

Estimating Stochastic Volatility and Jumps Using High-Frequency Data and Bayesian Methods*

Milan FIČURA (milan.ficura@vse.cz), *corresponding author*

Jiří WITZANY (jiri.witzany@vse.cz)

both authors: Faculty of Finance and Accounting, University of Economics, Prague

Abstract

We compare two approaches for estimation of stochastic volatility and jumps in the EUR/USD time series—the non-parametric power-variation approach using high-frequency returns and the parametric Bayesian approach (MCMC estimation of SVJD models) using daily returns. We have found that the estimated jump probabilities based on these two methods are surprisingly uncorrelated (using a rank correlation coefficient). This means that the two methods do not identify jumps on the same days. We further found that the non-parametrically identified jumps are in fact almost indistinguishable from the continuous price volatility at the daily frequency because they are too small. In most cases, the parametric approach using daily data does not in fact identify real jumps (i.e. discontinuous price changes) but rather only large returns caused by continuous price volatility. So if these unusually high daily returns are to be modelled, the parametric approach should be used, but if the goal is to identify the discontinuous price changes in the price evolution, the non-parametric high-frequency-based methods should be preferred. Among other results, we further found that the non-parametrically identified jumps exhibit only weak clustering (analyzed using the Hawkes process), but relatively strong size dependency. In the case of parametrically identified jumps, no clustering was present. We further found that after the beginning of 2012, the amount of jumps in the EUR/USD series greatly increased, but the results of our study still hold.

1. Introduction

Numerous studies have shown that asset price dynamics exhibit properties like fat tails of the return distribution, persistent long-term shifts in levels of volatility and abrupt changes in the levels of price and volatility (jumps). In order to tackle these empirical features, it has become common practice to decompose the variability of the price processes into two components—diffusive stochastic volatility and discontinuous price jumps (Craine *et al.*, 2000; Eraker, 2004; Witzany, 2013). This has proved to be useful in many financial applications including option pricing (Eraker *et al.*, 2003; Eraker, 2004; Fulop, Li and Yu, 2015), VaR estimation (Witzany, 2013), volatility forecasting (Andersen *et al.*, 2007; Lanne, 2006; Corsi *et al.*, 2010) and volatility risk premium modeling (Todorov, 2010; Chen and Poon, 2013; Fičura, 2014). The correct identification and modeling of asset price jumps may also provide profitable signals for quantitative trading (Novotny *et al.*, 2015).

There are currently two main approaches to decomposing the variability of the price process into its continuous and discontinuous component—a parametric

* This research was supported by Czech Science Foundation Grant P402/12/G097—Dynamical Models in Economics

approach based on the Bayesian estimation methods and a non-parametric approach based on the asymptotic theory of power variations and high-frequency data. Combined approaches using Bayesian methods together with high-frequency power-variation estimators have also been proposed (see Fičura and Witzany, 2015, and Maneesoonthorn *et al.*, 2016), greatly aiding the modeling of the self-exciting jump dynamics.

In the parametric approach, it is necessary to first define the underlying processes governing the evolution of the price, i.e. typically the processes of log returns and stochastic volatility, as well as the processes governing the occurrences, magnitudes and intensity of jumps in price and volatility. The parameters of the resulting SJVD (Stochastic-Volatility-Jump-Diffusion) models (see Shephard, 2005) and the evolution of the latent state variables (i.e. the unobservable processes) are then estimated usually through simulation-based Bayesian inference methods such as the MCMC (Markov chain Monte Carlo) algorithm (for different implementations, see Eraker *et al.*, 2003; Johannes and Polson, 2009; Nakajima, 2012; and Witzany, 2013), possibly in combination with Particle filters for the sequential estimation of the latent state variables (see, Stroud and Johannes, 2014 and Fulop, Li and Yu, 2015). An approach to sequential estimation of the parameters of SVJD models was further proposed by Golightly (2014). Among other applications, Liu and Li (2015) use Bayesian methods to test for persistence of stochastic volatility in models with jumps, while Szerszen (2009) uses them to estimate stochastic volatility models with Lévy α -stable jumps.

The non-parametric approach, on the other hand, does not define the parametric processes explicitly and it should theoretically be valid for a wide range of possible processes governing price evolution (i.e. it is model-free). This approach uses high-frequency returns and the asymptotic theory of power-variation measures (derived in Andersen *et al.*, 2003, and Barndorff-Nielsen and Shephard, 2004) to calculate quantities that converge (with increasing frequency for their calculation) either to the quadratic variation, which is a measure of the overall price variability, or to the integrated variance, which is a measure of the diffusive price variability. The most commonly used measures converging to the quadratic variance are the realized variance (Andersen and Bollerslev, 1998) and the realized kernels (Barndorff-Nielsen *et al.*, 2008). Proposed measures converging to the integrated variance are, among others, the bipower variation (Barndorff-Nielsen and Shephard, 2004), threshold bipower variation (Corsi, Pirino and Reno, 2010), multi-power variation measures (see Ysusi, 2006, and Shi, 2009) and the nearest-neighbor truncation measures (Andersen *et al.*, 2010). In order to estimate the statistically significant jumps in a practical setting (with non-continuously sampled returns), an approach is commonly used in which the difference between the realized variance and bipower variation is normalized with the use of integrated quarticity. For a comparison of different approaches, see Barndorff-Nielsen and Shephard (2004); Andersen *et al.* (2007); Corsi, Pirino and Reno (2010); Hanousek *et al.* (2013), and Andersen *et al.* (2014). An alternative method, estimating jumps at the exact time at which they occur without integrated quarticity, was also proposed by Lee and Mykland (2008).

As there are two approaches for estimating stochastic volatility and jumps, it is natural to ask whether they provide similar results. From the previously conducted studies, it is clear that the non-parametric approach typically identifies significantly

more jumps than the parametric approach does (as mentioned in Andersen *et al.*, 2007), which is to be expected because it is far easier to identify jumps from high-frequency returns than from the daily returns commonly used in parametric model estimation (for an intraday implementation of SVJD models, see Stroud and Johannes, 2014). Nevertheless, we can still ask if the two methods at least tend to identify jumps at the same times and if the jump processes exhibit similar properties.

To answer these questions, we estimated the stochastic volatility and jumps in the past history of the EUR/USD exchange rate (in the period between 3 February 2006 and 15 April 2014) using representative methods from both approaches. Specifically, we used a log-variance SVJD model (based on a model in Witzany, 2013) with self-exciting jumps and compared it with the bipower-variation approach for volatility and jumps estimation (developed in Andersen *et al.*, 2007), where significant jumps are identified with the shrinkage estimator constructed using the realized tripower quarticity.

Surprisingly, the results of our research indicate that the two methods do not tend to identify jumps at the same times at all. The jumps identified non-parametrically from high-frequency returns are usually not large enough to be distinguishable at the daily frequency through the Bayesian method and the alleged “jumps” identified using the parametric approach are actually not real jumps in most cases (i.e. discontinuous price changes), but just large daily returns caused by diffusive volatility (or possibly jumps in volatility).

So the study shows that although the daily SVJD models may be useful for the modeling of unusually large daily returns occurring in the financial time series, in most cases these returns seem not to be caused by discontinuous price changes (jumps), which, in general, cannot be identified accurately by these models. The high-frequency-based non-parametric methods, on the other hand, seem to be able to identify discontinuous price changes relatively well. As these have, however, only little to do with the unexpectedly large daily returns, the non-parametric methods may not be able to account for the occurrences of such returns in the asset price time series. So the optimal choice of the method to be used should depend on the purpose of the application. If the discontinuous price changes are the main feature of interest (as may be the case in some market-making and quantitative trading strategies), then the high-frequency nonparametric estimators may be the appropriate method to use. If, on the other hand, the unusually high daily returns are the feature that needs to be modeled (as may be the case in option pricing, VaR estimation or other applications in which the fat tails of the daily return distribution have to be accounted for), then the daily SVJD models may be preferable.

The rest of the paper is organized as follows: In Section 2 we present the general price process and the non-parametric estimators used for the estimation of the quadratic variation, integrated variance and jump variance. In Section 3 we present our parametric SVJD model with self-exciting jumps and the Bayesian methods used for the estimation of its parameters and the latent state variables. In Section 4 we perform empirical research on the EUR/USD exchange rate and on simulated time series in order to analyze main differences between the two methods. Section 5 concludes the paper.

2. Non-Parametric Estimation of Volatility and Jumps

Let us assume that the logarithmic price of an asset follows a general Stochastic-Volatility Jump-Diffusion process defined by the following stochastic differential equation:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + j(t)dq(t) \quad (1)$$

where $p(t)$ is the logarithm of the asset price, $\mu(t)$ is the instantaneous drift rate, $\sigma(t)$ is the instantaneous volatility, $W(t)$ is a Wiener process, $j(t)$ is a process determining the size of the jumps and $q(t)$ is a counting process whose differential determines the times of jump occurrences (i.e. the jump indicator).

