévy-driven OU process as in Barndorff-Nielsen and Shephard (2001).

3.3. Jump specification

In this subsection I model the jumps in both JDSV models. The approach adopted in this paper is to specify the functional form of $g(\cdot)$ and $k(\cdot)$ (and thus the jump dependence), but to leave the Poisson measure μ unspecified. Instead, in each of the models I estimate (parameterize) only cumulants associated with μ. This way the difference in the performance of the different models will not be influenced by a potentially wrong parametric choice for the distribution of the jumps, which can happen if a fully parametric approach is adopted instead.

3.3.1. The jump–diffusion JDSV model

I make the following assumption for $g(\cdot)$ and $k(\cdot)$ in the jump–diffusion JDSV model

H5. $g(\mathbf{x}) = \text{const1} \times x$ and $k(\mathbf{x}) = \text{const2} \times x^2$.

[12] The restrictions guaranteeing stationarity are given in Sections 3.2 and 3.4 below.

This assumption means that the jumps in the variance are proportional to the squared price jumps. This jump specification bears analogy with the GARCH modeling in discrete time where the conditional variance is determined by the past squared returns. It is important also to note that this specification of the jumps does not rule out jumps of infinite variation in the price (which will be the case for example if the jumps in the price are proportional to the jumps in the variance as in the Non-Gaussian OU-SV model).

In the estimation I treat as parameters the following cumulants associated with the Poisson measure μ: $m_c := \int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\frac{b_0}{\rho_1\rho_2}$, $m_d := \int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$ and $\upsilon := \int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$. The factor $\frac{b_0}{\rho_1\rho_2}$ in the expression for m_c is associated with the memory function, but makes m_c equal to the variance of the continuous component of the price and thus easier to interpret. The expression m_d is the variance of the discontinuous component. The cumulants m_c, m_d and υ are all we need to know for the measure μ in order to compute the moments used in the estimation of the jump–diffusion JDSV model (under the the jump specification **H5** of course).

3.3.2. The pure-jump JDSV model

The specification of the jumps in the pure-jump JDSV model that is used in the estimation is analogous to assumption **H5**

S6. $g(\mathbf{x}) = \text{const1} \times x$ and $k(\mathbf{x}) = \text{const2} \times x^2$.

Assumption **S6** does not rule out jumps of infinite variation, which as already mentioned is particularly important for the pure-jump models. Note also that combining **S5** and **S6** we have the implication $\int_{\mathbb{R}_0} x^3 G(dx) = 0$.

Under assumption **S6** we can see that the jumps in $\sigma^2(t)$ are proportional to the quadratic variation of the Lévy process driving the price. This is very similar to the COGARCH modeling (Klüppelberg et al., 2004; Brockwell et al., 2006) where the jumps in $\sigma^2(t)$ are proportional to the quadratic variation of the discontinuous component of the price. What makes the pure-jump JDSV model analyzed here different from the COGARCH intuitively is the fact that $\sigma^2(t)$ is an infinite moving average of the past squared Lévy jumps driving the price (i.e. under **S6** we have $k(x) \propto g^2(x)$), while in the COGARCH model $\sigma^2(t)$ is a moving average of the past squared price jumps.

For the estimation of the pure-jump JDSV model under the assumptions **S6** and **S7**, I parameterize the following two expressions related with the Poisson random measure μ

$$m := \int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x}) \int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})\frac{b_0}{\rho_1\rho_2}$$

and

$$\upsilon := \sqrt{\int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x}) \int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})}.$$

Similar to the jump–diffusion JDSV model, the scaling of m by the factor $\frac{b_0}{\rho_1\rho_2}$ makes it equal to the variance of the return over a unit interval and thus much easier to interpret. Given **S6**, m and υ synthesize all information about μ needed in the estimation (with the method-of-moment type estimator specified in Section 3.1).

