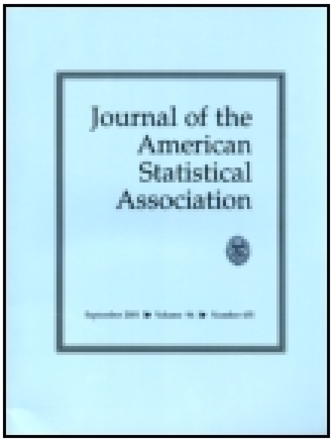

Journal of the American Statistical Association
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uasa20

Inverse Realized Laplace Transforms for Nonparametric Volatility Density Estimation in Jump-Diffusions

Viktor Todorov [a] & George Tauchen [a]

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

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

Accepted author version posted online: 14 May 2012.Published online: 24 Jul 2012.

To cite this article: Viktor Todorov & George Tauchen (2012) Inverse Realized Laplace Transforms for Nonparametric Volatility Density Estimation in Jump-Diffusions, Journal of the American Statistical Association, 107:498, 622-635, DOI: 10.1080/01621459.2012.682854

To link to this article: http://dx.doi.org/10.1080/01621459.2012.682854

Inverse Realized Laplace Transforms for Nonparametric Volatility Density Estimation in Jump-Diffusions

Viktor TODOROV and George TAUCHEN

This article develops a nonparametric estimator of the stochastic volatility density of a discretely observed Itô semimartingale in the setting of an increasing time span and finer mesh of the observation grid. There are two basic steps involved. The first step is aggregating the high-frequency increments into the realized Laplace transform, which is a robust nonparametric estimate of the underlying volatility Laplace transform. The second step is using a regularized kernel to invert the realized Laplace transform. These two steps are relatively quick and easy to compute, so the nonparametric estimator is practicable. The article also derives bounds for the mean squared error of the estimator. The regularity conditions are sufficiently general to cover empirically important cases such as level jumps and possible dependencies between volatility moves and either diffusive or jump moves in the semimartingale. The Monte Carlo analysis in this study indicates that the nonparametric estimator is reliable and reasonably accurate in realistic estimation contexts. An empirical application to 5-min data for three large-cap stocks, 1997–2010, reveals the importance of big short-term volatility spikes in generating high levels of stock price variability over and above those induced by price jumps. The application also shows how to trace out the dynamic response of the volatility density to both positive and negative jumps in the stock price.

KEY WORDS: High-frequency data; Ill-posed problems; Nonparametric density estimation; Regularization; Stochastic volatility.

1. INTRODUCTION

Continuous-time models are widely used in empirical finance to model the evolution of financial asset prices. The absence of arbitrage (under some technical conditions) implies that traded financial assets should be semimartingales, and typically, most, if not all, applications restrict attention to Itô semimartingales, that is, semimartingales with characteristics absolutely continuous in time. Thus, the standard asset pricing model for the log-financial price X_t is of the form

$$dX_t = \alpha_t dt + \sqrt{V_t} dW_t + dJ_t, \tag{1}$$

where α_t and $V_t > 0$ are processes with càdlàg paths, W_t is a Brownian motion, and J_t is a jump process.

In Section 2, we give the technical conditions for the components comprising X in (1), while here, we briefly describe their main roles in determining the dynamics of X. α_t captures risk premium (and possibly risk-free rate) and is well known to be present, but will not be the object of interest in this article. As will be seen below, it gets filtered out in our estimation process. The jump component, J_t, highlighted by Barndorff-Nielsen and Shephard (2004), among others, accounts, according to earlier empirical evidence, typically for 5%–15% of the total variance of the increments in X. Jumps reflect the fact that financial time series exhibit very sharp short-term moves incompatible with the continuous sample paths implied by diffusive models. Much of the evidence on jumps has been adduced using very-high-frequency data; for example, see Barndorff-Nielsen and Shephard (2006) and Ait-Sahalia and Jacod (2009); earlier efforts using coarsely sampled (daily) data were at best mildly successful in handling both jumps and diffusive price moves. Similar to α_t, J_t will be filtered out in our analysis. The volatility term, V_t, represents the dominant component of the variance of increments in X and thus is most widely studied; for example, see Mykland and Zhang (2009) and the many references therein. The volatility process is well known to be negatively correlated with the increments in the driving diffusion W_t. This negative correlation is the so-called "leverage" effect, a term proposed by Black (1976), and the effect has been extensively documented in a wide variety of studies using various statistical methods. Of course, price and volatility can have dependence beyond the "leverage" effect, as in the symmetric generalized autoregressive conditional heteroscedasticity (GARCH) processes; for example, see Klüppelberg, Lindner, and Maller (2004).

This article develops a method to estimate nonparametrically from high-frequency data (by way of a Laplace transform) the marginal law of the stochastic volatility process V_t as well as its conditional law for certain interesting events. For reasons just described, we develop the density estimation method within a very general setting where V_t and W_t can be dependent, and jumps and a drift term are present as well. Although the preceding discussion is for X being a financial asset price, the results in this article obviously apply to any statistical application where high-frequency observations of an Itô semimartingale (which includes many continuous-time models) are available.

On an intuitive level, our method can be described as follows. We use a cosine transformation of appropriately rescaled high-frequency returns data (which is akin to using a bounded influence function) that essentially separates out the jumps and the drift, thereby leaving (essentially) the diffusive piece scaled by $\sqrt{V_t}$. Averaging the transform over time yields the

Viktor Todorov is Associate Professor, Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208 (E-mail: *v-todorov@northwestern.edu*). George Tauchen is Professor, Department of Economics, Duke University, Durham, NC 27708 (E-mail: *george.tauchen@duke.edu*). The authors thank the editor, associate editor, and two anonymous referees for their detailed comments that led to significant improvements. The authors also thank Tim Bollerslev, Francis Diebold, Jia Li, and Andrew Patton, as well as many seminar participants. Todorov's work was partially supported by the National Science Foundation grant SES-0957330.

Journal of the American Statistical Association
June 2012, Vol. 107, No. 498, Theory and Methods
DOI: 10.1080/01621459.2012.682854

realized Laplace transform of volatility studied in Todorov and Tauchen (2012). This transform estimates the real-valued Laplace transform of the underlying spot volatility process, and it further achieves this without any need for staggering of price increments, explicit truncation, or other techniques involving tuning parameters commonly used for jump-robust measures of volatility.

Real Laplace transforms uniquely identify the distribution of nonnegative random variables, so a second step in the estimation is to invert the realized Laplace transform of volatility and thereby recover an estimate of the volatility density. The task of inverting real Laplace transforms also arises in analysis of certain physical phenomena. Historically, inversion of the real Laplace transform, where transform values are only available on the nonnegative real axis but not on the entire complex plane, was among the most notorious of all ill-posed problems. However, recent regularization algorithms developed by Kryzhniy (2003a,b), along with the availability of high-speed computing equipment (for nested numerical integrations), render the inversion a quick and easy task to compute in a matter of a few minutes using standard software such as MATLAB.

The role of regularization in this context is to guarantee statistical consistency when the volatility Laplace transform is recovered with a sampling error, as is the case here. Our nonparametric volatility density estimate is an integral on $\mathbb{R}_+$ of the realized Laplace transform multiplied by a (deterministic) regularized kernel. To analyze the asymptotic behavior of this integral, the local uniform asymptotics of the realized Laplace transform derived in Todorov and Tauchen (2012) does not suffice. Here, therefore, we derive the asymptotic behavior of the realized Laplace transform considered as a process in a weighted $L_2(\mathbb{R}_+)$ space, which requires, in particular, to bound the discretization error for increasing values of the argument of the Laplace transform. We further derive bounds on the magnitude of the error in the nonparametric density estimation due to the regularization. Combining these results, we are able to bound the mean squared error of our nonparametric volatility density estimate and show that it achieves optimal rate of convergence (for the assumed smoothness of the density).

We can compare our method with the deconvolution approach of Van Es, Spreij, and Van Zanten (2003) and Comte and Genon-Catalot (2006). The methods in these articles are developed for a process without jumps, that is, without J_t in (1), and with stochastic volatility V_t independent from the Brownian motion W_t in (1) (and further without drift in the case of Comte and Genon-Catalot 2006). In such a setup, the logarithm of the squared price increments is (approximately) a sum of signal (the log volatility) and noise, and one can apply a deconvolution kernel (Van Es, Spreij, and Van Zanten 2003) or a penalized projection (Comte and Genon-Catalot 2006) to generate a nonparametric estimate of the distribution of the log volatility. The rate of convergence of this estimation depends, of course, on the smoothness of the volatility density, with a logarithmic rate in the least favorable situation. In our case, similar rates apply on comparable smoothness conditions for the density but under the much weaker and empirically plausible assumptions on jumps and leverage for X in (1). On a more practical level, the method here avoids taking logarithms of high-frequency squared price increments, which could be problematic in some instances, because the log transformation inflates (despite the centering) the nontrivial number of zero returns due to discretization, and it further weighs more heavily the smaller increments that are more prone to microstructure noise effects.

We test our method on simulated data that mimic a typical dataset available in finance and find that the method can recover reasonably well the volatility density. We further provide guidance on the choice of the regularization parameter.

In an empirical application, we investigate the distribution of the spot volatility for three large-cap stocks. An earlier work by Andersen et al. (2001) investigated the distribution of daily realized volatility (which is the sum of the squared daily high-frequency increments) of financial series—exchange rates in their case. Here, we go one step further and recover the distribution of spot volatility. Spot volatility and realized volatility differ due to time aggregation as well as due to the presence of price jumps. The evidence suggests that the density of spot volatility is less concentrated around the mode, with more mass in the extreme tails, than that of a (jump-robust) realized volatility measure. This latter finding underscores the presence of short-term volatility spikes. We also invert the realized Laplace transform on days following a significant price jump and provide nonparametric evidence that volatility increases significantly after jumps, with diminishing impact over time. Overall, the nonparametric analysis sheds light on the importance of the variability of the stochastic volatility process V_t in accounting for big moves in asset prices, in addition to pure price jumps generated by J_t.

The rest of the article is organized as follows. Section 2 states the problem and the assumptions, and presents the main asymptotic results. Section 3 reports numerical experiments on the inversion of real-valued Laplace transform using our proposed method—both in the case when the latter is the true transform and in the case where it is recovered from high-frequency observations of a jump-diffusion process. Section 4 reports on the empirical application to high-frequency stock data. Section 5 concludes. The Appendix contains the proof of Theorem 1.

2. MAIN RESULTS

We start with describing our estimation method and deriving its asymptotic behavior.

2.1 Setup and Definitions

We first state the necessary assumptions that we will need. The observed process X is given by its dynamics specified in (1) and is defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ that satisfies the usual conditions. Our assumption for X is as follows:

Assumption A. For the process X specified in (1), assume

A1. For every $t \geq 0$, we have $\mathbb{E}\{|\alpha_t|^2 + |V_t|^2 + |J_t|^2\} \leq K$ for some positive constant K.

A2. J_t is a jump process of the form $J_t = \int_0^t \int_{\mathbb{R}} \kappa(x)\widetilde{\mu}(ds, dx) + \int_0^t \int_{\mathbb{R}} \kappa'(x)\mu(ds, dx)$, where μ is an integer-valued measure on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\nu(ds, dx)$; $\widetilde{\mu}(ds, dx) = \mu(ds, dx) - \nu(ds, dx)$; $\kappa(x)$ is a continuous function, with $\kappa(x) = x$ around zero, or identically zero around zero, and zero outside a compact set containing the zero; and $\kappa'(x) = x - \kappa(x)$.