The logarithmic return over a given time period between $t-1$ and t can be expressed as

$$r(t) = p(t) - p(t-1) = \int_{t-1}^t \mu(\tau) d\tau + \int_{t-1}^t \sigma(\tau) dW(\tau) + \sum_{t-1 \leq \tau < t} \kappa(\tau) \quad (2)$$

where $\kappa(t) = j(t)I[q(t)=1]$, $I(\cdot)$ is the indicator function and the sum of $\kappa(t)$ measures the impact of jumps during the given period of time (from $t-1$ to t).

The total variability of the price process over a given period of time can be expressed with its quadratic variation in the following form:

$$QV(t) = \int_{t-1}^t \sigma^2(s) ds + \sum_{t-1 \leq s < t} \kappa^2(s) \quad (3)$$

where the first term—representing the continuous component of price variability—is called *Integrated Variance*, and the second term—representing the discontinuous component of price variability—is called *Jump Variance*. So we can write:

$$QV(t) = IV(t) + JV(t) \quad (4)$$

where $IV(t)$ is the integrated variance and $JV(t)$ is the jump variance.

Both of the quantities, $IV(t)$ and $JV(t)$, are unobservable and they have to be estimated. One possible approach is to parametrically define the processes $\sigma(t)$, $j(t)$ and $q(t)$ and estimate their parameters and the latent state variables using Bayesian inference methods (as will be shown in Section 3). Another approach is to use the non-parametric high-frequency power-variation estimators which converge either to $QV(t)$ or to $IV(t)$.

In our study we estimate the quadratic variation with the *realized variance* (see Andersen and Bollerslev, 1998), which is defined for a given frequency as the sum of squared returns on some higher frequency. Denoting Δ as some intraday time interval and $r(t, \Delta)$ as the logarithmic return between $t-\Delta$ and t , we can define the realized variance as

$$RV(t, \Delta) = \sum_{j=1}^{1/\Delta} r^2(t-1+j\Delta, \Delta) \quad (5)$$

and it holds that $RV(t, \Delta) \rightarrow QV(t)$ as $\Delta \rightarrow 0$.

The realized variance should theoretically provide an unbiased and consistent estimate of the underlying quadratic variation, as long as the price follows a continuous semi-martingale and the high-frequency returns are not plagued by microstructure noise (see Andersen *et al.*, 2003). In practical applications, the cumulation of the microstructure noise (bid-ask bounce, etc.) present at ultra-high frequencies (tick, minute, etc.) causes the returns to be serially autocorrelated and the $RV(t, \Delta)$ estimator to become positively biased. To cope with this problem, it is possible to use more advanced estimators of the quadratic variance, such as sub-sampling schemes (see Zhang, Mykland and Ait-Sahalia, 2005), the autocorrelation-robust estimator of Hansen and Lunde (2004), or the realized kernels of Barndorff-Nielsen and Shephard (2008). A simpler (though less precise) approach is to estimate the realized variance on slightly lower frequencies at which no serial autocorrelation is present (Andersen *et al.*, 2005). In our case, standard $RV(t, \Delta)$ is used with returns sampled at the 15-minute frequency, at which the estimated autocorrelation was negligible.

In order to estimate the continuous component of the quadratic variation—i.e. the integrated variance—it is possible to use the realized bipower variation (Barndorff-Nielsen and Shephard, 2004) defined as follows:

$$BV(t, \Delta) = \frac{\pi}{2} \sum_{j=2}^{1/\Delta} |r(t-1+j\Delta, \Delta)| |r(t-1+(j-1)\Delta, \Delta)| \quad (6)$$

and it holds that $BV(t, \Delta) \rightarrow IV(t)$ as $\Delta \rightarrow 0$.

The contribution of the jump component can then be estimated as

$$RJV(t, \Delta) = RV(t, \Delta) - BV(t, \Delta) \quad (7)$$

where $RJV(t, \Delta)$ is the realized jump variance and $RJV(t, \Delta) \rightarrow \sum_{t-1 \leq s < t} \kappa^2(s)$ as $\Delta \rightarrow 0$.

As long as we are not able to sample the absolute returns at infinitely fine frequency, the estimates of jump variance based on bipower variation are always plagued by some noise. This noise may cause the values of the estimator to be negative. The underlying integrated variance can never be negative, so it is reasonable to discard the negative values (Andersen, Bollerslev and Diebold, 2007) as follows

$$RJV(t, \Delta) = \max\{RV(t, \Delta) - BV(t, \Delta), 0\} \quad (8)$$

Consequently, in order to ensure that the estimates of the integrated variance and jump variance sum into the estimate of the quadratic variation (realized variance), it is necessary to re-estimate the integrated variance as follows:

$$EIV(t, \Delta) = RV(t, \Delta) - \max\{RV(t, \Delta) - BV(t, \Delta), 0\} \quad (9)$$

where $EIV(t, \Delta)$ is the adjusted estimate of the integrated variance.

Nevertheless, even if we discard the negative values, the noise present in the bipower variation estimator (due to the finiteness of our sampling frequency) causes the $RJV(t, \Delta)$ to indicate a non-negative jump component on almost every day. It would be reasonable to pick only the significant jumps in the presence of noise, which can be done by using the so-called shrinkage estimator for the jump contribution (see Barndorff-Nielsen and Shephard, 2004, and Andersen, Bollerslev and Diebold, 2007).

The shrinkage estimator is based on the idea that appropriately normalized differences between realized variance and bipower variation should—in the absence of jumps— asymptotically converge to the standard normal distribution. For the normalization, it is necessary to define the so-called integrated quarticity:

$$IQ(t) = \int_{t-1}^t \sigma^4(s) ds \quad (10)$$

which can be consistently estimated (even in the presence of jumps) with the realized tripower quarticity:

$$TQ(t, \Delta) = \frac{\pi^{3/2}}{4\Delta} \Gamma\left(\frac{7}{6}\right)^{-3} \sum_{j=3}^{1/\Delta} |r(t-1+j\Delta, \Delta)|^{4/3} |r(t-1+(j-1)\Delta, \Delta)|^{4/3} |r(t-1+(j-2)\Delta, \Delta)|^{4/3}$$

i.e. it holds that $TQ(t, \Delta) \rightarrow IQ(t)$ as $\Delta \rightarrow 0$.

Using $RV(t, \Delta)$, $BV(t, \Delta)$ and $TQ(t, \Delta)$ we can define a variable $Z(t, \Delta)$ which has asymptotically a standard normal distribution as long as the underlying process does not contain jumps:

$$Z(t, \Delta) = \frac{[RV(t, \Delta) - BV(t, \Delta)]RV(t, \Delta)^{-1}}{\sqrt{[(\pi/2)^2 + \pi - 5] \max\{1, TV(t, \Delta)BV(t, \Delta)^{-2}\} \Delta}} \quad (12)$$

Large values of $Z(t, \Delta)$ thus indicate that a jump occurred during the day. So the statistically significant jumps at significance level α can be identified as

$$RJV(t, \Delta) = I\{Z(t, \Delta) > \Phi(\alpha)^{-1}\} [RV(t, \Delta) - BV(t, \Delta)] \quad (13)$$

where

$I\{\cdot\}$ is the indicator function and $\Phi(\alpha)^{-1}$ is the quantile function of the standard normal distribution. It is also worth noting that for $\alpha = 0.5$, equation (13) becomes equal to equation (8) and only the negative jumps are eliminated by the estimator.

As we want the jump estimator and the integrated variance estimator to sum into the realized variance, we need to re-estimate the integrated variance as

$$EIV(t, \Delta) = RV(t, \Delta) - I\{Z(t, \Delta) > \Phi(\alpha)^{-1}\} [RV(t, \Delta) - BV(t, \Delta)] \quad (14)$$

It is worth noting that although the aforementioned jump estimation method is commonly used in the literature, it may underestimate the jump component in finite

samples, especially if consecutive jumps (i.e. jumps in two consecutive high-frequency time-periods) occur in the time series, which may indeed be probable, especially due to the self-exciting behavior of the jumps. For discussion of this issue and possible solutions, see Corsi *et al.* (2010).

An additional problem is the small sample size (with 15-minute frequency we have just 96 daily returns), which may also decrease the precision of the jump estimates. A solution would be to use a bootstrapped version of the estimator as described in Dovonon *et al.* (2014).

Nevertheless, we chose to use the basic $RJV(t, \Delta)$ estimator as we believe it should be accurate enough for the purposes of this study.

3. Bayesian Estimation of Volatility and Jumps

In order to estimate the continuous stochastic volatility and jumps using the parametric approach, we need to specify the underlying latent processes for $\mu(t)$, $\sigma(t)$, $j(t)$ and $q(t)$ in equation (1) as well as the latent process $\lambda(t)$, which determines the intensity of jumps governing the possible jump clustering. We will further present a parametric SVJD model inspired by Witzany (2013), with the addition of Self-Exciting Hawkes jumps instead of independent Poisson jumps.

