Journal of Econometrics 164 (2011) 367–381

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

Realized Laplace transforms for estimation of jump diffusive volatility models☆

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:
Received 12 September 2010
Received in revised form
7 June 2011
Accepted 13 June 2011
Available online 21 June 2011

JEL classification:
C51
C52
G12

Keywords:
Jumps
High-frequency data
Laplace transform
Stochastic volatility models

ABSTRACT

We develop an efficient and analytically tractable method for estimation of parametric volatility models that is robust to price-level jumps. The method entails first integrating intra-day data into the Realized Laplace Transform of volatility, which is a model-free estimate of the daily integrated empirical Laplace transform of the unobservable volatility. The estimation is then done by matching moments of the integrated joint Laplace transform with those implied by the parametric volatility model. In the empirical application, the best fitting volatility model is a non-diffusive two-factor model where low activity jumps drive its persistent component and more active jumps drive the transient one.

1. Introduction

Stochastic volatility and price-level jumps are two well-recognized features of asset price dynamics that differ in economically significant ways. Given the substantial compensations for these two risks demanded by investors, as evident from option prices, it is of central importance to better understand their dynamic characteristics. When using coarser sampling of stock prices such as daily data, the separation of volatility from price-level jumps becomes relatively difficult, and, most important, it depends crucially on the correct modeling of all aspects in the asset model; misspecification of any one feature of the model for the asset dynamics can lead to erroneous evidence about the role and significance of each of those risks. On the other hand, the availability of high-frequency data provides a model-free way of disentangling the key features of the price dynamics.

Price level jumps have recently been studied very extensively.[1] There is substantial parametric and non-parametric evidence for rare, but very sharp price movements, as would be expected from a compound Poisson process. There is also some evidence for smaller, more vibrant price-level jumps, as predicted by models built on Lévy processes with so-called Blumenthal–Getoor indexes above zero.[2] This evidence is documented by Ait-Sahalia and Jacod (2009a) for liquid individual equity prices, by Todorov and Tauchen (2011a) for index options prices, see also the references therein.

The aim of this paper is to remain robust to these price jumps while developing an enhanced methodology for understanding the latent volatility dynamics. The robustness is achieved by using high-frequency data in conjunction with techniques that effectively filter out the irrelevant parts of returns.

☆ We would like to thank the editor, an Associate Editor, a referee as well as Torben Andersen, Tim Bollerslev, Andrew Patton, and many conference and seminar participants for numerous suggestions and comments. Todorov's work was partially supported by NSF grant SES-0957330.

* Corresponding author. Tel.: +1 919 660 1812.
E-mail addresses: v-todorov@northwestern.edu (V. Todorov), george.tauchen@duke.edu (G. Tauchen), ig9@duke.edu (I. Grynkiv).

0304-4076/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2011.06.016

[1] See, for example, Barndorff-Nielsen and Shephard (2004, 2006) and Ait-Sahalia and Jacod (2009b), among many others, with evaluations of jump tests in Huang and Tauchen (2005), and quite comprehensively in Theodosiou and Zikes (2009).

[2] All Lévy processes can be divided into equivalence classes according to the Blumenthal–Getoor index (Blumenthal and Getoor, 1961), which lies in [0,2). The index of the relatively quiescent variance gamma process is 0; that of the Cauchy process is 1.0; and the indices of more vibrant processes are closer to the upper bound of 2.0.

The newly developed estimation method is for parametric volatility models. Restriction to parametric models is motivated by the fact we only directly observe financial prices, not the volatility, so the information regarding volatility dynamics is embedded more deeply into the process. Extracting this information pathwise is, in effect, a de-convolution effort that can be done nonparametrically only at a rate equal in general to the fourth root of the sample size, which is slow. It might be thought that a way around this problem is to use volatility-sensitive financial derivatives, e.g., options, but that is not generally the case. As noted by Todorov and Tauchen (2011a), a market-based volatility index based on a portfolio of option prices such as the VIX is actually a two-level convolution of the underlying volatility process; one convolution is the formation of the forward integrated variance while the other is the integration implicit in market's computation of the risk-neutral expectation of the forward variance. The risk-neutral volatility expectation contains a rather non-trivial volatility risk premium, and hence we need a model for the latter before we use volatility derivatives in the estimation. Thus, it seems the only way to discriminate more sharply across models, without imposing assumptions on the pricing of risk, is to use a relatively efficient parametric method that exploits the full strength of available high-frequency data.

Candidate volatility models and estimators already pervade the literature, but the evidence to date on the empirically most credible model is not very conclusive. The list of potential models includes, for example, purely diffusive affine models, affine jump diffusions, and pure jump models driven by Lévy processes of various activity levels. The evidence, however, at best suggests only the general features of the appropriate model. It seems clear that volatility is comprised of at least two factors, one very slow moving and the other quickly mean reverting, and there is also some evidence for volatility jumps. On these points see the findings regarding volatility factors in Bollerslev and Zhou (2002) and Barndorff-Nielsen and Shephard (2002) who use high-frequency data, along with Andersen et al. (2002) and Chernov et al. (2003) for earlier evidence from low frequency data; for volatility jumps see Todorov and Tauchen (2011a) along with earlier evidence using low frequency data provided by Eraker et al. (2003). Though suggestive, these findings taken together do not define a well-specified parsimonious model that jointly captures the distributional and dynamic characteristics of volatility. This rather hazy view of volatility dynamics should not be surprising, given the above-noted fact that volatility dynamics are embedded so deeply into the financial price process.

Our strategy to sharpen estimation precision entails a new moment-based estimator applied to daily aggregates of trigonometric transforms of the high-frequency returns; specifically, we use the Realized Laplace transform of volatility proposed by Todorov and Tauchen (2011b). This computation is a simple sum of cosine transformations of the high-frequency increments, and it is an empirical measure that embodies more information regarding volatility than the now-conventional measures of daily variability. Todorov and Tauchen (2011b) show that the Realized Laplace transform is a *nonparametric*, consistent and asymptotically normal estimate of the *unobservable* integrated Laplace transform of volatility over a fixed interval of time. The above paper derives the asymptotic behavior of this nonparametric measure when joint long-span fill-in asymptotics is used. Importantly, this measure is robust to the presence of jumps in the price, a desideratum discussed above. In this paper we build on the above nonparametric results and show how they can be used for efficient and robust *parametric* estimation of the stochastic volatility process.

As seen in Section 3 below, the Realized Laplace transform conveys information about the (joint) Laplace transform of volatility, and similarly the candidate volatility models are most easily expressed in terms of their characteristic functions.[3]

Thus some aspects of our estimation methodology are in common with the long tradition of estimation based on the characteristic function (Parzen, 1962) and conditional characteristic functions as in Feuerverger and Mureika (1977), Singleton (2001), Jiang and Knight (2002), Yu (2004), Bates (2006) and Carrasco et al. (2007). For example, we face issues similar to those of the continuum of moment conditions and also the numerical difficulties associated with computing inverse integral transforms. However, an important difference with prior work comes from the fact that we use the high-frequency data to "integrate out" in a *model-free way* the components of the process not directly linked to volatility such as the driving martingales in the price-level and the price jumps. So, unlike previous work using the empirical characteristic function, we cannot work directly with the empirical joint Laplace transform of the process of interest (the latent volatility) but rather we use a daily integrated version of it.

A key feature of the proposed method is its analytical tractability. Moments of the integrated joint Laplace transform of the volatility can be computed via one-dimensional numerical integration as soon as the joint characteristic function of stochastic volatility is known in closed form. This is the case for a variety of models, including the class of affine jump-diffusion models of Duffie et al. (2000, 2003). Given the wide applicability of this class in financial applications, it is important to explore its full flexibility and verify whether specification in this class can capture the key characteristics of volatility risk present in the data. Our method provides a convenient and efficient way to do that.

The full description of the method in Section 2.2 requires a moderate amount of detail, but a précis is as follows: Using the analytical tractability of Laplace transforms, we can form at any lag a continuum of estimating equations defined on the two-dimensional nonnegative orthant $\mathbb{R}^2_+$ that span the information in the integrated joint Laplace transform of current and lagged volatility. Then, kernel-based averaging schemes are used to condense the continuum of equations to individual estimating equations, but over regions of the orthant instead of all of $\mathbb{R}^2_+$, as would be the case in Paulson et al. (1975), Knight and Yu (2002), and Jiang and Knight (2010), who previously consider kernel-based averaging. Because the estimating equations are additively separable functions of the data and parameters, we can undertake minimum distance estimation (GMM) with a weighting matrix that is a model-free fixed function of the data. In effect, then, the information from the various estimating equations formed by regional kernel averaging at different lags all gets weighted together in a data-optimal manner to form the chi-squared estimation criterion function.

The proposed method can be compared with previous work on parametric (or semi-parametric) estimation of continuous-time stochastic volatility models. First, there is an earlier statistics literature on estimation of diffusion models from high-frequency data, see e.g. Prakasa Rao (1988) and the many references therein. The key difference with our method is that these papers are all concerned with estimation of directly observed Markov models, while our focus is estimating processes with hidden states, e.g., volatility and jumps. Indeed, our Realized Laplace transform measure is aimed exactly at estimating the volatility hidden in the price.

[3] For any scalar random variable X the characteristic function is $c(\omega) = \mathbb{E}\left(e^{i\omega X}\right), \omega \in \mathbb{R}$, while for a non-negatively supported random variable Y the *real* Laplace transform is $\mathcal{L}(u) = \mathbb{E}\left(e^{-uX}\right), u \in \mathbb{R}_+$. The multivariate extensions are obvious and both transforms are one-to-one with their probability distribution function, but the domain of the real Laplace transform is, of course, smaller.

Compared next with the earlier literature on estimation of stochastic volatility asset pricing models based on coarser frequencies, e.g., days, there are two key differences. First, we disentangle volatility from price jumps nonparametrically and second we perform estimation as if volatility is directly observable. These aspects of our estimation method provide both efficiency and robustness gains towards the complicated problem of modeling price jumps and their relation with the volatility jumps. On the other hand, our method is only for estimation of the volatility dynamics, and therefore we will need to combine it with existing estimation methods in order to make an inference for the rest of the return dynamics, e.g., the price jumps. Nevertheless, being able to estimate efficiently and robustly the volatility specification should help in pinning down the rest of the return dynamics both robustly and more efficiently.[4]

More recently, Bollerslev and Zhou (2002), Barndorff-Nielsen and Shephard (2002), Corradi and Distaso (2006) and Todorov (2009) consider estimation of stochastic volatility models using either method of moments or QML estimation on a model-free estimate of the daily Integrated Variance constructed from high-frequency data.[5] In our Monte Carlo application we compare our procedure with the above alternative (in particular we compare with QML estimation on a measure of the Integrated Variance). We consider a variety of one-factor pure-continuous and pure-jump volatility models as well as a two-factor volatility model. In all of the simulated models the price contains jumps. We find that our inference based on the Realized Laplace transform performs well without any significant biases unlike the QML estimation based on the high-frequency estimate of the Integrated Variance. The latter has significant biases coming from the high-frequency estimation. Also, comparison with the Cramer–Rao efficiency bound for an *infeasible* observation scheme of *daily spot variance* further reveals that our method has good efficiency properties and is able to extract the relevant information about volatility in the high-frequency data.

In the empirical application to the S&P 500 index futures sampled at five minutes, the method appears to discriminate reasonably well across a broad class of volatility models, and it shows promise for generating interesting new insights about the dynamics of volatility. We find that volatility exhibits transient and persistent shifts, which in itself is not too surprising, but interestingly both components of volatility appear to move through jumps and possibly without a diffusive component. The persistent shifts in volatility happen through a process with low activity, mainly rare big jumps. On the other hand, the driving Lévy process for the transient factor is far more active—it has big spikes but also a lot of small jumps which capture day to day variation in volatility. The finding that volatility appears to follow a two-factor pure-jump process needs to be tempered by other evidence that the preferred model does encounter some problems reproducing shifts from very low to very high levels of volatility as occur in the data.

The rest of the paper is organized as follows. In Section 2 we develop our method for estimating the parametric volatility specification of a stochastic process observed at high-frequency. In Section 3 we present the various jump-diffusion models for the stochastic volatility that we estimate on simulated and observed data. Section 4 contains the Monte Carlo study. In this section we also compare our method with a feasible alternative based on QML estimation on Integrated Variance. Section 5 contains the empirical application based on high-frequency S&P 500 Index futures data. Section 6 concludes. All technical details are given in Appendix.