A3. For every $t,s \geq 0$, we have $\mathbb{E}(|V_t - V_s|^2|\mathcal{F}_{s\wedge t}) \leq K_{s\wedge t}|t-s|$ for some $\mathcal{F}_{s\wedge t}$-measurable random variable $K_{s\wedge t}$, with $\mathbb{E}|K_{s\wedge t}|^{2+\iota} < K$ for some positive constant K and arbitrary small $\iota > 0$, and $\mathbb{E}|J_t - J_s|^p \leq K|t-s|$ for every $p \in (\beta, 2]$, some $\beta \in [0, 2)$, and a positive constant K.

A4. V_t is a stationary and α-mixing process, with $\alpha_t^{\text{mix}} = O(t^{-1-\iota})$ when $t \to \infty$ for some arbitrary small $\iota > 0$, where

$$\alpha_t^{\text{mix}} = \sup_{A\in\mathscr{F}_0, B\in\mathscr{F}^t} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$
$$\mathscr{F}_0 = \sigma(V_s, s \leq 0), \text{ and}$$
$$\mathscr{F}^t = \sigma(V_s, s \geq t). \tag{2}$$

Assumption A1 imposes some mild integrability conditions on the different components of X. Some of them can be potentially relaxed, but nevertheless, they are very weak and satisfied in virtually all parametric models used in empirical finance. Assumption A2 specifies the jump process in X_t. We note that there is very little structure that is assumed for the jumps and, in particular, time variation in the jump compensator (in the form of both time-varying jump size and time-varying jump intensity) is allowed for. In Assumption A3, we impose restriction on the variability in the processes V_t and J_t. The part of A3 concerning V_t is very minimal and, in particular, is satisfied when V_t is an Itô semimartingale (as is the case in the popular affine jump-diffusion framework), but it also holds for certain long-memory specifications (A3 also strengthens slightly the integrability condition for V_t in A1). We also point out that Assumption A allows for jumps in V_t that can have arbitrary dependence with J_t, which will be of practical importance, as we will see in the empirical application. The restriction of A3 on the jump component J_t is that the so-called Blumenthal–Getoor index of the latter (which can be random) is bounded by β. We note that we allow for $\beta > 1$, which means that infinite variation jumps are included in our analysis as well. Finally, A4 is a (standard) mixing condition on the volatility process and is satisfied in wide classes of volatility models.

As stated already in the Introduction, our goal in this article is to recover nonparametrically the density of the spot volatility marginal law (with respect to the Lebesgue measure), which we denote with $f(x)$ (and assume to exist almost everywhere, and it further does not depend on t, as we are interested in the case when volatility is a stationary process). Our next assumption imposes the necessary conditions on $f(x)$.

Assumption B. The marginal law of the stationary process V_t has density $f(x)$ which is piecewise continuous and has piecewise continuous derivatives on $[0, \infty)$ with $f(0+)$ and $f'(0+)$ possibly infinite. We further have

B1. $f'(x) = O(x^{-q})$ as $x \to 0$, and $f'(x) = O(x^{-1/2-\iota})$ and $f(x) = O(1)$ as $x \to \infty$, for some nonnegative $q < 5/2$ and arbitrary small $\iota > 0$.

B2. $f(x)$ and $f'(x)$ are bounded on $\mathbb{R}_+$, with $f(x) = o(x^{-1-\iota})$ and $f'(x) = o(x^{-2-\iota})$ for some arbitrary small $\iota > 0$ when $x \to \infty$ and $f(0+) = 0$.

The degree of smoothness of the density naturally impacts the precision of the estimation, as in standard nonparametric density estimation (based on direct observations of the process), and the above assumption provides such conditions. Assumption B1 is quite weak and allows for the density of V_t to explode around zero. Assumption B2 strengthens B1 by ruling out explosions around zero and further requiring a rate of decay of the volatility density (and its derivative) at infinity.

Our strategy of estimating nonparametrically the volatility density from high-frequency observations of X_t is based on first recovering the Laplace transform of the volatility density and then inverting it. To this end, we denote the *real-valued* Laplace transform of the marginal distribution of the process V_t with

$$\mathcal{L}(u) = \mathbb{E}[\exp(-uV_t)], \ u \geq 0. \tag{3}$$

In Todorov and Tauchen (2012), we have proposed the following nonparametric estimate of the unobserved volatility Laplace transform from high-frequency observation of X_t on the discrete equidistant grid $0, \frac{1}{n}, \ldots, \frac{i}{n}, \ldots, T$:

$$\widehat{\mathcal{L}}(u) = \frac{1}{nT}\sum_{i=1}^{nT}\cos\left(\sqrt{2un}\Delta_i^n X\right),$$
$$\Delta_i^n X = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}, \ u \geq 0, \tag{4}$$

which we refer to as realized Laplace transform. As shown in Todorov and Tauchen (2012), for $T \to \infty$ and $n \to \infty$, we have locally uniformly in u

$$\widehat{\mathcal{L}}(u) \xrightarrow{\mathbb{P}} \mathcal{L}(u), \tag{5}$$

and there is an associated central limit theorem, but we will not make use of it here. We note, in particular, that $\widehat{\mathcal{L}}(u)$ is robust to the presence of jumps in X (the component J_t in (1)) as well as to any dependence between the volatility process V_t and the Brownian motion W_t.

The results that follow will continue to hold if the observation times are nonequidistant but still nonrandom (by conditioning, they can also be further extended to the case when the sampling is random but independent from the process X); that is, if on the interval $[0, T]$, we observe X on the discrete grid, then $0 = \tau(n, 0) < \tau(n, 1) < \cdots < \tau(n, i) < \cdots$ ($\{\tau(n, i) : i \geq 0, n \geq 1\}$ is a double sequence, with n indexing the sequence of discretization grids). We denote $\Delta(n, i) = \tau(n, i) - \tau(n, i-1)$, and with $\pi_T^n = \sup_{i:\tau(n,i)\leq T}\Delta(n, i)$, the mesh of the grid on $[0, T]$. Then, $\widehat{\mathcal{L}}(u)$ in the nonequidistant case gets generalized to $\widehat{\mathcal{L}}(u) = \frac{1}{T}\sum_{i:\tau(n,i)\leq T}\Delta(n, i)\cos(\sqrt{2u\Delta(n, i)^{-1}}\Delta_i^n X)$, and we need $\pi_T^n \to 0$ as $T \to \infty$ and $n \to \infty$ for the consistency in (5). Further, the limit results in Theorem 1 continue to hold with n^{-1} replaced with π_T^n.

The (real-valued) Laplace transform of a nonnegative random variable uniquely identifies its distribution (see, e.g., Feller 1971). However, recovering the distribution from the Laplace transform is an ill-posed problem (Tikhonov and Arsenin 1977), and hence, one needs a regularization to make the inversion problem a continuous operator on the space of Laplace transforms. Here, we adopt an approach proposed in Kryzhniy (2003a,b) and propose the following regularized inversion of the true Laplace transform $\mathcal{L}(u)$:

$$f_R(x) = \int_0^\infty \mathcal{L}(u)\Pi(R, xu)du, \tag{6}$$

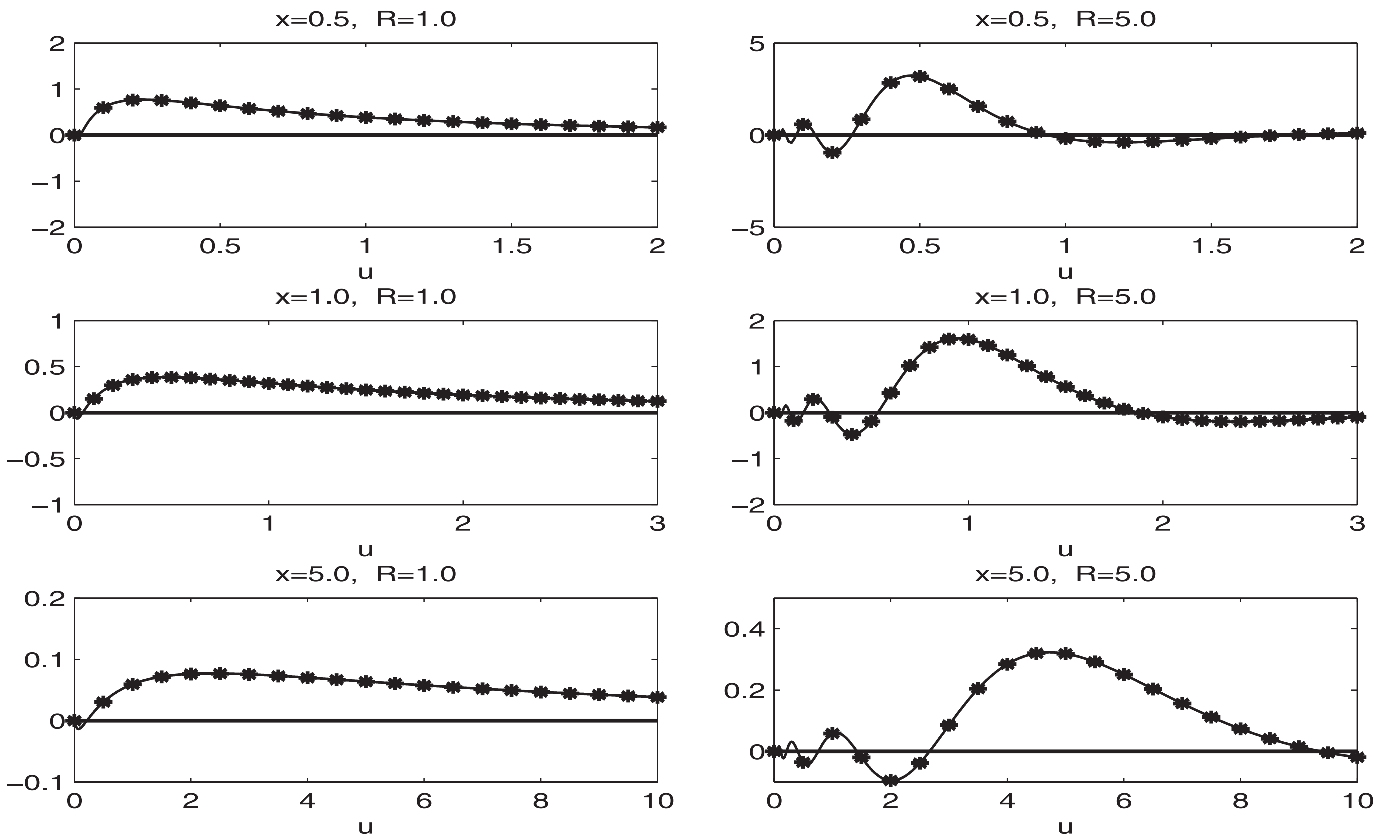

Figure 1. The function $\delta_{R,x}(u)$. The solid lines correspond to the theoretical value of $\delta_{R,x}(u)$ given in (8). The dots in the plots correspond to our numerical evaluation of $\delta_{R,x}(u)$ by evaluating the integral in (6) with $\mathcal{L}(u) = \exp(-ux)$.

where $R > 0$ is a regularization parameter and the kernel $\Pi(R, x)$ is defined as

$$\Pi(R,x) = \frac{4}{\sqrt{2}\pi^2}\left[\sinh\left(\frac{\pi R}{2}\right)\int_0^\infty \frac{\sqrt{u}\cos(R\ln(u))}{u^2+1}\sin(xu)du + \cosh\left(\frac{\pi R}{2}\right)\int_0^\infty \frac{\sqrt{u}\sin(R\ln(u))}{u^2+1}\sin(xu)du\right]. \quad (7)$$

As shown in Kryzhniy (2003a), $f_R(x) \to \frac{f(x-)+f(x+)}{2}$ for every $x > 0$ (pointwise) as $R \to \infty$, where we define $f(x+) = \lim_{y\downarrow x} f(y)$ and $f(x-) = \lim_{y\uparrow x} f(y)$. We further have

$$\left|\int_0^\infty (\mathcal{L}_1(u) - \mathcal{L}_2(u))\Pi(R, xu)du\right| \le K \sup_{u\in\mathbb{R}_+} |\mathcal{L}_1(u) - \mathcal{L}_2(u)|,$$

for any two Laplace transforms $\mathcal{L}_1(u)$ and $\mathcal{L}_2(u)$ (Kryzhniy 2003a, theorem 2) and a positive constant K, which shows that this is indeed a regularization of the ill-posed inversion problem (Tikhonov and Arsenin 1977).