In the model, we assume that the drift of the logarithmic return process is constant: $\mu(t) = \mu$. Equation (1) then changes to

$$dp(t) = \mu dt + \sigma(t)dW(t) + j(t)dq(t) \quad (15)$$

In order to model the instantaneous stochastic volatility $\sigma(t)$, we utilize the log-variance model which uses the mean-reverting Ornstein-Uhlenbeck process for the logarithm of the return variance. The process can be written as follows:

$$dh(t) = \kappa[\theta - h(t)]dt + \xi dW_V(t) \quad (16)$$

where $h(t)$ is the logarithm of the return variance, $h(t) = \ln[\sigma^2(t)]$ and $W_V(t)$, is a separate Wiener process governing the evolution of the stochastic volatility, which may be correlated with the Wiener process of the logarithmic price $W(t)$, but we will further assume no correlation, as the correlation does not seem to be present for currency markets (unlike stock markets) (see Franses and van Dijk, 2000, p. 18). Parameter θ represents the long-term level of volatility, κ determines the strength of the mean-reversion and ξ is the volatility of volatility, which is assumed to be constant. Additionally, it is possible to add jumps into the volatility process (see Eraker, 2004, and Fulop, Li and Yu, 2015), but we will omit these in the presented model.

Considering the process of jump sizes $j(t)$, we will model it as a series of normally distributed independent random variables: $j(t) \sim N(\mu_J, \sigma_J)$ (i.e. we will not include the possibility of time-varying jump magnitudes or the possible correlation between the jump magnitudes and the stochastic volatility).

Finally, the process of jump occurrences $q(t)$ will be modeled using a self-exciting Hawkes process with the exponential decay function (see Ait-Sahalia *et al.*, 2015, and Fulop, Li and Yu, 2015 for the application of the Hawkes processes for jump modeling, and Bacry *et al.*, 2015 for a detailed overview of the applications of these processes in finance). In the Hawkes process, the intensity of jumps $\lambda(t)$, defined by the relationship $\Pr[dq(t)=1] = \lambda(t)dt$, increases (by a fixed amount) every time a jump occurs and then decays exponentially back to its long-term level. The jump intensity process $\lambda(t)$ can be expressed using the following differential equation:

$$d\lambda(t) = \kappa_J [\theta_J - \lambda(t)]dt + \xi_J dq(t) \quad (17)$$

where θ_J is the long-term jump intensity, κ_J determines the rate of the exponential decay of the jump intensity towards its long-term level and ξ_J measures the immediate increase in jump intensity after a jump occurs (i.e. when $dq(t) = 1$).

By solving differential equation (17), it is possible to express the value of the jump intensity $\lambda(t)$ at any given point in time using the following relationship:

$$\lambda(t) = \theta_J + \int_{-\infty}^t \xi_J e^{-\kappa_J(t-s)} dq(s) = \theta_J + \sum_{dq(s)=1, s \leq t} \xi_J e^{-\kappa_J(t-s)} \quad (18)$$

It is worth noting that many recent studies (Eraker *et al.*, 2003; Eraker, 2004; and Fulop, Li and Yu, 2015) also add the possibility of jumps in the volatility process, which may be correlated with jumps in price. However, we do not include this feature in our model.

The final SVJD model used in our study is specified by equations (15), (16) and (17). In order to estimate the parameters of the model, we convert the model from continuous-time into discrete-time using the Euler discretization (with the notation where $t = 1, 2, \dots$ represents days). For the discretization of the jump intensity process, we have to assume that no more than one jump can occur during one day.

The discrete version of equation (15) governing the evolution of daily log-returns is

$$r(t) = \mu + \sigma(t)\varepsilon(t) + J(t)Q(t) \quad (19)$$

where $r(t)$ is the daily logarithmic return defined as $r(t) = p(t) - p(t-1)$ and $p(t)$ is the logarithm of the closing price at day t . Parameter μ represents the unconditional mean of the daily returns, $\sigma(t)$ is the daily conditional volatility, $\varepsilon(t) \sim N(0,1)$ is a standard normal random variable, $J(t) \sim N(\mu_J, \sigma_J)$ is a normally distributed random variable determining the size of the jumps and $Q(t) \sim \text{Bern}[\lambda(t)]$ is a variable following a Bernoulli process with intensity $\lambda(t)$.

The discrete version of equation (16) governing the daily conditional volatility is

$$h(t) = \alpha + \beta h(t-1) + \gamma \varepsilon_V(t) \quad (20)$$

where $h(t) = \ln[\sigma^2(t)]$ is the logarithm of the daily return variance, $\alpha = (1 - \beta)\theta$ is the long-term volatility, β is the autoregressive coefficient, γ is the volatility of volatility and $\varepsilon_V(t) \sim N(0,1)$ is a series of standard normal random variables uncorrelated with $\varepsilon(t)$.

Finally, the discrete version of equation (17) governing the jump intensity $\lambda(t)$ is

$$\lambda(t) = \alpha_J + \beta_J \lambda(t-1) + \gamma_J Q(t-1) \quad (21)$$

where $\lambda(t)$ is the jump intensity at day t , $a_J = (1 - b_J - \gamma_J)\theta_J$ determines the long-term jump intensity, β_J is the rate of the exponential decay of the jump intensity, and γ_J is the increase of jump intensity on the day following a jump occurrence.

For notational simplicity, we will further denote $V(t) = \sigma^2(t)$ as the daily conditional variance.

The final model in the discrete time setting has three equations (19, 20 and 21) with nine parameters to be estimated ($\mu, \alpha, \beta, \gamma, \theta_J, \beta_J, \gamma_J, \mu_J, \sigma_J$) and three vectors of latent state variables: **V**, **J** and **Q**.

To estimate the parameters of the model and the past values of the latent state variables, we use a Markov chain Monte Carlo (MCMC) algorithm constructed according to Witzany (2013) and based on the results in Jacquier *et al.* (2007) and Johannes and Polson (2009).

MCMC is a Bayesian estimation method that enables us to sample from high-dimensional multivariate densities by constructing a Markov chain that converges to the target joint density but uses only information about the conditional low-dimensional densities that are far easier to analytically express and sample from.

Let us assume that our goal is to estimate a vector of parameters denoted as $\Theta = (\theta_1, \dots, \theta_k)$ and that we are able to analytically express (and sample from) all of the univariate conditional densities $p(\theta_j | \theta_i, i \neq j, \text{data})$, but not the multivariate joint density $p(\Theta | \text{data})$. Then we can use the *Gibbs sampler* and construct a Markov chain that should converge to the multivariate joint density $p(\Theta | \text{data})$ and will thus generate asymptotically unbiased samples from it.

The procedure for the Gibbs sampler is as follows:

0. Assign a vector of initial values to $\Theta^0 = (\theta_1^0, \dots, \theta_k^0)$ and set $j = 0$.
1. Set $j = j + 1$.
2. Sample $\theta_1^j \sim p(\theta_1 | \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data})$
3. Sample $\theta_2^j \sim p(\theta_2 | \theta_1^j, \theta_3^{j-1}, \dots, \theta_k^{j-1}, \text{data})$
- ...
- $k + 1$. Sample $\theta_k^j \sim p(\theta_k | \theta_1^j, \theta_2^j, \dots, \theta_{k-1}^j, \text{data})$ and return to step 1.

According to the Clifford-Hammersley theorem (see Johannes and Polson, 2009) the univariate conditional distributions $p(\theta_j | \theta_i, i \neq j, \text{data})$ fully characterize the joint distribution $p(\Theta | \text{data})$ and it can also be proved that the Markov chain constructed according to the Gibbs sampler converges to the multivariate joint distribution $p(\Theta | \text{data})$. So, we only have to calculate enough iterations of the Gibbs sampler, discard the first ones and use the remaining ones to estimate all of the required statistical properties of the joint distribution $p(\Theta | \text{data})$.

The conditional densities $p(\theta_j | \theta_i, i \neq j, \text{data})$ are usually obtained by applying the Bayes theorem to the likelihood function and the prior density:

$$p(\theta_1 | \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data}) \propto L(\text{data} | \theta_1, \theta_2^{j-1}, \dots, \theta_k^{j-1}) \times \text{prior}(\theta_1 | \theta_2^{j-1}, \dots, \theta_k^{j-1}) \quad (22)$$

where $L(\cdot)$ denotes the likelihood function, $\text{prior}(\cdot)$ is the Bayesian prior density for the given parameter and \propto represents a proportionate relationship. In our application, we generally use uninformative priors (i.e. $\text{prior}(\theta_i) \propto 1$) and assume independence of the parameters.

In order to utilize the Gibbs sampler, it is necessary to normalize the right hand side (RHS) of equation (22) (i.e. to replace the proportionate relationship with equality). This can be done by integrating the RHS of the equation over θ_1 in order to get the density $p(\text{data} | \theta_2^{j-1}, \dots, \theta_k^{j-1})$ with which the RHS has to be divided.

If integration on the right hand side of equation (22) is not possible, then the Gibbs sampler cannot be used. Instead, we can use the *Metropolis-Hastings algorithm*, which is a rejection sampling algorithm that works as follows: a proposal value for each parameter is drawn from a proposal density and the new parameter value is then either accepted or rejected based on a given probability.