2. Parametric estimation via the realized Laplace transform

Assume we observe at discrete points in time a 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

$$\mathrm{d}X_t = \alpha_t \mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t + \int_{\mathbb{R}} \delta(t-, x)\tilde{\mu}(\mathrm{d}t, \mathrm{d}x), \tag{1}$$

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) $\nu(x)\mathrm{d}x$; $\delta(t, x)$: $\mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is càdlàg in t and $\tilde{\mu}(\mathrm{d}s, \mathrm{d}x) = \mu(\mathrm{d}s, \mathrm{d}x) - \nu(x)\mathrm{d}x\mathrm{d}s$.

Our interest in the paper is estimation of parametric models for the stochastic volatility process V_t which is robust to the specification of the rest of the components in the model, i.e., the drift term α_t and the price jumps. Importantly, given prior empirical evidence in e.g., Andersen et al. (2002) and Chernov et al. (2003), we will be interested in developing an estimation method that *does not* depend on the V_t being a Markov process (with respect to its own filtration).

We note that in (1) we have implicitly assumed that the process contains a diffusive component and we are interested in its stochastic volatility. Nonparametric tests in Todorov and Tauchen (2010) show that this assumption is satisfied for the S&P 500 index that we are going to use in our empirical study. More generally, however, if the observed price does not have a diffusive component, then the construction of the Realized Laplace Transform below needs to be modified in an important way. Otherwise the estimation will lead to inconsistent results. We do not consider this complication here and instead refer to Todorov and Tauchen (2011c) for the theoretical analysis.

2.1. Constructing the realized Laplace transform

We start with constructing the Realized Laplace Transform on which the proposed estimation is based. We assume that we observe the log-price process X at the equidistant times $0, \frac{1}{n}, \ldots, 1, 1+\pi, \frac{n+1}{n}+\pi, \ldots, \frac{i}{n}+\lfloor(i-1)/n\rfloor\pi, \ldots, T+\pi(T-1)$. Our unit interval will be the trading part of a day where we have high-frequency price records and the intervals $(1, 1+\pi), (2, 2+\pi), \ldots$ are the periods from close to open where we do not have high-frequency price observations. n denotes the number of high-frequency returns within the trading interval and T is the total span of the data.[6] For simplicity we will denote the log-price increment over a high-frequency interval as $\Delta_i^n X = X_{\frac{i}{n}+\lfloor(i-1)/n\rfloor\pi} - X_{\frac{i-1}{n}+\lfloor(i-1)/n\rfloor\pi}$. Our results in the paper will rely on joint asymptotic arguments: fill-in ($n \to \infty$) and long-span ($T \to \infty$).

Using the high-frequency data over a given trading day $[(t-1)(1+\pi), (t-1)(1+\pi)+1]$, we can estimate in a model-free way the empirical Laplace transform of the latent variance process over that interval. In particular, provided price jumps are of finite variation and under some further mild regularity restrictions (satisfied for all parametric models considered here), Todorov and Tauchen (2011b) show

[4] Similarly, Bollerslev and Zhou (2002), Barndorff-Nielsen and Shephard (2002) and Corradi and Distaso (2006) focus mainly on estimation of the stochastic volatility specification.

[5] Recently Dobrev and Szerszen (2010) consider MCMC estimation of the volatility dynamics of discrete-time stochastic volatility models using a measure of the Integrated Variance as an extra observation equation and provide a strong example of the gains from incorporating high-frequency data in the estimation.

[6] $\lfloor x\rfloor$ denotes the smallest integer not exceeding x and we will set it to zero if $x < 0$.

$$Z_t(u) = \frac{1}{n}\sum_{i=n(t-1)+1}^{nt}\cos\left(\sqrt{2u}\sqrt{n}\Delta_i^n X\right) = \int_{(t-1)(1+\pi)}^{(t-1)(1+\pi)+1} e^{-uV_s}ds + O_p\left(1/\sqrt{n}\right), \quad u \geq 0. \tag{2}$$

We refer to $Z_t(u)$ as the Realized Laplace transform of the variance over the day. The robustness with respect to jumps of $Z_t(u)$ follows from the following property of the cosine function: $|\cos(x) - \cos(y)| \leq K(|x-y| \wedge 1)$ for any $x, y \in \mathbb{R}$. We note that unlike existing jump–robust volatility measures where robustness is achieved either through explicit truncation of the price increments or through staggering by considering products of adjacent price increments, our measure has a "built-in" truncation through the choice of the function that transforms the increments. We also stress that the above result in (2) holds regardless of the volatility specification (as long as it is an Itô semimartingale) and in particular the result is robust to specifications in which the price and volatility jumps are strongly dependent as data would suggest.[7] A formal analysis of the jump–robustness properties of $Z_t(u)$ is given in Todorov and Tauchen (2011b).

Next, using long-span asymptotics combined with standard stationarity and ergodicity conditions for V_t, and provided $T/n \to 0$, Todorov and Tauchen (2011b) show that[8]

$$\begin{cases} \widehat{\mathcal{L}}_V(u) = \frac{1}{T}\sum_{t=1}^{T}\int_{(t-1)(1+\pi)}^{(t-1)(1+\pi)+1} e^{-uV_s}ds + o_p(1/\sqrt{T}), \\ \widehat{\mathcal{L}}_V(u,v;k) = \frac{1}{T-k}\sum_{t=k+1}^{T}\left[\int_{(t-1)(1+\pi)}^{(t-1)(1+\pi)+1} e^{-uV_s}ds \right. \\ \quad \left. \times \int_{(t-k-1)(1+\pi)}^{(t-k-1)(1+\pi)+1} e^{-vV_s}\right]ds + o_p(1/\sqrt{T}), \end{cases} \tag{3}$$

where we define

$$\widehat{\mathcal{L}}_V(u) = \frac{1}{T}\sum_{t=1}^{T} Z_t(u), \qquad \widehat{\mathcal{L}}_V(u,v;k) = \frac{1}{T-k}\sum_{t=k+1}^{T} Z_t(u)Z_{t-k}(v). \tag{4}$$

From here using a standard Law of Large Numbers, we easily have[9]

$$\widehat{\mathcal{L}}_V(u) \xrightarrow{\mathbb{P}} \mathcal{L}_V(u), \qquad \widehat{\mathcal{L}}_V(u,v;k) \xrightarrow{\mathbb{P}} \mathcal{L}_V(u,v;k), \quad u,v \geq 0, k \in \mathbb{Z}, \tag{5}$$

for

$$\mathcal{L}_V(u) = \mathbb{E}\left(e^{-uV_s}\right), \qquad \mathcal{L}_V(u,v;k) = \mathbb{E}\left(\int_{k(1+\pi)}^{k(1+\pi)+1} e^{-uV_s}ds \int_0^1 e^{-vV_s}ds\right). \tag{6}$$

$\mathcal{L}_V(u)$ is the Laplace transform of V_s and $\mathcal{L}_V(u,v;k)$ is just an *integrated* joint Laplace transform of the variance during two unit intervals which are k days apart; $\widehat{\mathcal{L}}_V(u)$ and $\widehat{\mathcal{L}}_V(u,v;k)$ are their sample counterparts. The connection between $\mathcal{L}_V(u,v;k)$ and the joint Laplace transform of the variance at two points in time can be directly seen from the following. First, we denote the joint Laplace transform of the vector $(V_{t_1}, \ldots, V_{t_K})$ as

$$\mathcal{L}_V(\mathbf{u};\mathbf{t}) = \mathbb{E}\left(e^{-\sum_{i=1}^{K} u_i V_{t_i}}\right), \quad \mathbf{u} = (u_1,\ldots,u_K) \geq 0, \quad \mathbf{t} = (t_1,\ldots,t_K) \geq 0. \tag{7}$$

Then by a change of variable, and using the fact that V_t is a stationary process and therefore $\mathcal{L}_V([u,v];[t,s]) = \mathcal{L}_V([u,v];[t-s,0])$ where $t \geq s$, we can write for $k \geq 1$

$$\begin{aligned} \mathcal{L}_V(u,v;k) &= \int_{\widetilde{k}-1}^{\widetilde{k}}\int_{\widetilde{k}-t}^{1}\mathcal{L}_V([u,v];[t+s,s])dsdt + \int_{\widetilde{k}}^{\widetilde{k}+1}\int_{0}^{\widetilde{k}+1-t}\mathcal{L}_V([u,v];[t+s,s])dsdt \\ &= \int_{\widetilde{k}-1}^{\widetilde{k}}(t-\widetilde{k}+1)\mathcal{L}_V([u,v];[t,0])dt + \int_{\widetilde{k}}^{\widetilde{k}+1}(\widetilde{k}+1-t)\mathcal{L}_V([u,v];[t,0])dt, \end{aligned} \tag{8}$$

where we have used the shorthand $\widetilde{k} = k(1+\pi)$. Thus, although we cannot estimate the joint Laplace transform of the stochastic variance process at two arbitrary points in time, we can get very close to it by the use of $\widehat{\mathcal{L}}_V(u,v;k)$. The potential loss of information that occurs in estimation based on $\mathcal{L}_V(u,v;k)$ instead of $\mathcal{L}_V([u,v];[t+k,t])$ is for volatility dynamics with very short persistence.

2.2. Estimation methodology

In the infeasible case when the variance process V_t is directly observed (at integer times), one can match the empirical and model-implied joint Laplace transform at a given lag K. As shown in Feuerverger and Mureika (1977), see also Carrasco et al. (2007), appropriate weighting of these moments can lead to asymptotic equivalence to the estimation equations

$$\frac{1}{T}\sum_{t=K+1}^{T}\nabla_\rho p(V_t|V_{t-1},V_{t-2},\ldots,V_{t-K};\rho) = \mathbf{0}, \tag{9}$$

where henceforth we denote with ρ the vector of parameters and $p(V_t|V_{t-1},V_{t-2},\ldots,V_{t-K};\rho)$ stands for the probability density of V_t conditional on the vector $(V_{t-1},V_{t-2},\ldots,V_{t-K})$. The estimation equations in (9) achieve the Cramer–Rao efficiency bound in the case when V_t is Markovian of order K.

Our case is more complicated as we *do not observe* V_t and hence for the estimation problem here we can "only" work with $\mathcal{L}_V(u,v;k)$ instead of $\mathcal{L}_V([u,v];[t+k,t])$. Using the above analysis, we propose to base inference on matching the model-implied $\mathcal{L}_V(u,v;k)$ with the sample estimate $\widehat{\mathcal{L}}_V(u,v;k)$. If the volatility is constant within a day, then exactly as above appropriate weighting of these moment conditions will yield the Cramer–Rao efficiency bound based on *daily direct observations* of V_t.

More specifically, our vector of moment conditions is given by

$$\mathbf{m}_T(\rho) = \left\{\int_{\mathcal{R}_{j,k}}\left[\widehat{\mathcal{L}}_V(u,v;k) - \mathcal{L}_V(u,v;k|\rho)\right] \times \omega(du,dv)\right\}_{j=1,\ldots,J,k=1,\ldots,K}, \quad \mathcal{R}_{j,k} \subset \mathbb{R}_+^2, \tag{10}$$

where we describe the construction of the regions $\mathcal{R}_{j,k}$ and the weight measure $\omega(du,dv)$ below. Then our estimator is minimum distance (GMM) with estimating equations in the vector given in

[7] The Monte Carlo analysis in Section 4 further shows that various realistic features of the return dynamics, such as price jumps, leverage effect and joint price and volatility jumps, have very limited finite-sample effects in the estimation.

[8] In fact, as shown in Todorov and Tauchen (2011b), the speed condition $T/n \to 0$ can be relaxed to the much weaker one $T/n^2 \to 0$ for the class of affine jump-diffusion stochastic volatility models that we estimate in this paper.

[9] We refer to Todorov and Tauchen (2011b) for the technical conditions required for this long-span asymptotics as well as the associated CLT with it.

(10) and some weight matrix $\widehat{\mathbf{W}}$ converging in probability to a positive definite matrix W:

$$\hat{\rho} = \underset{\rho}{\operatorname{argmin}}\ \mathbf{m}_T(\rho)'\widehat{\mathbf{W}}\mathbf{m}_T(\rho). \tag{11}$$

We set $\widehat{\mathbf{W}}$ to be an estimate of the optimal weight matrix defined by the asymptotic variance of the empirical moments to be matched, i.e., $\int_{\mathcal{R}_{j,k}} \widehat{\mathcal{L}}_V(u, v; k)\omega(\mathrm{d}u, \mathrm{d}v)$. Note that because of the separability of data from parameters in the moment vector $\mathbf{m}_T(\rho)$, we construct our optimal weight matrix $\widehat{\mathbf{W}}$ using only the data. This in particular means that all moments are weighted the same way regardless of the model that is estimated, provided the set of moments used in the estimation is kept the same of course.