It is easy to develop intuition about the regularization by using the connection between the regularized density and the true density derived in Kryzhniy (2003a)

$$f_R(x) = \int_0^\infty f(u)\delta_{R,x}(u)du,$$
$$\delta_{R,x}(u) = \frac{2\sqrt{ux}}{\pi(u^2-x^2)}\sin\left(R\ln(u/x)\right), \quad x > 0. \quad (8)$$

The function $\delta_{R,x}(u)$ is a smooth approximation of the Dirac delta function at the point x. The regularization parameter R determines the degree of smoothing and corresponds to the choice of the bandwidth in regular nonparametric kernel estimators where one has direct observations of the variable of interest (unlike here, where we do not directly observe V_t). Higher values of R means that $\delta_{R,x}(u)$ is closer to the Dirac delta, and hence, this implies less smoothing. However, these higher values can lead to a good result only if the precision of the input (here, the realized Laplace transform) is high; otherwise the oscillations in $\delta_{R,x}(u)$ will cause very noisy density estimates. The exact opposite holds for low values of R. This is further confirmed from Figure 1, where we plot the function $\delta_{R,x}(u)$ for several different values of x and a low and a high value of R, which we will actually use in our numerical work later on.

An alternative representation of $\delta_{R,x}(u)$ is as the regularized inverse of the function $\exp(-ux)$, that is, $\delta_{R,x}(u) = \int_0^\infty \exp(-tx)\Pi(R, ut)dt$. We can use this connection to check the impact on the estimation of the error due to the numerical integration involved in computing (6) and (7). We plot the resulting estimates of $\delta_{R,x}(u)$ with the dotted lines in Figure 1. As can be observed from the figure, the dotted lines plot on top of the solid lines (which correspond to the theoretical value of $\delta_{R,x}(u)$ in (8)), which indicates that the numerical error is negligible for the values of R used in the computations (which covers the range of R that we are going to use in practice).

The feasible analogue of $f_R(x)$, based on the realized Laplace transform, is naturally defined as

$$\widehat{f}_R(x) = \int_0^\infty \widehat{\mathcal{L}}(u)\Pi(R, xu)du, \quad (9)$$

and this will be our nonparametric estimate of the density of V_t from the discrete observations of the underlying process $X_0, X_{\frac{1}{n}}, \dots, X_T$. The local uniform asymptotics for $\widehat{\mathcal{L}}(u)$ developed in Todorov and Tauchen (2012) does not suffice to study $\widehat{f}_R(x)$ since the integral in (9) is defined on $\mathbb{R}_+$. To analyze the asymptotics of $\widehat{f}_R(x)$ (both pointwise and as a function), we need: (a) asymptotics of $\widehat{\mathcal{L}}(u)$ considered as a process on weighted $L_2(\mathbb{R}_+)$ space and (b) bounds for the order of magnitude of the regularization error. We turn to this next.

2.2 Inversion of Real-Valued Laplace Transforms

We have the following asymptotic result for the regularized estimated density $\widehat{f}_R(x)$.

Theorem 1. For the process X in (1), we assume that Assumption A holds and denote $\widetilde{f}(x) = \frac{f(x-)+f(x+)}{2}$. Let $n \to \infty$, $T \to \infty$, and $R \to \infty$.

(a) If Assumption B1 holds, then

$$(f_R(x) - \widetilde{f}(x))^2 = O\big(R^{-2(\frac{5-2q}{3}\wedge 1)} \times \log^2(R)\big), \quad (10)$$
$$\mathbb{E}(\widehat{f}_R(x) - f_R(x)) = O\big(\exp(\pi R/2) \times R^2 n^{-((1/\beta)\wedge 1-(\beta\wedge 1)/2)\wedge 1/4+\iota}\big),$$

and

$$\mathbb{E}(\widehat{f}_R(x) - f_R(x))^2 = O\big(\exp(\pi R)R^4\big(T^{-1} + n^{-(1/2)\wedge(1-\beta/2)+\iota}\big)\big), \forall \iota > 0. \quad (11)$$

(b) If Assumption B2 holds, then

$$\mathbb{E}\left\{\int_{\mathbb{R}_+} w(x)(\widehat{f}_R(x) - \widetilde{f}(x))^2 dx\right\} = O\left(\frac{\log^2(R)}{R^2} \bigvee \exp(\pi R)R^4\big(T^{-1} + n^{-1+\beta/2+\iota}\big)\right), \quad \forall \iota > 0, \quad (12)$$

where $w(x)$ is a bounded nonnegative-valued function, with $w(x) = o(x^2)$ for $x \to 0$.

If we further denote

$$\widetilde{f}_R(x) = \int_0^\infty \widehat{\mathcal{L}}(u)\widetilde{\Pi}(R, xu)du, \; \widetilde{\Pi}(R, xu) = \chi(R^2 x)\Pi(R, xu), \; \chi(u) = u \wedge 1, \quad (13)$$

then (under Assumption B2)

$$\mathbb{E}\left\{\int_{\mathbb{R}_+} (\widetilde{f}_R(x) - \widetilde{f}(x))^2 dx\right\} = O\left(\frac{\log^2(R)}{R^2} \bigvee \exp(\pi R)R^8\big(T^{-1} + n^{-1+\beta/2+\iota}\big)\right), \quad \forall \iota > 0. \quad (14)$$

The result of Theorem 1 implies that $\widehat{f}_R(x)$ is a consistent estimate of the volatility density $f(x)$ at the points of continuity of the latter and estimates the average of the right and left limits (which exist by Assumption B) at the points of discontinuity. The theorem goes one step further and provides bounds on the bias and the variance of the estimator. There are two sources of bias in the estimation. One, which is deterministic, arises from the regularization of the inversion, and naturally, depends only on the regularization parameter R. Its bound is given in (10). The second source of bias is stochastic and arises from the discretization error, that is, we do not directly observe the empirical volatility Laplace transform $\frac{1}{T}\int_0^T \exp(-uV_s)ds$ but we need to recover it from the high-frequency data. The magnitude of this bias is given by the first expression in (11), and naturally, depends only on the mesh of the observation grid, that is, $1/n$. Finally, the bound on the variance of the estimator is given in the second expression in (11). It depends both on the span of the data and the mesh of the observation grid. The leading component of $\mathbb{E}(\widehat{f}_R(x) - f_R(x))^2$ (provided n is increasing sufficiently fast relative to T and R is fixed) is given by

$$\frac{1}{T}\int_0^\infty\int_0^\infty\int_0^\infty \Pi(R, xu)\Pi(R, xv)\Sigma_{u,v}dudvdx,$$

where $\Sigma_{u,v}$ is the long-run variance-covariance kernel of $\frac{1}{T}\int_0^T(\exp(-uV_s) - \mathcal{L}(u))ds$, that is,

$$\Sigma_{u,v} = \int_0^\infty \{\mathbb{E}[(\exp(-uV_t) - \mathcal{L}(u))(\exp(-vV_0) - \mathcal{L}(v))] + \mathbb{E}[(\exp(-uV_0) - \mathcal{L}(u))(\exp(-vV_t) - \mathcal{L}(v))]\}dt.$$

In the most common case of Assumption B2 and provided $n \propto T^\alpha$ for some $\alpha > 0$, we can set $R = \gamma\frac{\log(T)}{\pi}$ for some positive $\gamma < \alpha(1 - \beta/2) \wedge 1$ and get the squared bias and the variance of the estimator of (almost) the same order of magnitude. Such choice of R will result in (only) logarithmic rate of convergence for our volatility density estimator. This is not surprising, given our weak assumption B2 for the density $f(x)$. The squared logarithmic rate of convergence for the log-volatility density, in a setting where X_t does not contain jumps and V_t is independent from W_t, is obtained via a deconvolution approach in Van Es, Spreij, and Van Zanten (2003), where it is connected with the optimal rate of deconvoluting a density in the presence of super-smooth noise derived in Fan (1991) (in the context of iid data). Van Es, Spreij, and Van Zanten (2003) assumed that $f(x)$ is twice continuously differentiable and obtained the optimal rate of convergence for their density estimator of squared logarithmic rate, while here, we assume only first-order derivatives for $f(x)$ and hence, we end up with (almost) logarithmic rate of convergence (which is the optimal deconvolution rate under this smoothness assumption for $f(x)$ in the presence of super-smooth noise; see Fan 1991).

Similarly, here, if we assume more smoothness conditions for the volatility density $f(x)$, we can show that the deterministic bias due to the regularization $f(x) - f_R(x)$ is of a smaller order of magnitude than the bound given in (10). This, in turn, will imply a faster rate of convergence for our volatility density estimator (provided R is chosen optimally). Thus, we have a natural link between the rate of convergence of our density estimator and the degree of smoothness of the unknown volatility density. This is similar to the results in Comte and Genon-Catalot (2006). In the setting of no drift and no jumps as well as independent W_t and V_t, Comte and Genon-Catalot (2006) showed that a penalized projection-type volatility density estimator can provide faster rates of convergence for smoother volatility densities.

Under the minimal smoothness requirement for the volatility density in Assumption B, the relative speed condition between T and n is relatively weak, that is, as pointed above, it is of the form

$n \propto T^{\alpha}$ for some $\alpha > 0$. Such a condition, in particular, is much weaker than the corresponding requirement in the problem of parametric estimation of diffusions from discrete observations; for example, see Prakasa Rao (1988). This, of course, is no surprise and is a mere reflection of the much smaller role (in relative terms) played by the discretization error in our nonparametric volatility density estimation. We also point out that, quite naturally, the bigger is the discretization error, the bigger is the bound on the activity index of jumps in X, β. This is because higher-activity jumps become harder to separate from diffusive innovations in the increments of X, which in particular implies that $|\cos(\sqrt{2un}\Delta_i^n X) - \cos(\sqrt{2un}(\Delta_i^n X - \Delta_i^n J))|$ is larger for a higher value of β (its order of magnitude is $n^{(\beta\wedge 1)/2-(1/\beta)\wedge 1+\iota}$ for $\iota > 0$ arbitrary small). We also point out that the discretization error is of smaller magnitude in (12) versus (11) (provided $\beta < 1$) because of the extra assumption on the behavior of the volatility density around zero in Assumption B2 (the volatility density approaches zero around the origin).

