Journal of Econometrics 178 (2014) 180–193

Contents lists available at ScienceDirect

Journal of Econometrics

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

Volatility activity: Specification and estimation☆

Viktor Todorov[a], George Tauchen[b,*], Iaryna Grynkiv[b]

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

[b] Department of Economics, Duke University, Durham, NC 27708, United States

ARTICLE INFO

Article history:
Available online 13 August 2013

JEL classification:
C51
C52
G12

Keywords:
Asymmetric volatility activity
High-frequency data
Laplace transform
Signed power variation
Specification testing
Stochastic volatility
Volatility jumps

ABSTRACT

The paper examines volatility activity and its asymmetry and undertakes further specification analysis of volatility models based on it. We develop new nonparametric statistics using high-frequency option-based VIX data to test for asymmetry in volatility jumps. We also develop methods for estimating and evaluating, using price data alone, a general encompassing model for volatility dynamics where volatility activity is unrestricted. The nonparametric application to VIX data, along with model estimation for S&P index returns, suggests that volatility moves are best captured by an infinite variation pure-jump martingale with a symmetric jump compensator around zero. The latter provides a parsimonious generalization of the jump-diffusions commonly used for volatility modeling.

1. Introduction

Volatility plays a central role in finance and economics, and there has been substantial research aimed at understanding volatility dynamics. A central obstacle in this effort is the fact that volatility itself is a latent process, with its dynamics hidden inside those of the observed price process. However, as now more high-frequency financial data becomes available, we can plausibly think in terms of continuous time, and our ability to precisely infer features of the volatility dynamics has considerably improved.

Existing classes of continuous time volatility models generally place quite sharp restrictions on the specification of volatility dynamics. For instance, the most common way of modeling volatility in continuous time is through the affine jump-diffusion class (see Duffie et al. (2000)); these models offer significant analytical tractability which facilitates estimation as well as various applications of the models, e.g., for option pricing. Although quite flexible, the affine jump-diffusion class of models imposes tight restrictions regarding the direction of volatility jumps and their activity.[1] The affine class of models only permits positive volatility jumps with the activity of volatility jumps further restricted to being of finite variation and commonly of finite activity as well.[2] Likewise, other classes of models that use nonnegative Lévy processes to model volatility dynamics (these processes are referred to as Lévy subordinators) restrict volatility jumps to being positive and of finite variation. Finally, the most commonly used volatility model in econometrics, the so-called exponential-diffusion, precludes any volatility jumps at all.[3] These models notwithstanding, actual volatility jumps can be much more general in terms of both their activity and their asymmetry, and there is no reason to expect the true features to agree with the model-implied versions.

Given the recent availability of very high-frequency data, these tight restrictions on both the direction and the activity level of volatility jumps in the extant models can be put to the test. In this paper we undertake such specification analysis by studying volatility activity using both parametric and nonparametric techniques. In Todorov and Tauchen (2011), we used the high-frequency (option-based) VIX index[4] data to estimate the activity

☆ Todorov's work was partially supported by an NSF grant.

* Corresponding author.
E-mail addresses: v-todorov@kellogg.northwestern.edu (V. Todorov), george.tauchen@duke.edu (G. Tauchen), ig9@duke.edu (I. Grynkiv).

[1] The activity level is characterized by a set of activity indexes, all taking values in the interval [0, 2], described in great detail below.

[2] In most applications, e.g., the double-jump model of Duffie et al. (2000), volatility jumps are modeled via a nonnegative compound Poisson process, which is of finite activity (and thereby has an activity index of zero).

[3] Another implication of the restrictive modeling of volatility jumps is that the model-implied volatility activity can only assume two possible values: either 1 (in the case where the driving process is a Lévy subordinator) or 2 (in the case where diffusion is used in the volatility modeling). The true level of volatility activity, however, can take any value in the interval [1, 2].

[4] The (volatility) VIX index is calculated by the Chicago Board of Options Exchange (CBOE) and is based on close-to-maturity S&P 500 index options. The

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.015

of volatility jumps and found evidence that the latter are of infinite variation with activity far exceeding that of finite variation typically used to model volatility jumps to date.[5]

In this paper we go one step further and answer the question: are the up and down moves in volatility equally active? In the context of a diffusion model, this is always the case. However, for jumps the activity of positive and negative moves might differ. One extreme case is of course when only positive volatility jumps are allowed in which case down jumps have trivially activity of zero. Here, we show that potential asymmetry in the volatility activity can be detected by utilizing the difference between power variation constructed from only positive and only negative high-frequency VIX increments. The relative magnitude of these power variations will be asymptotically different depending on the degree of asymmetry in the volatility activity. When implementing our nonparametric test we find evidence that the volatility jump activity is close to being symmetric. An implication of this result is also that negative volatility jumps are present.

The above nonparametric evidence about the volatility activity is based on the VIX volatility index, which is an option-based quantity. We next investigate whether these properties of volatility activity documented in options data can be "seen" with underlying asset price data alone.[6] The presence of positive and negative jumps in volatility, as well as their activity levels are features of the volatility dynamics that pertain to its *unobservable* path. Theoretically such features of the volatility can be inferred by relying solely on high-frequency price data on a fixed time interval. Such inference will, however, be associated with very slow rates of convergence and will be virtually infeasible with the frequencies that are available in practice.[7] Intuitively, the reason is that volatility is unobserved and changes in the volatility level are convoluted with the Gaussian shocks and in addition price jumps.

We adopt an alternative strategy in this paper by specifying a very general parametric volatility model and estimating it from the price data. The parametric model that we propose overcomes the above-mentioned limitations of existing models regarding volatility activity. In particular, jumps of arbitrary sign are allowed and their activity is left unrestricted. This is done by using infinite variation asymmetric Lévy processes as the driving martingale in an exponential Ornstein–Uhlenbeck (OU) volatility specification. In such a setting we address the question of whether the underlying asset data is informative enough to distinguish these model features (and hence limitations of the extant models).

The parametric structure helps the inference in several directions. First, regarding the volatility activity, which is associated with the small changes in volatility, the parametric model allows us to "borrow" information from volatility jumps which are medium-sized and hence "easier" to see with high-frequency *price* data. Second, the parametric specification allows us to pool information across time as opposed to working with fixed span and sampling more frequently. This can lead to significant increase in the precision of the estimation.[8]

Of course, the efficiency gains from the parametric specification come with the cost of robustness to potential model misspecification. To alleviate the concerns regarding the effect of model misspecification about the conclusions regarding asymmetric volatility activity, we do the following. First, we use a model that is general enough and in which importantly the different features of the volatility dynamics are captured by different parameters. In this regard, our modeling strategy follows Barndorff-Nielsen and Shephard (2001) and specifies the model by parameterizing separately the memory kernel and the marginal distribution of the volatility process. Further, in our parametric model separate parameters govern the tail behavior of volatility jumps and the behavior of the small positive and negative volatility jumps. The latter determines the volatility activity and the possible presence of asymmetric activity of volatility jumps. Second, we use the high-frequency price data to efficiently summarize the information about volatility in a robust way. So, in particular, we do not need to assume any particular model for the price jumps and their intensity as well as the challenging question of the relation between the shocks in volatility and those in the price level.

More specifically, our estimation method is based on integrating the high-frequency data into the realized Laplace transform of volatility proposed by Todorov and Tauchen (2012) to succinctly summarize the information about volatility in the data. The latter measure provides a nonparametric estimate of the empirical Laplace transform over a day (or any other period of time) and when aggregated over time can be used to measure the (integrated) joint volatility Laplace transform over different points in time. As is well-known, the joint Laplace transform preserves all the information about the volatility process dynamics. Our estimation is then based on minimizing the distance between the data-based and model-implied integrated joint volatility Laplace transforms. The latter unfortunately is unavailable in closed form and we evaluate it via simulation.

Our results from the parametric estimation show that the high-frequency price data alone contains information about the volatility activity. We find strong evidence against the standard pure-diffusive log stochastic volatility model. This model appears to be misspecified for modeling both small and big volatility moves. When a pure-jump specification is used with positive only jumps, like the non-Gaussian OU model of Barndorff-Nielsen and Shephard (2001) but with arbitrary jump activity, the fit improves significantly with strong evidence for jumps of infinite variation. Our best performing model is a pure-jump model with symmetric infinite variation small jumps and asymmetric big positive and negative jumps. These parametric results are in accordance with the nonparametric findings based on the VIX index.

The rest of the paper is organized as follows. Section 2 presents the setting in the paper. Section 3 defines the various notions of volatility activity. Section 4 shows how to use high-frequency data on the volatility VIX index to draw nonparametric inferences about the asymmetry of the volatility activity under some assumptions for the risk premium. In Section 5 we introduce our parametric Lévy-driven stochastic volatility model that allows for the general activity patterns found using the VIX index data and in Section 6 we present the estimation technique that we use for its estimation from high-frequency price data. Section 7 contains the empirical results from the parametric volatility estimation. Section 8 concludes. All technical results are given in Appendices A and B at the end of the paper.

prices of the options are weighted appropriately to replicate the risk-neutral future quadratic variation of the underlying S&P 500 index. Details on the calculations are available in the *white paper* on the CBOE website.

5 An interesting exception is the exponential continuous time GARCH model, recently studied theoretically by Haug and Czado (2007). Their setup allows for volatility jumps of infinite variation but within the confines of a particular parametric class of submodels.

6 The link between the activity properties of the VIX index and those of the unobserved latent volatility requires the volatility risk premia to depend on the same state variables which determine the stochastic volatility. While this assumption has been maintained in most (if not all) earlier empirical work, its violation would invalidate the link between the volatility activity and our nonparametric evidence about the VIX. This is another motivation for studying the volatility activity using only underlying price data.

7 The well-known presence of market microstructure noise further limits our ability to use "ultrahigh" frequencies such as seconds.

8 All parametric volatility models used to date imply that volatility activity does not change over time.

2. The model setup

Assume we observe at discrete points in time a price process X, defined on some filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ that has the following dynamics:

$$dX_t = \alpha_t dt + \sqrt{V_t} dW_t + \int_{\mathbb{R}} \delta(t-, x)\tilde{\mu}(dt, dx), \tag{1}$$

where α_t and V_t are càdlàg processes (and $V_t \geq 0$); W_t is a Brownian motion; μ is a homogeneous Poisson measure with compensator (Lévy measure) $ds\nu(x)dx$; $\delta(t, x) : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is càdlàg in t and $\tilde{\mu}(ds, dx) = \mu(ds, dx) - \nu(x)dxds$.

Our interest in the paper is in the specification of the stochastic process V_t, which we refer to as stochastic variance, and we will leave the rest of the components in the model, i.e., the drift term α_t and the price jumps, unspecified.

Restricting attention to the Markov setting (merely for ease of exposition, with an extension to a multifactor model being trivial), the most typical way to date used to model the volatility dynamics is by imposing it (or a transformation of it) to follow a Lévy-driven stochastic differential equation (SDE)

$$df(V_t) = f_1(V_t)dt + f_2(V_t)dB_t + f_3(V_{t-})dJ_t, \tag{2}$$

where B_t is a Brownian motion (having arbitrary connection with W_t) and J_t is a pure-jump Lévy martingale (having arbitrary connection with μ); f, f_1, f_2, f_3 are some functions guaranteeing the nonnegativity of the process V_t.

For the leading example of the affine jump-diffusion model of Duffie et al. (2000), the functions in (2) are respectively: $f(x) = x$, $f_1(x) = \kappa(\theta - x)$, $f_2(x) = \sigma\sqrt{x}$ and $f_3(x) = 1$ for κ, θ, σ some parameters. In this model J_t is restricted to positive jumps, necessarily of finite variation and in most cases even compound Poisson. The non-Gaussian OU model of Barndorff-Nielsen and Shephard (2001) further restricts $f_2(x) = 0$, i.e., no diffusion. Another well-known volatility model is the exponential Gaussian OU specification in which $f(x) = \log(x), f_1(x) = -\kappa\log(x), f_2(x) = \sigma$ and $f_3(x) = 0$ for κ and σ some parameters.