3.4. A two-factor affine jump–diffusion model

To compare the performance of the JDSV models I estimate also a standard one and two-factor affine jump–diffusion models.[13] The realized variance in the two-factor case has the same autocorrelation structure as in the JDSV models analyzed here for the choice of the CARMA(2,1) kernel. The two-factor affine jump–diffusion stochastic volatility model is given by

$$\mathrm{d}p(t) = \alpha\mathrm{d}t + \sqrt{V(t)}\mathrm{d}W(t) + \int_{\mathbb{R}_0^n} g(\mathbf{x})\tilde{\mu}(\mathrm{d}s, \mathrm{d}\mathbf{x}), \tag{16}$$

$$V(t) = V_1(t) + V_2(t),$$

$$\mathrm{d}V_i(t) = \kappa_i(\theta_i - V_i(t))\mathrm{d}t + \sigma_i\sqrt{V_i(t)}\mathrm{d}B_i(t), \quad i = 1, 2, \tag{17}$$

where $W(t)$, $B_1(t)$ and $B_2(t)$ are independent standard Brownian motions[14]; the function $g(\cdot)$ and the compensated random measure $\tilde{\mu}$ are as introduced in the JDSV models.

Like the jump–diffusion JDSV model, the jumps in the affine jump–diffusion model (16)–(17) are time-homogeneous[15]: $\Delta p(t) = g(\mathbf{x})$. The two variance factors follow square-root diffusion processes and take non-negative values. Note that unlike the JDSV models the stochastic volatility model (16)–(17) does not have jumps in the variance, i.e. $\Delta V(t) = 0$.

For the estimation of the affine jump–diffusion models, I parameterize the fourth cumulant of the jumps in the price. That is I estimate $\upsilon := \int_{\mathbb{R}_0^n} g^4(\mathbf{x})G(\mathrm{d}\mathbf{x})$. In addition, instead of working with σ_1 and σ_2, I estimate $\sigma_{1\upsilon} := \sigma_1\sqrt{\frac{\theta_1}{2\kappa_1}}$ and $\sigma_{2\upsilon} := \sigma_2\sqrt{\frac{\theta_2}{2\kappa_2}}$, which are the standard deviations of the two variance factors. Further, I set $\theta := \theta_1 + \theta_2 + \int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$ and estimate θ, since θ_1, θ_2 and $\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$ are not identified separately (with the moment conditions used here). Finally, to guarantee that the estimated parameters correspond to stationary variance factors V_1 and V_2, I impose the following restriction on $\sigma_{1\upsilon}$ and $\sigma_{2\upsilon}$[16]

$$\sigma_{1\upsilon} + \sigma_{2\upsilon} < \theta. \tag{18}$$

3.5. Monte Carlo study

Before estimating the different models using real data I conduct a Monte Carlo study to assess the finite sample properties of the estimation method. The model estimated in the Monte Carlo is the jump–diffusion JDSV model with CARMA(1,0) kernel and jump specification **H5**.[17] Table 1 contains details on the particular parametric distribution for the jumps used in the Monte Carlo as well as the parameter values in the different cases. Since the estimated model is an "one-factor" type model, I drop from the moment vector (given in Section 3.1) two of the moment conditions which identify the memory of the stochastic variance. These moments are the autocorrelation of RV at lag ten and the sum of the autocorrelations from lag twenty till lag forty. The estimated parameters are ρ_1, $-\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1$, $\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$ and $\sqrt{-0.5\int_{\mathbf{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1}$.[18] In the Monte Carlo I analyze

13 The estimation method is the same as the one used for the JDSV models and was discussed in Section 3.1.

14 The assumption for the independence of the Brownian motions allows for closed-form expressions for the moments needed in the estimation in the general two-factor case.

15 Of course in the general affine jump–diffusion models, the jumps could have time-varying intensity (but it should be affine in the factors), see e.g. (Duffie et al., 2003) for details.

16 Condition (18) follows from the restriction guaranteeing stationarity of a square-root process.

17 The reason for using a model in the jump–diffusion JDSV class in the Monte Carlo is that models in this class are more challenging to estimate as the estimation involves separation of continuous and discontinuous components. Moreover the estimation results later in this section show that the jump–diffusion JDSV models outperform the pure-jump JDSV models when variance jumps are proportional to squared price jumps.