We note further that the result in part (a) is pointwise, that is, for fixed x, while that in part (b) is for the mean integrated squared error (MISE) (note that under B1, $f(x)$ does not need to belong to $L^2(\mathbb{R})$). For the estimator $\widehat{f}_R(x)$, we provide in (12) a bound for its MISE weighted by a function $w(x)$ that is bounded and $o(x^2)$ as $x \to 0$ but is otherwise arbitrary. The role of $w(x)$ is to downweight the estimation error in $\widehat{f}_R(x)$ around zero. In (13), we propose a slight modification of $\widehat{f}_R(x)$, which we define as $\widetilde{f}_R(x)$. We have $\widetilde{f}_R(x) = \chi(R^2x)\widehat{f}_R(x)$ and the function $\chi(R^2x)$ serves to dampen our original density estimate around zero. This dampening in turn allows us to bound in (14) the MISE of $\widetilde{f}_R(x)$ with $w(x) = 1$, that is, without any downweighting of the estimation error around zero. In our applications below, we will use $\widehat{f}_R(x)$, but we will evaluate it starting from sufficiently small value of x that is above zero, guidance for which can be easily obtained from the quantiles of any nonparametric daily integrated volatility estimates.

Finally, the analysis here can be easily extended to recovering the volatility density conditioned on some set. One interesting example, which we will consider in our empirical application, is the occurrence of big price jumps. The analysis in Bollerslev and Todorov (2011) can be used to bound the discretization error in identifying the set of big jumps (under some conditions for J_t). Also, one can consider a setting of fixed span, that is, T fixed, in which $\widehat{f}_R(x)$ will recover the density of the empirical volatility distribution over the given interval of time. Of course, for this, we will need the smoothness assumption B to hold for the empirical volatility distribution. An example where this will be the case is when V_t is a non-Gaussian OU (Ornstein–Uhlenbeck) process (see Equation (20)) in which the driving Lévy process is compound Poisson (note that Assumption B allows for discontinuities in the density). Such models for volatility have been considered for example in Barndorff-Nielsen and Shephard (2001). More generally, however, the density of the empirical distribution will not be differentiable and one can instead recover the empirical cumulative distribution of volatility.

3. NUMERICAL EXPERIMENTS

We proceed next with numerical experiments to test our estimation method developed in the previous section. We first investigate how well our estimator can recover the volatility density in the infeasible scenario when the volatility Laplace transform is measured without error. We then consider the feasible scenario where the volatility Laplace transform is recovered from high-frequency price data via the realized Laplace transform.

3.1 Inverting Known Laplace Transforms

We use two distributions in our numerical analysis here, which will be the marginal laws of two popular volatility specifications that we will use in our Monte Carlo analysis below. The first is the Gamma distribution. We denote $Y \sim G(a, b)$ for a random variable with probability density

$$f^{\mathrm{G}}(x) = \frac{b^a x^{a-1}}{\Gamma(a)} \exp(-bx)\, 1_{\{x>0\}}, a, b > 0, \quad (15)$$

with the corresponding real-valued Laplace transform given by

$$\mathcal{L}^{\mathrm{G}}(u) = \left(\frac{1}{1+u/b}\right)^a. \quad (16)$$

The second distribution we use is the inverse-Gaussian. We denote $Y \sim IG(\mu, \nu)$ for a variable with probability density given by

$$f^{\mathrm{IG}}(x) = \sqrt{\frac{\nu}{2\pi x^3}} \exp\left[-\frac{\nu}{2\mu^2 x}(x-\mu)^2\right] 1_{\{x>0\}}, \quad (17)$$

with the corresponding real-valued Laplace transform

$$\mathcal{L}^{\mathrm{IG}}(u) = \exp\left\{(\nu/\mu)\left[1 - \sqrt{1+2\mu^2 u/\nu}\right]\right\}. \quad (18)$$

It is easy to check that the two distributions satisfy Assumption B2. In Table 1, we list all the different cases considered in this section and give the corresponding parameters. We look in particular at settings with small, average, and big dispersion around the mode of the density.

The Gamma and inverse-Gaussian distributions are the marginal laws of two volatility specifications, widely used in empirical finance. The first is the square-root diffusion process given by

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dB_t, \kappa, \sigma, \theta > 0, \sigma \le \sqrt{2\kappa\theta}. \quad (19)$$

The marginal distribution of the square-root diffusion process is the Gamma distribution with parameters $a = \frac{2\kappa\theta}{\sigma^2}$ and $b = \frac{2\kappa}{\sigma^2}$ in the parameterization of (15).

Table 1. Parameter setting for the Monte Carlo analysis

Case	Marginal distribution of V_t	Parameter values
G-L	Gamma	$\kappa = 0.02, \theta = 1.0, \sigma^2 = \frac{2\kappa\theta}{4.0}$
G-M	Gamma	$\kappa = 0.02, \theta = 1.0, \sigma^2 = \frac{2\kappa\theta}{2.5}$
G-H	Gamma	$\kappa = 0.02, \theta = 1.0, \sigma^2 = \frac{2\kappa\theta}{1.5}$
IG-L	Inverse-Gaussian	$\kappa = 0.02, \mu = 1.0, \nu = 3.0$
IG-M	Inverse-Gaussian	$\kappa = 0.02, \mu = 1.0, \nu = 1.0$
IG-H	Inverse-Gaussian	$\kappa = 0.02, \mu = 1.0, \nu = 0.5$

NOTE: Cases G-L, G-M and G-H correspond to the square-root diffusion process in (19) and the parameters of the Gamma distribution are given by $a = \frac{2\kappa\theta}{\sigma^2}$ and $b = \frac{2\kappa}{\sigma^2}$. Cases IG-L, IG-M and IG-H correspond to the non-Gaussian OU process in (20) with Inverse-Gaussian marginal distribution.

Table 2. Integrated squared error of density estimate: known Laplace transform

Case	$\int_{\mathbb{R}} f^2(x)dx$	Regularization parameter $R = 2.0$	$R = 3.0$	$R = 4.0$	$R = 5.0$
G-L	0.6247	5.40×10^{-2}	1.11×10^{-2}	1.60×10^{-3}	1.92×10^{-4}
G-M	0.5301	1.83×10^{-2}	2.10×10^{-3}	2.32×10^{-4}	8.65×10^{-4}
G-H	0.4765	4.10×10^{-3}	3.90×10^{-3}	6.99×10^{-2}	1.21×10^{-0}
IG-L	0.6518	5.01×10^{-2}	8.10×10^{-3}	1.10×10^{-3}	2.89×10^{-5}
IG-M	0.5959	8.40×10^{-3}	4.68×10^{-4}	2.31×10^{-5}	1.56×10^{-6}
IG-H	0.6911	5.30×10^{-3}	2.74×10^{-4}	5.15×10^{-6}	3.28×10^{-4}

NOTE: The ISE $\int_{Q_{0.005}}^{Q_{0.995}} (f_R(x) - f(x))^2 dx$ is approximated by a Riemann sum, with the length of the discretization interval being 0.005. The range of integration is $Q_{0.005}$–$Q_{0.995}$ for Q_α denoting the α-quantile. Each of the cases is explained in Table 1.

The second volatility specification is a non-Gaussian OU process given by

$$dV_t = -\kappa V_t dt + dL_t, \kappa > 0, \tag{20}$$

where L_t is a Lévy subordinator. Following Barndorff-Nielsen and Shephard (2001), we specify the non-Gaussian OU process via its marginal distribution, which will be the inverse-Gaussian (which is self-decomposable and hence this is possible; e.g., see Sato 1999), with parameterization given in the previous subsection. It can be shown, for example, see Todorov, Tauchen, and Grynkiv (2011), that the Lévy measure of L_t is given by

$$\frac{\kappa \nu \exp(-\nu^2 x/(2\mu^2))}{\sqrt{2}\Gamma(0.5)} \left[\frac{x^{-1.5}}{2} + \frac{\nu^2 x^{-0.5}}{2\mu^2}\right].$$

Further, both volatility specifications in (19) and (20) satisfy Assumption A. In Table 1, we report the parameter values of the volatility specifications corresponding to the different cases considered for their marginal distributions.

In Table 2, we report the integrated squared error (ISE) in recovering the volatility density from the exact Laplace transform (over the quantile range $Q_{0.005}$–$Q_{0.995}$). The precision across all cases is very high. When we use the exact Laplace transform, there is obviously no estimation error and all error is due to the regularization and the numerical integration. We consider a range of values for the regularization parameter R, and we can observe from Table 2 that R plays a big role with regard to the precision of the estimation. Small values of R result in bias due to oversmoothing (recall Figure 1), while very big values in R can result in a bigger error due to the numerical integration. We also point out that the optimal value of R depends on the volatility density, which is of course unknown.

3.2 Inverting Estimated Laplace Transforms

We turn next to the feasible case where the volatility density is not known and has to be estimated from high-frequency observations of X. In the simulations, the price process is given by (1), with volatility following either (19) or (20), $\alpha_t = 0$, J_t is a compound Poisson process with intensity 1/3 (i.e., one jump every three days on average), and normally distributed jump size with mean 0 and variance of 0.3. For simplicity, in the Monte Carlo setup, we set the volatility process independent from the price process. Simulation evidence in Todorov, Tauchen, and Grynkiv (2011) indicates that the leverage effect has negligible effect on the realized Laplace transform in finite samples (recall from (5) that leverage has no asymptotic effect on $\widehat{\mathcal{L}}(u)$). In all experiments, we set $\mathbb{E}(V_t) = 1$, which implies that jumps contribute approximately 10% of price variation, consistent with prior empirical evidence.

The unit of time in our simulation design is a day, and we assume the span is $T = 3000$ days with $n = 76$, which corresponds to sampling the price process every 5 min in a 6.5-hr trading day for approximately 12 years. In Table 3, we report the precision in recovering the volatility density from a single simulation from each of the scenarios for a range of values of the regularization parameter R. Comparing Table 2 with Table 3, not surprisingly, we can observe that the ISE is orders of magnitude higher when the Laplace transform has to be estimated from the data due to the estimation error. Nevertheless, provided the appropriate R is used, the error in estimation is reasonably small. The values of R for which the precision is highest in the case of the estimated Laplace transform are lower than the case when the Laplace transform is known. This is because the estimation error prevents us from using kernels with high "focusing" ability, that is, we need to smooth more to remove the effect of the estimation error.

In the case of estimated Laplace transforms, we have a U-shaped pattern in the ISE: too large and too small values of R correspond to bigger ISE. This effect can be most clearly observed in Figure 2, where we plot the estimated densities for three different values of R for the simulation scenario G-H. Too low R results in oversmoothing and relatively big estimation bias. Increasing R improves the precision. However, when R is

Table 3. Integrated squared error of density estimate: estimated Laplace transform

Case	$\int_{\mathbb{R}} f^2(x)dx$	Regularization parameter $R = 1.0$	$R = 1.5$	$R = 2.0$	$R = 3.0$	$R = 4.0$
G-L	0.6247	0.1882	0.1003	0.0488	0.0128	0.2473
G-M	0.5301	0.1066	0.0437	0.0162	0.0066	0.1275
G-H	0.4765	0.0482	0.0161	0.0102	0.0677	0.9009
IG-L	0.6518	0.2310	0.1390	0.0900	0.0840	0.3574
IG-M	0.5959	0.1130	0.0432	0.0163	0.0099	0.0698
IG-H	0.6911	0.0838	0.0268	0.0326	0.2980	4.3244

NOTE: The computations are based on a single simulation from the models given in Table 1 and the volatility Laplace transform estimate using the realized Laplace transform $\widehat{\mathcal{L}}(u)$ defined in (4). The ISE is approximated the same way as in the calculations for Table 2.