Consistency and asymptotic normality of the estimator in (11) follows from classical conditions required for GMM estimation, see e.g., Hansen (1982), as well as Theorem 2 in Todorov and Tauchen (2011b) that guarantee that (3) above holds.

The intuition behind our estimator is as follows. We split $\mathbb{R}^2_+$ into regions. Within the regions we weight the distance $\widehat{\mathcal{L}}(u, v; k) - \mathcal{L}(u, v; k|\rho)$ by the weight measure $\omega(\mathrm{d}u, \mathrm{d}v)$, while we let the data determine (through the optimal weight matrix $\widehat{\mathbf{W}}$) the relative importance of each pair region-lag. This approach can be viewed as a feasible alternative to the use of a continuum of moment conditions based on $\mathcal{L}(u, v; k|\rho)$ (over u, v and k) as in the case when the stochastic process of interest (here V_t) is fully observable considered in Carrasco et al. (2007). While for many models, e.g., the affine jump-diffusion class, the joint Laplace transform is known analytically, this is not the case generally for the integrated one, $\mathcal{L}(u, v; k|\rho)$, and hence the latter has to be evaluated by (one-dimensional) numerical integration which is quick and easy. On the other hand, using a continuum of points (u, v, k) in the estimation would (generally) involve high-dimensional numerical integrations which are unstable. This can be viewed as the price to be paid for "making" V_t from latent to "observable".

The weight measure $\omega(\mathrm{d}u, \mathrm{d}v)$ we consider here is of the form $\sum_i \delta_{(u_i,v_i)}\mathrm{e}^{-0.5(u_i^2+v_i^2)/c^2}$ for $\delta_{\mathbf{x}}$ denoting Dirac delta at the point $\mathbf{x}$, and we set $c = 0.50 \times u_{\max}$ where $u_{\max}$ is the maximum value of u and v that we use in the estimation. We explain how we set $u_{\max}$ later in our numerical applications: the goal is to pick $u_{\max}$ such that $[0, u_{\max}] \times [0, u_{\max}]$ contains "most of" the information in $\mathcal{L}(u, v; k|\rho)$. $\mathrm{e}^{-0.5(u^2+v^2)/c^2}$ weighs the information coming from the different points in the region that is included, with points closer to (0, 0) receiving more weight. Given our choice of c, the lowest weight corresponds to the density of a normal distribution at 2 standard deviations and this is exactly the region where the normal density has curvature (and hence weighs differently the different points (u, v)). This is similar to the use of the Gaussian kernel in empirical characteristic function based estimation in Jiang and Knight (2002) and Carrasco et al. (2007). From a practical point of view, using Dirac deltas would probably not lead to much loss of information as the joint Laplace transform is typically rather smooth.

The regions that we look at in the estimation are of the form

$$\mathcal{R}_{j,k} = \{(u\ v) \in [\underline{b}_{j,k}u_{\max}\ \overline{b}_{j,k}u_{\max}] \times [\underline{b}'_{j,k}u_{\max}\ \overline{b}'_{j,k}u_{\max}]\},$$
$$j = 1, \ldots, J,\ k = 1, \ldots, K, \tag{12}$$

where $\underline{b}_{j,k}$, $\underline{b}'_{j,k}$, $\overline{b}_{j,k}$ and $\overline{b}'_{j,k}$ are all in $[0, 1]$, and in the numerical applications in the next sections we will explain how to choose them. The general idea is to cover all useful information in the (integrated) joint Laplace transform, making sure at the same time that the regions contain sufficiently different information so that we do not end up with a perfectly correlated set of moment conditions in the GMM.

This approach stands in contrast to existing kernel averaging approaches that enforce the same kernel averaging scheme – almost always Gaussian – over the entire continuum. Our approach uses a GMM weight matrix on top of kernel-weighting over regions, and it also differs from a strategy like that of Carrasco and Florens (2000, 2002) which entails a significant computational burden, at least in the present context, to determine a model-based kernel that is optimal conditional on model validity. Here the weighting over the continuum of estimating equations is model-free, and we force all models to confront the same vector of data-determined conditions.

3. Parametric volatility specifications

The proposed estimation procedure based on the Realized Laplace transform is particularly easy to implement in volatility models for which the joint Laplace transform is known in closed form (or up to numerical integration). The general affine jump-diffusion models proposed in Duffie et al. (2000, 2003) are a leading example. They have been used widely in many finance applications. Therefore, in our Monte Carlo study as well as the empirical application we illustrate the proposed estimation technique to estimate multi-factor affine jump-diffusion models for the unobserved market variance process. Using earlier evidence from daily estimation, e.g., Andersen et al. (2002) and Chernov et al. (2003), we will focus on two-factor specifications

$$\begin{aligned} &V_t = V_{1t} + V_{2t}, \quad t > 0, \\ &\mathrm{d}V_{it} = \kappa_i(\theta_i - V_{it})\mathrm{d}t + \sigma_i\sqrt{V_{it}}\mathrm{d}W_{it} + \mathrm{d}L_{it}, \quad i = 1, 2, \end{aligned} \tag{13}$$

where L_{it} are Lévy subordinators with Lévy measures $\nu_i(\mathrm{d}x)$. The list of the particular model specifications that we estimate and compare performance are

- *Pure-continuous volatility model*: one or two factor specification of (13) with $L_{it} \equiv 0$.
- *Pure-jump volatility model*: one or two factor specification of (13) with $\sigma_i \equiv \theta_i \equiv 0$ and jump measure of L specified with (15) below.
- *Continuous-jump volatility model*: one-factor is pure-continuous and the other is pure-jump with jump measure of L specified with (15) below.

The pure-continuous volatility models are just a superposition of the standard square-root diffusion processes. The pure-jump volatility factors are also known as non-Gaussian OU models, see e.g., Barndorff-Nielsen and Shephard (2001a). In those models the volatility factor moves only through positive jumps and it reverts afterwards back to its unconditional mean level till another jump arrives (infinite activity, but finite variation, jumps are allowed). The marginal distribution is infinitely divisible (see e.g., Sato (1999)) and hence by the Lévy–Khinchine theorem can be represented (identified uniquely) by its Lévy measure. Here we follow an approach proposed by Barndorff-Nielsen and Shephard (2001a) and model the process by specifying the marginal distribution of the volatility factor and we back out from it the model for the driving Lévy process. This has the advantage that the parameters controlling the memory of the volatility process are separated from those controlling its distribution.

In particular we work with pure-jump volatility factors whose marginal distribution is that corresponding to the increments of a tempered stable process (Carr et al., 2002; Rosiński, 2007). The latter is known to be a very flexible distribution. Its corresponding Lévy density is given by

$$\nu_{V_i}(x) = c_i\frac{\mathrm{e}^{-\lambda_i x}}{x^{1+\alpha_i}}1_{\{x>0\}}, \quad c_i > 0, \lambda_i > 0, \alpha_i < 1. \tag{14}$$

Table 1
Parameter setting for the Monte Carlo.

Case	Paramaters			
One-factor pure-continuous models				
A	$\kappa_1 = 0.50$	$\theta_1 = 1.0$	$\sigma_1 = 0.5$	
B and B–L	$\kappa_1 = 0.15$	$\theta_1 = 1.0$	$\sigma_1 = 0.2$	
C	$\kappa_1 = 0.03$	$\theta_1 = 1.0$	$\sigma_1 = 0.1$	
One-factor pure-jump models				
D	$\kappa_1 = 0.50$	$\alpha_1 = 0.5$	$c_1 = 0.7979$	$\lambda_1 = 2.0$
E and E–L	$\kappa_1 = 0.15$	$\alpha_1 = 0.5$	$c_1 = 0.7979$	$\lambda_1 = 2.0$
Two-factor pure-jump model				
F	$\kappa_1 = 0.03$	$\alpha_1 = 0.5$	$c_1 = 0.7596$	$\lambda_1 = 5.0$
	$\kappa_2 = 1.00$	$\alpha_2 = 0.5$	$c_2 = 0.2257$	$\lambda_2 = 1.0$

Note: In cases A, B and C, V_t is independent from W_t while in case B–L we have $\text{corr}(W_t, W_{it}) = -0.5 \times t$. In cases D, E and F, volatility and price jumps are independent. In case E–L, price jumps are equal in magnitude to volatility jumps.

The parameters α_i control the small jumps in the volatility factors, while λ_i control the big jumps. A value $\alpha_i < 0$ corresponds to finite activity jumps and $\alpha_i \geq 0$ to infinite activity. Intuitively, the activity of the volatility jumps determines the vibrancy of the volatility factor trajectories. There are two special cases of (14): the case $\alpha = 0.0$ corresponds to Gamma marginal distribution (which is also the marginal distribution of the square-root diffusion) and the case $\alpha = 0.5$ corresponds to Inverse Gaussian marginal distribution. Using (34) in Appendix, we have that the Lévy density of the driving Lévy process for our pure-jump volatility factors with the specified marginal distribution by (14) is given by

$$\nu_{L_i}(x) = \left(\frac{\alpha_i c_i \kappa_i e^{-\lambda_i x}}{x^{\alpha_i + 1}} + \frac{c_i \lambda_i \kappa_i e^{-\lambda_i x}}{x^{\alpha_i}} \right) 1_{\{x>0\}}. \tag{15}$$

4. Monte Carlo study

We next test the performance of our estimation method on simulated data from the following models for the stochastic volatility: one-factor square-root diffusion, one-factor non-Gaussian OU model with Inverse Gaussian marginal distribution, and two-factor superposition of the above non-Gaussian OU model. The different simulated models are summarized in Table 1. In all models the mean of V_t is set to 1 (variance reported in daily percentage units). The different cases differ in the volatility persistence, the volatility of volatility and the presence of volatility jumps.[10] Also, in each of the scenarios, except for case E–L, we set the price jumps to be of Lévy type with the following Lévy density (i.e., jump compensator)

$$\nu_X(x) = 0.2 \times \frac{e^{-x^2}}{\sqrt{\pi}}, \tag{16}$$

which corresponds to compound Poisson jumps with normally distributed jump size. The selected values of the parameters in (16) imply variance due to price jumps is 0.1, which is consistent with earlier non-parametric evidence. In case E–L price and volatility jumps arrive at the same time and are equal in magnitude (but the price jumps might be with negative sign as well) and this case allows us to explore the small sample effect of dependence between price and volatility jumps as well as the robustness of our results to presence of infinite-activity price jumps.

In each simulated scenario we have $T = 5000$ days and we sample $n = 80$ times during the day, which mimics our available data in the empirical application, and for simplicity we also set $\pi = 0$, i.e., there is no overnight period. The Monte Carlo results

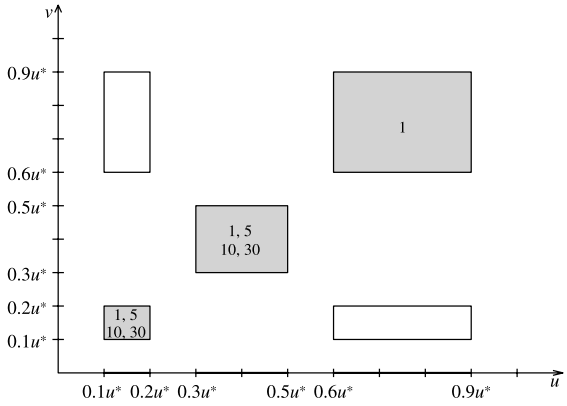

Fig. 1. Regions in R^2_+ for fixed lag used in estimation. The shaded regions are the ones used in the set of moment conditions MC1 (together with their projections on the u axis), with the numbers in them corresponding to the lags.

are based on 1000 replications. Finally, each estimation is done via the MCMC approach of Chernozhukov and Hong (2003) to classical estimation, with length of the MCMC chain of 15,000.[11] The weight matrix $\widehat{W}$ is computed using Parzen kernel with lag length of 70 days. The results are summarized in Tables 2 and 3.