18 The last parameter is the volatility of $\sigma^2(t)$, but only for the CARMA(1,0) case and this is why this transformation of $\int_{\mathbf{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$ is not used when reporting the estimation results for the JDSV models in Section 3.6.

Table 1
Details on Monte Carlo. $\mathrm{d}p(t) = \sigma(t-)\mathrm{d}W(t) + \int_{\mathbb{R}_0} g(x)\tilde{\mu}(\mathrm{d}t, \mathrm{d}x)$, $\sigma^2(t) = \int_{-\infty}^{t}\int_{\mathbb{R}_0} e^{\rho(t-s)}k(x)\mu(\mathrm{d}t, \mathrm{d}x)g(x) = k_1x$, $k(x) = k_2x^2$, $G(\mathrm{d}x) = c\frac{e^{-\lambda|x|}}{|x|^{1+\alpha}}\mathrm{d}x$.

Case	Parameter values						Implied moments	
	α	λ	c	k_1	k_2	ρ	$\mathbb{E}\left([p,p]_{(t,t+1]}\right)$	$CV\left([p,p]_{(t,t+1]}\right)$
Low persistence, low volatility	0.1	0.5785	1.0283	0.2177	0.0426	−0.10	0.5	0.5347
Low persistence, high volatility	0.1	0.1928	1.0283	0.3770	0.1279	−0.10	0.5	0.9261
High persistence, low volatility	0.1	0.2000	0.2487	0.0961	0.0025	−0.03	0.5	0.6377
High persistence, high volatility	0.1	0.0667	0.2487	0.1665	0.0075	−0.03	0.5	1.1045

Note: The Poisson measure used in the Monte Carlo is of a symmetric tilted stable process (also known as CGMY process). The simulation is done using the series representation method (see Rosiński, 2007 and Todorov and Tauchen, 2006 for details). $CV\left([p,p]_{(t,t+1]}\right)$ stands for coefficient of variation of the daily quadratic variation.

Table 2
Monte Carlo results.

Parameter	True value	Mean	RMSE	5-th percentile	Median	95-th percentile
Panel A. Low persistence, low volatility						
$-\rho_1$	0.1000	0.0998	0.0120	0.0810	0.0996	0.1191
$-\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1$	0.4500	0.4433	0.0200	0.4127	0.4430	0.4759
$\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$	0.0500	0.0499	0.0052	0.0414	0.0497	0.0590
$\sqrt{-0.5\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1}$	0.2236	0.2105	0.0252	0.1773	0.2092	0.2482
Panel B. Low persistence, high volatility						
$-\rho_1$	0.1000	0.0981	0.0127	0.0776	0.0982	0.1182
$-\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1$	0.4500	0.4348	0.0358	0.3825	0.4338	0.4910
$\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$	0.0500	0.0480	0.0065	0.0385	0.0479	0.0582
$\sqrt{-0.5\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1}$	0.3873	0.3401	0.0715	0.2568	0.3379	0.4304
Panel C. High persistence, low volatility						
$-\rho_1$	0.0300	0.0279	0.0088	0.0140	0.0279	0.0419
$-\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1$	0.4500	0.4387	0.0354	0.3870	0.4377	0.4970
$\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$	0.0500	0.0472	0.0115	0.0299	0.0471	0.0654
$\sqrt{-0.5\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})/rho_1}$	0.2236	0.1984	0.0411	0.1514	0.1955	0.2558
Panel D. High persistence, high volatility						
$-\rho_1$	0.0300	0.0275	0.0085	0.0142	0.0277	0.0409
$-\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1$	0.4500	0.4271	0.0594	0.3410	0.4227	0.5238
$\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$	0.0500	0.0453	0.0126	0.0262	0.0456	0.0640
$\sqrt{-0.5\int_{\mathbb{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1}$	0.3873	0.3150	0.1055	0.2095	0.3072	0.4495