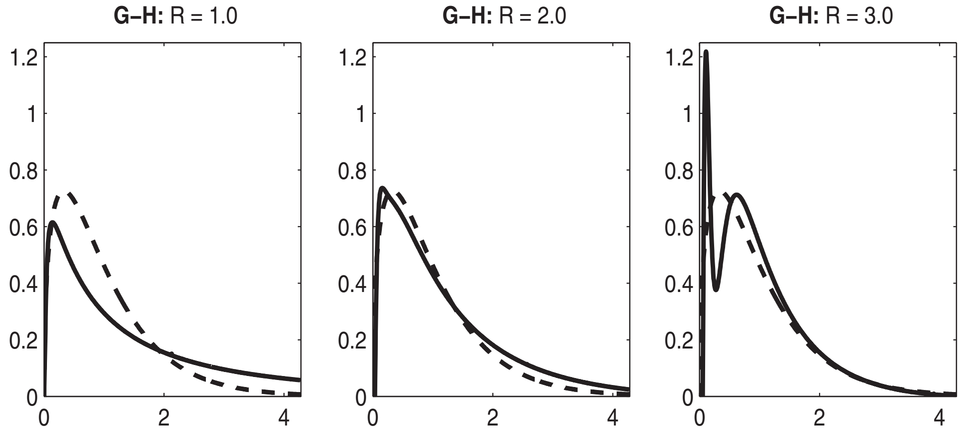

Figure 2. The effect of regularization. The figure shows the recovered density of V_t for a simulated series from model G-H, whose integrated squared error is reported in Table 3. The solid lines correspond to the estimated density, and the dashed lines to the true one.

too big, the estimation noise gets "blown up" and this leads to the oscillations in the estimated density ("inherited" from the more-focused kernel), which can be observed in the last plot in Figure 2.

3.3 Monte Carlo Analysis

We now turn to a Monte Carlo analysis using the above specified setup and 1000 replications. Based on the analysis in the previous subsections, the crucial question is how to pick R, as the optimal value of R depends on the unknown volatility density. From Figure 2, we know that when R is "too high" for the precision with which we can recover the Laplace transform of volatility from the high-frequency data, then the recovered density starts to oscillate. Therefore, a reasonable and very easy rule is to set R as the largest value, which results in a minimum number of violations of the quasiconcavity conditions (see, e.g., Koenker and Mizera 2010) of the recovered volatility density. In particular, for the case plotted in Figure 2, this will lead us to picking the middle value of $R = 2.0$.

We implement this rule in the Monte Carlo analysis. We note that this can lead to a different value of R for the different realizations of the simulated processes. The results from the Monte Carlo analysis are reported in Table 4. As can be observed from the results in the table, we have relatively good precision, with which we can recover the volatility densities across the different simulation scenarios. In general, also, the MISE is comparable with the minimal ISE from the single realizations of the process reported in Table 3. The hardest case of all simulation scenarios is the IG-H, which corresponds to inverse-Gaussian with very high volatility of volatility. The estimation error involved in this case is relatively big, necessitating small values of R to attenuate its effect on the inversion, which in turn leads to some bias.

Table 4. Monte Carlo analysis results: MISE

Case	$\int_{\mathbb{R}} f^2(x)dx$	MISE	Case	$\int_{\mathbb{R}} f^2(x)dx$	MISE
G-L	0.6247	0.0249	IG-L	0.6518	0.0532
G-M	0.5301	0.0184	IG-M	0.5959	0.0321
G-H	0.4765	0.0192	IG-H	0.6911	0.0998

NOTE: The computations are based on 1000 replica of the models given in Table 1, with $T = 3000$ and $n = 76$. The MISE is a sample average over the replications of the ISE, which in turn is approximated the same way as in the calculations for Table 2. The choice of the regularization parameter R is over the discrete grid $1 : 0.25 : 3.5$ and is the largest number of this set with least violations of quasiconcavity of the density estimate.

4. EMPIRICAL APPLICATION

We next illustrate the use of the developed nonparametric technique in a short empirical application. We analyze three large-cap stocks that are part of the S&P 100 Index: one in the technology sector (IBM), one in utilities and services (Johnson & Johnson, abbreviated by its ticker JNJ), and one in the financial sector (Bank of America, abbreviated by its ticker BAC). The sample period is from April 1997 till December 2010, and we sample every 5-min during the trading hours on each trading day (which is our unit of time) and this results in 76 high-frequency return observations per day (we omit the price at the opening and at the closing to attenuate potential special effects with start and end of trading). We exclude days in which there was no trading of the stock for more than half of the day. This resulted in a total of 3423 days for IBM, 3421 for BAC, and 3420 days for JNJ in our sample. The 5-min sampling frequency is coarse enough so that the effect of microstructure noise is negligible. Using the truncated variation defined later in (22), we estimate that jumps contribute the nontrivial 11.6%, 11.1%, and 12.7% of the total price variation for IBM, BAC, and JNJ stocks, respectively.

Before turning to the actual estimation, we "standardize" the high-frequency returns, when using them in the calculation of the realized Laplace transform, to account for the well-known presence of a diurnal deterministic within-day pattern in volatility; for example, see Andersen and Bollerslev (1997). To this end, V_t in (1) is replaced by $\widetilde{V}_t = V_t \times d(t - \lfloor t \rfloor)$, where V_t is our original stationary volatility process satisfying Assumption A and $d(s)$ is a positive differentiable deterministic function on $[0, 1]$ that captures the diurnal pattern. Then, we standardize each high-frequency increment $\Delta_i^n X$ with $1/\sqrt{\widehat{d}_i}$ for

$$\widehat{d}_i = \frac{\widehat{g}_i}{\widehat{g}}, \widehat{g}_i = \frac{n}{T} \sum_{t=1}^{T} \left|\Delta_{i_t}^n X\right|^2 1\left(\left|\Delta_{i_t}^n X\right| \leq \alpha n^{-\varpi}\right),$$

$$\widehat{g} = \frac{1}{n} \sum_{i=1}^{n} \widehat{g}_i, i = 1, \ldots, nT, \alpha > 0, \varpi \in (0, 1/2),$$

where $i_t = t - 1 + i - [i/n]n$, for $i = 1, \ldots, nT$ and $t = 1, \ldots, T$. We set $\varpi = 0.49$ and $\alpha = 3 \times \sqrt{BV_t}$, with BV_t denoting the bipower variation of Barndorff-Nielsen and Shephard (2004, 2006), defined as

$$BV_t = \frac{\pi}{2} \sum_{i=n(t-1)+1}^{nt} \left|\Delta_{i-1}^n X\right| \left|\Delta_i^n X\right|. \tag{21}$$

Intuitively, $\widehat{d_i}$ estimates the deterministic component of the stochastic variance and then we "standardize" the high-frequency increments with it. Todorov and Tauchen (2012) derived the asymptotic effect of this "cleaning" procedure, but since $\widehat{d_i}$ estimates quite precisely the deterministic pattern, naturally, this effect is rather small.

We plot the estimated densities of the spot volatility obtained from the estimation method based on the regularized inverse of the realized Laplace transform in (9) for each of the three stocks in the three top panels in Figure 3. For ease of interpretation, we present the density estimates for $\sqrt{V_t}$ (and not V_t) in percentage terms, as this is the standard way of quoting volatility on the market. We can contrast these spot volatility density estimates with estimates of the density of a jump-robust measure of the daily integrated volatility obtained by inverting its direct empirical Laplace transform. We measure the integrated volatility using the truncated variation, originally proposed in Mancini (2009), and applied here as

$$TV_t = \sum_{i=n(t-1)+1}^{nt} \left|\Delta_i^n X\right|^2 1\left(\left|\Delta_i^n X\right| \le \alpha n^{-\varpi}\right), \tag{22}$$

for the same choice of $\varpi = 0.49$ and $\alpha = 3 \times \sqrt{BV_t}$ as for the estimation of the diurnal component of volatility $\widehat{d_i}$. Under Assumption A, TV_t is a consistent estimate for the unobservable integrated volatility (and this is the reason for using it as a benchmark for the volatility density), that is, we have for each fixed $t \ge 1$ (see, e.g., Jacod 2008),

$$TV_t \xrightarrow{\mathbb{P}} \int_{t-1}^{t} \sigma_s^2\, ds, \text{ as } n \to \infty. \tag{23}$$

The dashed lines in the top three panels in Figure 3 show the implied density for the daily integrated volatility by using our method to invert the empirical Laplace transform of the $\sqrt{TV_t}$ series, while the three lower panels show standard kernel-density estimates of $\sqrt{TV_t}$ obtained from the same observations.

There are several conclusions to be made from Figure 3. First, there is significant volatility of volatility: the spot volatility can take values as high as five to six times its modal value. Thus, volatility dynamics (and in particular short-lived sharp changes in it) plays an important role in generating tail events in individual stock returns, in addition to the price jumps. We recall that the realized Laplace transform is robust to the presence of price jumps. Therefore, our nonparametric separation of volatility from jumps identifies a rather nontrivial role of stochastic volatility in generating extreme events in asset prices. Comparing the spot with integrated volatility, we can observe a common pattern in the three stocks. The mode of the spot volatility density is slightly to the left of that for the daily integrated volatility, and spot volatility has somewhat fatter distribution than that of