We will refer to the Lévy process $L_t = t + \sigma B_t + J_t$ as the driving Lévy process of the stochastic volatility. Then, the volatility model in (2) is uniquely identified by the functions f, f_1, f_2, f_3 and the driving Lévy process. We recall that for a generic Lévy process L_t with finite first moment, the Lévy–Khinchin theorem implies (see, e.g., Sato (1999))

$$\mathbb{E}\left(e^{iuL_t}\right) = \exp\left(iu\gamma - \sigma^2u^2/2 + \int_{\mathbb{R}}(e^{iux} - 1 - iux)\nu(dx)\right), \quad u \in \mathbb{R}, \tag{3}$$

where γ is the drift (which is one for the driving Levy process of the volatility specification in (2)), σ^2 is the variance of the Gaussian part, and $\nu(dx)$ is the Lévy measure. These three quantities identify uniquely the Lévy process and the corresponding infinitely divisible distribution of the process at a fixed point in time. Therefore, in the following we will identify a Lévy process, or an infinitely divisible distribution, by its so-called characteristic triplet (γ, σ, ν); see, e.g., Sato (1999). We note that the definition of the characteristic triplet is always unique up to the choice of a truncation function needed to ensure that integrals with respect to the counting jump measure are always well defined. Our implicit choice here is that this truncation function is the identity (and hence the requirement for the existence of first moment of the Lévy process).

3. Volatility activity

We next analyze the activity of the different components of the driving martingale of the stochastic volatility: positive jumps, negative jumps and diffusion. We first state the various notions of volatility activity and then show in the setting of the volatility specification in (2) how they relate to the characteristics of the driving Lévy process. Henceforth, for a generic semimartingale process Y, we use $\Delta Y_s = Y_s - Y_{s-}$ to denote its jumps.[9] We define the activity of the positive volatility jumps *pathwise* as

$$RA_V = \inf\left\{p : \sum_{s\leq T} |\Delta V_s|^p 1_{\{\Delta V_s > 0\}} < \infty\right\}, \quad \forall\, T > 0. \tag{4}$$

Similarly, the activity of negative volatility jumps is defined as

$$LA_V = \inf\left\{p : \sum_{s\leq T} |\Delta V_s|^p 1_{\{\Delta V_s < 0\}} < \infty\right\}, \quad \forall\, T > 0. \tag{5}$$

The activity levels RA_V and LA_V can take in general any value in the range $[0, 2]$, are random quantities, and depend on T. When V_t follows the Lévy-driven SDE in (2), since the process V_t is càdlàg (and hence locally bounded), RA_V and LA_V are equal to RA_L and LA_L, i.e., the corresponding activities of the driving Lévy process L_t. Hence, using the Lévy property, we have $RA_V \equiv \inf\{p : \int_{\mathbb{R}^+} x^p\nu(dx) < \infty\}$ and $LA_V \equiv \inf\{p : \int_{\mathbb{R}^-}(-x)^p\nu(dx) < \infty\}$ which are constant and do not depend on T.

For the affine jump-diffusion model we further have trivially $LA_V \equiv 0$, as negative jumps are simply not allowed in this class of models, and $RA_V < 1$, as the positive jumps are necessarily of finite variation. Furthermore, when jumps are of finite activity, as in the popular double-jump stochastic volatility model of Duffie et al. (2000), we even have $RA_V \equiv 0$.

The overall jump activity of both positive and negative jumps, defined in Ait-Sahalia and Jacod (2009) as a simple generalization of the Blumenthal–Getoor index, is simply $JA_V = \max\{RA_V, LA_V\}$. We should also point out that the jump activities are determined by the "small" jumps, since big jumps are always of finite number over finite time intervals.

The stochastic volatility models used to date have constrained the volatility jump activities to be as high as 1 and in most cases to be even exactly zero. It is clear from the above discussion, however, that the "universe" of available processes for modeling volatility risk is much wider, and in Section 5 we will introduce a general volatility model that can allow RA_V and LA_V to take any value in the interval $(0, 2)$.

Finally, the overall activity of the volatility process, proposed in Todorov and Tauchen (2010), is simply defined as

$$TA_V = \inf\left\{p : \operatorname*{plim}_{n\to\infty} \sum_{i=1}^{nT} |\Delta_i^n V|^p < \infty\right\}, \quad \Delta_i^n V = V_{\frac{i}{n}} - V_{\frac{i-1}{n}}, \quad \forall\, T > 0. \tag{6}$$

In words, TA_V is the smallest power for which the power variation of the volatility process does not explode. The overall activity is determined by the "dominating" component of the volatility process among the drift, the diffusion and the positive and negative jumps. The components of volatility order in terms of their activity from least to most active as follows: finite variation jumps, drift, infinite variation jumps and diffusion. For example when the volatility process contains a diffusion component, then its activity is always at its highest level of 2. When the volatility model is of pure-jump type then the volatility activity is determined by RA_V and LA_V.

[9] Y_{s-} denotes the limit from the left of the process, which always exists as the realizations of the process have càdlàg paths.

4. Nonparametric evidence on volatility activity

We continue with providing nonparametric evidence for the volatility activity. We use the VIX index quoted by the CBOE for this. We recall that the VIX index which we henceforth denote with IV_t (abbreviation for option *implied volatility*) is a basket of out-of-the-money European-style options on the S&P 500 index and provides a nonparametric measure for the *risk-neutral expected future quadratic variation*, i.e., with our notation in (1) (for X being the underlying S&P 500 index), we have

$$IV_t = \mathbb{E}^{\mathbb{Q}}\left(\int_t^{t+\tau} V_s ds + \int_t^{t+\tau}\int_{\mathbb{R}} \delta^2(s-,x)\mu(ds,dx)\middle|\mathcal{F}_t\right), \tag{7}$$

where τ corresponds to 1 calendar month and $\mathbb{Q}$ denotes the risk-neutral measure. The risk-neutral probability measure differs from the statistical one by the risk premia. Assuming multifactor structure for V_t, in Theorem 1 of Todorov and Tauchen (2011) we have shown that under the additional assumption that risk premia is a sole function of the state variables that determine the stochastic variance,[10] we have that IV_t is some smooth function of the current state of the multivariate volatility factor. For ease of exposition here we will assume that V_t follows (2), i.e., that V_t is a Markov process, and therefore we have $IV_t = g(V_t)$ for $g(\cdot)$ some smooth function. Then, using again Theorem 1 of Todorov and Tauchen (2011), we have

$$RA_V \equiv RA_{IV}, \qquad LA_V \equiv LA_{IV}, \qquad JA_V \equiv JA_{IV}, \qquad TA_V \equiv TA_{IV}. \tag{8}$$

This analysis shows that under the above-mentioned assumptions about the volatility risk premia, we can infer the volatility activity, i.e., the quantities RA_V, LA_V, JA_V and TA_V, from the corresponding quantities associated with IV_t. In Todorov and Tauchen (2011), using high-frequency data on the VIX index we estimated $JA_{IV} \equiv TA_{IV}$ in the range 1.73–1.83; see Table 5 of that paper. This implies that V_t is pure-jump process of infinite variation. Now, we will go one step further and will investigate the question of where the activity is coming from. In other words, we will be interested in statistical inference about the relation between RA_{IV} and LA_{IV}. Given the above-mentioned evidence for IV being pure-jump, we will do so in a pure-jump setting (and assuming high-frequency data on IV to be available). To this end we denote with Y_t a generic pure-jump semimartingale process.[11]

Our strategy for detecting asymmetry in the jump activity of Y will be based on the different behavior that signed power variation has depending on $RA_Y \lessgtr LA_Y$. We define the signed p-power variation as[12]

$$\begin{aligned} V_T^+(Y,p) &= \sum_{i=1}^{nT} |\Delta_i^n Y|^p 1\left(\Delta_i^n Y > 0\right), \\ V_T^-(Y,p) &= \sum_{i=1}^{nT} |\Delta_i^n Y|^p 1\left(\Delta_i^n Y < 0\right). \end{aligned} \tag{9}$$

We further denote as $V_T(Y,p) = V_T^+(Y,p) + V_T^-(Y,p)$ the total pth variation. Then we have the following result.

Theorem 1. *Assume that on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ we observe a process Y at times $0, \frac{1}{n}, \frac{2}{n}, \ldots, T$ for T fixed and $n \uparrow \infty$, where the process Y has the following dynamics:*

$$Y_t = Y_0 + \int_0^t \alpha_s' ds + \int_0^t \int_{\mathbb{R}} \sigma_{s-} x \tilde{\mu}'(ds,dx), \tag{10}$$

for α_s' and σ_s processes with càdlàg paths, and $\mu'(ds,dx)$ is a homogeneous Poisson measure with compensator $ds\nu'(x)dx$ of the form

$$\begin{aligned} &\nu'(x) = \frac{c^+}{|x|^{\beta^+ +1}} 1_{\{x>0\}} + \frac{c^-}{|x|^{\beta^- +1}} 1_{\{x<0\}} + \nu''(x), \\ &\quad c^+ > 0,\ c^- > 0, \\ &|\nu''(x)| \leq \frac{A}{|x|^{\beta' +1}}, \\ &\quad \text{for } |x| \leq x_0 \text{ for some } x_0 > 0 \text{ and } A > 0,\ \beta' < \max\{\beta^+, \beta^-\}. \end{aligned} \tag{11}$$

If σ_s is an Itô semimartingale, with locally bounded coefficients, and $\sigma_s > 0$ for $\forall s \in [0,T]$, and further $\max\{\beta^+,\beta^-\} > 1$, we have for any $0 < p < \max\{\beta^+,\beta^-\}$ as $n \uparrow \infty$

$$\begin{aligned} \Delta_n^{1-p/\beta} V_T^+(Y,p) &\xrightarrow{\mathbb{P}} K^+(p,\beta) \int_0^T |\sigma_s|^p ds, \\ \Delta_n^{1-p/\beta} V_T^-(Y,p) &\xrightarrow{\mathbb{P}} K^-(p,\beta) \int_0^T |\sigma_s|^p ds, \end{aligned} \tag{12}$$

where we use the shorthand $\beta = \max\{\beta^+,\beta^-\}$ and further $K^+(p,\beta)$ and $K^-(p,\beta)$ are the pth positive and negative moments of stable distribution with Lévy density $\frac{c^+}{|x|^{1+\beta}} 1_{\{x>0,\beta\equiv\beta^+\}} + \frac{c^-}{|x|^{1+\beta}} 1_{\{x<0,\beta\equiv\beta^-\}}$.

We note that $\nu''(x)$ in (11) is signed measure and therefore the above assumption accommodates Lévy measures with exponentially decaying tails like those of the tempered stable process that we use for our parametric model in Section 5. Implicit in the above theorem is that $\int_{\mathbb{R}} |x|\nu'(x) < \infty$ and this assumption can be spared. The above setting covers most models specified via Lévy-driven SDEs which is the dominant way of modeling continuous time processes.[13] In this setup, $RA_Y = \beta^+$ and $LA_Y = \beta^-$. To simplify notation therefore, in what follows (as in the theorem), we will use the shorthand $\beta = JA_Y$, $\beta^+ = RA_Y$, $\beta^- = RA_Y$.

When $\beta^+ > \beta^-$, then the constants $K^+(p,\beta)$ and $K^-(p,\beta)$ are determined by the pth signed moments of a β^+-stable *spectrally positive* process.[14] When $\beta^+ < \beta^-$, then $K^+(p,\beta)$ and $K^-(p,\beta)$ are determined by the pth signed moments of a β^--stable spectrally *negative* process. Finally, when $\beta^+ = \beta^-$, then $K^+(p,\beta)$ and $K^-(p,\beta)$ are determined by the pth signed moments of a β-stable distribution with asymmetry parameter controlled by the ratio of the local scales of the positive and negative Lévy measures.

We note that $K^+(p,\beta) \neq K^-(p,\beta)$ either when $\beta^+ \neq \beta^-$ or when $\beta^+ = \beta^-$ but $c^+ \neq c^-$. Both cases stem from the asymmetry of the driving Lévy measure of Y around zero.

We have $K^+(p,\beta) < K^-(p,\beta)$ when the limiting stable process is spectrally positive and the opposite when it is spectrally

[10] Parametric models in empirical work to date typically impose the models under statistical and risk-neutral measure to be of the same class, and further that jump intensity is a sole function of V_t; see, e.g., Singleton (2006) and references therein. This implies automatically the above requirement for the risk premia.