Note: The table reports the results of the Monte Carlo experiment. The number of Monte Carlo replications is 1000 and each simulation contains 3000 days of 288 high-frequency observations.

scenarios of high and low level of volatility of the stochastic variance and of high and low persistence in the stochastic variance, which makes a total of four different cases. In all the cases I set the variance of the continuous price component to be 0.45 and that of the jump price component to be 0.05, following the non-parametric empirical findings in Andersen et al. (2007) and Huang and Tauchen (2005). Finally, to match the FX high-frequency data that I am going to use in the empirical application, I set the sample size to 3000 and the number of intraday observations to 288. The results from the Monte Carlo study are summarized in Table 2. Below I outline the key findings.

(**1**) The moments of the realized power variation statistics used in the estimation can efficiently separate the total variance into two parts — one due to the continuous price component and the other one due to the discontinuous price component. Both estimates, $-\int_{\mathbb{R}_0^n} k(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1$ and $\int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$, are slightly downward biased, more so the variance of the continuous price component. This is consistent with the downward bias in the mean of the stochastic volatility that is found in the Monte Carlo studies of Andersen and Sørensen (1996) and Bollerslev and Zhou (2002).

(**2**) The autocorrelations of the realized variance used in the estimation are able to identify ρ_1. This parameter estimate is again slightly downward biased.

(**3**) The volatility of the stochastic variance parameter, $\sqrt{-0.5\int_{\mathbf{R}_0^n} k^2(\mathbf{x})G(\mathrm{d}\mathbf{x})/\rho_1}$, is the hardest parameter to estimate. This is reflected in the relatively bigger (as compared with the other parameters) bias and RMSE. Similar conclusion is made also in Andersen and Sørensen (1996).

(**4**) Increasing either the persistence of the stochastic variance or the volatility of the stochastic variance worsens the results, i.e. the biases and the RMSE-s of all parameters increase.

Overall, the Monte Carlo study shows satisfactory performance of the estimator in empirically realistic situations.

3.6. Estimation and discussion of the results

I finally proceed with the estimation. For the empirical application I use continuously compounded five-minute returns on the German Deutsch Mark/US Dollar (DM/\$) spot exchange rate series. It spans the period from December 1, 1986 till June 30, 1999. From the data set were removed missing data, weekends, fixed holidays and similar calendar effects, with details explained in Andersen et al. (2001). The total number of days left in the data set is 3045, each of which consists of 288 five-minute continuously compounded returns. The results from the estimation of the different models are reported in Tables 3–4.

The first important question, which I try to answer using the estimation results, is whether price jumps matter for the ability of the models to fit the data. Looking at the results for the jump–diffusion JDSV model in Table 3 and the affine

Table 3
Parameter estimates for the JDSV models.

Panel A: Jump–Diffusion JDSV models	CARMA(1,0) no price jumps	CARMA(1,0) with price jumps	CARMA(2,1) no price jumps	CARMA(2,1) with price jumps
b_0			0.0609(0.0402)	0.1533(0.0634)
$-\rho_1$	0.2736(0.0458)	0.1838(0.0383)	0.0079(0.0097)	0.0221(0.0162)
$-\rho_2$			1.9516(0.2297)	1.4960(0.1845)
m_c	0.3881(0.0143)	0.2752(0.0131)	0.4147(0.0244)	0.4285(0.0268)
m_d		0.1229(0.0141)		0.0435(0.0155)
v	0.0007(0.0002)	0.0047(0.0011)	0.2080(0.1068)	0.1564(0.0485)
J test	46.3400(6)[0.0000]	32.4440(5)[0.0000]	21.2280(4)[0.0003]	5.1458(3)[0.1614]

Panel B: Pure-Jump JDSV Models	CARMA(1,0)	CARMA(2,1)
b_0		241.4600(53609.3)
$-\rho_1$	0.5680(0.1320)	0.5266(0.4335)
$-\rho_2$		3.0439(17.0097)
m	0.4296(0.0129)	0.4102(0.0164)
v	0.0447(0.0097)	0.0004(0.0929)
J test	51.2000(6)[0.0000]	45.958(4)[0.0000]