Our choice of $u_{\max}$ is such that $\widehat{\mathcal{L}}'_V(u)$ is around -0.01.[12] This resulted in $\widehat{\mathcal{L}}_V(u_{\max}) \approx 0.005$ which in turn corresponded to $u_{\max}$ of around 8 for the simulated models. Therefore, in the estimation, we choose $u_{\max}$ by the simpler rule $u_{\max} = \widehat{\mathcal{L}}_V^{-1}(0.005)$ which satisfies our target in terms of the derivative of the estimated Laplace transform. The moment conditions that we use are

(a) regions $[0.1u_{\max}\ 0.2u_{\max}] \times [0\ 0]$, $[0.3u_{\max}\ 0.5u_{\max}] \times [0\ 0]$ and $[0.6u_{\max}\ 0.9u_{\max}] \times [0\ 0]$,
(b) squares $[0.1u_{\max}\ 0.2u_{\max}]^2$, $[0.3u_{\max}\ 0.5u_{\max}]^2$ and $[0.6u_{\max}\ 0.9u_{\max}]^2$ for lags $k = 1$,
(c) square $[0.1u_{\max}\ 0.2u_{\max}]^2$ for lag $k = 5, 10, 30$,
(d) square $[0.3u_{\max}\ 0.5u_{\max}]^2$ for lag $k = 5, 10, 30$.

Fig. 1 displays the above regions with the lag lengths k entered within the block and the one-dimensional regions are the heavily shaded segments along the abscissa; in subsequent diagnostic work we use other lag lengths and include the "off-diagonal" blocks in Fig. 1 in the estimation. We refer to the set of moment conditions immediately above as MC1. This results in 12 moment conditions, and as we confirm later in the Monte Carlo, this moment vector captures well, in a relatively parsimonious way, the information in the data about the distribution and memory of volatility. Finally, in each of the two-dimensional regions above we evaluate the integrated joint Laplace transform only in the four edges. This is done to save on computational time and does not have a significant effect on the estimation.

For the one-factor models in our Monte Carlo, we can compare the efficiency of our estimation method with the *infeasible* case when we observe directly the latent variance process at daily frequency. The Cramer–Rao efficiency bound for the latter observational scheme is easily computable in the one-factor volatility setting with details provided in Appendix. Note that our benchmark is a *daily* variance and not a *continuous record* of the latter. A continuous record of V_t would imply that the parameter σ in the square-root model and α and c in the one-factor non-Gaussian OU model can be inferred from a *fixed* span of data without estimation error. Instead, our goal with this comparison

[10] We do not consider a pure-jump alternative to case C, i.e., a persistent one-factor pure-jump model as the numerical integrations needed for the Cramer–Rao bound are relatively unstable since the integrands have too much oscillation.

[11] Of course any other optimization method can be used for finding the parameter estimates than our MCMC-based approach.

[12] Note that $L'_V(u) \equiv \mathbb{E}(Ve^{-uV})$ is strictly decreasing in u.

Table 2
Monte Carlo results: one-factor models.

Par	True value	RLT-based estimation			QML estimation			CRB
		Median	MAD	SE	Median	MAD	SE	
Case A								
κ_1	0.5000	0.4734	0.0267	0.0211	0.2999	0.2001	0.0093	0.0187
θ_1	1.0000	1.0106	0.0113	0.0120	1.0057	0.0087	0.0121	0.0142
σ_1	0.5000	0.4921	0.0098	0.0126	0.4068	0.0932	0.0062	0.0063
Case B								
κ_1	0.1500	0.1474	0.0107	0.0132	0.1855	0.0355	0.0085	0.0086
θ_1	1.0000	1.0097	0.0119	0.0165	1.0040	0.0114	0.0159	0.0188
σ_1	0.2000	0.2014	0.0059	0.0072	0.2488	0.0488	0.0035	0.0021
Case B–L								
κ_1	0.1500	0.1505	0.0089	0.0137	0.1808	0.0308	0.0081	0.0086
θ_1	1.0000	1.0115	0.0146	0.0165	0.9983	0.0111	0.0160	0.0188
σ_1	0.2000	0.2044	0.0062	0.0081	0.2430	0.0430	0.0035	0.0021
Case C								
κ_1	0.0300	0.0336	0.0040	0.0049	0.1121	0.0821	0.0199	0.0035
θ_1	1.0000	1.0059	0.0305	0.0413	0.9971	0.0284	0.0553	0.0465
σ_1	0.1000	0.1019	0.0035	0.0049	0.2050	0.1050	0.0079	0.0010
Case D								
κ_1	0.5000	0.4893	0.0218	0.0296	0.3032	0.1968	0.0118	0.0032
α_1	0.5000	0.5042	0.0140	0.0270	–	–	–	0.0095
c_1	0.7979	0.7845	0.0477	0.0943	0.7664	0.0315	0.0163	0.0342
λ_1	2.0000	1.9283	0.1252	0.1896	1.8361	0.1639	0.0743	0.0810
Case E								
κ_1	0.1500	0.1511	0.0077	0.0124	0.1396	0.0106	0.0087	*0.0002*
α_1	0.5000	0.5149	0.0317	0.0469	–	–	–	*0.0058*
c_1	0.7979	0.7541	0.1008	0.1647	0.7491	0.0488	0.0274	*0.0294*
λ_1	2.0000	1.8996	0.2164	0.3026	1.7513	0.2487	0.1255	*0.1042*
Case E–L								
κ_1	0.1500	0.1519	0.0078	0.0125	0.1424	0.0084	0.0088	*0.0002*
α_1	0.5000	0.5128	0.0268	0.0380	–	–	–	*0.0058*
c_1	0.7979	0.7580	0.0803	0.1263	0.7487	0.0492	0.0270	*0.0294*
λ_1	2.0000	1.8793	0.1897	0.2377	1.7181	0.2819	0.1221	*0.1042*

Note: MAD stands for median absolute value around the true value; SE stands for standard error; CRB stands for Cramer–Rao bound; the reported CRBs for Case E are italicized because the associated numerical integrations for the inverse Fourier transform described in Appendix were very delicate, mainly for κ_1; the difficulties with κ_1 indirectly affects the values for the other three parameters because of the matrix inversion.

here is to gauge the potential loss of efficiency due to the use of our moments based on $\mathcal{L}(u, v; k)$ instead of working directly with the *infeasible* daily transitional density of the latent variance.

In the one-factor models, we also compare our estimator with a feasible alternative using the high-frequency data that has been widely used to date. It is based on performing inference on the Integrated Variance defined as

$$IV_t = \int_{(t-1)(1+\pi)}^{(t-1)(1+\pi)+1} V_s \mathrm{d}s, \quad t \in \mathbb{Z}. \tag{17}$$

In many models, and in particular the ones we use in our Monte Carlo, see e.g., Meddahi (2003) and Todorov (2009), the Integrated Variance follows an ARMA process whose coefficients are known functions of the structural parameters for the volatility (for the simulated one-factor models it is ARMA(1,1), see Appendix for the details). Then, one way of estimation based on the Integrated Variance is to match moments like mean, variance and covariance, see e.g., Bollerslev and Zhou (2002) and Corradi and Distaso (2006) in the continuous setting and Todorov (2009) in the presence of jumps. An alternative, following Barndorff-Nielsen and Shephard (2002), that we use here to compare our method with, is to do a Gaussian Quasi-Maximum Likelihood for the sequence $\{IV_t\}_{t\in\mathbb{Z}}$.[13] The details of the necessary computations are given in the Appendix.

Integrated Variance is of course unobserved, but it can be substituted with a model-free estimate from the high-frequency data. One possible such estimate that we use here is the Truncated Variance, proposed originally by Mancini (2009), defined as

$$TV_t(\alpha, \varpi) = \sum_{i=n(t-1)+1}^{nt} |\Delta_i^n X|^2 1_{\{|\Delta_i^n X| \le \alpha n^{-\varpi}\}}, \quad \alpha > 0, \quad \varpi \in (0, 1/2), \tag{18}$$

where here we use $\varpi = 0.49$, i.e., a value very close to 1/2 and we further set $\alpha = 3 \times \sqrt{BV_t}$ for BV_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):

$$BV_t = \frac{\pi}{2} \sum_{i=n(t-1)+2}^{nt} |\Delta_{i-1}^n X||\Delta_i^n X|. \tag{19}$$

Under certain regularity conditions, see e.g., Jacod (2008), the Truncated Variance is a model-free consistent and asymptotically normal estimator (for the fill-in asymptotics) of the (unobservable) Integrated Variance defined in (17). The asymptotic justification for the joint fill-in and long-span asymptotics of the QML estimator in the presence of price jumps can be done exactly as in Todorov (2009).

Table 2 summarizes the results from the Monte Carlo for the different one-factor models with Table 3 reporting the rejection rates for the corresponding test of overidentifying restrictions. In the case of the QML estimation of the pure-jump model we fix the parameter α at its true value, since the QML estimation cannot identify such richly specified marginal distribution of

[13] Barndorff-Nielsen and Shephard (2002) apply the method in the context of no jumps, and so use the Realized Variance. Also, unlike our use of QML here, Barndorff-Nielsen and Shephard (2002) take into account the error in the Realized Variance in measuring the Integrated Variance (the error does not matter in the fill-in asymptotic limit but can have a finite sample effect). We do not take this error into account to be on par with our proposed method which similarly does not account for the error in measuring $L_V(u, v; k)$ from high-frequency data.

Table 3
Monte Carlo results: J-test.

Case	df	Nominal Size		Case	df	Nominal Size	
		1%	5%			1%	5%
A	9	4.29	8.95	D	8	2.60	9.80
B	9	8.00	17.00	E	8	1.90	5.50
B–L	9	3.70	17.90	E–L	8	3.10	8.80
C	9	0.87	4.13	F	4	6.50	19.98

the volatility.[14] We can see from the table that our proposed method behaves quite well. It is virtually unbiased for almost all parameters—the only exception perhaps is the parameter λ_1 in the pure-jump models which is the hardest parameter to estimate (recall it controls the big jumps in volatility). However its bias is still very small, especially when compared with the precision of its estimation.

Comparing the standard errors of our estimator with Cramer–Rao bounds for the infeasible scheme of daily observations of V_t, which is used as a benchmark, we see that the performance of our proposed method is generally very good. There are, however, a few notable deviations from full efficiency. One of the reasons for this (for some of the parameters) is in the observational scheme: our estimator is based on *integrated* volatility measure whereas the Cramer–Rao bounds are for daily observations of *spot* volatility. For example, we see that for the square-root diffusion the standard errors for σ_1 are somewhat bigger than the efficient bound. Note, however, that the efficiency bound for this parameter can be driven to 0 by considering more frequent (than daily) observations of the spot volatility. The same observation can be made also for the parameters c_1, α_1 and κ_1 in the pure-jump model.

The second reason for the deviation from efficiency is in the use of high-frequency data in estimating the integrated joint Laplace transform of volatility in our estimation method. This effect can be seen by noting, for example, that the wedge between the standard errors of our estimator and the Cramer–Rao efficiency bounds widens by going from the square-root diffusive models to the pure-jump ones. The effect of discretization error in the latter class of models should be bigger because of the volatility jumps.

Looking at cases B–L and E–L, we can note that the dependence between the price and volatility innovations, either via correlated Brownian motions or dependent price and volatility jumps, has virtually no finite-sample distortion on the estimation. The same conclusion can be made regarding the presence of price jumps in X that are infinitely active as in the case E–L. Overall, the only effect of the presence of price jumps and the frequency of sampling on our RLT-based estimation is in the standard errors.

Comparing our estimator with the feasible alternative of QML on the Truncated Variance, we can see that overall the former behaves much better than the latter. The main problem of the QML estimation is that it is significantly biased for the parameters controlling memory and persistence of volatility across all considered cases. The standard errors of the QML estimators are smaller for the low persistence cases (even from the efficiency bounds) and much higher for the high persistence case, but this is hard to interpret because of the very significant biases in the estimation.

What is the reason for the significant biases in the QML estimation based on the Truncated Variance? Using a Central Limit Theorem for the fill-in asymptotics we have approximately

$$TV_t(\alpha, \varpi) \approx \int_{(t-1)(1+\pi)}^{(t-1)(1+\pi)+1} V_s \mathrm{d}s + \frac{1}{\sqrt{n}}\epsilon_t, \tag{20}$$

Table 4
Monte Carlo results for RLT-based estimation of case F.

Par	True value	Median	MAD	SE
κ_1	0.0300	0.0306	0.0072	0.0118
α_1	0.5000	0.5031	0.0340	0.0695
c_1	0.7596	0.7808	0.0963	0.1809
λ_1	5.0000	4.8609	0.6966	1.1049
κ_2	1.0000	1.0000	0.0995	0.1581
α_2	0.5000	0.4531	0.1097	0.1855
c_2	0.2257	0.2482	0.0551	0.1077
λ_2	1.0000	1.0620	0.2151	0.3740

Notation as in Table 2. The results are based on 800 Monte Carlo replications.

where conditional on the volatility process, ϵ_t is a Gaussian error (whose volatility depends on V_s and does not shrink as the sampling frequency n increases). Then note that the objective function of the QML estimator involves squares of the Integrated Variance. Substituting $TV_t(\alpha, \varpi)$ for IV_t in the objective function therefore introduces an error in the latter whose expectation is not zero (as it will involve squares of ϵ_t) and this in turn generates the documented biases. This error of course will decrease as we sample more frequently, i.e., as $n \to \infty$, but it clearly has a very strong effect on the precision of the QML estimator for the frequency we are interested in. A possible solution to this problem of the QML estimation based on the Truncated Variance is to recognize the approximation in (20) and derive an expression for the variance of ϵ_t as done for example in Barndorff-Nielsen and Shephard (2002) in the context of no price jumps.