Specifically, Step 2 in the Gibbs sampler is replaced by the following two-step procedure:

- A. Draw θ_1^j from the proposal density $q(\theta_1 | \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data})$.
- B. Accept θ_1^j with the probability $\alpha = \min(R, 1)$, where R denotes the acceptance ratio defined as

$$R = \frac{p(\theta_1^j | \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data}) q(\theta_1^{j-1} | \theta_1^j, \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data})}{p(\theta_1^{j-1} | \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data}) q(\theta_1^j | \theta_1^{j-1}, \theta_2^{j-1}, \dots, \theta_k^{j-1}, \text{data})} \quad (23)$$

which is in practice evaluated by sampling $u \sim U(0, 1)$ from a uniform distribution and accepting θ_1^j if and only if $u < R$.

It can be shown that the resulting Markov chain converges to the joint distribution $p(\Theta | \text{data})$ (see Johannes and Polson, 2009).

There are many different versions of the Metropolis-Hastings algorithm which often differ in the proposal density assumed. A very popular version is the *Random-Walk Metropolis-Hastings* with the following proposal density:

$$\theta_1^j \sim \theta_1^{j-1} + N(0, c) \quad (24)$$

where c is a meta-parameter which may influence the computational efficiency of the algorithm and the practice is to set it so that approximately 50% of the proposals get accepted.

A great advantage of the Random-Walk Metropolis-Hastings algorithm is that its proposal distribution is symmetric, which means that the probability of going from θ_1^{j-1} to θ_1^j is the same as the probability of going from θ_1^j to θ_1^{j-1} . Because of that, the proposal densities in equation (23) cancel out. Consequently, by utilizing relationship (22) and assuming non-informative priors, the acceptance ratio reduces to the likelihood ratio

$$R = \frac{L(\text{data} | \theta_1^j, \theta_2^{j-1}, \dots, \theta_k^{j-1})}{L(\text{data} | \theta_1^{j-1}, \theta_2^{j-1}, \dots, \theta_k^{j-1})} \quad (25)$$

So, in theory, as long as we are able to analytically express the likelihood function, we can use this algorithm to estimate the joint posterior density of the parameters.

In our case, we want to estimate a vector of a few model parameters Θ and a large number of latent state variables X . Since we know from the Bayes theorem that

$$p(\Theta, X | \text{data}) \propto p(\text{data} | \Theta, X) * p(X, \Theta) \quad (26)$$

we can estimate iteratively the parameters and the latent state variables

$$\begin{aligned} p(\Theta | X, \text{data}) &\propto p(\text{data} | \Theta, X) * p(X | \Theta) * p(\Theta) \\ p(X | \Theta, \text{data}) &\propto p(\text{data} | \Theta, X) * p(\Theta | X) * p(X) \end{aligned} \quad (27)$$

In order to achieve better computational efficiency, we combine in our application different versions of the MCMC algorithm for different variables. Specifically, we use the Gibbs sampler to estimate the parameters $\mu, \mu_j, \sigma_j, \alpha, \beta, \gamma$ and latent state variables \mathbf{Q}, \mathbf{J} and \mathbf{V} (for \mathbf{V} using an accept-reject Gibbs sampler based on Kim, Shephard and Chib, 1998) and Random-Walk Metropolis Hastings to estimate the Hawkes process parameters $\theta_j, \beta_j, \gamma_j$.

The full estimation algorithm (based on Witzany, 2013) proceeds as follows:

1. Sample reasonable initial values $\mu^{(0)}, \mu_j^{(0)}, \sigma_j^{(0)}, \alpha^{(0)}, \beta^{(0)}, \gamma^{(0)}, \theta_j^{(0)}, \beta_j^{(0)}, \gamma_j^{(0)}, \mathbf{V}^{(0)}, \mathbf{J}^{(0)}, \mathbf{Q}^{(0)}$. Denoting s^2 as the estimate of unconditional variance $\text{Var}(r)$, the following initial values were used: $\mu^{(0)} = 0, \mu_j^{(0)} = 0, \sigma_j^{(0)} = 2 * s, \alpha^{(0)} = \log(s^2) * (1 - 0.9), \beta^{(0)} = 0.9, \gamma^{(0)} = 0.3, \theta_j^{(0)} = 0.05, \beta_j^{(0)} = 0.8, \gamma_j^{(0)} = 0.01$.

The initial stochastic variances $\mathbf{V}^{(0)}$ were set equal to the exponential moving average of r^2 and jump occurrences $\mathbf{Q}^{(0)}$ were set equal to zero.

2. For $i = 1, \dots, T$ sample jump sizes $J_i^{(g)} \propto \varphi\left(J; \mu_J^{(g-1)}, \sigma_J^{(g-1)}\right)$ if $Q_i^{(g-1)} = 0$ and $J_i^{(g)} \propto \varphi\left(r_i; \mu^{(g-1)} + J, \sqrt{V_i^{(g-1)}}\right) \varphi\left(J; \mu_J^{(g-1)}, \sigma_J^{(g-1)}\right)$ if $Q_i^{(g-1)} = 1$.
3. For $i = 1, \dots, T$ sample jump occurrences $Q_i^{(g)} \in \{0, 1\}$, $\Pr[Q = 1] = p_1 / (p_0 + p_1)$, where $p_0 = \varphi\left(r_i; \mu^{(g-1)}, \sqrt{V_i^{(g-1)}}\right) (1 - \lambda^{(g-1)})$ and $p_1 = \varphi\left(r_i; \mu^{(g-1)} + J, \sqrt{V_i^{(g-1)}}\right) \lambda^{(g-1)}$.
4. Sample new stochastic log-variances $h_i^{(g)} = \log\left(V_i^{(g)}\right)$ for $i = 1, \dots, T$ using the Gibbs Sampler with the accept-reject procedure developed in Kim, Shephard and Chib (1998), i.e. we calculate the series $y_i = r_i - \mu^{(g-1)} - J_i^{(g)} Q_i^{(g)}$ and sample $h_i^{(g)}$ from a proposal distribution $\varphi(h_i; \mu_i, \sigma)$, where $\mu_i = \phi_i + \frac{\sigma^2}{2} [y_i^2 \exp(-\phi_i) - 1]$, $\phi_i = \frac{[\alpha(1-\beta) + \beta(\log V_{i+1} + \log V_{i-1})]}{(1+\beta^2)}$ and $\sigma = \frac{\gamma}{\sqrt{1+\beta^2}}$.

The proposal is accepted with probability f^* / g^* (otherwise a new proposal is drawn), where

$$\log f^* = -\frac{h_i}{2} - \frac{y_i^2}{2} [\exp(-h_i)] \quad \text{and} \quad \log g^* = -\frac{h_i}{2} - \frac{y_i^2}{2} [\exp(-\phi_i)(1+\phi_i) - h_i \exp(-\phi_i)].$$

5. Sample new stochastic volatility autoregression coefficients $\alpha^{(g)}, \beta^{(g)}, \gamma^{(g)}$ from $h_i = \log(V_i^{(g)})$ for $i = 1, \dots, T$ using the Bayesian linear regression model (Lynch, 2007), i.e.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}\mathbf{y}, \quad \hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}, \quad \text{where} \quad \mathbf{X} = \begin{pmatrix} 1 & \dots & 1 \\ h_1 & \dots & h_{T-1} \end{pmatrix}' \quad \text{and} \quad \mathbf{y} = (h_2 \dots h_T)',$$

so we sample $(\gamma^{(g)})^2 \propto IG\left(\frac{n-2}{2}, \frac{\hat{\mathbf{e}}'\hat{\mathbf{e}}}{2}\right)$ and $(\alpha^{(g)}, \beta^{(g)})' \propto \varphi\left[(\alpha, \beta)'; \hat{\boldsymbol{\beta}}, (\gamma^{(g)})^2 (\mathbf{X}'\mathbf{X})^{-1}\right]$.

6. Sample $\mu^{(g)}$ based on the normally distributed time series $r_i - J_i^{(g)} Q_i^{(g)}$ with variances $V_i^{(g)}$: $p\left(\mu^{(g)} \mid \mathbf{r}, \mathbf{J}^{(g)}, \mathbf{Q}^{(g)}, \mathbf{V}^{(g)}\right) \propto \varphi\left(\mu; \sum_{i=1}^T \frac{r_i - J_i^{(g)} Q_i^{(g)}}{V_i^{(g)}} \middle/ \sum_{i=1}^T \frac{1}{V_i^{(g)}}, \sum_{i=1}^T \frac{1}{V_i^{(g)}}\right)$.
7. Sample $\theta_J, \beta_J, \gamma_J$ using Random-Walk Metropolis-Hastings and the likelihood function $L\left(\mathbf{Q}^{(g)} \mid \theta_J, \beta_J, \gamma_J\right) = \prod_{i=1}^T \lambda_i^{Q_i} (1 - \lambda_i)^{1-Q_i}$.