Note: The table reports the parameter estimates for the JDSV models with choices for the memory function $f(\cdot)$: CARMA(1,0) and CARMA(2,1) and jump specifications given in assumption **H5** and **S6** respectively. The data used in the estimation is $DM/\$$ exchange rate over the period from December 1 1986 till June 30 1999, for a total of 3045 daily observations, each of which consists of 288 five-minute returns. The power variation statistics were computed using the intraday five-minute returns. The asymptotic variance–covariance matrix is calculated using Bartlett weights with a lag-length of eighty. The numbers in parenthesis after the J tests are the degrees of freedom and those in square brackets are the corresponding p-values.

Table 4
Parameter estimates for the affine jump–diffusion SV models.

	One-factor no price jumps	One-factor with price jumps	Two-factor no price jumps	Two-factor with price jumps
θ	0.3881(0.0223)	0.3980(0.0136)	0.4140(0.0269)	0.4429(0.0280)
κ_1	0.2707(5.7096)	0.1872(0.0389)	0.0084(0.7818)	0.0122(0.0119)
σ_{1v}	0.0360(0.8432)	0.1561(0.0151)	0.1118(4.3350)	0.1696(0.0344)
κ_2			1.9496(0.1807)	1.4957(0.2052)
σ_{2v}			0.2299(0.0872)	0.2733(0.0626)
v		0.0278(0.0072)		0.0272(0.0085)
J test	46.2800(6)[0.0000]	32.4400(5)[0.0000]	21.1680(4)[0.0003]	6.9470(3)[0.0736]

Note: The table reports the parameter estimates for the affine jump–diffusion stochastic volatility models given in (16)–(17), applied to the $DM/\$$ exchange rate. $\theta = \theta_1 + \theta_2 + \int_{\mathbb{R}_0^n} g^2(\mathbf{x})G(\mathrm{d}\mathbf{x})$, $\sigma_{1v} = \sigma_1\sqrt{\frac{\theta_1}{2\kappa_1}}$ and $\sigma_{2v} = \sigma_2\sqrt{\frac{\theta_2}{2\kappa_2}}$. The data spans the period from December 1 1986 till June 30 1999, for a total of 3045 daily observations, each of which consists of 288 five-minute returns. The power variation statistics were computed using the intraday five-minute returns. The asymptotic variance–covariance matrix is calculated using Bartlett weights with a lag-length of eighty. The numbers in parenthesis after the J tests are the degrees of freedom and those in square brackets are the corresponding p-values.

jump–diffusion model in Table 4 one can see that inclusion of jumps in the price significantly improves the fit of the model. This holds true regardless of the choice of the memory function for the jump–diffusion JDSV model and the number of factors in the affine jump–diffusion model. Also, this conclusion for the importance of the price jumps is robust to the different ways of modeling the stochastic variance (as a sum of square-root processes or moving average of positive jumps). The big difference in the performance of the models with or without price jumps indicates that the moments used in the estimation significantly penalize for their omission. On the other hand, both estimated pure-jump JDSV models provide very bad fit to the data. Therefore, pure-jump JDSV model in which the jumps in the variance are proportional to the squares of the driving Lévy process in the price does not seem to provide good description of the data.

The second question of interest is how well is the persistence of the realized variance described by a CARMA(2,1) kernel. Comparing the performance of the JDSV models with CARMA(1,0) kernel with those with CARMA(2,1) kernel, one can see that the CARMA(2,1) kernel provides a much better fit for the autocorrelation structure of RV. The same pattern emerges when comparing the one-factor affine jump–diffusion model with the two-factor one. This finding is in line with results in most of the empirical studies of multi-factor stochastic volatility models. What is important to note here is that the CARMA(2,1) kernel does as good job as a two-factor stochastic volatility model in capturing the persistence in the realized variance without the need of introducing multiple factors. On Fig. 1 I compare the autocorrelation of the realized variance with that implied by the parameter estimates of the jump–diffusion JDSV model with CARMA(2,1) kernel, which contains price jumps. In the estimation, as already discussed in Section 3.1, I use only the first forty autocorrelations of the realized variance. Fig. 1 shows that the CARMA(2,1) kernel provides pretty good fit and matches well the observed autocorrelations even beyond lag forty.