We can use similar reasoning as above to explain why our proposed RLT-based estimation does not suffer from the above problem. A Central Limit Theorem, see Todorov and Tauchen (2011b), implies

$$Z_t(u) \approx \int_{(t-1)(1+\pi)}^{(t-1)(1+\pi)+1} \mathrm{e}^{-uV_s}\mathrm{d}s + \frac{1}{\sqrt{n}}\tilde{\epsilon}_t, \tag{21}$$

where conditional on the volatility process, $\tilde{\epsilon}_t$ is a Gaussian error (whose volatility depends on V_s and u and does not shrink as n increases) with $\mathbb{E}(\tilde{\epsilon}_t\tilde{\epsilon}_s) = 0$ for $t \neq s$ and $t, s \in \mathbb{Z}$. Note that our estimation is based on the idea that $\widehat{\mathcal{L}}_V(u, v; k)$ contains all the information for V_t in the high-frequency data and hence uses only these moment conditions in the estimation without any further nonlinear transformations. Using the approximation (21), we see that $\widehat{\mathcal{L}}_V(u, v; k)$ are (approximately) unbiased for $\mathcal{L}_V(u, v; k)$.[15]

Turning to the test for overidentifying restrictions, we can see from Table 3 that overall the test performance is satisfactory although in some of the cases there is moderate over-rejection. The worst performance is for cases B and B–L in which the finite sample size of the test is bigger with 10% from its nominal level. Such finite sample over-rejections though are consistent with prior evidence for GMM reported in Andersen and Sørensen (1996) particularly when the number of moment conditions is large (as is the case for scenarios B and B–L).

Finally the results for case F are given in Table 4. We see that our estimator behaves well in this richly parameterized model. Some of the parameters have small biases (particularly α_2) but they are insignificant compared with the magnitude of the associated standard errors. The hardest parameters to estimate are those of the transient factor, which is also twice as volatile as the persistent factor.

5. Empirical application

5.1. Initial data analysis

We continue next with our empirical application where we use 5-min level data on the S&P 500 futures index covering the

[14] Of course a method of moment estimator based on the Truncated Variation could have identified that moment.

[15] For more formal discussion about the asymptotic bias in $Z_t(u)$, see Todorov and Tauchen (2011b).

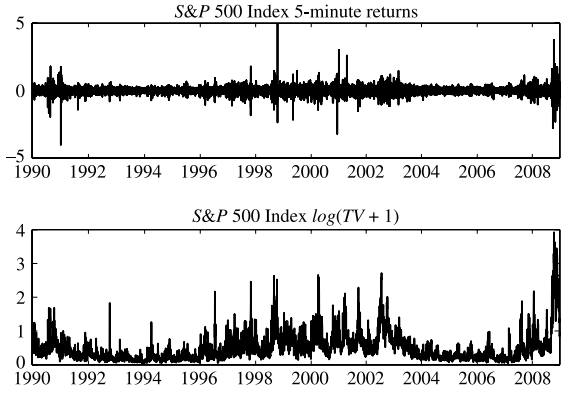

Fig. 2. S&P 500 index data.

period January 1, 1990, to December 31, 2008. Each day has 80 high-frequency returns. The data frequency is sparse enough so that microstructure-noise related issues are of no series concern here.[16] The ratio of the overnight returns variance to the mean realized variance over the trading day is 0.3437, and we therefore set π to this number. On Fig. 2 we plot the raw high-frequency data used in our estimation as well as (a log transformation of) the Truncated Variation, which as explained in Section 4 is a model-free measure for the daily Integrated Variance. The high-frequency returns have clearly distinguishable spikes, which underscores the importance of using volatility measure robust to jumps as is the case for the Realized Laplace Transform. Also the bottom panel of the figure suggests a complicated dynamic structure of the stochastic volatility with both persistent and transient volatility spikes present.

Before turning to the estimation we need to modify slightly our analysis because of the well-known presence of a diurnal deterministic within-day pattern in volatility, see e.g., Andersen and Bollerslev (1997). To this end, V_t in (1) needs to be replaced by $\widetilde{V}_t = V_t \times f(t - \lfloor t/(1+\pi)\rfloor)$ where V_t is our original stationary volatility process and $f(s)$ is a positive differentiable deterministic function on $[0, 1]$ that captures the diurnal pattern. Then we correct our original Realized Laplace transform for the deterministic pattern in the volatility by replacing Z_t with

$$\widetilde{Z}_t(u) = \frac{1}{n}\sum_{i=n(t-1)+1}^{nt} \cos\left(\sqrt{2u}\sqrt{n}\,\widehat{f}_i^{-1/2} 1_{\{\widehat{f}_i\neq 0\}}\Delta_i^n X\right), \quad \widehat{f}_i = \frac{\widehat{g}_i}{\widehat{\widehat{g}}},$$
$$\widehat{g}_i = \frac{n}{T}\sum_{t=1}^{T} |\Delta_{i_t}^n X|^2 1(|\Delta_{i_t}^n X| \le \alpha n^{-\varpi}), \qquad \widehat{\widehat{g}} = \frac{1}{n}\sum_{i=1}^{n}\widehat{g}_i, \tag{22}$$
$$i = 1, \ldots, nT, \alpha > 0, \varpi \in (0, 1/2),$$

where $i_t = t-1+i-[i/n]n$, for $i = 1, \ldots, nT$ and $t = 1, \ldots, T$. As for the construction of the Truncated Variance in Section 4, we set $\alpha = 3\times\sqrt{BV_t}$ and $\varpi = 0.49$. We further put tilde to all estimators of Section 2 in which Z_t is replaced with $\widetilde{Z}_t$. In this setting of diurnal patterns in volatility (5) will still hold when we replace $\widehat{\mathcal{L}}_V(u)$ and $\widehat{\mathcal{L}}_V(u, v; k)$ with $\widetilde{\mathcal{L}}_V(u)$ and $\widetilde{\mathcal{L}}_V(u, v; k)$.[17] Intuitively, $\widehat{f}_i$ estimates the deterministic component of the stochastic variance, and then in $\widetilde{Z}_t(u)$ we "standardize" the high-frequency increments by it.

[16] For example, the autocorrelations in the 5-min returns series are very small and insignificant.

[17] One can further quantify the effect from the correction for diurnal pattern on the standard errors of $\widetilde{\mathcal{L}}_V(u, v; k)$, see Todorov and Tauchen (2011b), but this effect is relatively small and therefore we ignore it in the subsequent work.

Table 5
Estimation results: one-factor models.

Pure-continuous		Pure-jump	
Parameter	Estimate	Parameter	Estimate
κ_1	0.0146	κ_1	2.6834
	(0.0027)		(0.1926)
θ_1	2.2437	α_1	0.5593
	(0.1078)		(0.0247)
σ_1	0.2082	c_1	0.2047
	(0.0132)		(0.0178)
		λ_1	1.9459
			(0.2669)
J Test (df)	324.36(6)		209.64(5)
(P-Val)	(p = 0.00)		(p = 0.00)

Note: The set of moment conditions MC0 defined in the text is used in the estimation. Standard errors for the parameter estimates are reported in parentheses.

5.2. Estimation results

We proceed with the estimation of the different volatility models discussed in Section 3. In our estimation, we set $u_{\max} = \widehat{\mathcal{L}}_V^{-1}(0.1)$ which results in a derivative $\widehat{L}_V'(u)$ at $u_{\max}$ of around -0.01 and a value of $u_{\max}$ close to 8 exactly as in the Monte Carlo. Later we check the robustness of our findings with respect to the choice of $u_{\max}$. In the estimation we use the same set of moment conditions as in the Monte Carlo but we drop initially the last three moments in (d) of MC1, which results in 9 moment conditions. We refer to this reduced set of moments as MC0. As for the estimation on the simulated data, for all results here the optimal weighting matrix is estimated using a Parzen kernel with lag length of 70. Also, we always impose the stationarity restriction $\sigma_i \le \sqrt{2\theta_i\kappa_i}$ for $i = 1, 2$ for the square-root processes.[18]

The results for the one-factor volatility models are given in Table 5. Not surprisingly these models cannot fit the data very well as evidenced by the extremely large values of the J test. The pure-jump model performs far better than the pure-continuous model and this is because it is more flexible in the type of marginal distribution for the volatility it can generate. We also note that the estimated mean reversion parameters in the two models are very different as both models struggle to match the initial fast drop in the autocorrelation of the volatility caused by the many short-term volatility spikes evident from the bottom panel of Fig. 2.

We turn next to the two-factor stochastic volatility models. The estimation results for these models are given in Table 6. As we see from the table, the J tests for the two-factor models drop significantly as these models have a better chance to capture simultaneously the short-lived spikes in volatility together with its more persistent shifts. At the same time the performance of the models differ significantly: two-factor pure-continuous and continuous-jump models have still significant difficulties in matching the moments from the data, unlike the two-factor pure-jump model. Where is that difference in performance coming from? Looking at the mean-reversion parameter estimates, we see that they are quite similar across models: one is very persistent (capturing the persistent shifts in volatility) and the other one is very fast mean-reverting (capturing the short-term volatility spikes). Also, the implied mean of the volatility across models is very similar.

Where the models start to differ, which explains their different success, is the ability to generate volatility of volatility in the different factors. First, the pure-continuous model cannot

[18] The estimation is done as in the Monte Carlo via MCMC but the length of the chain is 500,000 draws.

Table 6
Estimation results: two-factor models.

Pure-continuous		Continuous-jump		Pure-jump	
Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
κ_1	0.0186	κ_1	0.0176	κ_1	0.0122
	(0.0024)		(0.0106)		(0.0045)
θ_1	0.6100	θ_1	0.4152	α_1	0.3630
	(0.0423)		(0.0702)		(0.1459)
σ_1	0.1504	σ_1	0.1209	c_1	0.1826
	(0.0052)		(0.0148)		(0.0666)
				λ_1	0.1836
					(0.1852)
κ_2	1.3776	κ_2	2.3605	κ_2	2.2167
	(0.1336)		(0.4000)		(0.1779)
θ_2	0.4434	α_2	0.5403	α_2	0.5692
	(0.0168)		(0.0548)		(0.0864)
σ_2	1.1052	c_2	0.1050	c_2	0.1143
	(0.0523)		(0.0246)		(0.0286)
		λ_2	0.3420	λ_2	0.6039
			(0.1110)		(0.1509)
J Test (df)	97.87(3)		39.64(2)		1.61(1)
(P-Val)	(p = 0.000)		(p = 0.000)		(p = 0.205)

Note: The set of moment conditions MC0 defined in the text is used in the estimation. Standard errors for the parameter estimates are reported in parentheses.

generate enough volatility of volatility both in the persistent and the transient volatility components. This explains its very bad performance. This fact is most clearly seen by noting that for both factors, the parameters are on the boundary of the stationarity restriction (which generates the highest possible volatility of the factors). When we move from the pure-continuous to the continuous-jump model, we can see a significant improvement of the fit: the J test drops approximately by half. It is interesting to note that in this model the persistent volatility factor is the square-root diffusion and the pure-jump factor captures the transient day-to-day moves in volatility. Now the volatility of the transient factor can increase. Indeed, its coefficient of variation (standard deviation over mean) rises from 1.00 to 2.01. However, as for the pure-continuous model, the persistent factor is on the boundary of the stationarity condition as the model is struggling to reproduce the pattern of the persistent shifts in "observed" volatility. This shows also that our set of moment conditions identify not only the unconditional distribution of the volatility and its persistence, but it also extracts from the data information about the volatility of the persistent and transient shifts in volatility.