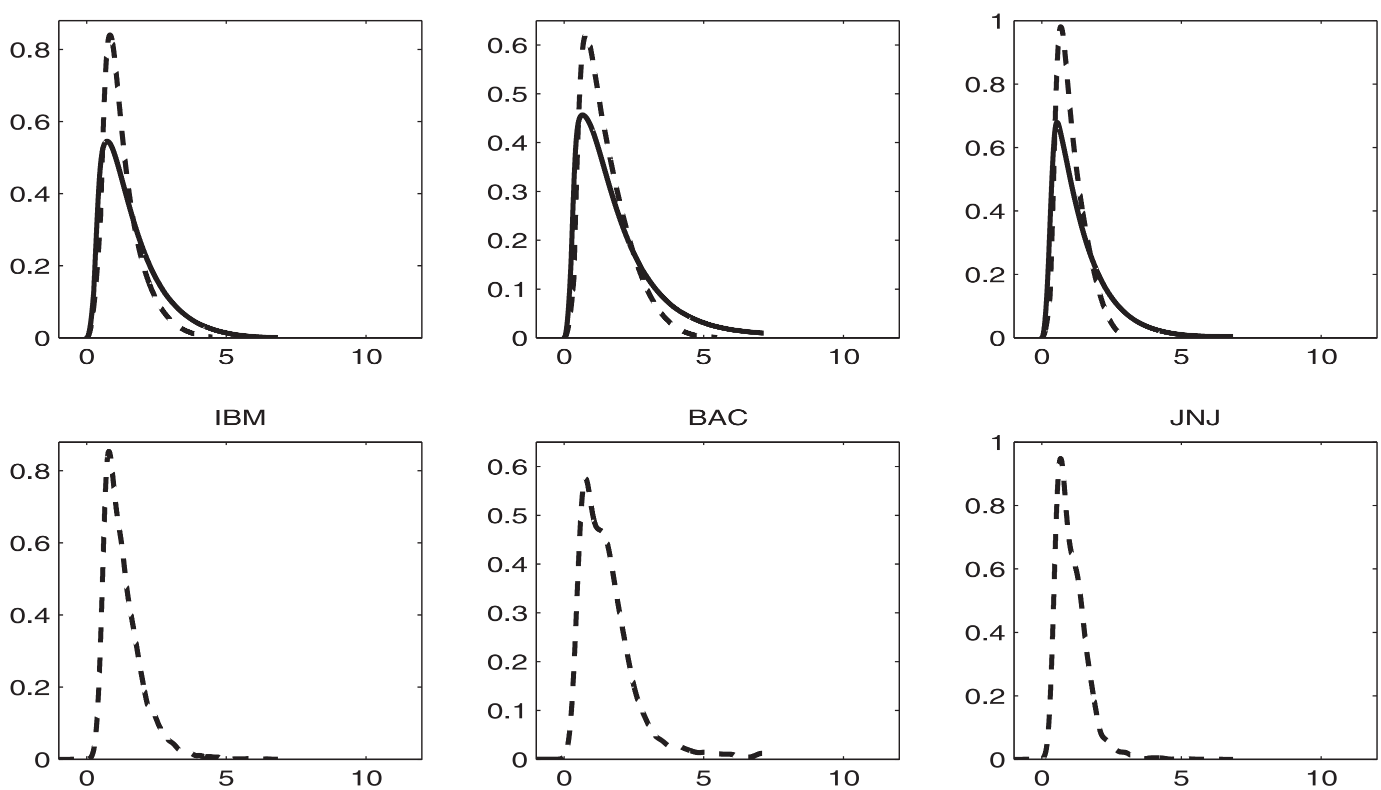

Figure 3. Nonparametric spot and integrated volatility density estimates. The solid lines correspond to our nonparametric density estimate of $\sqrt{V_t}$, while the dashed lines are nonparametric density estimates of daily $\sqrt{TV_t}$. The dashed lines in the top plots are based on inverting the empirical Laplace transform of TV_t using our estimator in (9), while the ones in the bottom plots are standard kernel estimates with Silverman's automatic bandwidth of $h = 0.79 \times \text{IQR} \times T^{-1/5}$, where IQR denotes the interquartile range.

the integrated volatility. The reason for this is the presence of short-term volatility moves in the form of volatility jumps and the mean reversion. The daily integrated volatility "averages out" the sharp moves in volatility. Overall, our nonparametric evidence here points to stochastic volatility, with significant volatility of volatility possibly generated by volatility jumps.

We investigate further the hypothesis of volatility jumps and their interaction with the price jumps by computing conditional density estimates. In particular, we will use the methodology developed here to gather nonparametric evidence regarding the effect of price jumps on stochastic volatility. In standard volatility models, volatility is a diffusion process and thus by construction stochastic volatility does not jump when the price jumps (volatility and price jumps though can still be dependent in such setting as volatility can drive the jump intensity). More recent parametric work has allowed for volatility jumps, as in the non-Gaussian OU model of Barndorff-Nielsen and Shephard (2001), although in some specifications, volatility and price jumps are constrained to be uncorrelated.

To investigate the effect of price jumps on volatility, we do the following. We identify the days in the sample where relatively large jumps occurred (we will be precise about what constitutes a large jump below) and then construct the realized Laplace transform of volatility on a given (fixed) number of days after the day with the large jump. We then use our nonparametric procedure to invert the estimated Laplace transform and recover the density of volatility a fixed number of days after the occurrence of price jumps. More formally, for some "big" fixed $\tau > 0$, we define for any integer $k \geq 1$ the set of days with a "big" positive, respectively negative, jump as

$$I_n^{\pm\tau}(k) = \Big\{t = k+1,\dots,T : \{i = (t-k-1)n+1,\dots,(t-k)n\} \bigcap \big\{i = 1,\dots,nT : \Delta_i^n X \gtrless \pm(\alpha n^{-\varpi} \vee \tau)\big\} \neq \emptyset\Big\},$$

where α and ϖ are the same as the ones used in the construction of TV_t. $I_n^{\pm\tau}(k)$ is the set of days in the sample where k days ago, a big positive or negative jump has occurred. We can then construct the realized Laplace transform on the sets $I_n^{\pm\tau}(k)$, that is,

$$\widehat{\mathcal{L}}^{\pm\tau}(k) = \frac{1}{|I_n^{\pm\tau}(k)|} \sum_{t\in I_n^{\pm\tau}(k)} \left[\sum_{i=(t-1)n+1}^{tn} \cos\left(\sqrt{2un}\,\Delta_i^n X\right)\right], \tag{24}$$

where $|I_n^{\pm\tau}(k)|$ denotes the size of the set $I_n^{\pm\tau}(k)$. Finally, we can invert $\widehat{\mathcal{L}}^{\pm\tau}(k)$ using (9). Our goal is to produce a nonparametric estimate of the densities of $\sqrt{V_t}1_{\{t\in I^{-\tau}(k)\}}$ and $\sqrt{V_t}1_{\{t\in I^{+\tau}(k)\}}$ where the set $I^{\pm\tau}(k)$ is defined via

$$I^{\pm\tau}(k) = \Big\{t : [\lfloor t\rfloor - k - 1; \lfloor t\rfloor - k] \bigcap \big\{s > 0 : \Delta X_s \gtrless \pm\tau\big\} \neq \emptyset\Big\}. \tag{25}$$

In our actual application, we set τ to $\tau = \sqrt{\mathbb{E}(TV_t)} \times 5/\sqrt{n}$, which is five-standard deviation move for the continuous part of the high-frequency return (the mean of the truncated variation is estimated by the corresponding sample average). This results in approximately 100 jumps of each sign in our sample to estimate the realized Laplace transforms in (24). We further set

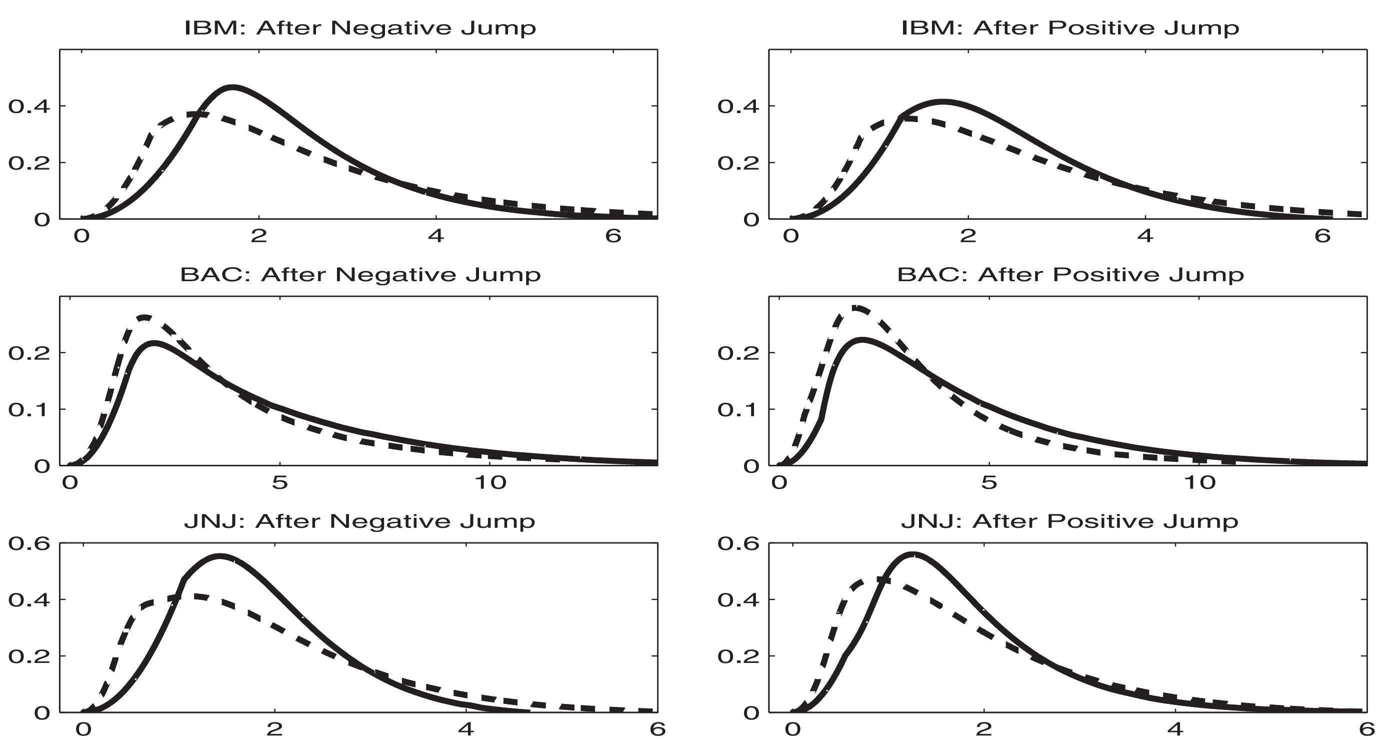

Figure 4. Nonparametric volatility density estimates after a price jump. The solid (dashed) line is a nonparametric density estimate of $\sqrt{V_t}$ over the days in the sample that follow a day (respectively 22 days) after a positive (left side) or negative "big" jump (right side) in the price. The threshold for the "big" jumps is set to five standard deviations of the continuous part of the high-frequency return based on the sample mean of TV_t of each individual series.

k to 1 and 22, which amounts to looking at volatility 1 day and 1 calendar month after a "big" jump.

The result of the calculations are presented in Figure 4. Comparing the estimated volatility densities in Figure 4 with the unconditional ones in Figure 3, we can observe a very pronounced shift of the mode toward the right, that is, volatility unambiguously increases after a "big" jump. Interestingly, the density of the volatility 1 month after the jump starts moving toward the unconditional one (compared with the density the day following the jump). The interpretation is that a big price jump "feeds" into higher future volatility, with the effect diminishing over time due to the mean reversion in volatility. Of course, one should be careful in interpreting the above evidence, as high volatility might be generating the big price jumps, which in turn are followed by higher volatility. This will be the case where price jump intensity depends on volatility. Nevertheless, our analysis clearly shows that price jumps are very closely related with the stochastic volatility dynamics, and in particular, in terms of parametric volatility modeling, we need models that allow for this connection, as, for example, in the non-Gaussian OU model of Barndorff-Nielsen and Shephard (2001). Another interesting common pattern across the three stocks is that there is a significant spread in the estimated volatility densities in Figure 4. This suggests that the size of the price jumps plays a big role in determining the size of the impact it has on the future stochastic volatility. Thus, the connection between volatility and price jumps is size-dependent. Further, comparing the left and the right side of Figure 4, we can observe that the volatility densities after a positive and a negative jump are rather similar for these three stocks. Finally, it is interesting to point out that among the three stocks, the one whose volatility reacts strongest to the occurrence of price jumps is BAC. This is consistent with the view that stocks in the financial sector are most sensitive to financial distress in the form of extreme market events.