[11] We denote it with Y_t to avoid confusion with our underlying price process X_t in (1).

[12] Signed power variation for $p = 2$ and in the context of jump-diffusions was introduced by Barndorff-Nielsen et al. (2010) who study its asymptotic properties and use it as a measure of downside risk. Our application here is for the pure-jump semimartingale, which asymptotically behaves very differently from the jump-diffusion, and further our only goal is of detecting asymmetry in the volatility activity. We note also that in the context of diffusion, the constants $K^{\pm}(p,\beta)$ in (12) below will be equal to each other, i.e., the diffusion cannot generate asymmetric volatility activity unlike pure-jump processes.

[13] Examples include the non-Gaussian OU processes of Barndorff-Nielsen and Shephard (2001) as well as the exponential Lévy-driven OU processes that we introduce in Section 5.

[14] A *spectrally positive* jump process is a jump process with positive only jumps and similarly a *spectrally negative* jump process is a jump process with negative only jumps.

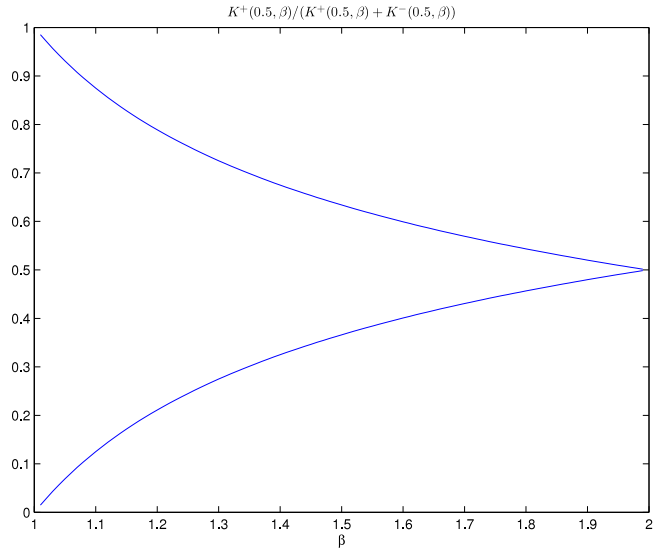

Fig. 1. The limit $\frac{K^+(0.5,\beta)}{K^+(0.5,\beta)+K^-(0.5,\beta)}$ as a function of β. The lower line corresponds to $\beta^+ > \beta^-$ and the upper line to $\beta^+ < \beta^-$.

negative. However, the limiting result in (12) is not convenient to use directly for our inference about the activity asymmetry for two reasons: (1) the limits in (12) are time-varying (because of σ_t), and (2) the above limit results involve scaling of the power variations that include the unknown β. Both these problems can be overcome by looking instead at the ratio $\frac{V_T^+(Y,p)}{V_T^+(Y,p)+V_T^-(Y,p)}$. Its limiting behavior follows directly from Theorem 1:

$$\frac{V_T^+(Y,p)}{V_T^+(Y,p)+V_T^-(Y,p)} \xrightarrow{\mathbb{P}} \frac{K^+(p,\beta)}{K^+(p,\beta)+K^-(p,\beta)}, \tag{13}$$

and can therefore reveal whether there is potential asymmetric jump activity of the discretely observed process Y.

For example for $p = 0.5$ and $\beta = 1.5$, if the limiting process is spectrally positive (no negative jumps), the limiting value of the ratio is 0.366. If the limiting process is spectrally negative (no positive jumps), the limiting value in (13) is 0.634 and finally if the process is symmetric, then its value is 0.5. More generally, the limit in (13) depends on β and on Fig. 1 we plot it for the case where $\beta^+ \neq \beta^-$.[15] As seen from the figure, for $\beta = \max\{\beta^+, \beta^-\}$ approaching 2, the limit in (13) converges to 0.5. That is, the activity asymmetry shrinks. Intuitively, this is because when β converges to 2, the pure-jump process "converges" to a diffusion. For the latter, the limiting value of the ratio is 0.5, corresponding to symmetric activity.

We calculated the signed power variation ratio in (13) for the 5 min high-frequency VIX index data used in Todorov and Tauchen (2011) which covers the period September 2003 to December 2008.[16] We set $p = 0.5$ (other powers produced very similar results). The left panel of Fig. 2 shows the resulting series. As seen from the graph, the ratio is surprisingly close to 0.5. Indeed the sample mean is 0.4862 and the sample median is 0.4819. Moreover, the first-order autocorrelation of the series is -0.03 and statistically insignificant, further confirming that deviations from the mean are just estimation error.

To contrast this behavior of the VIX index, we simulated from our exponential OU volatility model that we introduce in (15)–(16) in the next section (the model allows for asymmetric volatility activity). Fig. 3 corresponds to the case where the driving martingale is spectrally positive, Fig. 4 to the case where it is spectrally negative and Fig. 5 to the symmetric driving martingale case. As seen from the left panels of Figs. 3–5, the asymmetric volatility activity can be readily captured by our ratio of signed power variations. The means of $V_T^+(Y,p)/V_T(Y,p)$ in the simulated series are respectively 0.385 and 0.634, which are very close to their asymptotic limits (0.366 and 0.634 respectively) and indicate volatility asymmetry. On the other hand, the left panel of Fig. 5, consistently with theory, signals that the activity of the process is symmetric.

To further gauge the magnitude of the deviations from the symmetric limit of 0.5 for the signed power variation in our VIX data set, we computed in a long Monte Carlo study the quantiles of the daily signed power variation statistic for an exponential OU volatility model with exactly the same parameters as for Figs. 3–5 but with activity parameters set to $\beta^+ = \beta^- = 1.73$, which is approximately the estimate for the VIX activity from the data. Since the volatility of the VIX series is quite persistent, it is approximately constant over a short period of time like a day, and the distribution of the signed power variation is hence like that of a (locally) stable Lévy process, and this holds true also for the simulated model. Therefore, the simulated quantiles of the statistic should provide a reasonable proxy for the actual quantiles provided the VIX data has locally symmetric jump measure.[17] Comparing the quantiles of the signed power variation in the simulated and observed data, we see that in approximately only 4% of the days in the VIX data set, the signed power variation ratio is outside the 2.5%–97.5% quantile range determined from the simulation. This provides further nonparametric evidence that the small volatility changes are of pure-jump type and approximately symmetric.

While the behavior of $V_T^{\pm}(Y,p)$ for $p < \beta$ reveals the small scale behavior of the jumps and in particular their activity asymmetry, the behavior of the ratio $V_T^+(Y,p)/V_T(Y,p)$ for $p > \beta$ would reveal the symmetry of the big jumps of the process. Indeed, we have for $p > \beta$ (see, e.g., Theorem 2.2 in Jacod (2008))

$$\begin{aligned} V_T^+(X,p) &\xrightarrow{\mathbb{P}} \sum_{s\le T} |\Delta X_s|^p 1(\Delta X_s > 0), \\ V_T^-(X,p) &\xrightarrow{\mathbb{P}} \sum_{s\le T} |\Delta X_s|^p 1(\Delta X_s < 0). \end{aligned} \tag{14}$$

Of course, we note that for $p > \beta$, $V_T^+(Y,p)/V_T(Y,p)$ will have a *random* limit unlike for the case $p < \beta$ where the limit is a constant. Nevertheless using the sample mean of this statistic, averaged over the days in the sample, can provide information for the symmetry of the "big" jumps (for p high the relative importance of the small jumps in the pth variation is minimal). We plot in the right panels of Figs. 2–5 the daily ratios for $p = 2.5$ (which of course is above β). As seen from the figures, the spectral positivity results in a ratio which on average is close to 1 while exactly the opposite holds for the spectrally negative process. On the other hand for the symmetric martingale case and the VIX data the ratio is close to 0.5. Of course, the mere fact that $V_T^+(Y,p)/V_T(Y,p)$ for the VIX index can take values that are far below 1 is another manifestation of the presence of negative volatility jumps. At the same time, we stress that the evidence for the symmetry of the "big" jumps in the VIX

[15] For the case $\beta^+ < \beta^-$, the limit in (13) is 1 minus the limit for the case $\beta^+ > \beta^-$.

[16] We refer the reader to that paper for details and various summary measures of the VIX index data set.

[17] Furthermore, since only the stable part of the Lévy density matters in the limit (see (13) above), in the simulation the quantiles remained nearly unchanged when changing $\lambda^{\pm}$ and $c^{\pm}$. This further implies that the quantiles of the signed power variation are not very sensitive to the assumed parametric model (at least as far as it is realistically calibrated with the actual data) in the simulation.

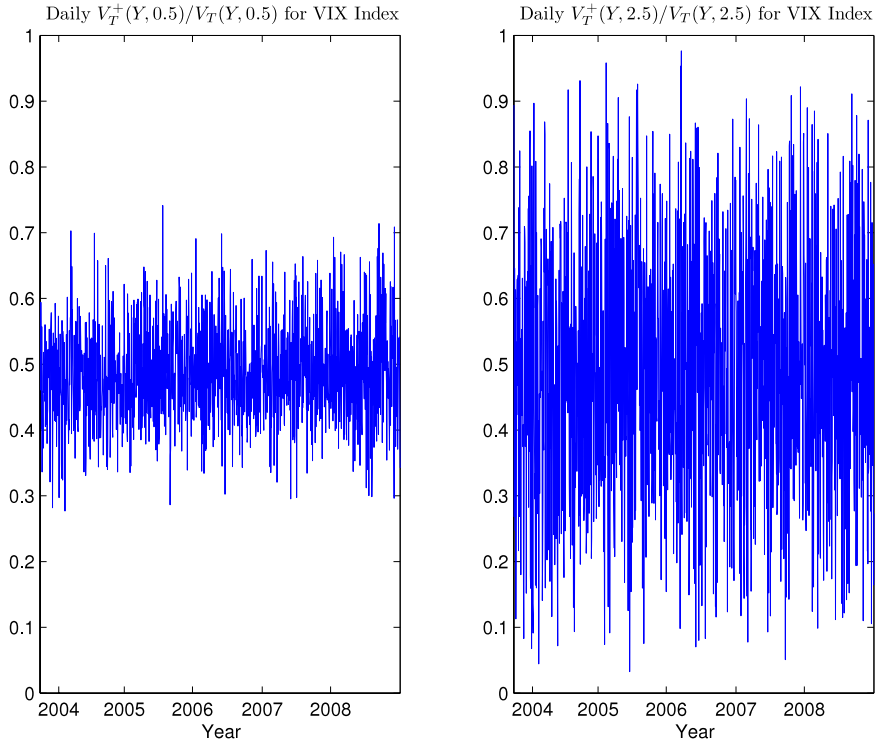

Fig. 2. The data set is the 5 min VIX index spanning the period 2003–2008 for a total of 1212 days.

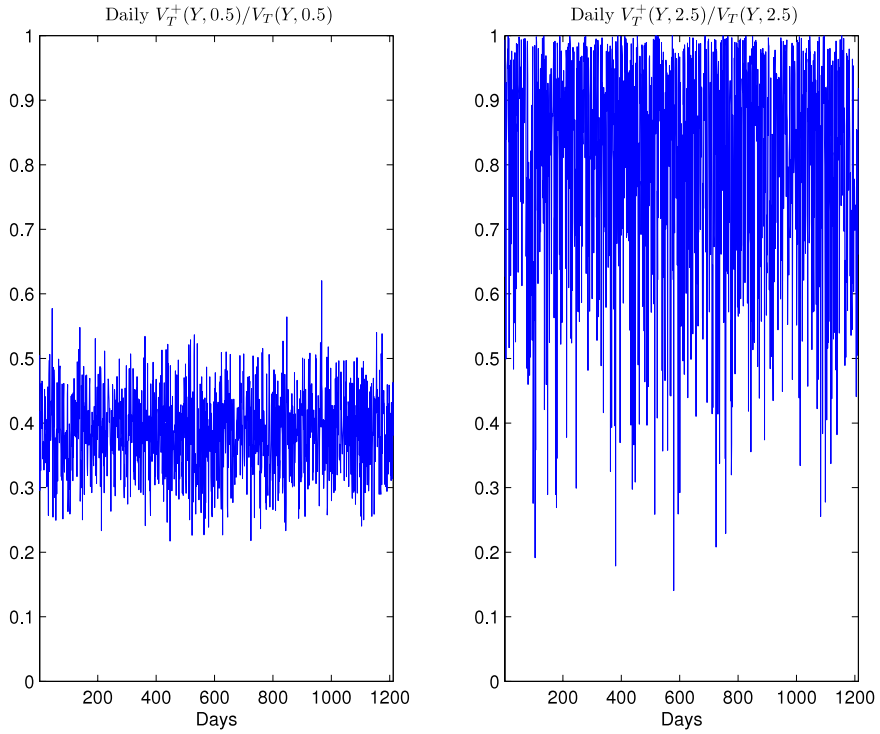

Fig. 3. The results are based on a simulated Exp-OU model (15)–(16). The simulated series is sampled at 5 min intervals for a total of 1212 days mimicking the VIX index set. The parameters of the scenario are: $\beta^+ = 1.5, c^+ = 0.2, \lambda^+ = 3.0, c^- = 0, \kappa = 0.01$, which corresponds to a spectrally positive tempered stable marginal of the volatility process. The limiting value of the ratio on the left panel is 0.366 and the sample mean is 0.385.

index does not automatically translate into symmetric big jumps in V_t as the presence of risk premia distorts the mapping.[18]

On the basis of the nonparametric evidence from the high-frequency VIX index data, we can draw several conclusions for the volatility process V_t: (1) the process is pure-jump of infinite variation, (2) volatility activity is approximately symmetric, and (3) negative jumps are present.