Finally, when we model the two volatility factors to be of pure-jump type, we see that the J test falls to a level that corresponds to $p = 0.205$, i.e., such a specification does not "struggle" any more to fit the moments in the set MC0. We discuss briefly the parameter estimates of this best performing two-factor pure-jump model. First, the implied mean of the persistent V_1 is 0.7580 while that of the transient V_2 is 0.2913. This implies that the estimated unconditional mean of the diffusive volatility is 1.05 (recall we quote in daily percentage units). Note that the Realized Laplace Transform captures only the diffusive volatility and is robust to the price jumps. It has "built-in" truncation and we did not have to remove the "big" price increments in its construction to make it robust to jumps as is done in the Truncated Variation. We will later compare the estimated model's implications for the Integrated Variance with those observed in the data (via the model-free Truncated Variation).

The half life of the persistent factor is 56.81 and of the transient is 0.32. This provides a good fit for the persistent and transient shocks in the volatility observed in the bottom panel of Fig. 2. The coefficient of variation for the persistent factor is 2.1395 while that for the transient factor is 1.5627. Interestingly, the data "requires" quite a volatile persistent factor in addition to the already present volatile transient factor.

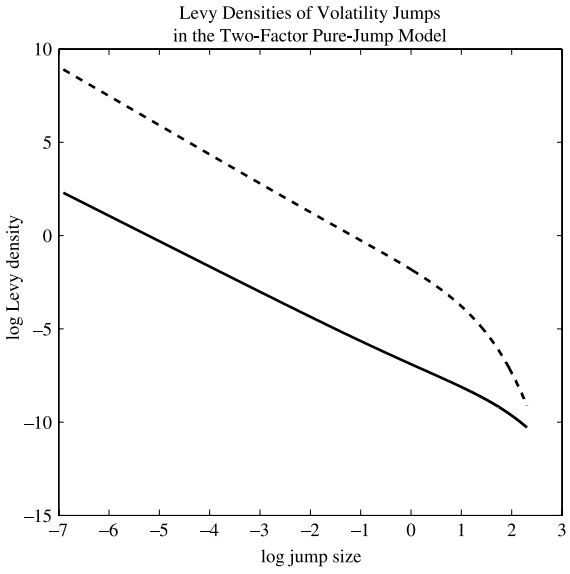

Fig. 3. The Lévy densities of the volatility jumps are defined in (15) and the parameter estimates used in the calculation are reported in the last column of Table 6. The solid line corresponds to V_{1t} and the dotted line to V_{2t}.

On Fig. 3 we contrast the implied Lévy densities of the driving Lévy processes of the two factors. As seen from the figure, the Lévy density of the transient factor is above that of the persistent one except for the very big jumps (not shown on the figure). The transient factor contains more jumps than the persistent one and their effect on the future value of volatility quickly dies out. The slight increase in the wedge between the two densities for the smaller jump sizes is a manifest of $\widehat{\alpha}_2 > \widehat{\alpha}_1$, i.e., the transient volatility factor is more vibrant.[19]

To sum up, the estimation results suggest a very persistent component of volatility which moves mainly through big jumps and fast mean-reverting component which is much more vibrant and captures day-to-day moves in volatility as well as occasional spikes in it. This behavior of the components of the volatility can be clearly seen from Fig. 4 which plots a simulation of the two factors over a period of length as that of our sample.

[19] Note that since both driving jump processes of the volatility factors are infinitely active, their Lévy densities explode at 0.

Table 7
Two-factor pure-jump model diagnostics.

	Alternative u-cutoffs		Alternative moments			
Par	$\widehat{\mathcal{L}}_V(u_{\max}) = 0.15$	$\widehat{\mathcal{L}}_V(u_{\max}) = 0.05$	MC1	MC2	MC3	MC4
κ_1	0.0119	0.0123	0.0118	0.0228	0.0145	0.0229
	(0.0053)	(0.0043)	(0.0043)	(0.0041)	(0.0045)	(0.0041)
α_1	0.5252	0.1994	0.3067	0.5059	0.1120	0.5059
	(0.1205)	(0.1504)	(0.1584)	(0.1698)	(0.1212)	(0.1690)
c_1	0.1490	0.2594	0.1981	0.1501	0.2148	0.1501
	(0.0526)	(0.0787)	(0.0631)	(0.0850)	(0.0512)	(0.0846)
λ_1	0.0835	0.3374	0.1993	0.1233	0.2303	0.1233
	(0.1905)	(0.2094)	(0.1707)	(0.2865)	(0.1271)	(0.2855)
κ_2	2.0505	2.3683	2.6130	2.7780	2.6168	2.7780
	(0.1763)	(0.1640)	(0.1584)	(0.1381)	(0.4359)	(0.1380)
α_2	0.4450	0.5675	0.5888	0.4525	0.5899	0.4525
	(0.1791)	(0.0879)	(0.1331)	(0.1677)	(0.04570)	(0.1671)
c_2	0.1274	0.1283	0.1066	0.1331	0.1361	0.1331
	(0.0481)	(0.0346)	(0.0383)	(0.0528)	(0.0246)	(0.0526)
λ_2	0.5877	0.7693	0.6094	0.7979	0.9915	0.7979
	(0.2128)	(0.2174)	(0.1948)	(0.2644)	(0.1464)	(0.2637)
J Test (df)	1.00(1)	3.23(1)	14.00(4)	10.59(4)	26.81(5)	23.44(5)
P-Val	$p = 0.310$	$p = 0.072$	$p = 0.007$	$p = 0.032$	$p = 0.000$	$p = 0.000$

Note: The alternative sets of moment conditions MC1, MC2, MC3, and MC4 are defined in the text. Standard errors for the parameter estimates are reported in parentheses.

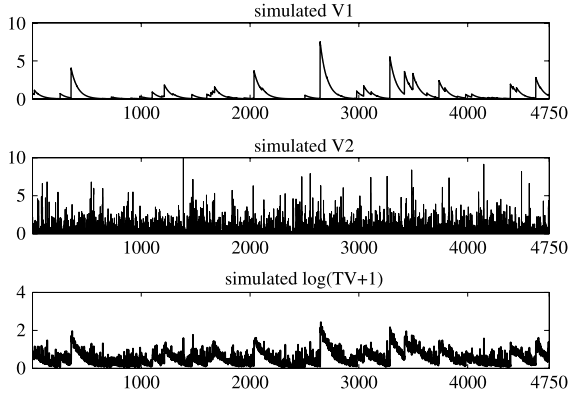

Fig. 4. Simulated volatility from the estimated two-factor pure-jump model with parameter estimates reported in the last column of Table 6. The units on the horizontal axes are simulated (trading) days.

We next test the performance of the best performing model, i.e., the two-factor pure-jump model, in two ways. First, we add additional moments of the integrated joint Laplace transform and also try alternative cutoff levels $u_{\max}$. The list of the alternative sets of moment conditions we test the model are listed below

(1) MC1: Defined in Section 4,
(2) MC2: We replace (d) in MC1 with region $[0.6u_{\max}0.9u_{\max}]^2$ for lags $k = 5, 10, 30$,
(3) MC3: We replace (d) in MC1 with region $[0.1u_{\max}0.2u_{\max}] \times [0.6u_{\max}0.9u_{\max}]$ for lags $k = 1, 5, 10, 30$,
(4) MC4: We replace (d) in MC1 with region $[0.6u_{\max}0.9u_{\max}] \times [0.1u_{\max}0.2u_{\max}]$ for lags $k = 1, 5, 10, 30$.

The results of all the above robustness checks are reported in Table 7. Looking at the overall fit first, i.e., the J test, we can see that under the alternative truncation rules for picking $u_{\max}$ as well as the first two alternative moment sets MC1 and MC2, the model fit is relatively good. The worst of those cases is the set of moment conditions MC1 where the value of the J test corresponds to a p-value of only 0.7% but this is consistent with the slight finite-sample overrejections of the J test reported in the Monte Carlo experiment. On the other hand, the model struggles with the moment sets MC3 and MC4. In these two regions, u is high and v is low and vice versa (recall the definition in 4). Intuitively, the high/low level of u puts relatively more importance to very low/high levels of V_t (for this moment condition) and the same holds true for the link between v and V_{t-k}. Given the above interpretation of the added moments in these two moment sets, the model clearly appears to struggle in matching simultaneously the persistence in volatility and the frequency and speed with which volatility moves from very low to very high levels and vice versa.

Turning next to the parameter estimates of the model across the different estimation setups reported in Table 7, we can see that the parameters controlling the persistence of the volatility factors and the marginal distribution of the transient volatility factor are relatively stable. On the other hand, the parameters controlling the persistent volatility factor vary somewhat over the different cases, most notably the estimates of α_1. Apart from the fact that the latter is hard to identify (which is reflected in its relatively big standard error), this is suggestive of some model misspecification of the persistent component of the volatility.

Our second test for the model performance is to verify how successfully it can fit the moments of the Truncated Variation (which is directly observed). The latter has not been used in the estimation as our inference is based only on the Realized Laplace Transform, and hence this provides a stringent test for the model performance. In Table 8, we compare the first and the second moment as well as the autocorrelation of $\log[TV_t(\alpha, \varpi) + 1]$ implied by the model with that in our data. The transformation $\log(1 + x)$ behaves like x for small values of x but is more robust to the outliers, and hence this transformation of the Truncated Variation is much more reliably estimated from the high-frequency data. This is the reason why we use it in our analysis here.[20] As seen from the table, the model can very comfortably match the moments of the Truncated Variation estimated from the data.[21] This is due to the fact that our estimation procedure selected the model not only by its fit to the mean, variance and persistence of volatility, but rather by its ability to fit the whole transitional density of the volatility.

[20] This is similar to the transformations of measures of realized variation used in e.g., Andersen et al. (2003) when constructing reduced-form based volatility forecasts.

[21] We also redid the calculations in Table 8 by "removing" the deterministic component of volatility in constructing $TV_t(\alpha, \varpi)$ using $\widehat{f}_i$ in (22). This resulted in very small changes in the last column of Table 8.

Table 8
Two-factor pure-jump model diagnostics: implied moments of truncated variation.

Moment	Model-implied	95% CI from data
$\mathbb{E}\,(\log[TV_t(\alpha,\varpi)+1])$	0.5488	[0.4253 0.5613]
$\mathbb{E}\,(\log[TV_t(\alpha,\varpi)+1])^2$	0.5255	[0.2739 0.5564]
AC 1 of $\log[TV_t(\alpha,\varpi)+1]$	0.8763	[0.7948 0.9051]
AC 5 of $\log[TV_t(\alpha,\varpi)+1]$	0.8260	[0.6240 0.8420]
AC 10 of $\log[TV_t(\alpha,\varpi)+1]$	0.7857	[0.5423 0.8114]
AC 30 of $\log[TV_t(\alpha,\varpi)+1]$	0.6518	[0.2324 0.6921]

Note: The model implied moments are computed from a long simulation of the estimated two-factor pure-jump model with parameters reported in the last column of Table 6. Standard errors for the confidence intervals in the last column are computed using the Parzen kernel with lag length of 70.

6. Conclusion

In this paper we propose an efficient method for estimation of parametric models for the volatility of general Itô semimartingales sampled at high-frequencies. The estimation is based on the model-free Realized Laplace Transform of volatility proposed in Todorov and Tauchen (2011b) and is robust to the presence of jumps in the price dynamics. The technique is particularly tractable and easy to apply in volatility models with joint characteristic function known in closed form up to a relatively easy numerical integration. The latter is the case for the class of the general affine jump-diffusion models of Duffie et al. (2000, 2003), which are widely used in financial applications. A Monte Carlo assessment documents good robustness and efficiency properties of the estimator in empirically plausible settings.

The empirical application illustrates the ability of the proposed estimator to extract important information in the data regarding the dynamic properties of volatility. Our method identifies two components of volatility, which is consistent with earlier empirical evidence. However, the method has the power to discriminate among different models for the dynamic properties of the two volatility components, and in particular indicates they are of pure jump type. In the preferred volatility model, the transient volatility component has occasional big spikes but also a lot of small jumps that capture day-to-day variations. On the other hand the persistent volatility factor moves mainly through big jumps—its dynamics are somewhat similar to a regime switching type model but with gradual decay of the high volatility regime.

Finally, our estimation is robust to price level jumps in a way that avoids extra tuning parameters, so inferences regarding volatility are not influenced by an incorrect specification for the price jumps. The general dynamics of volatility and price jumps are well known to be quite different, which, in the literature, motivates modeling of volatility separately from price jumps. In subsequent applications, of course, one needs to reassemble the pieces and therefore naturally be interested also in the price jumps, as they form an important part (between five to fifteen percent) of the total variability risk associated with the asset. Inference for the jump part of the price in the general case is complicated as different parameters of the jump specification can be estimated at various rates even in the relatively simple i.i.d. setting as shown in Ait-Sahalia and Jacod (2008). In the case of finite activity jumps, however, one can adopt the approach of Bollerslev and Todorov (forthcoming) for estimation of jump tails. We leave the systematic study of the problem of parametric estimation of the jump component of the price for future work.