The model that provides best fit to the moments used in the estimation is the CARMA(2,1) jump–diffusion JDSV model containing price jumps. The level of the J-statistic is well below commonly accepted critical values. In fact, the only other model which cannot be rejected at a 5% significance level is the two-factor affine jump–diffusion model with price jumps. I compare these two models closer.[19] Both of them generate the

[19] Note that these two models are non-nested. They both have the same number of parameters and are estimated with the same number of moment conditions. Therefore the difference in their J statistics can be used to compare their performance, see Andrews (1999) for example.

Table 5
Moment condition tests.

	CARMA(2,1) Jump–Diffusion JDSV model	Two-factor affine Jump–Diffusion model
Autocorrelation at lag 1	2.1654	2.5694
Autocorrelation at lag 4	2.3476	2.1448
Autocorrelation at lag 7	1.7215	1.2226
Autocorrelation at lag 10	1.3912	0.8910
Aver. autocorrelation for lags 20–40	0.0161	−1.2865
$\mathbb{E}(RV_\delta(t))$	1.8370	3.5623
$\mathbb{E}(RV_\delta^2(t))$	1.8791	3.8200
$\mathbb{E}(FV_\delta(t))$	−1.7197	7.1224
$\mathbb{E}(r_1^4(t))$	−0.1585	5.0830

Note: The table reports the diagnostic t-statistics for each of the moment conditions underlying the estimation results for the CARMA(2,1) Jump–Diffusion JDSV model and the two-factor Affine Jump–Diffusion model, both containing price jumps, reported in the last columns of Tables 3 and 4 respectively.

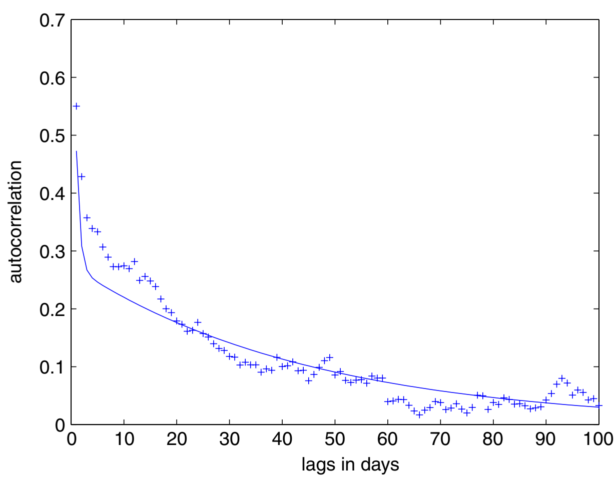

Fig. 1. The figure shows the fit of the CARMA(2,1) kernel. The empirical autocorrelation of the Realized Variance is marked with $+$. The solid line is the autocorrelation implied from the jump–diffusion JDSV model given in (1)–(2) with CARMA(2,1) memory function and jump specification given in assumption **H5**. The parameters were set at the estimated values reported in the last column of Table 3.

same autocorrelation structure. Therefore the difference in their performance comes from the jump specification and the fact that in the JDSV model the variance is solely driven by jumps while in the affine jump–diffusion model it is driven by a sum of square-root diffusion processes. Following Tauchen (1985), in Table 5 for each of the two models I report the t-statistics associated with the moments used in the estimation. These statistics should give an idea which moments the models have difficulty in matching.[20] Firstly, as expected both models produce comparable fit for the autocorrelations used in the estimation. The difference comes in the fit to the rest of the moments. As seen from Table 5, the CARMA(2,1) jump–diffusion JDSV model has no difficulty in fitting the variance of RV, the mean of FV and the fourth moment of the daily returns. In contrast, the affine jump–diffusion model does not provide a good fit to those moments. In fact, looking at the parameter estimates of the affine jump–diffusion model, reported in the last column of Table 4, it can be seen that the volatility parameters for the two square-root processes driving the stochastic variance are on the boundary of the stationarity condition (18) (in the estimation I impose this restriction). That is, the square-root diffusion processes, used for modeling the stochastic variance, could not produce enough volatility in the stochastic variance. This is in contrast with JDSV models where volatility in the stochastic variance is very easy to generate.