8. Sample $\mu_j^{(g)}, \sigma_j^{(g)}$ based on the normally distributed series $\mathbf{J}^{(g)}$ and uninformative priors $p(\mu) \propto 1$ and $p(\log \sigma^2) \propto 1$ (which is equivalent to $p(\sigma^2) \propto 1/\sigma^2$), i.e. we sample from

$$p\left(\mu_j^{(g)} \mid \mathbf{J}^{(g)}, \sigma_j^{(g-1)}\right) \propto \varphi\left(\mu_j^{(g)}; \frac{\sum_{i=1}^T J_i^{(g)}}{T}, \frac{\sigma_j^{(g-1)}}{\sqrt{T}}\right),$$

$$p\left[\left(\sigma_j^{(g)}\right)^2 \mid \mathbf{J}^{(g)}, \mu_j^{(g)}\right] \propto IG\left[\left(\sigma_j^{(g)}\right)^2; \frac{T}{2}, \frac{\sum_{i=1}^T \left(J_i^{(g)} - \mu_j^{(g)}\right)^2}{2}\right].$$

4. Empirical Results

The empirical research was performed on a time series of the EUR/USD exchange rate consisting of 2,083 trading days over the period between 3 February 2006 and 15 April 2014. All of the non-parametric high-frequency estimators were calculated on 15-minute returns provided by Forexhistorydatabase.com. The data correspond to the GMT+2 time-zone and consist of ninety-six 15-minute returns for each trading day (Monday to Friday). We further removed all days with less than twenty 15-minute returns from the sample, as these were viewed as non-standard and, due to the small sample sizes, could potentially cause bad behavior of the non-parametric estimators. All of the calculations were performed in Matlab.

Before applying the models to real price data, we compared the performance of the two methods in simulated time series. We simulated the series of price evolution using our SVJD model with parameters calibrated to the EUR/USD history (equal to the third row of *Table 5*), with the modification that we put the mean jump size parameter equal to zero ($\mu_j = 0$) and simulated the series in 12 different variants with different values of σ_j (i.e. the jump volatility parameter). The simulation was performed with 15-minute time steps and the length of the simulated series was set to 5,000 trading days. The purpose of the performed simulation was to assess the ability of the two analyzed methods (i.e. non-parametric estimation and the Bayesian approach) to identify jumps of different magnitudes.

The performance of the two jump estimators was compared using the accuracy ratio (AR) which can be defined as follows. If i is a random day on which a jump occurred and j a random day on which a jump did not occur, p_1 denotes the probability that the values of the jump estimator (JE) are $JE(i) > JE(j)$ and p_2 the probability that $JE(i) < JE(j)$, then the accuracy ratio can be defined as $AR = p_1 - p_2$.

As can be seen in *Figure 1*, the non-parametric power-variation approach is able to identify simulated jumps with far greater accuracy than the Bayesian approach, especially if the magnitude of the simulated jumps is relatively low. This is to be expected, as it is far easier to identify small discontinuities in the price evolution from high-frequency returns than from the daily returns to which

Figure 1 Accuracy Ratios for the Parametric and Non-Parametric Jump Estimators in Simulated Time Series with Different Values of SigmaJ

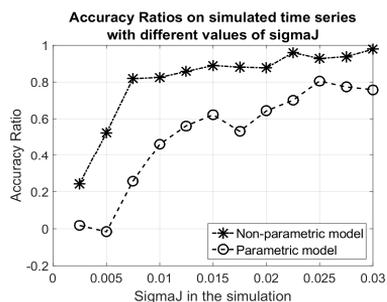


Table 1 Basic Statistics for the Realized Variance and Bipower Variation Series

	mean	st.dev.	skew	kurt	JB-test	acf(1)	LB-test
RV	4.14E-05	4.60E-05	4.58	37.57	111014.48	0.746	15410.05
BV	3.72E-05	4.17E-05	4.23	30.96	74071.31	0.785	17128.07
ln(RV)	-10.4376	0.7922	0.34	3.24	43.98	0.762	19601.80
ln(BV)	-10.5535	0.8002	0.38	3.19	54.00	0.792	20793.67
RV^(1/2)	0.0059	0.0026	1.89	9.06	4426.47	0.786	19555.14
BV^(1/2)	0.0056	0.0025	1.86	8.49	3815.76	0.815	20825.31

the Bayesian approach is applied. The main contribution of the simulation is to tell us the degree of accuracy that we can expect from the two methods based on the sizes of the relevant jumps. As we can see, for jumps with sigma smaller or equal to 0.5%, the accuracy ratio of the Bayesian approach is virtually equal to zero, indicating that the method cannot distinguish jump days from non-jump days if the volatility of the jumps is too small.

Knowing the simulation performance of the two methods, we applied them to the empirical EUR/USD log returns. *Table 1* shows the main summary statistics of realized variance and bipower variation as well as of their logarithmic and root transformations. For the Jarque-Bera test and the Ljung-Box test with lag 20, we report only the test statistic for comparison. All of the series exhibit strong, statistically significant autocorrelation and non-normality. The non-normality is least pronounced in the case of the logarithmically transformed series, which supports the application of our log-variance model in the parametric approach.

In the next step, we estimated the contribution of the jump component to the quadratic variation using equation (13) with probability levels (alphas) of 50%, 90%, 95%, 99%, 99.9% and 99.99%. *Table 2* shows some basic statistics of the identified jumps.

From the first and second column we can see that jumps are clearly present in the time series. The last column (m.jump) shows the estimated mean absolute jump size under the assumption that, at most, one jump can happen during one day. It is calculated as the mean of the square roots of the estimated jump component. The value of m.jump is relatively low, meaning that the jumps are rather small and it

Table 2 Statistics of Jumps Identified Non-Parametrically with Different Alphas

	number	rate	mean	st.dev.	skew	kurt	JB-test	m.jump
JV_50%	1620	77.77%	5.96E-06	9.41E-06	6.114	62.765	251194.00	0.0021
JV_90%	812	38.98%	9.42E-06	1.19E-05	5.121	42.320	55856.44	0.0028
JV_95%	596	28.61%	1.05E-05	1.23E-05	4.758	37.629	32027.99	0.0029
JV_99%	330	15.84%	1.27E-05	1.47E-05	4.374	30.062	11122.59	0.0032
JV_99.9%	132	6.34%	1.51E-05	1.30E-05	2.678	12.737	679.27	0.0036
JV_99.99%	63	3.02%	1.83E-05	1.47E-05	2.591	11.246	248.97	0.0040

Table 3 Autocorrelations of the Magnitudes of the Estimated Jump Components

	JV_50%	JV_90%	JV_95%	JV_99%	JV_99.9%	JV_99.99%
acf(1)	0.198	0.332	0.264	0.223	0.039	-0.246
LB-test	919.942	985.041	514.019	228.486	8.826	25.331
LB-pval	0	0	0	0	0.985	0.189

Table 4 Parameters of the Hawkes Process Applied to the Jumps Identified Non-Parametrically with Different Alphas (i.e. Probability Levels)

	ThetaJ		BetaJ		GammaJ	
JV_90%	0.4014	(0.0297)	0.9937	(0.0041)	0.0041	(0.0024)
JV_95%	0.2861	(0.0107)	0.6175	(0.3722)	0.0286	(0.0206)
JV_99%	0.1582	(0.0100)	0.8746	(0.0784)	0.0250	(0.0134)
JV_99.9%	0.0643	(0.0059)	0.8159	(0.5456)	0.0136	(0.0150)
JV_99.99%	0.0302	(0.0038)	0.8195	(1.0868)	0.0008	(0.0137)

may be difficult to identify them accurately (especially using the Bayesian approach, which exhibited close to zero predictive accuracy in the simulations when the volatility of the jumps was smaller than 0.5%).

We further analyze the jump size dependency by first removing the non-jump days and then calculating the autocorrelation of the resulting time series. We can see from the autocorrelation on lag one (acf(1)) and the Ljung-Box test with lag 20 that for jumps estimated with $\alpha \leq 99\%$, there exists some dependency in their absolute sizes. (see *Table 3*)

In order to analyze jump clustering, we first calculated the autocorrelation of the time series of the logarithms of the numbers of days between jump occurrences, but the results indicated no statistically significant autocorrelation.

As an alternative approach, we estimated the parameters of the Hawkes process on the non-parametrically identified jump time series (while assuming that the daily jump component is always generated by only a single jump). The estimation was done by numerically maximizing the likelihood function with the Matlab `fmincon` function. *Table 4* shows the results.

As can be seen in *Table 4*, the parameters indicate some tendency to jump clustering, as the BetaJ parameter is relatively high for all of the time series. Nevertheless, from the confidence intervals (estimated using the Hessian), we can see that

Table 5 Mean Bayesian Estimates of the Parameters of the SVJD Model

	mui	muiJ	sigmaJ	alpha	beta	gamma	thetaJ	betaJ	gammaJ
B_2_Mean	0.0002	0.0012	0.0057	-0.0450	0.9957	0.0698	0.0307	0.4133	0.0441
B_2_Std	(0.0001)	(0.0033)	(0.0011)	(0.0260)	(0.0025)	(0.0080)	(0.0255)	(0.2510)	(0.0282)
B_3_Mean	0.0001	0.0020	0.0079	-0.0475	0.9954	0.0686	0.0205	0.4414	0.0423
B_3_Std	(0.0001)	(0.0036)	(0.0015)	(0.0274)	(0.0026)	(0.0102)	(0.0194)	(0.2688)	(0.0278)
B_4_Mean	0.0002	0.0065	0.0109	-0.0401	0.9962	0.0653	0.0132	0.4231	0.0455
B_4_Std	(0.0001)	(0.0071)	(0.0023)	(0.0251)	(0.0024)	(0.0086)	(0.0133)	(0.2647)	(0.0286)

the GammaJ parameter is statistically insignificant for most of the time series (with the possible exception of JV_90% and JV_99%). We also calculated the Bayesian confidence intervals at which even the BetaJ parameter becomes insignificant.