5. CONCLUSION

In this article, we propose a nonparametric method for estimation of the spot volatility density from high-frequency data with increasing time span in a jump-diffusion model. The method consists of aggregating the high-frequency returns data into a function known as realized Laplace transform, which provides a consistent estimate of the unobservable real-valued volatility Laplace transform. On a second stage, the estimated volatility Laplace transform is inverted using a regularized kernel method to obtain an estimate of the density of spot volatility. We derive bounds on the MISE in the density estimation and provide guidance on the feasible choice of the regularization parameter. An empirical application for three large-cap stocks indicates the importance of short-term high-frequency movements in volatility that get smoothed over in forming estimates of daily realized variation, and it shows how to trace out the dynamic response of the spot volatility density to large price jumps of either sign as it relaxes back to the steady-state unconditional density.

APPENDIX: PROOF OF THEOREM 1

The proof of Theorem 1 consists of two parts. The first part is the analysis of the deterministic bias $f_R(x) - \widetilde{f}(x)$ (respectively $\int_{\mathbb{R}}(f_R(x) - \widetilde{f}(x))^2dx$), and the second part deals with the estimation error $\widehat{f}_R(x) - f_R(x)$ (respectively $\int_{\mathbb{R}} w(x)(\widehat{f}_R(x) - f_R(x))^2dx$ and $\int_{\mathbb{R}}(\widetilde{f}_R(x) - f_R(x))^2dx$). In what follows, we will denote with K a positive constant that does not depend on R, x, T, and n.

A.1 The Deterministic Bias $f_R(x) - \widetilde{f}(x)$

Using the representation of $f_R(x)$ in Kryzhniy (2003a) as an integral with respect to the true density $f(x)$, we have

$$\begin{aligned} f_R(x) &= \frac{2}{\pi}\int_0^\infty f(xu)\sqrt{u}\frac{\sin(R\ln(u))}{u^2-1}du \\ &= \frac{2}{\pi}\int_{-\infty}^\infty f(x\exp(z))\exp(3z/2)\frac{\sin(Rz)}{\exp(2z)-1}dz. \end{aligned} \tag{A.1}$$

Using the above, we can make the decomposition (recall the definition of $\widetilde{f}(x)$ in the theorem) for some constant $\delta > 1/R$

$$f_R(x) - \widetilde{f}(x) = \sum_{i=1}^4 A_i(x), \tag{A.2}$$

$$\begin{aligned} A_1(x) &= \frac{2}{\pi}\int_{-\delta}^0 [f(x\exp(z)) - f(x-)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz \\ &\quad + \frac{2}{\pi}\int_0^\delta [f(x\exp(z)) - f(x+)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz, \\ A_2(x) &= \frac{2}{\pi}\int_{-\infty}^{-\delta} [f(x\exp(z)) - f(x-)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz, \\ A_3(x) &= \frac{2}{\pi}\int_\delta^\infty [f(x\exp(z)) - f(x+)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz, \end{aligned}$$

and

$$\begin{aligned} A_4(x) &= f(x-)\left(\frac{1}{\pi}\int_{-\infty}^0 \frac{\exp(z/2)\sin(Rz)}{\sinh(z)}dz - \frac{1}{2}\right) \\ &\quad + f(x+)\left(\frac{1}{\pi}\int_0^\infty \frac{\exp(z/2)\sin(Rz)}{\sinh(z)}dz - \frac{1}{2}\right). \end{aligned}$$

Starting with $A_1(x)$, we can split the range of integration $(-\delta, 0)$ to $(-\delta, -1/R)$ and $(-1/R, 0)$ (and similarly for the positive side), and we use different arguments to bound each of the integrals. First, since $f(x)$ is piecewise differentiable, we can apply the Taylor expansion and trivially get

$$\begin{aligned} &\left|\int_{-1/R}^0 [f(x\exp(z)) - f(x-)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz\right| \\ &\quad + \left|\int_0^{1/R} [f(x\exp(z)) - f(x+)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz\right| \\ &\leq \sup_{u\in[\exp(-1/R)x,\exp(1/R)x]} |f'(u)u|\frac{K}{R}. \end{aligned} \tag{A.3}$$

Next, using integration by parts, we have

$$\begin{aligned} &\int_{-\delta}^{-1/R} [f(x\exp(z)) - f(x-)]\frac{\exp(3z/2)\sin(Rz)}{\exp(2z)-1}dz \\ &\quad = \frac{1}{R}\int_{-\delta}^{-1/R} g'(z)\cos(Rz)dz \\ &\quad\quad - \frac{1}{R}(g(-1/R)\cos(-1) - g(-\delta)\cos(-R\delta)), \end{aligned} \tag{A.4}$$

where $g(z) = \frac{(f(x\exp(z)) - f(x-))\exp(3z/2)}{\exp(2z)-1}$ and $g'(z)$ is its derivative. Using the fact that $f(x)$ has a piecewise continuous derivative, we get

$$|g(-1/R)| + |g(-\delta)| \leq K \sup_{u\in[\exp(-\delta)x,x]} |f'(u)u|. \tag{A.5}$$

We further have

$$\begin{aligned} g'(z) &= \frac{f'(x\exp(z))x\exp(5z/2)}{\exp(2z)-1} \\ &\quad + \frac{1.5\exp(3z/2)[f(x\exp(z)) - f(x-)]}{\exp(2z)-1} \\ &\quad - \frac{2e^{7z/2}[f(x\exp(z)) - f(x-)]}{(\exp(2z)-1)^2}, \end{aligned}$$

from which it follows that

$$|g'(z)| \le K \sup_{u\in[\exp(-\delta)x,x]} |f'(u)u| \frac{1}{1-\exp(z)}, z \in (-\delta, -1/R). \quad \text{(A.6)}$$

Therefore, $\int_{-\delta}^{-1/R} g'(z)\cos(Rz)dz = O(\log(R))$ for a given $x > 0$. Similar analysis, of course, can be made on the interval $(1/R, \delta)$, and thus altogether, we have

$$A_1(x) = O\left(\frac{\log(R)}{R}\right), \forall x > 0. \quad \text{(A.7)}$$

We continue next with $A_2(x)$. We can decompose it as

$$A_2(x) = \frac{2}{\pi}\left[A_2^{(1)}(x) + A_2^{(2)}(x) - A_2^{(3)}(x)\right],$$
$$A_2^{(1)}(x) = \int_{-\infty}^{-\delta-\lfloor\frac{R\log(R)}{3\pi}\rfloor\frac{2\pi}{R}} f(x\exp(z))\exp(3z/2)\frac{\sin(Rz)}{\exp(2z)-1}dz,$$
$$A_2^{(2)}(x) = \int_{-\delta-\lfloor\frac{R\log(R)}{3\pi}\rfloor\frac{2\pi}{R}}^{-\delta} f(x\exp(z))\exp(3z/2)\frac{\sin(Rz)}{\exp(2z)-1}dz,$$
$$A_2^{(3)}(x) = f(x-)\int_{-\infty}^{-\delta} \exp(3z/2)\frac{\sin(Rz)}{\exp(2z)-1}dz.$$

First, using the behavior of $f(x)$ around zero from Assumption B1 (and the identity $f(y) = f(x) + \int_x^y f'(u)du$ for each interval $[x, y]$ on which $f(u)$ and $f'(u)$ are continuous), we trivially have

$$A_2^{(1)}(x) = O\left(R^{-(5/3-2q/3)\wedge 1}\right), \forall x > 0, \quad \text{(A.8)}$$

with q being the constant in Assumption B1. We further have

$$A_2^{(2)}(x) = \sum_{k=1}^{\lfloor\frac{R\log(R)}{3\pi}\rfloor} h(-\delta - 2\pi k/R)\int_{-\delta-2\pi k/R}^{-\delta-2\pi(k-1)/R} \exp(3z/2)\sin(Rz)dz$$
$$+ \sum_{k=1}^{\lfloor\frac{R\log(R)}{3\pi}\rfloor} \int_{-\delta-2\pi k/R}^{-\delta-2\pi(k-1)/R} (h(z) - h(-\delta - 2\pi k/R))$$
$$\times \exp(3z/2)\sin(Rz)dz, \quad \text{(A.9)}$$

where we denote $h(z) = \frac{f(x\exp(z))}{\exp(2z)-1}$. Using Assumption B1, we have $|h(z)|/\exp(-z((q-1)\vee 0)) \le K(x)$ for z sufficiently small (and fixed x). Therefore, using direct integration, we have for every $x > 0$,

$$\sum_{k=1}^{\lfloor\frac{R\log(R)}{3\pi}\rfloor} h(-\delta - 2\pi k/R)\int_{-\delta-2\pi k/R}^{-\delta-2\pi(k-1)/R} \exp(3z/2)\sin(Rz)dz$$
$$= O(R^{-(5/3-2q/3)\wedge 1} \times \log(R)).$$

Similarly, using $|h'(z)|/\exp(-z((q-1)\vee 0)) \le K(x)$ for z sufficiently small (and fixed x), we have for every $x > 0$,

$$\sum_{k=1}^{\lfloor\frac{R\log(R)}{3\pi}\rfloor} \int_{-\delta-2\pi k/R}^{-\delta-2\pi(k-1)/R} (h(z) - h(-\delta - 2\pi k/R))\exp(3z/2)\sin(Rz)dz$$
$$= O(R^{-(5/3-2q/3)\wedge 1} \times \log(R)).$$

Thus, altogether, we have

$$A_2^{(2)}(x) = O\left(R^{-(5/3-2q/3)\wedge 1} \times \log(R)\right), \forall x > 0. \quad \text{(A.10)}$$

Finally, using integration by parts, it is easy to show that

$$A_2^{(3)}(x) = O(1/R), \forall x > 0. \quad \text{(A.11)}$$

Similar analysis can be made, of course, for $A_3(x)$, and thus altogether, we have

$$|A_2(x)| + |A_3(x)| = O\left(R^{-(5/3-2q/3)\wedge 1} \times \log(R)\right), \forall x > 0. \quad \text{(A.12)}$$

We are left with $A_4(x)$. We can write

$$A_4(x) = \frac{f(x-)}{\pi}\int_{-\infty}^{0} \frac{(\exp(z/2)-1)\sin(Rz)}{\sinh(z)}dz$$
$$+ \frac{f(x+)}{\pi}\int_{0}^{\infty} \frac{(\exp(z/2)-1)\sin(Rz)}{\sinh(z)}dz$$
$$+ \widetilde{f}(x)\left(\frac{1}{\pi}\int_0^\infty \frac{\sin(Rz)}{\sinh(z)}dz - \frac{1}{2}\right). \quad \text{(A.13)}$$

Using integration by parts for the first two integrals on the right side of the above equality and the identity $\int_0^\infty \frac{\sin(Rz)}{\sinh(z)}dz = \frac{\pi}{2}\tanh\left(\frac{\pi R}{2}\right)$ for the third one, we trivially get

$$A_4(x) = O(1/R), \forall x > 0. \quad \text{(A.14)}$$

Combining the results in (A.7), (A.12), and (A.14), we have the second claim in (10).