[18] As discussed above, this is unlike the case for the small jumps which determine the jump activity, for which there is one-to-one mapping as given in (8).

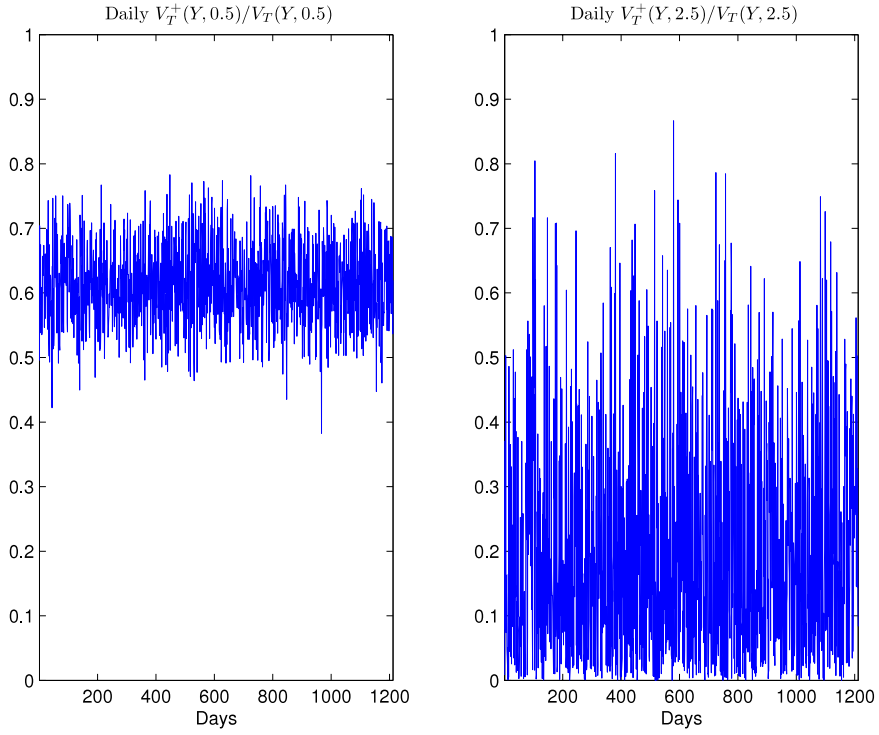

Fig. 4. The results are based on a simulated Exp-OU model (15)–(16). The simulated series is sampled at 5 min intervals for a total of 1212 days mimicking the VIX index set. The parameters of the scenario are: $c^+ = 0$, $\beta^- = 1.5$, $c^- = 0.2$, $\lambda^- = 3.0$, $\kappa = 0.01$, which corresponds to a spectrally negative tempered stable marginal of the volatility process. The limiting value of the ratio on the left panel is 0.634 and the sample mean is 0.615.

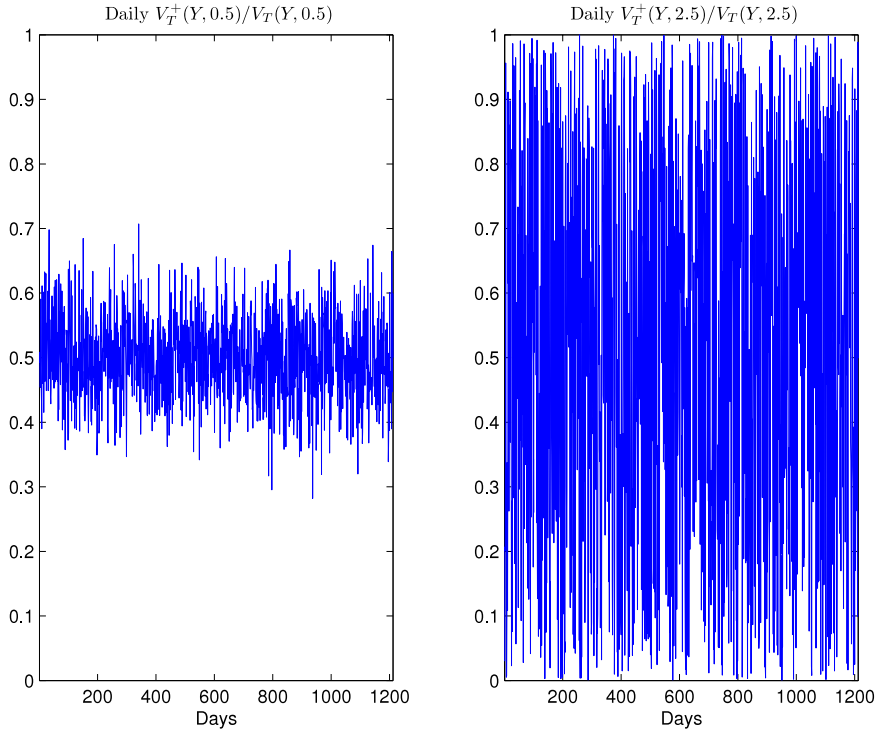

Fig. 5. The results are based on a simulated Exp-OU model (15)–(16). The simulated series is sampled at 5 min intervals for a total of 1212 days mimicking the VIX index set. The parameters of the scenario are: $\beta^+ = \beta^- = 1.5$, $c^+ = c^- = 0.2$, $\lambda^+ = \lambda^- = 3.0$, $\kappa = 0.01$, which corresponds to a symmetric tempered stable marginal of the volatility process. The limiting value of the ratio on the left panel is 0.5 and the sample mean is 0.5.

5. General exponential Lévy-driven volatility models

We will next try to find evidence for the volatility activity solely on the basis of the underlying price data and compare with the above option-based nonparametric evidence. The questions that we seek to answer are the following. Is there evidence for volatility being of pure-jump type? Are there negative jumps in volatility? Are the volatility jumps of infinite variation? Is there symmetry in the volatility activity as our evidence based on the VIX index data would suggest? Is there enough information in the price data alone

to answer the above questions in a statistically affirmative way? Also, we should bear in mind, the link between the activity of the VIX index and the underlying stochastic volatility depends on the assumption that risk premia are uniquely identified by the state variables determining V_t. This has been a standard assumption in earlier empirical work but if it fails then the evidence from the VIX index of the previous section will have limited implications for the properties of the stochastic volatility process.

Our strategy for answering the above questions will be to extract from the price data nonparametrically the information about the volatility process and then compare it with that of a general parametric volatility model. Here we introduce the model and analyze its properties and in the next section we present the estimation method. As already discussed in the introduction, the popular affine jump-diffusion volatility model does not allow for negative jumps and further restricts the activity of positive jumps to being of finite variation. Therefore, it cannot be used for the purposes of our analysis and we propose an alternative model that makes no restrictions regarding the jumps, either their sign or their activity level. We will introduce the model in the Markov setting, with the generalization to a multifactor setting being obvious. Our model is given by

$$V_t = \exp(\mu + v_t), \qquad dv_t = -\kappa dt + dL_t, \tag{15}$$

where $\kappa > 0$ and L_t is a Lévy process. This is simply a general exponential OU process (the generalization being in the driving Lévy process). The exponential transformation allows for negative jumps in V_t as well as arbitrary volatility activity. The factor v_t has a marginal distribution which is infinitely divisible with characteristic triplet $(0, \sigma, \nu_v)$ where (recall the definition of the characteristic triplet in (3))

$$\nu_v(dx) = \left(c^+ \frac{e^{-\lambda^+ x}}{x^{1+\beta^+}} 1_{\{x>0\}} + c^- \frac{e^{-\lambda^- |x|}}{|x|^{1+\beta^-}} 1_{\{x<0\}} \right) dx,$$

$$c^\pm \geq 0,\ \lambda_\pm > 0,\ \beta^\pm \in [0, 2). \tag{16}$$

We note that v_t is an OU martingale. This way of modeling the volatility process is similar to an approach proposed by Barndorff-Nielsen and Shephard (2001) for modeling a non-Gaussian OU volatility by specifying its marginal distribution and then "back out" from it the model for the driving Lévy process. This approach has the advantage that the parameters controlling the memory of the volatility process are separated from those controlling its distribution.

In our parametric specification, the distribution of v_t is a mixture of a normal distribution with variance σ^2 and that of a pure-jump tempered stable martingale (evaluated at time 1) with Lévy measure ν_v (Carr et al. (2002); Rosiński (2007)). The latter is known to be a very flexible distribution accommodating as special cases many known ones such as the inverse Gaussian and the Gamma distribution. A very attractive feature of this parametric model for the jumps is that over small scales the increments of the volatility V_t behave like those of a stable process but at the same time, unlike for the stable process, all moments of the volatility increments exist.

Using Barndorff-Nielsen and Shephard (2001); Sato (1999), we have that L_t is characterized by the characteristic triplet $\left(0, \sqrt{2\kappa}\sigma, \nu_L\right)$ where[19]

$$\nu_L(dx) = \left\{ \beta^+ \kappa c^+ \frac{e^{-\lambda^+|x|}}{|x|^{1+\beta^+}} + \lambda^+ \kappa c^+ \frac{e^{-\lambda^+|x|}}{|x|^{\beta^+}} \right\} 1_{\{x>0\}} dx + \left\{ \beta^- \kappa c^- \frac{e^{-\lambda^-|x|}}{|x|^{1+\beta^-}} + \lambda^- \kappa c^- \frac{e^{-\lambda^-|x|}}{|x|^{\beta^-}} \right\} 1_{\{x<0\}} dx. \tag{17}$$

[19] The explicit link between ν_v and ν_L is $\nu_L(x) = -\kappa(\nu_V(x) + x\nu_V'(x))$ and follows from Theorem 17.5 in Sato (1999).

The jumps in the volatility process V_t are associated with the jumps in the driving Markov process v_t via

$$\frac{\Delta V_t}{V_{t-}} = e^{\Delta v_t} - 1. \tag{18}$$

In words, the percentage jumps in volatility are given by $e^{\Delta v_t} - 1$. Our interest is in the properties of these jumps and therefore we derive their Lévy measure explicitly. For this, we introduce the transformation $\psi(x) = e^x - 1$. We denote with $\nu_{\psi(L)}$ the image of the Lévy measure ν_L (of the jumps of the driving Lévy process L_t in (15)) under the transform $x \to \psi(x)$. It is easy to derive then the following:

$$\nu_{\psi(L)}(dx) = \left(\beta^+ \kappa c^+ \frac{|x+1|^{-\lambda^+ - 1}}{|\log(x+1)|^{\beta^+ + 1}} + \lambda^+ \kappa c^+ \frac{|x+1|^{-\lambda^+ - 1}}{|\log(x+1)|^{\beta^+}} \right) 1_{\{x \in (0, +\infty)\}} dx + \left(\beta^- \kappa c^- \frac{|x+1|^{\lambda^- - 1}}{|\log(x+1)|^{\beta^- + 1}} + \lambda^- \kappa c^- \frac{|x+1|^{\lambda^- - 1}}{|\log(x+1)|^{\beta^-}} \right) 1_{\{x \in (-1, 0)\}} dx. \tag{19}$$

Using the fact that on each (finite) time interval the volatility V_t is bounded,[20] it is easy to see from (18) that for our general model in (15)–(16), the jump activity indexes are constant and are given by $RA_V \equiv \beta^+ 1_{\{c^+ > 0\}}$ and $LA_V \equiv \beta^- 1_{\{c^- > 0\}}$. In words, the parameters $\beta^\pm$ completely determine the volatility jump activity in our model and can take values in the whole possible range of jump activity $[0, 2)$. Further, the total jump activity is then determined by $TA_V = \max\{JA_V,\ 2_{\{\sigma \neq 0\}},\ 1\}$, i.e., it depends on whether a diffusive component is present in the driving Lévy process L_t or not.