[20] Of course, these statistics should be interpreted with care because the inconsistency of parameter estimates in general would lead to the inconsistency of even correctly specified moment conditions.

I end this section with a short discussion on the parameter estimates for the best performing model from the ones compared here, i.e., the CARMA(2,1) jump–diffusion JDSV model with jump specification given in **H5**. The parameter estimates of this model are reported in the last column of Panel A in Table 3. The estimation results confirm our expectation regarding the autoregressive coefficients of the CARMA(2,1) kernel. The first autoregressive root has a half-life of approximately thirty days and is therefore relatively slow mean reverting. On the other hand, the second autoregressive root has a half-life of approximately half a day and thus has much faster mean reversion. Fig. 1 illustrates that this kernel does provide a good fit of the empirical autocorrelation function. Turning to the parameters controlling the jumps, it can be seen that the jump component in the price has around 9% share in the total variance of the return process. This level is similar to the proportion of jumps found in financial asset prices in the studies of Andersen et al. (2007) and Huang and Tauchen (2005). These studies use the non-parametric jump detection tests developed in Barndorff-Nielsen and Shephard (2004) to disentangle the jumps from the continuous price component. Thus, overall I conclude that the jump-driven JDSV model with CARMA(2,1) kernel containing jumps in the price with specification given in **H5** is able to fit well the moments used in the estimation and captures the main empirical features observed in the high-frequency data.

4. Conclusion

In this paper a general semiparametric class of jump-driven stochastic volatility models is introduced and their econometric properties are analyzed. These models have the distinctive feature that the state variables determining the time-variation of the continuous and discontinuous component of the price, when time-varying, are representable as moving averages of positive jumps. This allows for writing the integrated variance itself as another moving average of the same positive jumps. Using that, I derive moments of the return process and use them in estimation of the models in the JDSV class via GMM. The empirical results in the paper confirm the ability of models in the JDSV class to capture salient features of the high-frequency financial data, better than traditionally used stochastic volatility models.

Finally, there are two directions in which the analysis of the current paper should be extended. First, in the estimated models the leverage effect is either ruled out or linked in a one-to-one relationship with the skewness of the returns, which is potentially rather restrictive. Therefore, more general specifications of the jumps in the price and the variance (and consequently their dependence) need to be considered in order to capture this feature of the data. The modeling of the jumps proposed here provides convenient framework for this. For example, following an analogy with the asymmetric GARCH in discrete-time, one possibility is to set the variance jumps proportional to the squared price jumps but with coefficient of proportionality being different for the negative and positive price jumps. However, for identifying such more general jump structures in estimation, moments of other realized (multi)power variation statistics are needed and this is the second direction in which the current work should be extended.

Acknowledgements

I am grateful to the editors and two anonymous referees for their helpful comments. I would also like to thank Tim Bollerslev, Ron Gallant, Han Hong and George Tauchen for many discussions and encouragement along the way. In addition, I am indebted to Tim Bollerslev for providing me with the high-frequency FX data for the empirical application and for helpful comments on the second draft of the paper. I have benefited from suggestions made by seminar participants at the Duke Econometrics and Finance Lunch Group, the Opening Workshop of the SAMSI Program on Financial Mathematics, Statistics and Econometrics, Durham, September 2005, the Conference on Stochastics in Science in Honor of Ole Barndorff-Nielsen, Guanajuato, Mexico, March 2006 and the Conference on Realized Volatility, Montreal, April 2006.