Appendix

A.1. Laplace transforms for affine jump-diffusion models

In general, if V_t is a superposition of independent factors, i.e., if $V_t = \sum_{j=1}^{k} V_{jt}$, then we have

$$\mathcal{L}_V(\mathbf{u};\mathbf{t}) = \prod_{j=1}^{k} \mathcal{L}_{V_j}(\mathbf{u};\mathbf{t}). \tag{23}$$

Therefore, for our inference methods, we will need a formula for $\mathcal{L}_{V_j}(\mathbf{u};\mathbf{t})$ for each of the individual factors.

We do the calculations first for a general affine jump-diffusion volatility factor, and then we specialize to its two special forms: pure-continuous (square-root diffusion) and pure-jump (non-Gaussian OU model) models. For simplicity in the subsequent calculations we refer to the factor as V, i.e., we drop the subscript. Also in what follows we denote with lower case l the log-Laplace transforms (both marginal and joint).

We denote

$$\psi(u) = \int_{\mathbb{R}} (\mathrm{e}^{ux} - 1)\nu(\mathrm{d}x), \quad u \in \mathbb{C} \quad \text{with } \Re(u) \le 0, \tag{24}$$

where recall $\nu(\mathrm{d}x)$ is the Lévy measure of the Lévy subordinator L_t. $\psi(u)$ is the characteristic exponent of L_1. We note that $l_L(u) \equiv \psi(iu)$ for $u \in \mathbb{R}$, where recall $l_L(u)$ denotes the log-Laplace transform.

Set $f(t,v) = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}uV_T}|V_t = v\right)$. Then $f(t,v)$ solves the following partial integro-differential equation

$$\frac{\partial f}{\partial t} + \kappa(\theta - v)\frac{\partial f}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 f}{\partial v^2} + \int_{\mathbb{R}} (f(v+z) - f(z))\nu(\mathrm{d}z), \tag{25}$$

with terminal condition $f(0,v) = \mathrm{e}^{\mathrm{i}uv}$. Guessing a solution of the form

$$f(t,v) = \mathrm{e}^{\alpha(u,T-t)+\beta(u,T-t)v}, \tag{26}$$

reduces the problem to the following system of ODE-s:

$$\begin{cases} \alpha' = \kappa\theta\beta + \psi(\beta), & \alpha(u,0) = 0, \\ \beta' = -\kappa\beta + \dfrac{\sigma^2}{2}\beta^2, & \beta(u,0) = \mathrm{i}u, \end{cases} \tag{27}$$

where α' and β' denote derivatives with respect to t. Thus, finally for $u \in \mathbb{R}$ and $T \ge t$, we have:

$$\begin{aligned} &\mathbb{E}\left(\mathrm{e}^{\mathrm{i}uV_T}|\mathcal{F}_t\right) = \exp\left(\alpha(u,T-t) + \beta(u,T-t)V_t\right) \\ &\alpha(u,s) = \kappa\theta\int_0^s \beta(u,z)\mathrm{d}z + \int_0^s \psi(\beta(u,z))\mathrm{d}z, \\ &\beta(u,s) = \frac{\mathrm{e}^{-\kappa s}\kappa\mathrm{i}u}{\kappa - \mathrm{i}u\sigma^2(1-\mathrm{e}^{-\kappa s})/2}. \end{aligned} \tag{28}$$

A.1.1. Square-root diffusion

Specializing (28) for this case, we get

$$\mathcal{L}_V([u,v];[t,s]) = \left(1 + \frac{u}{c(|t-s|)}\right)^{-2\kappa\theta/\sigma^2} \times \mathcal{L}_V\left(\frac{u\mathrm{e}^{-\kappa|t-s|}}{1 + u/c(|t-s|)} + v\right), \tag{29}$$

where

$$c(z) = \frac{2\kappa}{\sigma^2(1-\mathrm{e}^{-\kappa z})}. \tag{30}$$

The marginal of the square-root diffusion is a Gamma process, see e.g., Cont and Tankov (2004, p. 476), and we have

$$\mathcal{L}_V(u) = \left(\frac{1}{1 + u\sigma^2/(2\kappa)}\right)^{2\kappa\theta/\sigma^2}. \tag{31}$$

A.1.2. Non-Gaussian OU process

Specializing (28) for this case (or even by direct computation), we get

$$\mathcal{L}_V([u,v];[t,s]) = \mathcal{L}_V\left(ue^{-\kappa|t-s|}+v\right) \times \exp\left(\int_0^{|t-s|} l_L\left(ue^{-\kappa z}\right)dz\right). \quad (32)$$

For the non-Gaussian OU model there is a very convenient link (for the purposes of volatility modeling) between the Laplace transform of the driving Lévy subordinator L_t and that of the process V_t. In particular, we have (see e.g., Barndorff-Nielsen and Shephard (2001a))

$$l_L(u) = u\kappa \times l_V'(u), \quad u \geq 0. \quad (33)$$

Hence, once we specify the the Laplace transform of the marginal, we can determine that of the driving Lévy subordinator L_t, and from here easily calculate the joint Laplace transform $\mathcal{L}_V(u,v;t,s)$.

Further the Lévy densities of V_t and L_t, ν_V and ν_L respectively, are linked via (Barndorff-Nielsen and Shephard, 2001a; Sato, 1999)

$$\nu_L(x) = -\kappa(\nu_V(x) + x\nu_V'(x)). \quad (34)$$

Example. *Non-Gaussian OU model with tempered stable marginal distribution.*

The log-Laplace transform of V, i.e., the log-Laplace transform of the tempered stable process, is

$$l_V(u) = \begin{cases} c\Gamma(-\alpha)\left[(\lambda+u)^\alpha - \lambda^\alpha\right], & \text{if } \alpha \in (0,1), \\ -c\log(1+u/\lambda), & \text{if } \alpha = 0, \end{cases} \quad (35)$$

where $\Gamma(-\alpha) = -\frac{1}{\alpha}\Gamma(1-\alpha)$ for $\alpha \in (0,1)$ and Γ is the standard Gamma function. From here, using (33), we easily get

$$l_L(u) = \begin{cases} c\Gamma(-\alpha)\alpha\kappa u(\lambda+u)^{\alpha-1}, & \text{if } \alpha \in (0,1), \\ -\dfrac{c\kappa u}{\lambda+u}, & \text{if } \alpha = 0. \end{cases} \quad (36)$$

$$\int_0^{|t-s|} l_L(ue^{-\kappa z})dz = \begin{cases} c\Gamma(-\alpha)\left[(\lambda+u)^\alpha - (\lambda+ue^{-\kappa|t-s|})^\alpha\right], \\ \quad \text{if } \alpha \in (0,1), \\ -c\left[\log(\lambda+u) - \log(\lambda+ue^{-\kappa|t-s|})\right], \quad \text{if } \alpha = 0. \end{cases} \quad (37)$$

A.2. Details on the simulation of volatility models in the Monte Carlo

To keep notation simple we continue to remove the subscript for the volatility factors. The simulation of the square-root diffusion is done by a standard Euler scheme. The simulation of the non-Gaussian OU processes is done via the following scheme (recall that we need the volatility process V_t on the grid $0, \frac{1}{n}, \frac{2}{n}, \ldots, T$) using

$$V_{\frac{i}{n}} = e^{-\kappa/n}\left(V_{\frac{i-1}{n}} + \int_{\frac{i-1}{n}}^{\frac{i}{n}} e^{\kappa(s-(i-1)/n)}dL_s\right) \approx e^{-\kappa/n}\left(V_{\frac{i-1}{n}} + \sum_{j=1}^{m} e^{\kappa\frac{j-1}{nm}}\left(L_{\frac{i-1}{n}+\frac{j}{nm}} - L_{\frac{i-1}{n}+\frac{j-1}{nm}}\right)\right), \quad i = 1,\ldots,nT. \quad (38)$$

In our case $T = 5000$, $n = 80$ and $m = 80$, which corresponds to discretization of around 4 s.

The simulation of the driving Lévy subordinator in the Monte Carlo is done as follows. We make use of the following representation of L_t for $\alpha \geq 0$ which follows immediately from (34) (see also e.g., Barndorff-Nielsen and Shephard (2001b))

$$L_t \stackrel{d}{=} L_{1t} + L_{2t}, \quad L_{1t} \perp L_{2t},$$

L_{1t} is Lévy process with Lévy measure $\kappa c\alpha\frac{e^{-\lambda x}}{x^{1+\alpha}}1_{\{x>0\}}$,

$$L_{2t} = \sum_{j=1}^{N_t} Y_j,$$

$N_t \sim$ Poisson process with intensity $t \times \kappa c\lambda^\alpha\Gamma(1-\alpha)$ and $Y_j \sim G(1-\alpha,\lambda)$, (39)

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 $a, b > 0$.

For $\alpha = 0.5$, L_{1t} has Inverse-Gaussian distribution, denoted as $IG(\mu,\nu)$, with parameters $\mu = \frac{1}{2}\frac{\Gamma(0.5)\kappa ct}{\sqrt{\lambda}}$ and $\nu = \frac{1}{2}(\kappa ct)^2$ $(\Gamma(0.5))^2$. The Laplace transform of a variable Y with $Y \sim IG(\mu,\nu)$ is given by

$$\mathbb{E}(e^{-uY}) = \exp\left((\nu/\mu)\left[1-\sqrt{1+2\mu^2u/\nu}\right]\right).$$

To simulate $Y \sim IG(\mu,\nu)$, do the following: $x \sim N(0,1)$ and $u \sim U(0,1)$ and denote $z = \mu + \frac{\mu^2x^2}{2\nu} - \frac{\mu}{2\nu}\sqrt{4\mu\nu x^2+\mu^2x^4}$. Then

$$Y = \begin{cases} z & \text{if } u < \mu/(\mu+z), \\ \mu^2/z & \text{else}. \end{cases}$$

Finally, the simulation of the driving Lévy subordinators in the estimated two-factor pure-jump volatility model in Section 5 in which $\alpha_i \neq 0.5$ is done using (39) together with a shot-noise decomposition of the Lévy measure of L_{1t} in (39) with 500,000 shot noise terms on average in each discretization period.

A.3. ARMA representations for integrated variance in one-factor models

Easy but rather tedious calculations show that, see e.g., Todorov (2011), the ARMA(1,1) representation of IV in the one-factor pure-continuous and pure-jump model is given by

$$(IV_t - \mu) - e^{-\kappa}(IV_{t-1} - \mu) = \epsilon_t + \phi\epsilon_{t-1}, \quad (40)$$

where ϵ_t is white noise, i.e., $\mathbb{E}(\epsilon_t\epsilon_s) = 0$ for $t \neq s$. In both cases we have

$$\phi = \frac{1+e^{-2\kappa}-2\eta e^{-\kappa}-\sqrt{(1+e^{-2\kappa}-2e^{-\kappa}\eta)^2-4(\eta-e^{-\kappa})^2}}{2(\eta-e^{-\kappa})},$$

$$\eta = \frac{e^{-\kappa}(e^{-\kappa}-1)^2}{2(e^{-\kappa}+\kappa-1)}. \quad (41)$$

For the rest of the parameters in the ARMA representation we have:

- pure-continuous model

$$\mu = \theta, \quad \mathrm{Var}(\epsilon_t) = \frac{\sigma^2\theta}{2\kappa}\frac{e^{-\kappa}}{\kappa^2\phi}[(e^{-\kappa}-1)^2 - 2(e^{-\kappa}+\kappa-1)], \quad (42)$$

- pure-jump model

$$\mu = c\Gamma(1-\alpha)\lambda^{\alpha-1}, \quad \mathrm{Var}(\epsilon_t) = \frac{ce^{-\kappa}\lambda^{\alpha-2}(\alpha\Gamma(2-\alpha)+\Gamma(3-\alpha))}{2\kappa^2\phi} \times [(e^{-\kappa}-1)^2 - 2(e^{-\kappa}+\kappa-1)]. \quad (43)$$

The QML estimators are found by maximizing the Gaussian likelihood

$$-\frac{1}{2T}\sum_{t=1}^{T}\frac{\widehat{\epsilon}_t^2}{\mathrm{Var}(\epsilon_t)} - \frac{1}{2}\log(\mathrm{Var}(\epsilon_t)), \quad (44)$$

$$\mathbb{E}\left(\frac{\int_{\mathbb{R}_+}\mathrm{Re}\left[\mathrm{e}^{-\mathrm{i}uV_{t+1}}(\partial/\partial\rho)\psi(u,V_t,\rho)\right]\mathrm{d}u\int_{\mathbb{R}_+}\mathrm{Re}\left[\mathrm{e}^{-\mathrm{i}uV_{t+1}}(\partial/\partial\rho')\psi(u,V_t,\rho)\right]\mathrm{d}u}{\left\{\int_{\mathbb{R}_+}\mathrm{Re}\left[\mathrm{e}^{-\mathrm{i}uV_{t+1}}\psi(u,V_t,\rho)\right]\mathrm{d}u\right\}^2}\right) \tag{53}$$

Box I.

where for a given parameter vector, $\widehat{\epsilon}_t$ is determined recursively from the data by $\widehat{\epsilon}_t = (IV_t - \mu) - \mathrm{e}^{-\kappa}(IV_{t-1} - \mu) - \phi\widehat{\epsilon}_{t-1}$ with $\widehat{\epsilon}_0 = 0$ (note that the MA coefficient is smaller than 1 in absolute value). IV_t is estimated from the high-frequency data via (18)–(19).