We further note from (19) that the two parameters λ^- and λ^+ control the behavior of the jump tails (of $\psi(L)$), respectively at -1 and $+\infty$. This ensures that our parametric model is flexible enough, so that it does not allow for a "transfer" of information from the relatively big to small jumps and vice versa.[21] λ^- governs the behavior of the "big" negative jumps that can reduce the volatility to zero, while λ^+ controls the big positive jumps. Using the connection between the Lévy density and the tail probability derived in Rosinski and Samorodnitsky (1993), Theorem 2.1, we have for our model in (15)–(16) (provided positive jumps are present, which as we will see is the case of interest empirically)

$$\mathbb{P}(V_t \geq x) \sim \frac{c^+}{\lambda^+} \frac{x^{-\lambda^+}}{|\ln(x)|^{\beta^+ + 1}}, \quad \text{for } x \uparrow \infty. \tag{20}$$

Finally, our model (15)–(16) possesses also some analytical tractability. In particular, using the characteristic function of a tempered stable process (see, e.g., Cont and Tankov (2004), Proposition 4.2), we have

$$\log\left[\mathbb{E}(e^{uv_t})\right] = \psi(u), \quad u \leq \lambda^+, \qquad \beta^\pm \neq 0, 1, \tag{21}$$

[20] This is because the process is a semimartingale. Note that the bound is not uniform, i.e., it is for each volatility realization.

[21] This can be contrasted with more tightly parameterized jump models such as the stable, variance gamma or compound Poisson processes often used as building blocks in asset pricing models where the jump activity is either fixed or is a function of the parameter governing the jump tails.

where

$$\psi(u) = u^2\sigma^2/2 + c^+\Gamma(-\beta^+)\Big[(\lambda^+ - u)^{\beta^+} - (\lambda^+)^{\beta^+} + u\beta^+(\lambda^+)^{\beta^+-1}\Big] + c^-\Gamma(-\beta^-) \times \Big[(\lambda^- + u)^{\beta^-} - (\lambda^-)^{\beta^-} - u\beta^-(\lambda^-)^{\beta^--1}\Big]. \tag{22}$$

Thus, moments of the volatility process are known in closed form. This will be convenient for the implementation of our simulation-based estimation where we can perform variance-targeting techniques by restricting the model-implied mean to being within its nonparametric confidence bound inferred from the data. We will provide details on this in Section 7. The result in (21) reveals an interesting continuity in $\beta^\pm$ which is important to bear in mind when interpreting the estimation results. For $\sigma \equiv 0$, V_t is a pure-jump process with jump activity $JA_V = \max\{\beta^+, \beta^-\}$. When $\max\{\beta^+, \beta^-\} \to 2$ and, further, $c^\pm$ decrease such that $c^\pm\Gamma(-\beta^\pm)$ stays constant, then the pure-jump case "degenerates" to the continuous volatility case.

The result in (21) can be further extended to the case of joint moments of the volatility process over different points in time. We have for any $t, \tau \geq 0$ and $u, v \leq \lambda^+$

$$\begin{aligned}\log\left[\mathbb{E}(e^{uv_t + vv_{t-\tau}})\right] &= \psi(ue^{-\kappa\tau} + v) + \left(1 - e^{-2\kappa\tau}\right)u^2\sigma^2/2 \\ &+ c_+\Gamma(-\beta^+)\Big[(\lambda^+ - u)^{\beta^+} - (\lambda^+ - ue^{-\kappa\tau})^{\beta^+} \\ &+ u\beta^+(\lambda^+)^{\beta^+-1}(1 - e^{-\kappa\tau})\Big] \\ &+ c_-\Gamma(-\beta^-)\Big[(\lambda^- + u)^{\beta^-} - (\lambda^- + ue^{-\kappa\tau})^{\beta^-} \\ &- u\beta^-(\lambda^-)^{\beta^--1}(1 - e^{-\kappa\tau})\Big].\end{aligned} \tag{23}$$

The above result implies, in particular, that for $\tau \to \infty$, the asymptotic decay of the stochastic volatility auto-covariance function is exponential which is analogous to the standard affine jump-diffusion volatility models. The behavior for small lags, however, can differ not only from the affine jump-diffusion class, but also within our model depending on whether the model is pure-jump, jump-diffusion or pure-diffusion.

The analytical expressions for the moment conditions in (21)–(23) suggest that a simple method of moments condition can be developed. However, such estimation will be probably inefficient as the set of moment conditions are restricted to higher powers (the lowest possible power is 1) and those will be governed mainly from the parameters determining the volatility tail behavior, i.e., $\lambda^\pm$. Instead, in the next section we will use a method based on the conditional Laplace transform of volatility which can more efficiently extract the information about the different volatility characteristics that are in the data.

We conclude this section with a brief discussion about the multifactor extension. Given that the model is very richly parameterized in the one-factor setting, it is obvious that we need to impose some parametric restrictions in order to be able to identify the parametric structure. One parsimonious multifactor extension that we adopt in the empirical section is to let only the scale parameters σ and $c^\pm$ of the diffusive and jump parts respectively differ across the factors. In this way, the sum of the factors has exactly the same distribution as the individual factors with the only change being in the scales of the diffusion and the jumps.

6. Parametric volatility estimation using high-frequency data

We next briefly describe our estimation of the parametric volatility model (15)–(16) using high-frequency price data. The main challenge is extracting in an efficient and robust way the information for the latent volatility process from the discrete price observations. We do this here by using the realized Laplace transform proposed in Todorov and Tauchen (2012) defined over a day $[t-1, t]$ as

$$\begin{aligned}Z_t(u) &= \frac{1}{n}\sum_{i=n(t-1)+1}^{nt}\cos\left(\sqrt{2u}\sqrt{n\widehat{f}_i}^{-1/2}1_{\{\widehat{f}_i\neq 0\}}\Delta_i^n X\right), \\ \widehat{f}_i &= \frac{\widehat{g}_i}{\widehat{\overline{g}}}, \\ \widehat{g}_i &= \frac{n}{T}\sum_{t=1}^{T}|\Delta_{i_t}^n X|^2 1(|\Delta_{i_t}^n X| \leq 3n^{-0.49}), \\ \widehat{\overline{g}} &= \frac{1}{n}\sum_{i=1}^{n}\widehat{g}_i, \quad i = 1, \ldots, nT,\end{aligned} \tag{24}$$

where $i_t = t - 1 + i - [i/n]n$, for $i = 1, \ldots, nT$ and $t = 1, \ldots, T$. As shown in Todorov and Tauchen (2012), under the restriction of jumps in (1) being of finite variation and additional mild regularity conditions, we have $Z_t(u) = \int_{t-1}^t e^{-uV_s}ds + O_p\left(1/\sqrt{n}\right)$, for $u \geq 0$. Therefore, defining $\widehat{\mathcal{L}}_V(u, v; k) = \frac{1}{T-k}\sum_{t=k+1}^T Z_t(u)Z_{t-k}(v)$, we have for $u, v \geq 0$

$$\widehat{\mathcal{L}}_V(u, v; k) = \frac{1}{T-k}\sum_{t=k+1}^{T}\int_{t-1}^{t}e^{-uV_s}ds\int_{t-k-1}^{t-k}e^{-vV_s}ds + o_p(1/\sqrt{T}), \quad T \uparrow \infty,\ n \uparrow \infty,\ T/n \downarrow 0. \tag{25}$$

Under standard stationarity and ergodicity conditions, satisfied by our parametric model in (15)–(16), $\widehat{\mathcal{L}}_V(u, v; k)$ is a consistent and asymptotic normal estimator of

$$\mathcal{L}_V(u, v; k) = \mathbb{E}\left(\int_{t-1}^{t}e^{-uV_s}ds\int_{t-k-1}^{t-k}e^{-vV_s}ds\right).$$

We refer to $\mathcal{L}_V(u, v; k)$ as the integrated joint Laplace transform of volatility. As is well-known (see, e.g., Carrasco et al. (2007)), minimizing the distance between the empirical and model-implied joint Laplace transforms of a stochastic process can lead to efficient estimation of the underlying model for its dynamics. Therefore, we follow Todorov et al. (2011) and estimate our parametric volatility model via the following minimum-distance estimator:

$$\begin{aligned}\hat{\rho} &= \operatorname*{argmin}_{\rho}\ \mathbf{m}_T(\rho)'\widehat{\mathbf{W}}\mathbf{m}_T(\rho), \\ \mathbf{m}_T(\rho) &= \left\{\int_{\mathcal{R}_{j,k}}\left[\widehat{\mathcal{L}}_V(u, v; k) - \mathcal{L}_V(u, v; k|\rho)\right]\omega(du, dv)\right\}_{j=1,\ldots,J,k=1,\ldots,K},\end{aligned} \tag{26}$$

where ρ denotes the parameter vector; $\mathcal{R}_{j,k} \subset \mathbb{R}_+^2$ and ω is a weight function on $\mathbb{R}_+^2$; $\widehat{\mathbf{W}}$ is an estimate of the optimal weight matrix defined by the asymptotic variance of the empirical moments to be matched, i.e., that of $\int_{\mathcal{R}_{j,k}}\widehat{\mathcal{L}}_V(u, v; k)\omega(du, dv)$. Consistency and asymptotic normality of our minimum-distance estimator follow from classical conditions required for identification and CLT results for the moment vector. The choice of the regions and weight function follows Todorov et al. (2011). We set $u_{\max} = \mathcal{L}^{-1}(0.01, 0, ; 0)$ and use the following regions:

$$\begin{aligned}\mathcal{R}_{1,k} &= \{(u\ v) \in [0.1u_{\max}\ 0.2u_{\max}]^2\}, \quad k = 0, 1, 3, 10, 30, \\ \mathcal{R}_{2,k} &= \{(u\ v) \in [0.3u_{\max}\ 0.5u_{\max}]^2\}, \\ \mathcal{R}_{3,1} &= \{(u\ v) \in [0.6u_{\max}\ 0.9u_{\max}]^2\}, \quad k = 0, 1 \\ \mathcal{R}_{4,k} &= \{(u\ v) \in u_{\max}[0.1\ 0.2] \times u_{\max}[0.6\ 0.9]\}, \\ \mathcal{R}_{5,k} &= \{(u\ v) \in u_{\max}[0.6\ 0.9] \times u_{\max}[0.1\ 0.2]\}, \quad k = 0, 1.\end{aligned}$$

Within region, we use the weight function $\sum_i \delta_{(u_i,v_i)} e^{-0.5(u_i^2+v_i^2)/c^2}$ for $\delta_{\mathbf{x}}$ denoting Dirac delta at the point $\mathbf{x}$, $c = 0.50 \times u_{\max}$ and (u_i, v_i) being the edges of the regions. This choice of regions and weight within them provides a compromise between efficiency, computational speed and numerical stability. Monte Carlo work in Todorov et al. (2011) provides evidence that the above choice of lags and regions of integration $\mathcal{R}_{j,k}$ allows one to get very close to the Cramer–Rao efficiency bound corresponding to the infeasible scenario of direct daily observations of the variance process V_t.

The model-implied moments are not known in closed form and we evaluate them via simulation. Here we provide some details on this rather nontrivial step. A key feature of the parametric model (15)–(16), that we make use of in the simulation, is that the stationary distribution of the volatility process is known. Therefore, the simulation is done by generating an independent replica of the process over the interval $[0, K]$ where K is the highest number of lags used in the estimation (here 30). This way of estimating the moments implied by the model is more efficient than the alternative of simulating a single very long realization of the volatility process due to the strong persistence of the volatility process.