References

Ait-Sahalia, Y., 2004. Disentangling diffusion from jumps. Journal of Financial Economics 74, 487–528.
Andersen, T., Bollerslev, T., Diebold, F., 2007. Roughing it up: Disentangling continuous and jump components in measuring, modeling and forecasting asset return volatility. Review of Economics and Statistics 89, 701–720.
Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2001. The distribution of exchange rate volatility. Journal of the American Statistical Association 96, 42–55.
Andersen, T., Sørensen, B., 1996. GMM Estimation of a stochastic volatility model: a monte casrlo study. Journal of Business and Economic Statistics 14, 328–352.
Andrews, D., 1999. Consistent moment selection procedures for generalized method of moments estimation. Econometrica 67, 543–564.
Barndorff-Nielsen, O.E., Shephard, N., 2001. Non- Gaussian Ornstein-Uhlenbeck-based models and some of their applicaions in financial economics. Journal of the Royal Statistical Society: Series B 63, 167–241.
Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realised volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society: Series B 64, 253–280.
Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics 2, 1–37.
Barndorff-Nielsen, O.E., Shephard, N., 2006. Impact of jumps on returns and realised variances: econometric analysis of time-deformed Lévy processes. Journal of Econometrics 131, 217–252.
Bollerslev, T., Zhou, H., 2002. Estimating stochastic volatility diffusionusing conditional moments of integrated volatility. Journal of Econometrics 109, 33–65.
Brockwell, P., 2001. Lévy-driven CARMA processes. Annals of the Institute of Statistical Mathematics 53, 113–124.
Brockwell, P., Chadraa, E., Lindner, A., 2006. Continuous time GARCH processes. Annals of Applied Probability 16, 790–826.
Carr, P., Geman, H., Madan, D., Yor, M., 2003. Stochastic volatility for Lévy processes. Mathematical Finance 13, 345–382.
Carr, P., Wu, L., 2004. Time-changed Lévy processes and option pricing. Journal of Financial Economics 71, 113–141.
Chernozhukov, V., Hong, H., 2003. An MCMC approach to classical estimation. Journal of Econometrics 115, 293–346.
Das, S., Sundaram, R., 1999. Of smiles and smirks: a term structure perspective. Journal of Financial and Quantitative Analysis 34, 211–239.
Duffie, D., Filipović, D., Schachermayer, W., 2003. Affine processes and applications in Finance. Annals of Applied Probability 13 (3), 984–1053.
Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269–1300.
Garcia, R., Lewis, M., Pastorello, S., Renault, E., 2006. Estimation of objective and risk-neutral distributions based on moments of integrated volatility. Working Paper, Universite de Montreal.
Griffin, J., Steel, M., 2006. Inference with Ornstein-Uhlenbeck processes for stochastic volatility. Journal of Econometrics 134, 605–644.
Huang, X., Tauchen, G., 2005. The relative contributions of jumps to total variance. Journal of Financial Econometrics 3, 456–499.
Klüppelberg, C., Lindner, A., Maller, R., 2004. A continuous time GARCH process driven by a Lévy process: stationarity and second order behavior. Journal of Applied Probability 41, 601–622.
Meddahi, N., 2002. Theoretical comparision between integrated and realized volatility. Journal of Applied Econometrics 17, 479–508.
Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics 63, 3–50.
Roberts, G., Papaspiliopoulos, O., Dellaportas, P., 2004. Bayesian inference for non-Gaussian Ornstein–Uhlenbeck stochastic volatility processes. Journal of Royal Statistical Society B 66, 369–393.
Rosiński, J., 2007. Tempering stable processes. Stochastic Processes and Their Applications 117, 677–707.
Tauchen, G., 1985. diagnostic testing and evaluation of maximum likelihood models. Journal of Econometrics 30, 415–443.
Todorov, V., Tauchen, G., 2006. Simulation methods for Lévy -Driven CARMA stochastic volatility models. Journal of Business and Economic Statistics 24, 450–469.