In the parametric approach for jump identification, we estimated the posterior distributions of the parameters and latent state variables of the SVJD model described in Section 3 by using a MCMC algorithm with 20,000 iterations (of which the first 5,000 were discarded). We first performed the estimation without any prior information, which caused pathological behavior of the jump size parameters (sigmaJ converging to zero). Thus we decided to use a prior inverse gamma distribution for the sigmaJ parameter in the Gibbs sampler with the scale parameter $\alpha = 4$ and with the mean alternatively equal to 2, 3 and 4 times the unconditional standard deviations of the log-returns (denoted as B_2, B_3 and B_4), thereby forcing the possible jumps in the time series to be relatively large compared to the overall volatility. Table 5 shows the Bayesian posterior means and standard deviations of the parameter estimates.

As can be seen in Table 5, the parameter beta is very close to 1, indicating potential non-stationarity of the log-variance time series. Considering the jumps, we can see that the parameters of the Hawkes process (thetaJ, betaJ and gammaJ) do not indicate any jump clustering (betaJ is low and statistically insignificant). As for the sigmaJ parameter determining the jump volatility, its value varies between 0.57% and 1.09%, depending on the applied prior. It is thus apparent that the jumps are rather small in their absolute value and the accuracy of the Bayesian approach (regarding the identification of jumps) may be poor (as can be seen in the simulation results in Figure 1).

Figure 2 compares the Bayesian stochastic volatility estimates (for the B_3 case) with the non-parametric estimates of integrated variance. We can see that both of the series exhibit similar behavior with the stochastic variance only being less noisy.

Finally, we compare the jumps estimated using the Bayesian approach with the jumps estimated using the non-parametric power-variation approach. To identify jumps using the Bayesian approach, averages of the sampled jump occurrences (for every single day) from the last 15,000 sweeps of the MCMC algorithm were computed (Figures 3 and 4). The jump estimates thus constructed correspond to the posterior probabilities of jump occurrences for every single day. Table 6 shows the numbers of days in which the estimated probabilities that a jump occurred are greater than a given percentage (10% to 99%). The rows correspond to the three

Figure 2 Comparison of the SV (B_3) Estimate and the Integrated Variance Estimate (EIV) Based on the Differences between BV and JV_99.9%

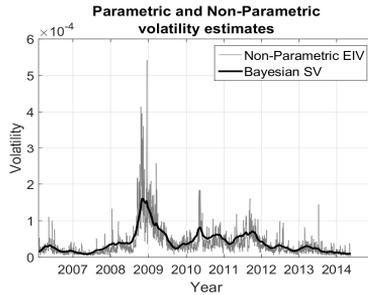


Figure 3 Bayesian Probabilities of Jump Occurrences for B_2 (left) and B_3 (right)

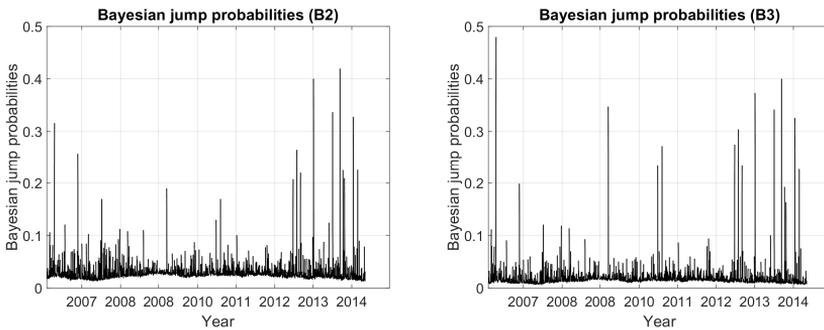
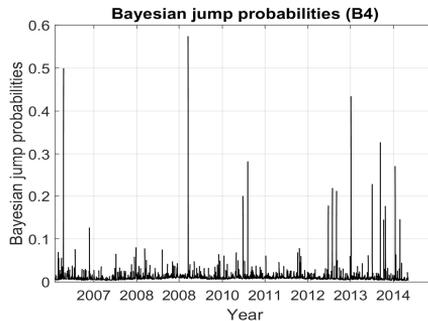


Figure 4 Bayesian Probabilities of Jump Occurrences (B_4)



MCMC runs with different priors regarding the jump volatility. As can be seen in *Table 6*, the method identified only one possible jump with Bayesian probability of occurrence greater than 50% (in the last row, i.e. the one with the highest prior jump volatility) and zero jumps at the higher confidence levels (90% to 99%). This is in stark contrast to the non-parametric approach, which identified many jumps in the time series with relatively high levels of confidence (as can be seen in *Table 2*).

Although the Bayesian approach identified a significantly smaller number of jumps in the time series than the non-parametric approach did, we can still ask

Table 6 Parametrically Identified Jumps at Different Bayesian Probability Levels (bp)

	bp > 10%	bp > 20%	bp > 30%	bp > 50%	bp > 90%	bp > 95%	bp > 99%
B_2_jumps	25	12	5	0	0	0	0
B_3_jumps	20	12	7	0	0	0	0
B_4_jumps	15	9	4	1	0	0	0

Table 7 Rank Correlations of the Estimated Jump Probabilities and with Other Variables

	bp2	bp3	bp4	Z(t,Δ)	RV	RV-BV	R2(day)	max(r2(15))
bp2	1.0000	0.9551	0.9462	0.0285	0.5382	0.2032	0.8529	0.4865
bp3	0.9551	1.0000	0.9459	0.0256	0.5584	0.2072	0.8395	0.4949
bp4	0.9462	0.9459	1.0000	0.0263	0.5795	0.2131	0.8433	0.5065
Z(t,Δ)	0.0285	0.0256	0.0263	1.0000	0.0223	0.8698	0.0365	0.2607

if the two methods tend to identify jumps at the same times (i.e. if, when one of the methods assigns a relatively high jump probability to a given day, the second method assigns a relatively high jump probability to that day as well). In order to find out, we calculated the Spearman rank correlation coefficients between the Bayesian probabilities of jump occurrences and the values of variable Z from equation (12).

As can be seen in *Table 7*, the rank correlation coefficients between the Bayesian jump probabilities (bp2, bp3 or bp4) and variable Z are very low, indicating that the two methods do not tend to identify jumps at the same times. This is strange, considering that the two methods should both identify the same phenomenon, i.e. discontinuous price changes.

In order to explain these results we calculated rank correlations of the jump estimators with some other quantities related to price volatility—namely the realized variance (RV), the difference between RV and BV ($RV - BV$), the squared daily return ($R2_Day$) and the squared maximum 15-minute return during the day ($\max(r2(15m))$). As we can see from the results, the Bayesian jump probabilities are strongly rank-correlated with the squared daily returns and somewhat correlated also with the realized variance and the maximum 15-minute returns. Conversely, the value of the Z -statistics is almost uncorrelated with the squared daily returns and surprisingly even with the daily realized variance.

As the Z -statistics exhibit correlation only with variables calculated from the intraday returns ($RV - BV$ and $\max(r2(15m))$) and virtually no correlation with the daily squared returns, it seems that the real jumps are in fact indistinguishable at the daily frequency (because they are too small to significantly influence the daily returns).

In order to examine this issue in more detail, we used the following approach: We estimated the daily integrated variance using equation (14) with different values of α and we then applied the HAR model (Corsi, 2004) to these series in order to calculate one-day predictions of the continuous price volatility. We then used these predictions to normalize the log-returns for every single day and calculated the distribution of the normalized returns on jump days and non-jump days to see the extent to which they are different.

Figure 5 Distribution of Normalized Returns on Jump Days and Non-Jump Days for Non-Parametrically Identified Jumps with $\alpha = 0.99$ (left) and $\alpha = 0.999$ (right)

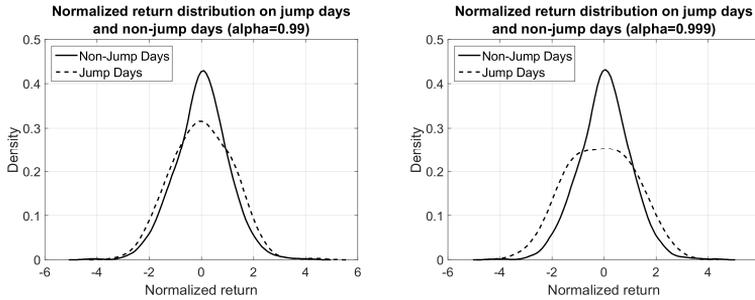
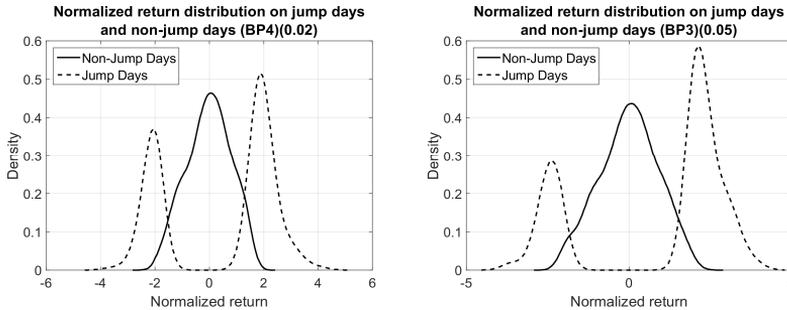


Figure 6 Distribution of Normalized Returns on Jump Days and Non-Jump Days for Parametrically Identified Jumps Using $bp \geq 0.02$ (B_4) (left) and $bp \geq 0.05$ (B_3) (right)



As we can see in *Figure 5*, on the majority of days the presence of jumps does not dramatically increase the daily returns compared to their size in the days without jumps. The vast majority of non-parametrically identified jumps are thus indistinguishable at the daily frequency.