We define $Z \overset{\mathcal{L}}{\sim} PTS(\beta, c, \lambda)$ (PTS stands for positive-jump tempered stable) for a random variable with

$$\mathbb{E}\left(e^{iuZ}\right) = \exp\left(c\int_{\mathbb{R}^+}\left(e^{iux} - 1 - iux\right)\frac{e^{-\lambda x}}{x^{\beta+1}}dx\right),$$
$$c > 0,\ \lambda > 0,\ \beta < 2. \tag{27}$$

This is just the distribution of the positive-jump part of L_1, for L_t the Lévy process in (15)–(16). In Appendix B we provide details for the simulation of this process. The simulation of v_t for $\beta \geq 0$ is then done in the following way:

1. Simulate v_0 from its stationary distribution which is a mixture of normal distribution and tempered stable distribution:
$$v_0 \overset{\mathcal{L}}{\sim} \sigma_l \times \mathcal{N}(0,1) + Z_0^+ - Z_0^-,$$
$$Z_0^+ \overset{\mathcal{L}}{\sim} PTS(\beta_l^+, c_l^+, \lambda_l^+),\ Z_0^- \overset{\mathcal{L}}{\sim} PTS(\beta_l^-, c_l^-, \lambda_l^-). \tag{28}$$
2. Simulate $\{v_t\}_{t\in(0,K]}$ where K is the highest number of lags used in the estimation on the discrete grid $\frac{1}{n}, \frac{2}{n}, \ldots, K$. This is done via the following discretization of the dynamics in (15):
$$v_{\frac{i}{n}} \approx e^{-\kappa/n}\left(v_{\frac{i-1}{n}} + \sum_{j=1}^{m} e^{\kappa\frac{j-1}{nm}}\left(L_{\frac{i-1}{n}+\frac{j}{nm}} - L_{\frac{i-1}{n}+\frac{j-1}{nm}}\right)\right),$$
$$i = 1, \ldots, nK, \tag{29}$$
where
$$\begin{cases} L_{\frac{i-1}{n}+\frac{j}{nm}} - L_{\frac{i-1}{n}+\frac{j-1}{nm}} \overset{\mathcal{L}}{\sim} \frac{\sqrt{2\kappa}\sigma}{\sqrt{nm}} \times \mathcal{N}(0,1) \\ \quad + Z_1^+ - Z_1^- + Z_2^+ - Z_2^-, \\ Z_1^+ \overset{\mathcal{L}}{\sim} PTS\left(\beta^+, \frac{\beta^+\kappa c^+}{nm}, \lambda^+\right), \\ Z_1^- \overset{\mathcal{L}}{\sim} PTS\left(\beta^-, \frac{\beta^-\kappa c^-}{nm}, \lambda^-\right) \\ Z_2^+ \overset{\mathcal{L}}{\sim} PTS\left(\beta^+ - 1, \frac{\lambda^+\kappa c^+}{nm}, \lambda^+\right), \\ Z_2^- \overset{\mathcal{L}}{\sim} PTS\left(\beta^- - 1, \frac{\lambda^-\kappa c^-}{nm}, \lambda^-\right). \end{cases} \tag{30}$$

We set $m = 1$ and we do 10,000 Monte Carlo replications of $\{v_t\}_{t\in[0,K]}$ on the discrete grid $0, \frac{1}{n}, \frac{2}{n}, \ldots, K$.

7. Empirical results from parametric volatility estimation

We next turn to the results from the estimation of the parametric volatility model. We use 5 min level data on the S&P 500 futures index covering the period January 1, 1990, to December 31, 2008, for a total of 4750 days. Each day has 80 high-frequency returns. We start with initial analysis of the data.

7.1. Initial data analysis

Using the high-frequency data, we form a nonparametric measure for the daily integrated variance, $\int_{t-1}^{t} V_s ds$. We use the truncated variance, proposed originally by Mancini (2009), which we implement here in the following way:

$$TV_{[t-1,t]}(\alpha, \varpi) = \sum_{i=n(t-1)+1}^{nt} |\Delta_i^n X|^2 1_{\{|\Delta_i^n X| \leq \alpha n^{-\varpi}\}},$$
$$\alpha > 0,\ \varpi \in (0, 1/2), \tag{31}$$

where here we use $\varpi = 0.49$, i.e., a value very close to $1/2$, and we further set $\alpha = 3 \times \sqrt{BV_{[t-1,t]}}$ for $BV_{[t-1,t]}$ denoting the bipower variation of Barndorff-Nielsen and Shephard (2004) over the day (which is another consistent estimator of the integrated variance in the presence of jumps):

$$BV_{[t-1,t]} = \frac{\pi}{2}\sum_{i=n(t-1)+2}^{nt} |\Delta_{i-1}^n X||\Delta_i^n X|. \tag{32}$$

Using the truncated variance, we estimate the mean of V_t for our data to be 0.8704 with a 95% confidence interval of [0.6057 1.1350].[22] The mean of the total quadratic variation (i.e., including the contribution from the jumps) in our sample is 1.0021 which implies a contribution of price jumps in the total price variation consistent with those of previous empirical work.

To reduce computational time in the parametric estimation, we implement variance targeting. We do this by imposing a prior support restriction on the parameter vector that ensures that the model-implied mean volatility is in the estimated nonparametric 95% confidence interval of [0.6057 1.1350] (recall that $\mathbb{E}(V_t)$ is known in closed form in our model).

7.2. Estimation results

We continue next with presenting the results from the parametric estimation. As is well-known, to capture volatility persistence we need a two-factor volatility structure; see, e.g., Andersen et al. (2002) and Chernov et al. (2003). Therefore, we estimate two-factor extensions of the model (15)–(16) which differ in the way in which the volatility (jump) activity is modeled. These specifications are:

- Exp-OU diffusion: **D**: $\beta_i^{\pm} = c_i^{\pm} = \lambda_i^{\pm} = 0, i = 1, 2$.
- Exp-OU positive jumps: **RJ**: $\sigma_i = \beta_i^- = c_i^- = \lambda_i^- = 0, i = 1, 2$ and $\beta_2^+ = \beta_1^+, c_2^+ = \phi c_1^+, \lambda_2^+ = \lambda_1^+$.
- Exp-OU symmetric activity finite variation jumps: **SJ-FV**: $\sigma_1 = \sigma_2 = 0$, $\beta_1^+ = \beta_1^- = \beta_2^+ = \beta_2^-, c_2^+ = c_2^- = \phi c_1^+ = \phi c_1^-$, $\lambda_2^+ = \lambda_1^+, \lambda_2^- = \lambda_1^-$ and $\beta_i^{\pm} < 1$ for $i = 1, 2$.
- Exp-OU symmetric activity jumps: **SJ**: $\sigma_1 = \sigma_2 = 0$, $\beta_1^+ = \beta_1^- = \beta_2^+ = \beta_2^-, c_2^+ = c_2^- = \phi c_1^+ = \phi c_1^-, \lambda_2^+ = \lambda_1^+, \lambda_2^- = \lambda_1^-$.

In specification **D**, volatility does not contain jumps, while in specification **RJ**, volatility moves only through positive jumps whose activity is unrestricted and negative jumps are absent. In model **SJ** there are both positive and negative jumps, with the small ones being symmetric while the big ones can be asymmetric. Finally, **SJ-FV** is a restricted version of **SJ** in which jumps are constrained to be of finite variation. We recall that all estimated specifications "face" the same moments from the high-frequency data.

[22] The asymptotic standard error was computed using a Parzen kernel with a lag length of 70.

Table 1
Estimation results for parametric volatility models.

Parameter	**D**	**RJ**	**SJ-FV**	**SJ**
μ	−0.5968 (0.0672)	−0.9302 (0.0997)	−0.5198 (0.0799)	−0.6986 (0.0699)
κ_1	0.0099 (0.0040)	0.0191 (0.0019)	0.0322 (0.0018)	0.0172 (0.0011)
κ_2	2.4703 (0.2873)	2.5693 (0.7257)	2.4405 (0.2835)	2.5758 (0.2469)
β^+		1.9259 (0.1833)	0.9787 (0.0397)	1.8958 (0.0468)
c^+		0.0455 (0.0033)	0.1914 (0.0061)	0.0335 (0.0011)
λ^+		1.3201 (0.2718)	1.1148 (0.0464)	2.5894 (0.1315)
β^-			$\equiv \beta^+$	$\equiv \beta^+$
c^-			$\equiv c^+$	$\equiv c^+$
λ^-			1.3354 (0.0958)	3.0836 (0.1964)
σ_1	0.7891 (0.3233)			
σ_2	0.5993 (0.0460)			
ϕ		0.7206 (0.0574)	1.5003 (0.0186)	0.7884 (0.0253)
J test (df)	45.80 (6)	17.74 (4)	72.48 (3)	13.28 (3)

Note: The estimation is based on the minimum-distance estimator in (26) with 11 moment conditions given in Section 6. The optimal weight matrix was computed using the Parzen kernel and a lag length of 70. Standard errors for the parameter estimates are reported in parentheses.

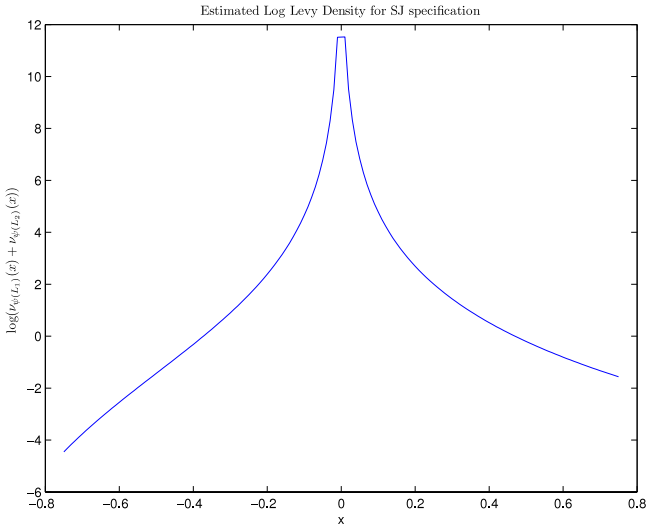

Fig. 6. The figure plots the estimated log Lévy density $\nu_{\psi(L_1)} + \nu_{\psi(L_2)}$ for the model specification **SJ** using the parameter estimates reported in the last column of Table 1.

The results from the parametric estimation are given in Table 1. We start our discussion with the standard diffusive log-volatility model, i.e., our **D** model specification. The results are given in the first column of the table. The overall fit of the model is relatively poor as signaled by the big values of the corresponding J statistics. This model is found to have good performance in full-parametric estimation based on daily frequency in Chernov et al. (2003). We recall that our estimation method is based on separating price jumps from volatility nonparametrically using the high-frequency data and then fitting the parametric volatility model. This robustifies the inference and sharpens the precision of the estimation compared with estimation based on daily frequency where the model needs to "separate" stochastic volatility from jumps. When the high-frequency data is used in an efficient way as is done here, the fit of the diffusive log-volatility model is found to be relatively poor.

The second column in Table 1 presents next the results for the **RJ** specification. This model is similar to the non-Gaussian OU model of Barndorff-Nielsen and Shephard (2001) where volatility is driven by positive jumps only. The generalization here is that the jump activity is allowed to be in the interval $[0, 2)$ while in the above-mentioned model jumps are restricted to being of finite variation (i.e., activity always smaller than 1). As we see from the overall fit, the **RJ** specification outperforms significantly the pure-diffusive one. Looking at the persistence parameters estimates for the two models **D** and **RJ**, we see that they are quite similar, indicating that both models capture the well-known memory features of volatility. The improved fit of **RJ** comes with two additional parameters — one which controls the Lévy measure around zero, i.e., the jump activity, and another one that controls the tail of the Lévy measure. The diffusive log-volatility model constrains parametrically the big and small volatility moves by one parameter, σ, and this parametric link is clearly not supported by the data. Also, looking at the estimate of the parameter β^+, it is interesting to note that our parametric estimation implies volatility jumps of infinite variation with activity level similar to that found by nonparametric methods using VIX index data; see Section 4.