A.4. Computing the Cramer–Rao lower bound

We wish to compute the Cramer–Rao lower bound for the parameter vector ρ of model defined by (13) in the main text on the assumption the volatility process V_t is observed at integer values of t. Let $p(V_{t+1}|V_t,\rho)$ denote the model-implied density so the task is to compute the information matrix

$$\mathbb{E}\left(\left\{\frac{\partial}{\partial\rho}\log[p(V_{t+1}|V_t,\rho)]\right\}\left\{\frac{\partial}{\partial\rho}\log[p(V_{t+1}|V_t,\rho)]\right\}'\right). \tag{45}$$

A.4.1. Pure-continuous volatility models, cases A–C

For these cases the parameter vector is $\rho = (\kappa\,\theta\,\sigma)'$ and the conditional density of $2\,c(\rho)\,V_{t+1}$ is non-central chi-squared where

$$c(\rho) = \frac{2\kappa}{\sigma^2(1-\mathrm{e}^{-\kappa\Delta})}, \tag{46}$$

the degrees of freedom parameter is

$$df(\rho) = \frac{4\kappa\theta}{\sigma^2}, \tag{47}$$

and non-centrality parameter

$$\nu_t(\rho) = 2c(\rho)\mathrm{e}^{-\kappa\delta}V_t. \tag{48}$$

Thus the gradient term for (45) is

$$\begin{aligned}&\frac{\partial}{\partial\rho}\log[p(V_{t+1}\,|V_t,\rho\,)]\\ &\quad=\frac{\partial}{\partial\rho}\log\left\{2c(\rho)\,n\left[2c(\rho)V_{t+1}\,|\mathrm{d}f(\rho),\nu_t(\rho)\,\right]\right\},\end{aligned} \tag{49}$$

where $n(\cdot|\mathrm{d}f,\nu)$ is the noncentral chi squared density. We use numerical derivatives, which appeared quite stable and accurate, to compute the gradient term immediately above and then Monte Carlo to compute the expectation in (45).

A.4.2. Pure-jump volatility models, cases D and E

For these cases $\rho = (\kappa\,\alpha\,c\,\lambda)'$. We need to work from the conditional characteristic function because the density $p(V_{t+1}\,|V_t,\rho\,)$ is not available in convenient closed form. Define the conditional characteristic function

$$\psi(u,V_t,\rho) = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}uV_{t+1}}\,|V_t\right), \tag{50}$$

and from the Fourier inversion formula the transition density is up to a constant

$$\int_{\mathbb{R}_+}\mathrm{Re}\left[\mathrm{e}^{-\mathrm{i}uV_{t+1}}\psi(u,V_t,\rho)\right]\mathrm{d}u, \tag{51}$$

with gradient

$$\int_{\mathbb{R}_+}\mathrm{Re}\left[\mathrm{e}^{-\mathrm{i}uV_{t+1}}(\partial/\partial\rho)\psi(u,V_t,\rho)\right]\mathrm{d}u. \tag{52}$$

Recalling that $(\partial/\partial\rho)\log(p) = (1/p)(\partial/\partial\rho)p$, then whenever the derivatives exist and the magnitude of the characteristic function (and its derivatives) is dominated by an integrable function not dependent upon ρ (45) becomes (see Box I).

For Cases D and E the gradient of the characteristics functions is

$$\begin{aligned}(\partial/\partial\kappa)\psi(u,V_t,\rho) &= \psi(u,V_t,\rho)(V_t+\alpha c\,\Gamma(-\alpha)\\ &\quad\times(\lambda-\mathrm{i}u\mathrm{e}^{-\kappa\delta})^{(\alpha-1)})\mathrm{i}u\delta\mathrm{e}^{-\kappa\delta};\\ (\partial/\partial\alpha)\psi(u,V_t,\rho) &= -\psi(u,V_t,\rho)c\,\Gamma(-\alpha)(\Psi(1-\alpha)\\ &\quad+1/\alpha)((\lambda-\mathrm{i}u)^\alpha-(\lambda-\mathrm{i}u\mathrm{e}^{(-\kappa\delta)})^\alpha)\\ &\quad+\log(\lambda-\mathrm{i}u)(\lambda-\mathrm{i}u)^\alpha\\ &\quad-\log(\lambda-\mathrm{i}u\mathrm{e}^{(-\kappa\delta)})(\lambda-\mathrm{i}u\mathrm{e}^{(-\kappa\delta)})^\alpha;\\ (\partial/\partial c)\psi(u,V_t,\rho) &= \psi(u,V_t,\rho)\Gamma(-\alpha)((\lambda-\mathrm{i}u)^\alpha\\ &\quad-(\lambda-\mathrm{i}u\mathrm{e}^{-\kappa\delta})^\alpha);\\ (\partial/\partial\lambda)\psi(u,V_t,\rho) &= \psi(u,V_t,\rho)c\,\Gamma(-\alpha)\alpha((\lambda-\mathrm{i}u)^{(\alpha-1)}\\ &\quad-(\lambda-\mathrm{i}u\mathrm{e}^{(-\kappa\delta)})^{(\alpha-1)});\end{aligned}$$

To compute (53) the inner integrals with respect to u are done numerically while the outer integration over the joint distribution of V_t, V_{t+1} is done by Monte Carlo.

The integrand $\mathrm{Re}\left[\mathrm{e}^{-\mathrm{i}uV_{t+1}}(\partial/\partial\rho)\psi(u,V_t,\rho)\right]$ in (53) may exhibit highly oscillatory behavior, which make numerical integration difficult. Filipovic et al. (2010) employ special routines QAWF and QAWO from the GNU scientific library to numerically compute the Fourier integral of a characteristic function. However, those routines assume that the main source of oscillations in the integrand is the sine (or cosine) factor coming from the Fourier transform. Wild oscillations of the characteristic function and its gradient, which in our case happen for several values of V_t, V_{t+1}, and ρ, may make the QAWF and QAWO algorithms fail.

In such cases, we employ the adaptive Gauss–Kronrod integration. This numerical integration procedure handles the oscillatory behavior of integrands quite well and is implemented both in Matlab and the GNU scientific library.

In Case D, when $\kappa = 0.5$, we can use the QAWF and QAWO algorithms to compute the inner integrals in (53). However, in Case E, when κ decreases to 0.15, the frequency of oscillations in the characteristic function and its derivatives increases too much so that QAWF and QAWO stop working. Case E, therefore, is computed using the QAGIU routine, which implements the adaptive Gauss–Konrod integration.

References

Ait-Sahalia, Y., Jacod, J., 2008. Fisher's information for discretely sampled Lévy processes. Econeomtrica 76, 727–761.

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

Ait-Sahalia, Y., Jacod, J., 2009b. Testing for jumps in a discretely observed process. 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.

Andersen, T., Sørensen, B., 1996. GMM estimation of a stochastic volatility model: a Monte Casrlo study. Journal of Business and Economic Statistics 14, 328–352.

Andersen, T.G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

Barndorff-Nielsen, O., Shephard, N., 2001a. 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., 2001b. Normal modified stable processes. Theory of Probability and Mathematical Statistics 65, 1–19.

Barndorff-Nielsen, O., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society Series B 64, 253–280.

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., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4, 1–30.
Bates, D., 2006. Maximum likelihood estimation of latent affine models. Review of Financial Studies 909–965.
Blumenthal, R., Getoor, R., 1961. Sample functions of stochastic processes with independent increments. Journal of Mathematics and Mechanics 10, 493–516.
Bollerslev, T., Todorov, V., 2011. Estimation of jump tails. Econometrica (forthcoming).
Bollerslev, T., Zhou, H., 2002. Estimating stochastic volatility diffusion using conditional moments of integrated volatility. Journal of Econometrics 109, 33–65.
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.
Carrasco, M., Florens, J.P., 2000. Generalization of gmm to a continuum of moment conditions. Econometric Theory 797–834.
Carrasco, M., Florens, J.P., 2002. GMM Estimation using the empirical characteristic function. Technical report, Institut National de la Statistique et des Etudes Economiques.
Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.
Chernozhukov, V., Hong, H., 2003. An MCMC approach to classical estimation. Journal of Econometrics 115, 293–346.
Cont, R., Tankov, P., 2004. Financial Modelling with Jump Processes. Chapman and Hall, Boca Raton, Florida, USA.
Corradi, V., Distaso, W., 2006. Semiparametric comparision of stochastic volatility models using realized measures. Review of Economic Studies 73, 635–667.
Dobrev, D., Szerszen, P., 2010. The information content of high-frequency data for estimating equity return models and forecasting risk. Technical report.
Duffie, D., Filipović, D., Schachermayer, W., 2003. Affine processes and applications in finance. Annals of Applied Probability 13 (3), 984–1053.
Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.
Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269–1300.
Feuerverger, A., Mureika, R., 1977. The empirical characteristic function and its application. Annals of Statistics 5, 88–97.
Filipovic, D., Mayerhofer, E., Schneider, P., 2010. Transition density approximations for multivariate affine jump diffusion processes. Working paper.
Hansen, L., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.
Huang, X., Tauchen, G., 2005. The relative contributions of jumps to total variance. Journal of Financial Econometrics 3, 456–499.
Jacod, J., 2008. Asymptotic properties of power variations and associated functionals of semimartingales. Stochastic Processes and their Applications 118, 517–559.
Jiang, G., Knight, J., 2002. Estimation of continuous-time processes via empirical characteristic function. Journal of Business and Economic Statistics 20, 198–212.
Jiang, G.J., Knight, J.L., 2010. ECF estimation of Markov models where the transition density is unknown. Econometrics Journal 13, 245–270.
Knight, J.L., Yu, J., 2002. Empirical characteristic function in time series estimation. Econometric Theory 18, 691–721.
Mancini, C., 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics 36, 270–296.
Meddahi, N., 2003. ARMA representation of integrated and realized variances. The Econometrics Journal 6, 334–355.
Parzen, E., 1962. On estimation of a probability density function and mode. Annals of Mathematical Statistics 33, 1065–1076.
Paulson, A.S., Holcomb, E.W., Leitch, R.A., 1975. The estimation of the parameters of the stable laws. Biometrika 62, 163–170.
Prakasa Rao, B., 1988. Statistical inference from sampled data for stochastic processes. In: In: Contemporary Mathematics, vol. 80. Amer. Math. Soc., Providence.
Rosiński, J., 2007. Tempering stable processes. Stochastic Processes and their Applications 117, 677–707.
Sato, K., 1999. Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge, UK.
Singleton, K., 2001. Estimation of affine asset pricing models using the empirical characteristic function. Journal of Econometrics 102, 111–141.
Theodosiou, M., Zikes, F., 2009. A comprehensive comparison of alternative tests for jumps in asset prices. Technical report, Imperial College London, Business School.
Todorov, V., 2009. Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. Journal of Econometrics 148, 131–148.
Todorov, V., 2011. Econometric analysis of jump-driven stochastic volatility models. Journal of Econometrics 160, 12–21.
Todorov, V., Tauchen, G., 2010. Activity signature functions for high-frequency data analysis. Journal of Econometrics 154, 125–138.
Todorov, V., Tauchen, G., 2011a. Volatility jumps. Journal of Business and Economic Statistics 29, 356–371.
Todorov, V., Tauchen, G., 2011b. The realized Laplace transform of volatility. Working paper.
Todorov, V., Tauchen, G., 2011c. Realized Laplace transforms for pure-jump semimartingales. Working paper.
Yu, J., 2004. Empirical characteristic function estimation and its applications. Econometric Reviews 23, 93–123.