As we can see in *Figure 6*, for the parametrically identified jumps the distribution of normalized returns on jump days is bi-modal and most of the returns on those days are in their absolute magnitude far greater than in the case of the non-jump days. This is to be expected, as the Bayesian method uses primarily the size of the daily returns in order to determine if a jump occurred. The shape of the distribution in *Figure 6* also supports the possible utilization of bi-modal jump distribution in the SVJD model (see Božovic, 2008).

Finally, we can ask if the Bayesian method really identifies jumps (i.e. discontinuous price changes) in the time series or if the alleged jumps are just large price movements caused by continuous variability in conjunction with some other phenomenon, i.e. jumps in volatility (see Eraker *et al.*, 2003) or dependencies in the conditional mean (see Stádník, 2014).

In order to answer that question, we compare the distribution of the Z -statistic on the days when the Bayesian jump probabilities are relatively high with its distribution on the days when the Bayesian jump probabilities are relatively low.

Figure 7 Distribution of the Values of the Z-Statistic on Days when $bp \geq 0.02$ and $bp < 0.02$ for B_4 (left) and on Days when $bp \geq 0.05$ and $bp < 0.05$ for B_3 (right)

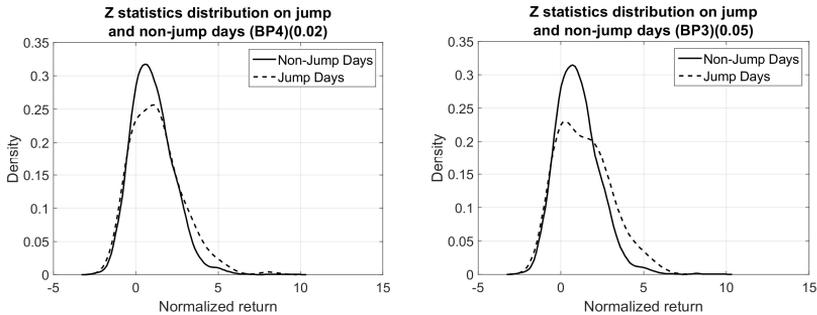
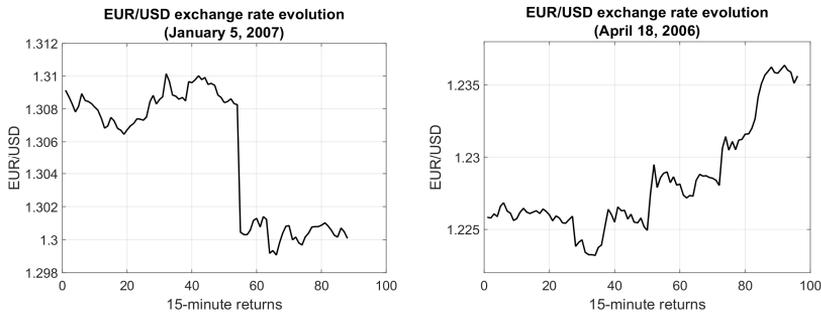


Figure 8 Days with the Largest Value of the Z-Statistics ($Z = 9.48$) (left) and the Largest Bayesian Probability of Jump Occurrences ($bp = 48\%$) (B_3) (right)



As can be seen in *Figure 7*, high Bayesian probabilities of jump occurrence coincide to a limited degree with higher values of the Z-statistic. Nevertheless, most of the days with high Bayesian probabilities of jumps probably do not contain realized jumps. We further performed a Welch *t*-test for the differences between the two sample means, which gave us mixed results (see *Table 8*).

As a practical illustration, *Figure 8* plots the day with the highest value of the Z statistics (9.48) and the day with the highest Bayesian jump probability for B_3 (48%). As we can see, the return on the day with $Z = 9.48$ was not enormous, but it contained a single large 15-minute return. This could not be seen at the daily frequency and the Bayesian method actually assigned a mere 2.13% jump probability to that day. On the other hand, if we look at the returns to which the Bayesian approach assigned 48% probability of a jump occurrence, we can see that there were no visible discontinuities in the price evolution on that day and that the price merely trended the whole day, achieving a large daily return gradually during the whole period (in fact the value of the Z-statistic on that day was only 0.2662).

Considering the distribution of jumps in *Figures 3* and *4*, we can see that the frequency of jump occurrences increases greatly after the beginning of 2012. In order to determine whether the results of our study hold even for this more recent

Figure 9 Bayesian Probabilities of Jump Occurrences (after 1 January 2012) (B₄)

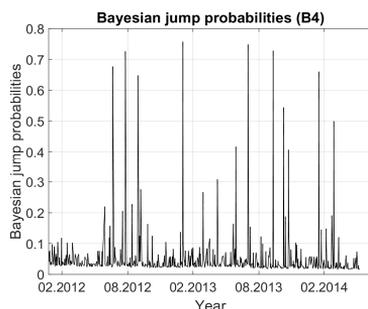


Table 8 Welsh *t*-Tests for the Difference between the Two Sample Means (bp \geq x% and bp $<$ x%) (1 = statistically significant difference, 0 = statistically insignificant difference)

	bp \geq 2%	bp \geq 5%	bp \geq 10%
B_2_jumps	0	1	0
B_3_jumps	1	1	0
B_4_jumps	1	0	0

Table 9 Bayesian Estimates of the Parameters of the SVJD Model (after 1 January 2012)

	mui	muiJ	sigmaJ	alpha	beta	gamma	thetaJ	betaJ	gammaJ
B_2_Mean	-0.0001	0.0049	0.0045	-0.1283	0.9885	0.0593	0.0551	0.4286	0.0409
B_2_Std	(0.0002)	(0.0037)	(0.0009)	(0.1274)	(0.0117)	(0.0260)	(0.0264)	(0.2636)	(0.0278)
B_3_Mean	-0.0001	0.0038	0.0061	-0.1620	0.9854	0.0634	0.0531	0.4528	0.0408
B_3_Std	(0.0002)	(0.0033)	(0.0011)	(0.1430)	(0.0131)	(0.0260)	(0.0264)	(0.2681)	(0.0274)
B_4_Mean	-0.0001	0.0013	0.0076	-0.1273	0.9886	0.0491	0.0535	0.4428	0.0433
B_4_Std	(0.0002)	(0.0103)	(0.0016)	(0.1325)	(0.0122)	(0.0288)	(0.0305)	(0.2721)	(0.0286)

Table 10 Rank Correlation between the Jump Estimators (after 1 January 2012)

	bp2	bp3	bp4	Z(t, Δ)	RV	RV-BV	R2(day)	max(r2(15))
bp2	1.0000	0.9295	0.8911	0.0603	0.4170	0.1799	0.6244	0.3411
bp3	0.9295	1.0000	0.9743	0.0456	0.4858	0.1937	0.8060	0.3896
bp4	0.8911	0.9743	1.0000	0.0265	0.5012	0.1787	0.8511	0.3971
Z(t, Δ)	0.0603	0.0456	0.0265	1.0000	0.0945	0.8864	0.0254	0.3474

period, we re-estimated the Bayesian model on the data for the period after 1 January 2012. From the results in *Tables 9* and *10* we can see that although jumps are much more common in this period, the main results still hold to a certain degree (although the correlation between the Bayesian jump probabilities and variable *Z* is slightly larger for this period; for the estimated jump probabilities see *Figure 9*).

5. Conclusion

We compared two different approaches to stochastic volatility and jump estimation. The first approach uses power-variation measures calculated from high-frequency returns in order to non-parametrically estimate the continuous and discontinuous components of price variability. We used the realized variance to estimate the quadratic variation, bipower variation to estimate the integrated variance and the shrinkage estimator to estimate the jump component of the quadratic variation at different probability levels.

The second approach specifies the underlying volatility and jump processes parametrically and then estimates their parameters and the latent state variables using Bayesian inference methods. We used a log-variance Stochastic-Volatility Jump-Diffusion model with self-exciting jumps governed by the Hawkes process as our underlying parametric model and we estimated its parameters through a MCMC algorithm.