The third column of Table 1 contains the results for the specification **SJ-FV**. We recall that in this specification jumps can be of arbitrary sign, but are restricted to being of finite variation as in the non-Gaussian OU model of Barndorff-Nielsen and Shephard (2001). As seen from the corresponding J-test, the fit of the model is poor although it contains more parameters than specification **RJ**. This can be associated with the restriction on the jump activity. In the last estimated specification **SJ** we remove this restriction and this yields the best fit across all estimated specifications to the moments from the high-frequency data. As for the specification **RJ**, the estimated jump activity is well above 1, indicating volatility jumps of infinite variation, exactly as our nonparametric analysis using the VIX index indicated.

We recall from our discussion in Section 5 that the diffusive specification **D** can be seen as a limiting case of **SJ** (when $\beta^{\pm}$ approaches 2). The estimation results here suggest that the preferred model is one in which volatility activity is below 2. Comparing the J-tests of **RJ** and **SJ** specifications we see that the difference in their performance is not very big. This indicates that our nonparametric statistics from the high-frequency price data do not penalize heavily for the omission of negative volatility jumps.

Finally, on Fig. 6, we plot the implied Lévy density of the volatility jumps via the parameter estimates of the **SJ** specification. As seen from the figure, the positive and negative big jumps are only mildly asymmetric.

In conclusion, we point out that even our best performing model specification **SJ** is formally rejected by the J-test. This suggests that further improvements are potentially possible. One possible extension is to relax the connection between the Lévy measures of the two volatility factors. This could be done with a richer data set more informative about the different behavior of the two volatility factors over small and large scales. We leave such further extensions and estimation of the models for future work.

8. Conclusion

The paper studies asymmetric volatility activity, i.e., the possibility that small up and down moves in volatility have different intensities. Nonparametric evidence in Todorov and Tauchen (2011) suggests that volatility is pure-jump, i.e., evolving only through jumps, and when this is the case the asymmetric volatility activity maps into asymmetric Lévy density around zero of the volatility jumps. We show in this paper how asymmetric volatility activity can be detected using high-frequency VIX index data, under certain assumptions for the volatility risk premium, and our empirical analysis based on this data suggests approximate symmetry of up and down volatility moves (with up moves slightly more active). We then look at the same question about the volatility activity but

using only high-frequency price data together with a general parametric volatility model that overcomes the restrictions on jumps (and in particular their activity) in the extant models. We show that the price data alone contains information regarding volatility activity and our parametric evidence is generally supportive of the findings using the VIX index data. The price data also clearly suggests that small and big volatility moves should be modeled separately.

Appendix A. Proof of Theorem 1

We prove only the case $\beta^+ > \beta^-$ and the other cases are analyzed in exactly the same way (the case $\beta^+ = \beta^-$ and $c^+ = c^-$ have been already shown in Todorov and Tauchen (2010)). In what follows C will denote a positive constant that does not depend on n and can change from line to line.

First, we do a localization similar to that done in Jacod (2008) by assuming that α'_s is bounded in absolute value and σ_s is bounded both from below and above by some positive constant. Proving the result for the general case (when α'_s and σ_s are only locally bounded) can be done as in Lemma 4.6 of Jacod (2008).

We start with introducing some notation. For some $\tau > 0$, we define

$$Y_t(\tau) = \int_0^t \alpha'_s ds + \int_0^t \int_{|x|\le\tau} \sigma_{s-} x \widetilde{\mu}'(ds, dx). \tag{33}$$

For arbitrary process Z we define $Z_{t,n} = Z_t - Z_{(i-1)\Delta_n}$ for $t \in [(i-1)\Delta_n, i\Delta_n]$. With this notation we can write on an extension of the original probability space

$$\begin{aligned} Y_{t,n}(\tau) &= \int_{(i-1)\Delta_n}^t \alpha'_s ds \\ &+ \int_{(i-1)\Delta_n}^t \int_{|x|\le\tau} \left(\sigma_{s-} - \sigma_{(i-1)\Delta_n-}\right) x \widetilde{\mu}'(ds, dx) + \sigma_{(i-1)\Delta_n-} S^+_{t,n} \\ &+ \sigma_{(i-1)\Delta_n-} \left(L^-_{t,n} + L^{3,+}_{t,n} - L^{1,+}_{t,n} - L^{2,+}_{t,n}\right), \\ &t \in [(i-1)\Delta_n, i\Delta_n], \end{aligned} \tag{34}$$

where S^+_t, $L^{1,+}_t$, $L^{2,+}_t$, $L^{3,+}_t$ and L^-_t are pure-jump Lévy processes with zero drift and Lévy measures respectively: (1) $\frac{c^+}{|x|^{1+\beta^+}} 1_{\{x>0\}}$ for S^+_t, (2) $-2\nu''(x)1(x : \nu''(x) < 0, x \in (0,\tau])$ for $L^{1,+}_t$, (3) $\frac{c^+}{|x|^{1+\beta^+}} 1(x > \tau)$ for $L^{2,+}_t$, (4) $|\nu''(x)|1(0 < x \le \tau)$ for $L^{3,+}_t$, (5) $\left(\frac{c^-}{|x|^{1+\beta^-}} + \nu''(x)\right) 1_{\{x<0\}}$ for L^-_t. The processes $S^+_t, L^{1,+}_t, L^{2,+}_t$ and $L^{3,+}_t$ can have dependence between them.

We carry out the decomposition

$$\begin{aligned} &\Delta_n^{1-p/\beta^+} V^+_T(Y,p) - K^+(p,\beta^+) \sum_{i=1}^{nT} |\sigma_{(i-1)\Delta_n-}|^p \\ &= A^{(1)}_T + A^{(2)}_T + A^{(3)}_T + A^{(4)}_T, \end{aligned} \tag{35}$$

$$\begin{aligned} A^{(1)}_T &= \Delta_n^{1-p/\beta^+} \sum_{i=1}^{nT} \left(|\Delta^n_i Y|^p - |\Delta^n_i Y(\tau)|^p\right) 1(\Delta^n_i Y > 0), \\ A^{(2)}_T &= \Delta_n^{1-p/\beta^+} \sum_{i=1}^{nT} |\Delta^n_i Y(\tau)|^p \left[1(\Delta^n_i Y > 0)\right. \\ &\left. -1(\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n} > 0)\right], \\ A^{(3)}_T &= \Delta_n^{1-p/\beta^+} \sum_{i=1}^{nT} \left\{|\Delta^n_i Y(\tau)|^p \right. \\ &\left. - |\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n}|^p\right\} 1(\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n} > 0), \\ A^{(4)}_T &= \Delta_n^{1-p/\beta^+} \sum_{i=1}^{nT} |\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n}|^p 1(\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n} > 0) \\ &- \Delta_n K^+(p,\beta^+) \sum_{i=1}^{nT} |\sigma_{(i-1)\Delta_n-}|^p. \end{aligned} \tag{36}$$

Since $\Delta_n \sum_{i=1}^{nT} |\sigma_{(i-1)\Delta_n-}|^p$ converges pathwise to $\int_0^T |\sigma_s|^p ds$, to prove the result in (12), it suffices to prove the asymptotic negligibility of each of the terms $A^{(j)}_T$ for $j = 1, \ldots, 4$.

We start with $A^{(1)}_T$. Using the fact that $p \in (0,2)$, and the boundedness of the jumps of $Y(\tau)$ as well as that of the processes α'_s and σ_s, we have

$$\begin{aligned} |A^{(1)}_T| &\le C\Delta_n^{1-p/\beta^+} \sup_{i=1,\ldots,nT} |\Delta^n_i Y(\tau)| \\ &\times \left(\sum_{s\le T} (|\Delta Y_s|^p + |\Delta Y_s|) 1(|\Delta Y_s| > \tau) + C\right), \end{aligned} \tag{37}$$

and since $\sum_{s\le T} (|\Delta Y_s|^p + |\Delta Y_s|) 1(|\Delta Y_s| > \tau)$ is bounded in probability and so is $\sup_{i=1,\ldots,nT} |\Delta^n_i Y(\tau)|$ (as the jumps of $Y(\tau)$ are bounded), we have

$$A^{(1)}_T \xrightarrow{\mathbb{P}} 0. \tag{38}$$

We turn next to $A^{(2)}_T$. First, using the Burkholder–Davis–Gundy inequality, we have

$$\mathbb{E}|\Delta^n_i Y(\tau)|^{\beta^+ + \iota} \le C\Delta_n \int_{|x|\le\tau} |x|^{\beta^+ + \iota} \nu'(dx) < C\Delta_n, \quad \forall \iota > 0. \tag{39}$$

Second, we have

$$\begin{aligned} &\left|1(\Delta^n_i Y > 0) - 1(\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n} > 0)\right| \\ &\le 1\left(|\Delta^n_i Y - \sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n}| > |\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n}|\right). \end{aligned}$$

Therefore, we can bound $|A^{(2)}_T| \le |A^{(2,a)}_T| + |A^{(2,b)}_T|$ where

$$A^{(2,a)}_T = \Delta_n^{1-p/\beta^+} \sum_{i=1}^{nT} |\Delta^n_i Y(\tau)|^p 1\left(|\sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n}| < \Delta^\alpha_n\right), \tag{40}$$

$$\begin{aligned} A^{(2,b)}_T &= \Delta_n^{1-p/\beta^+} \sum_{i=1}^{nT} |\Delta^n_i Y(\tau)|^p 1\left(|\Delta^n_i Y \right. \\ &\left. - \sigma_{(i-1)\Delta_n-} S^+_{i\Delta_n,n}| > \Delta^\alpha_n\right), \end{aligned} \tag{41}$$

for some α that satisfies $\min\{1/\beta^-, 1/\beta'\} > \alpha > 1/\beta^+$ (this is possible because of the relation between $\beta^\pm$ and β' assumed in the theorem). For $A^{(2,a)}_T$, using that the density of the stable law is bounded, we have by an application of Hölder's inequality and the bound in (39)

$$\mathbb{E}|A^{(2,a)}_T| \le C\Delta_n^{-p/\beta^+ + p/(\beta^+ + \iota) + (1-p/(\beta^+ + \iota))(\alpha - 1/\beta^+)}, \quad \forall \iota > 0. \tag{42}$$

For $A^{(2,b)}_T$ we first derive several auxiliary bounds. Using the boundedness of α'_s and Burkholder–Davis–Gundy inequality, we have

$$\begin{aligned} &\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} \alpha'_s ds\right| \le C\Delta_n, \\ &\mathbb{E}\left(|L^{1,+}_{i\Delta_n,n}|^{\beta'+\iota} + |L^{3,+}_{i\Delta_n,n}|^{\beta'+\iota}\right) \le C\Delta_n, \quad \forall \iota > 0. \end{aligned} \tag{43}$$

Using that σ_t is an Itô semimartingale (and upon making an additional localization to bound the coefficients describing the dynamics of σ_t), we next have

$$\begin{aligned} &\mathbb{E}\left|\int_{(i-1)\Delta_n}^{i\Delta_n} \int_{|x|\le\tau} (\sigma_{s-} - \sigma_{(i-1)\Delta_n-}) x \tilde{\mu}(ds, dx)\right|^{\beta^+ + \iota} \\ &\le C\Delta_n^{1+\beta^+/2+\iota/2}, \quad 0 < \iota < 2 - \beta^+. \end{aligned} \tag{44}$$

We also have for arbitrary constant $\chi > 0$ and arbitrary small $0 < \iota < 1$ (recall that the Lévy density of $L^{2,+}_t$ is zero around zero)

$$\mathbb{P}\left(|\sigma_{(i-1)\Delta_n-} L^{2,+}_{i\Delta_n,n}| > \chi\Delta^\alpha_n\right) \le C\Delta_n^{1-\iota}. \tag{45}$$

Similarly, for $0 < \iota < 1/\alpha - \max\{\beta^-, \beta'\}$ and arbitrary constant $\chi > 0$, we have

$$\mathbb{P}\left(|\sigma_{(i-1)\Delta_n-}L^-_{i\Delta_n,n}| > \chi\Delta_n^\alpha\right) \le C\Delta_n^{1-\alpha(\max\{\beta^-,\beta'\}+\iota)}. \tag{46}$$