We applied the two methods to simulated time series, finding that the non-parametric approach identifies jumps with far greater accuracy than the parametric approach, especially when the jumps are small. We then applied the two methods to the past history of the EUR/USD exchange rate during the period between 3 February 2006 and 15 April 2014.

The most surprising result of our study is the finding that the probabilities of jump occurrences estimated using the two methods are virtually uncorrelated (using Spearman's rank-correlation coefficient), which means that the two methods do not identify jumps at the same times. This is strange, considering that the methods should theoretically identify the same phenomenon, i.e. discontinuous price changes.

Our further research showed that in the vast majority of cases the jumps identified using the non-parametric approach from high-frequency returns are not large enough to be distinguishable at the daily frequency—indeed, they are almost uncorrelated with the squared daily returns as well as with the realized variance.

The parametrically identified jumps, on the other hand, are strongly correlated with the daily squared returns, but in most cases they do not represent real jumps at all (i.e. they do not coincide with discontinuous price changes) but only large daily returns generated by the continuous price variability, amplified possibly by some other phenomenon—jumps in volatility or the potentially stochastic behavior of the conditional mean. Therefore, it may be interesting to include these features in the SVJD model and examine the problem further.

Among other results, we found that non-parametrically identified jumps exhibit only weak (statistically insignificant) clustering and that parametrically identified jumps do not exhibit clustering at all. Conversely, we found some autocorrelation in the jump magnitudes for the non-parametrically identified jumps, which is a feature that may further be included in the parametric model as well. From the results, it also seems that it may be useful to use a bi-modal distribution for the jump component in the SVJD models, as it corresponds more closely to the distribution of jumps that are identifiable using that method.

We further found that the intensity of jumps greatly increases after the beginning 2012, but the main results of our study still seem to more or less hold.

REFERENCES

- Ait-Sahalia Y, Cacho-Diaz J, Leaven R (2015): Modeling Financial Contagion Using Mutually Exciting Jump Processes. *Journal of Financial Econometrics*, 117(3):585–606.
- Andersen TG, Bollerslev T (1998): Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, 39(4):885–905.
- Andersen TG, Bollerslev T, Christoffersen PF, Diebold FX (2005): Volatility Forecasting. *NBER Working Paper*, no. 11188.
- Andersen TG, Bollerslev T, Diebold FX (2007): Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*, 89(4):701–720.
- Andersen TG, Bollerslev T, Diebold FX, Labys P (2003): Modeling and Forecasting Realized Volatility. *Econometrica*, 71(2):579–625.
- Andersen TG, Dobrev D, Schaumburg E (2014): A Robust Neighborhood Truncation Approach to Estimation of Integrated Quarticity. *Econometric Theory*, 30(1):3–59.
- Bacry E, Mastromatteo I, Muzy J-F (2015): Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1).
- Barndorff-Nielsen OE, Hansen PR, Lunde A, Shepard N (2008): Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76:1481–1536.
- Barndorff-Nielsen OE, Shepard N (2004): Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, 2(1):1–48.
- Božovic M (2008): The Role of Jumps in Foreign Exchange Rates. *Universitat Pompeu Fabra, Barcelona, Working paper*, November, pp. 1–50.
- Chen K, Poon S-H (2013): Variance Swap Premium under Stochastic Volatility and Self-Exciting Jumps. University of Manchester, *Manchester Business School, Working paper*, pp. 1–50.
- Corsi F (2004): A Simple Long Memory Model of Realized Volatility. *University of Lugano, Institute of Finance, Working paper*, August, pp. 1–31.
- Corsi F, Pirino D, Reno R (2010): Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting. *Sant'Anna School of Advanced Studies, LEM Working Paper Series*, no. 11, pp. 1–33.
- Craine R, Lochstoer LA, Syrtveit K (2000): Estimation of a Stochastic-Volatility Jump-Diffusion Model. *Economic Analysis Review*, 15(1):61–87.
- Dovonon P, Goncalves S, Hounyo U, Meddahi N (2014): Bootstrapping High-Frequency Jump Tests. *Toulouse School of Economics, Discussion paper*, September 2014:1–41.
- Eraker B (2004): Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *Journal of Finance*, 59(3):1367–1403.
- Eraker B, Johannes M, Polson NG (2003): The Impact of Jumping Equity Index Volatility and Returns. *Journal of Finance*, 58(3):1269–1300.
- Fičura M (2014): Analysis of factors influencing the size of the volatility risk premium of the EUR/USD exchange rate. *Business & IT*, 4(2).
- Fičura M, Witzany J (2015): *Using high-frequency power-variation estimators in the Bayesian estimation of Stochastic-Volatility Jump-Diffusion models*. The 9th International Days of Statistics and Economics, Prague, September, 2015, pp. 423–434 (ISBN 978-80-87990-09-3). Available at: https://msed.vse.cz/msed_2015/article/170-Ficura-Milan-paper.pdf
- Franses PH, Dijk D van (2000): *Non-Linear Time Series Models in Empirical Finance*. Australia, Cambridge University Press (ISBN 0-511-01100-8).
- Fulop A, Li J, Yu J (2015): Self-Exciting Jumps, Learning, and Asset Pricing Implications. *Review of Financial Studies*, 28(3):876–912.
- Golightly A (2009): Bayesian Filtering for Jump-Diffusions with Application to Stochastic Volatility. *Journal of Computational and Graphical Statistics*, 18(2):384–400.

- Hanousek J, Kocenda E, Novotny J (2013): Price Jumps on European Stock Markets. *William Davidson Institute Working Paper*, no. 1059. Available at SSRN: <http://ssrn.com/abstract=2370586> or <http://dx.doi.org/10.2139/ssrn.2370586>
- Hansen PR, Lunde A (2004): *An Unbiased Measure of Realized Variance*. London School of Economics: Available at: http://www.lse.ac.uk/fmg/documents/events/seminars/capitalMarket/2004/P_Hansen.pdf
- Jacquier E, Johannes M, Polson N (2007): MCMC Maximum Likelihood for Latent State Models. *Journal of Econometrics*, 137:615–640.
- Jacquier E, Polson N, Rossi P (1994): Bayesian Analysis of Stochastic Volatility Models. *Journal of Business & Economic Statistics*, 12(4):69–87.
- Johannes M, Polson N (2009): MCMC Methods for Financial Econometrics. In: Ait-Sahalia Y, Hansen LP (Eds.): *Handbook of Financial Econometrics*, pp.1–72 (North-Holland).
- Lanne M (2006): Forecasting Realized Volatility by Decomposition. *European University Institute, Economics Working Papers*, no. ECO2006/20, pp. 1–26.
- Lee SS, Mykland PA (2008): Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *Review of Financial Studies*, 21(6):2535–2563.
- Liu X-B, Li Y (2014): Bayesian Testing Volatility Persistence in Stochastic Volatility Models with Jumps. *Quantitative Finance*, 14(8):1415–1426.
- Lynch SM (2007): *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Springer, August 2007, pp. 359 (ISBN-13: 978-0387712642).
- Maneesoonthorn W, Forbes CS, Martin GM (2016): Inference on Self-Exciting Jumps in Prices and Volatility Using High-Frequency Measures. *Cornell University Library, arXiv preprint*, March 2016, arXiv:1401.3911v3, pp. 1–47.
- Nakajima J (2012): Bayesian Analysis of GARCH and Stochastic Volatility: Modeling Leverage, Jumps and Heavy-Tails for Financial Time Series. *Japanese Economic Review*, 63(1):81–103.
- Novotny J, Petrov D, Urga G (2015): Trading price jump clusters in foreign exchange markets. *Journal of Financial Markets*, 24:66–92.
- Shephard N (2005): *Stochastic Volatility: Selected Readings (Advanced Texts in Econometrics)*. Oxford, Oxford University Press, May 2005 (ISBN-13: 978-0199257195).
- Shi P (2009): *Correcting Finite Sample Biases in Conventional Estimates of Power Variation and Jumps*. Duke University, Durham NC, Final Paper for Econ201FS, pp. 1–22.
- Stádník B (2014): The Riddle of Volatility Clusters. *Verslas: Teorija ir Praktika / Business: Theory and Practice*, 15(2):140–148.
- Stroud J, Johannes M (2014): Bayesian modeling and forecasting of 24-hour high-frequency volatility: A case study of the financial crisis. *Cornell University Library, arXiv preprint*, January 2014, arXiv:1211.2961, pp. 1–48
- Szerszen PJ (2009): Bayesian Analysis of Stochastic Volatility Models with Levy Jumps: Application to Risk Analysis. *Federal Reserve Board, Washington, D.C., Division of Research & Statistics and Monetary Affairs, Finance and Economics Discussion Series*.
- Todorov V (2010): Variance Risk Premium Dynamics: The Role of Jumps. *Review of Financial Studies*, January 2010, 23(1):345–383.
- Witzany J (2013): Estimating Correlated Jumps and Stochastic Volatilities. *Prague Economic Papers*, 22(2):251–283.
- Ysusi C (2006): Detecting Jumps in High-Frequency Financial Series Using Multipower Variation. *Banco de Mexico, Working paper*, no.2006-10, pp. 1–35.
- Zhang L, Mykland PA, Ait-Sahalia Y (2005): A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, 100(472):1394–1411.