Application of Hölder's inequality, together with the bounds in (39) and (43)–(46), then gives

$$\mathbb{E}|A_T^{(2,b)}| \le C\Delta_n^{-p/\beta^+ + p/(\beta^+ + \iota) + (1-p/(\beta^+ + \iota))(1-\alpha(\max\{\beta^-,\beta'\}+\iota'))},$$
$$\forall \iota,\ \iota' > 0. \tag{47}$$

From here, upon choosing appropriately α, we have altogether for some $\eta > 0$

$$\mathbb{E}|A_T^{(2,a)}| + \mathbb{E}|A_T^{(2,b)}| \le C\Delta_n^\eta, \tag{48}$$

which implies asymptotic negligibility of $A_T^{(2)}$. We next turn to $A_T^{(3)}$. We make use of the following algebraic inequality:

$$\|x|^p - |y|^p| \le C|x-y|^p + C|x-y||x|^{p-1}1_{\{p>1\}}, \tag{49}$$

to bound $|A_T^{(3)}| \le C(A_T^{(3,a)} + A_T^{(3,b)})$ where

$$\begin{aligned} A_T^{(3,a)} &= \Delta_n^{1-p/\beta^+}\sum_{i=1}^{nT}|\Delta_i^n Y(\tau) - \sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n}|^p, \\ A_T^{(3,b)} &= \Delta_n^{1-p/\beta^+}\sum_{i=1}^{nT}|\sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n}|^{p-1}|\Delta_i^n Y(\tau) \\ &\quad - \sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n}|1_{\{p>1\}}. \end{aligned} \tag{50}$$

For $A_T^{(3,a)}$, using the bounds in (43)–(46), we have

$$\mathbb{E}|A_T^{(3,a)}| \le C\Delta_n^{\left(\frac{p}{\max\{\beta^-,\beta'\}}\wedge 1\wedge p\right)-p/\beta^+-\iota}, \quad \forall \iota > 0. \tag{51}$$

This guarantees asymptotic negligibility of $A_T^{(3,a)}$ because $\max\{\beta^-, \beta'\} < \beta^+$. For $A_T^{(3,b)}$, we apply Hölder's inequality (with powers $p/(p-1)$ and p for the terms $|\sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n}|^{p-1}$ and $|\Delta_i^n Y(\tau) - \sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n}|$ respectively) and use the bounds in (43)–(46) to get

$$\mathbb{E}|A_T^{(3,b)}| \le C\Delta_n^{\left(\frac{p}{\max\{\beta^-,\beta'\}}\wedge 1\wedge p\right)\frac{1}{p}-\frac{1}{\beta^+}-\iota}, \quad \forall \iota > 0. \tag{52}$$

We are left with $A_T^{(4)}$. Using the definition of $K^+(p,\beta^+)$ in the theorem, we have (note $\sigma_{(i-1)\Delta_n-}$ is positive)

$$\begin{aligned} &\mathbb{E}\left\{|\sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n}|^p 1\left(\sigma_{(i-1)\Delta_n-}S^+_{i\Delta_n,n} > 0\right)\Big|\mathcal{F}_{(i-1)\Delta_n}\right\} \\ &= \Delta_n^{p/\beta^+}K^+(p,\beta^+)|\sigma_{(i-1)\Delta_n-}|^p. \end{aligned}$$

We also have $\mathbb{E}_{i-1}|S^+_{i\Delta_n,n}|^q \le C\Delta_n^{q/\beta^+}$ for any $q < \beta^+$, and together with the boundedness of σ_s we have the convergence to zero of $A_T^{(4)}$ by an application of Theorem VIII.2.27 in Jacod and Shiryaev (2003). □

Appendix B. Simulation of tempered stable process

In this section we provide details on the simulation of the increments of a tempered stable process which is needed in evaluating the model-implied volatility joint Laplace transform in Section 6. It obviously suffices to consider only the simulation of the process evaluated at time 1 and for its positive-jump part, i.e., the variable $Z \overset{\mathcal{L}}{\sim} PTS(\beta, c, \lambda)$ (recall our notation in (27)), and we do so henceforth. If we denote with L_t^{pts} a Lévy process with law at time $t = 1$ given by $PTS(\beta, c, \lambda)$, then to simulate the Lévy process over arbitrary interval t, we just use the fact that $L_t^{\text{pts}} \overset{\mathcal{L}}{\sim} PTS(\beta, c\times t, \lambda)$.

The simulation is based on the acceptance–rejection algorithm proposed by Baeumer and Meerschaert (2009) with the stable distribution used as the proposal density and the exponential distribution used to temper it appropriately. The method is exact when $\beta < 1$ and approximate when $\beta \ge 1$. Kawai and Masuda (2010) provide evidence that the method is efficient (in terms of computational time) and the approximation error can be easily controlled without significant loss of computational time.

The method works very well for small time intervals (which is the case for our application), because over small time intervals, the leading component of the tempered stable is its stable part (i.e., one can show, with the notation as in the previous paragraph, $\frac{L_t^{\text{pts}}}{t^{1/\beta}} \overset{\mathcal{L}}{\longrightarrow} S(\beta,\tilde{c})$ where $S(\beta,\tilde{c})$ is a spectrally positive β-stable distribution with $\tilde{c}$ a scaling parameter which is a function of the parameters of the original tempered stable process). Hence, the rejection rate of the algorithm can be kept relatively low. Alternative is the method of Rosiński (2007) based on the shot-noise decomposition of the Lévy measure which involves simulation of the actual jumps. To achieve reasonable accuracy, however, this method requires simulation of too many jumps particularly for β close to 2 which is computationally costly.

Given the discussion at the beginning of the section, to keep the rejection rates in the simulation relatively low, for the simulation of the random variables Z_0^+ and Z_0^-, we use the fact that the distributions are infinitely divisible and we split them into n subincrements (where recall that n is the number of high-frequency subintervals in a unit interval — this is equivalent to viewing $Z_0^\pm$ as a Lévy process evaluated at time 1 and simulating its value by summing its n high-frequency subincrements).

The simulation of the tempered stable distribution via the acceptance–rejection method of Baeumer and Meerschaert (2009) is done in the following way depending on the value of β:

- Case $\beta < 0$:

$$Z = \sum_{j=1}^N Y_j - c\lambda^{\beta-1}\Gamma(1-\beta),$$
$$N \sim \text{Poisson distribution with intensity}$$
$$c\lambda^\beta\Gamma(-\beta) \text{ and } Y_j \sim G(-\beta,\lambda), \tag{53}$$

  where $G(a,b)$ stands for the Gamma distribution with probability density: $\frac{b^a x^{a-1}}{\Gamma(a)}e^{-bx}1_{\{x>0\}}$ for $a, b > 0$.
- Case $\beta = 0$:

$$Z \overset{\mathcal{L}}{\sim} G(c,\lambda) - \frac{c}{\lambda}, \tag{54}$$

  where $G(a,b)$ stands for the Gamma distribution as defined above.
- Case $\beta \in (0,1)$:
  (a) Simulate U uniform on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and W exponential with mean 1, and set $\theta = \arctan(\tan(\beta\pi/2))$,
  (b) set

$$V = [-c\Gamma(-\beta)\cos(\beta\pi/2)]^{1/\beta}\frac{\sin(\beta U+\theta)}{(\cos(U)\cos(\theta))^{1/\beta}} \times\left[\frac{\cos[(1-\beta)U-\theta]}{W}\right]^{(1-\beta)/\beta}, \tag{55}$$

  (c) simulate Y exponential with mean $1/\lambda$,
  (d) if $Y < V$, go back to (a); otherwise return $Z = V - c\Gamma(1-\beta)\lambda^{\beta-1}$.
- Case $\beta = 1$: set a tuning parameter $a = 5\times[-c\Gamma(-\beta)\cos(\beta\pi/2)]^{1/\beta}$.
  (a) Simulate U uniform on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and W exponential with mean 1, and set $\theta = \arctan(\tan(\beta\pi/2))$,

(b) set

$$V = c\left\{\log\left(\frac{\pi c}{2}\right) + \left(\frac{\pi}{2} + U\right)\tan(U) - \log\left(\frac{\frac{\pi}{2}W\cos(U)}{\frac{\pi}{2} + U}\right)\right\}, \quad (56)$$

(c) simulate Y exponential with mean $1/\lambda$,
(d) if $Y < V + a$, go back to (a); otherwise return $Z = V + c\,(1 + \log(\lambda))$.

- Case $\beta \in (1, 2)$: set a tuning parameter $a = 5 \times [-c\Gamma(-\beta)\cos(\beta\pi/2)]^{1/\beta}$.
  (a) Simulate U uniform on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and W exponential with mean 1, and set $\theta = \arctan(\tan(\beta\pi/2))$,
  (b) set

$$V = [-c\Gamma(-\beta)\cos(\beta\pi/2)]^{1/\beta} \times \frac{\sin(\beta U + \theta)}{(\cos(U)\cos(\theta))^{1/\beta}}\left[\frac{\cos[(1-\beta)U - \theta]}{W}\right]^{(1-\beta)/\beta}, \quad (57)$$

  (c) simulate Y exponential with mean $1/\lambda$,
  (d) if $Y < V + a$, go back to (a); otherwise return $Z = V - c\Gamma(1-\beta)\lambda^{\beta-1}$.

Baeumer and Meerschaert (2009) show that for $a \uparrow \infty$ (a being the tuning parameter in the last two cases, i.e., when $\beta \geq 1$), the Kolmogorov–Smirnov distance between the generated random variable and the targeted tempered stable one converges to zero. Here, we have set the tuning parameter a in the case $\beta \geq 1$ such that the probability that the stable variable $V < -a$ is negligible. Numerical experiments, not reported in the paper, showed that for plausible values of λ (close to the estimates reported in Section 7), the resulting approximation error is very small.

References

Ait-Sahalia, Y., Jacod, J., 2009. Estimating the degree of activity of jumps in high frequency financial data. Annals of Statistics 37, 2202–2244.

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

Baeumer, B., Meerschaert, M., 2009. Tempered stable levy motion and transit superdiffusion. Journal of Computational and Applied Mathematics 223, 2438–2448.

Barndorff-Nielsen, O., Shephard, N., 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. Journal of the Royal Statistical Society Series B 63, 167–241.

Barndorff-Nielsen, O., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics 2, 1–37.

Barndorff-Nielsen, O.E., Kinnebrock, S., Shephard, N., 2010. Measuring downside risk — realised semivariance. In: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics, Essays in Honor of Robert F. Engle. Oxford University Press.

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

Carrasco, M., Chernov, M., Florens, J., Ghysels, E., 2007. Efficient estimation of general dynamic models with a continuum of moment conditions. Journal of Econometrics 140, 529–573.

Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Cont, R., Tankov, P., 2004. Financial modelling with jump processes. In: Boca Raton. Chapman and Hall, Florida, USA.

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

Haug, S., Czado, C., 2007. An exponential continuous time GARCH process. Journal of Applied Probability 44 (4), 1–18.

Jacod, J., 2008. Asymptotic properties of power variations and associated functionals of semimartingales. Stochastic Processes and their Applications 118, 517–559.

Jacod, J., Shiryaev, A.N., 2003. Limit Theorems for Stochastic Processes, second ed. Springer-Verlag, Berlin.

Kawai, R., Masuda, H., 2010. On Simulation of Tempered Stable Random Variates. Working paper.

Mancini, C., 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics 36, 270–296.

Rosiński, J., 2007. Tempering stable processes. Stochastic Processes and their Applications 117, 677–707.

Rosinski, J., Samorodnitsky, G., 1993. Distributions of subadditive functionals of sample paths of infinite divisible processes. Annals of Probability 21, 996–1014.

Sato, K., 1999. Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge, UK.

Singleton, K., 2006. Empirical Dynamic Asset Pricing. Princeton University Press.

Todorov, V., Tauchen, G., 2010. Activity signature functions for high-frequency data analysis. Journal of Econometrics 154, 125–138.

Todorov, V., Tauchen, G., 2011. Volatility jumps. Journal of Business and Economic Statistics 29, 356–371.

Todorov, V., Tauchen, G., 2012. The realized Laplace transform of volatility. Econometrica 80, 1105–1127.

Todorov, V., Tauchen, G., Grynkiv, I., 2011. Realized Laplace transforms for estimation of jump diffusive volatility models. Journal of Econometrics 164, 367–381.