For the claim in (12), we will show here that

$$\int_0^\infty \left(f_R(x) - \widetilde{f}(x)\right)^2 dx = O\left(\frac{\log^2(R)}{R^2}\right). \quad \text{(A.15)}$$

We make use of the decomposition of $f_R(x) - \widetilde{f}(x) = \sum_{i=1}^4 A_i(x)$. Then, using the bounds in (A.3), (A.5), (A.6), and (A.13) and Assumption B2, we have

$$\int_0^\infty (A_1^2(x) + A_4^2(x))dx = \left(\frac{\log^2(R)}{R^2}\right). \quad \text{(A.16)}$$

For the two terms $A_2(x)$ and $A_3(x)$, we use a different decomposition than the one we used above for showing (10). In particular, we now decompose $A_2(x)$ using integration by parts as

$$A_2(x) = \frac{1}{R}\int_{-\infty}^{-\delta} g'(z)\cos(Rz)dz - \frac{1}{R}(g(-\delta)\cos(-R\delta)$$
$$- g(-\infty)\cos(-\infty)),$$

where recall that we denote $g(z) = \frac{[f(x\exp(z))-f(x-)]\exp(3z/2)}{\exp(2z)-1}$ and $g'(z)$ as its derivative, and we do exactly the same decomposition for $A_3(x)$. Then, using the boundedness of $f(x)$ as well as $f'(x)$, with $f(x) = o(x^{-1-\iota})$ and $f'(x) = o(x^{-2-\iota})$ for $x \to \infty$, together with Fubini's theorem (integrating first over x), we can easily get

$$\int_0^\infty \left(A_2^2(x) + A_3^2(x)\right)dx = O\left(\frac{\log^2(R)}{R^2}\right). \quad \text{(A.17)}$$

Finally, for the deterministic bias component of the bound in (14), we can directly use (A.15), combined with

$$\int_0^\infty \left(\chi(R^2x) - 1\right)^2 dx \le K/R^2,$$

to get

$$\int_0^\infty \left(\widetilde{f}_R(x) - f(x)\right)^2 dx = O\left(\frac{\log^2(R)}{R^2}\right). \quad \text{(A.18)}$$

A.2 The Estimation Error $\widehat{f}_R(x) - f_R(x)$

We first decompose the error in estimating the volatility Laplace transform as follows:

$$\widehat{\mathcal{L}}(u) - \mathcal{L}(u) = B_1(u) + B_2(u) + B_3(u) + B_4(u) + B_5(u),$$
$$B_1(u) = \frac{1}{T}\int_0^T \left[\exp(-uV_s) - \mathcal{L}(u)\right]ds,$$
$$B_2(u) = \frac{1}{T}\int_0^T \left[\exp(-uV_{\lfloor sn\rfloor/n}) - e^{-uV_s}\right]ds,$$

$$B_3(u) = \frac{1}{nT}\sum_{i=1}^{nT}\left[\cos\left(\sqrt{2un}\Delta_i^n X\right) - \cos\left(\sqrt{2un}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\sqrt{V_s}dW_s\right)\right],$$

$$B_4(u) = \frac{1}{nT}\sum_{i=1}^{nT}\left[\cos\left(\sqrt{2un}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\sqrt{V_s}dW_s\right) - \cos\left(\sqrt{2un}\sqrt{V_{\frac{i-1}{n}}}\Delta_i^n W\right)\right],$$

and

$$B_5(u) = \frac{1}{nT}\sum_{i=1}^{nT}\left[\cos\left(\sqrt{2un}\sqrt{V_{\frac{i-1}{n}}}\Delta_i^n W\right) - e^{uV_{\lfloor sn\rfloor/n}}\right].$$

Starting with the first term, we can apply lemma VIII.3.102 in Jacod and Shiryaev (2003), which provides bounds on conditional expectations of centered mixing processes, and Assumption A4 to get

$$\mathbb{E}(B_1(u)) = 0 \quad \text{and} \quad \mathbb{E}|B_1(u)|^2 \le \frac{K}{T}\int_0^\infty \alpha_s^{\text{mix}}ds. \tag{A.19}$$

We note that $\int_0^\infty \alpha_s^{\text{mix}}ds$ is finite due to our assumption on the rate of decay of the mixing coefficients in Assumption A4.

We continue next with $B_2(u)$. First, using the Taylor expansion, we can get the inequality $|\exp(-x^p) - \exp(-y^p)| \le K|x-y|$ for $x, y \in \mathbb{R}_+$ and some constant K and $p \ge 1$. Then, using the classical inequality

$$||x|^r - |y|^r| \le |x-y|^r, \text{ for some} r \in (0,1],$$

we can get $|\exp(-x) - \exp(-y)| \le K|x-y|^r$ for $x, y \in \mathbb{R}_+$ and any $r \le 1$. Applying the last inequality with $x = V_{\lfloor sn\rfloor/n}$ and $y = V_s$, and then using Assumption A3, together with lemma VIII.3.102 in Jacod and Shiryaev (2003) and Assumption A4, we get

$$\mathbb{E}|B_2(u)| \le \frac{K}{T}u^{1/2-2\iota}\int_0^T \mathbb{E}|V_{\lfloor sn\rfloor/n} - V_s|^{1/2-2\iota}ds \le Ku^{1/2-2\iota}n^{-1/4+\iota},$$

$$\begin{aligned}\mathbb{E}|B_2(u)|^2 &= \frac{1}{T^2}\mathbb{E}\left\{\int_0^T\int_0^T\left[\exp(-uV_{\lfloor tn\rfloor/n}) - \exp(-uV_t)\right]\times\left[\exp(-uV_{\lfloor sn\rfloor/n}) - \exp(-uV_s)\right]dsdt\right\}\\ &\le \frac{Ku^\iota n^{-\iota/2}}{T^2}\int_0^T\int_0^T 1_{\{|t-s|\ge 1/n\}}\left(\alpha^{\text{mix}}_{|t-s|-1/n}\right)^{1-\iota}dtds\\ &\le \frac{Ku^\iota n^{-\iota/2}}{T},\end{aligned} \tag{A.20}$$

for sufficiently small $\iota > 0$. Turning next to $B_3(u)$, we first can derive the following bounds, using the trigonometric identity for $\cos(a) - \cos(b)$ and the inequality $|\sin(x)| \le |x|$, as well as Assumptions A1 and A3 and the Hölder inequality:

$$\mathbb{E}\left|\cos\left(\sqrt{2un}(\Delta_i^n X - \Delta_i^n J)\right) - \cos\left(\sqrt{2un}\int_{\frac{i-1}{n}}^{\frac{i}{n}}\sqrt{V_s}dW_s\right)\right|^p \le Ku^{\frac{p(1-\iota)}{2}}n^{-\frac{p(1-\iota)}{2}}, \forall\iota\in(0,1),\ p=1,2,$$

$$\mathbb{E}\left|\cos\left(\sqrt{2un}\Delta_i^n X\right) - \cos\left(\sqrt{2un}(\Delta_i^n X - \Delta_i^n J)\right)\right| \le \begin{cases} Ku^{\frac{\beta+\iota}{2}}n^{-1+\frac{\beta+\iota}{2}}, \forall\iota\in(0,1-\beta), \text{ if } \beta<1,\\ Ku^{(1-\iota)/2}n^{1/2-1/\beta+\iota/2}, \forall\iota\in(0,\frac{2-\beta}{\beta}), \text{ if } \beta\ge 1,\end{cases}$$

$$\mathbb{E}\left|\cos\left(\sqrt{2un}\Delta_i^n X\right) - \cos\left(\sqrt{2un}(\Delta_i^n X - \Delta_i^n J)\right)\right|^2 \le Ku^{(\beta+\iota)/2}n^{\beta/2+\iota/2-1}, \forall\iota\in(0,2-\beta).$$

Using the above bounds, the inequality $(x+y)^2 \le 2x^2 + 2y^2$, and the Cauchy–Schwartz inequality, we get altogether

$$\mathbb{E}|B_3(u)| \le K(u^{1/2-\iota}\vee 1)n^{[(\beta\wedge 1)/2-1\wedge(1/\beta)]\vee(-1/2)+\iota}$$

and

$$\mathbb{E}|B_3(u)|^2 \le K(u^{1-\iota}\vee 1)n^{\beta/2-1+\iota}, \tag{A.21}$$

where $\iota > 0$ is arbitrary small. Turning to $B_4(u)$, we can distinguish two cases: when Assumption B1 holds, and when the stronger condition B2 holds. In the case of B1 only, we can use the Itô isometry and the Hölder inequality to get

$$\mathbb{E}|B_4(u)| \le Ku^{1/2-\iota}n^{-1/4+\iota}$$

and

$$\mathbb{E}|B_4(u)|^2 \le Ku^{1-\iota}n^{-1/2+\iota}, \forall\iota>0, \text{ under B1}. \tag{A.22}$$

When the stronger condition B2 holds, we can make use of the following identity:

$$\sqrt{x} - \sqrt{y} = \frac{1}{2\sqrt{x}}(x-y) + \frac{\left(\sqrt{x}-\sqrt{y}\right)^2}{2\sqrt{x}} \Rightarrow \left|\sqrt{x}-\sqrt{y}\right| \le \frac{|x-y|}{\sqrt{x}}, \forall x>0,\ y\ge 0.$$

Applying the above, together with the Itô isometry; the Hölder inequality; the fact that $f(x) \le Kx$ for x around zero (because of the boundedness of $f'(x)$ and $f(0+) = 0$), which implies that $\mathbb{E}|V_t|^{\iota-2} \le K$ for any $\iota \in (0,2)$; and Assumption A3, we can get the stronger bound

$$\mathbb{E}|B_4(u)| \le Ku^{1/2-\iota}n^{-1/2+\iota}$$

and

$$\mathbb{E}|B_4(u)|^2 \le Ku^{1-\iota}n^{-1+\iota}, \forall\iota>0, \text{ under B2}. \tag{A.23}$$

Finally, for $B_5(u)$, we can use successive conditioning and get

$$\mathbb{E}(B_5(u)) = 0 \quad \text{and} \quad \mathbb{E}(B_5(u))^2 \le \frac{K}{nT}. \tag{A.24}$$

Now, we are ready to show the result in (11) and its equivalent for the ISE, and we note for this that the constant K in (A.20)–(A.24) does not depend on u. In what follows, we will make use of

$$\mathbb{E}\left(\widehat{\mathcal{L}}(u) - \mathcal{L}(u)\right)^2 \le K\sum_{i=1}^5 \mathbb{E}(B_i(u))^2,$$

and the bounds in (A.19)–(A.24). Recalling the definition of $\widehat{f}_R(x)$ in (9), we have

$$\widehat{f}_R(x) - f_R(x) = \int_{\mathbb{R}_+}\Pi(R,xu)\left(\widehat{\mathcal{L}}(u) - \mathcal{L}(u)\right)du.$$

Using integration by parts twice, we have for $x > 0$,

$$\int_0^\infty \frac{\sqrt{u}\sin(R\ln(u))}{u^2+1}\sin(xu)du = \frac{1}{x^2}\int_0^\infty \sin(xu)\Xi(u,R)du,$$

$$\begin{aligned}\Xi(u,R) = &-\frac{0.25u^{-3/2}\sin(R\ln(u))}{u^2+1} - \frac{u^{-3/2}\sin(R\ln(u))}{(u^2+1)^{3/2}}\\ &-\frac{3u^{1/2}\sin(R\ln(u))}{(u^2+1)^2} + \frac{4u^{5/2}\sin(R\ln(u))}{(u^2+1)^3}\\ &-\frac{4Ru^{1/2}\cos(R\ln(u))}{(u^2+1)^2} - \frac{R^2u^{-3/2}\sin(R\ln(u))}{u^2+1}.\end{aligned}$$

From here, using the identity $|\sin(x)/x| \le 1$, we have $|\int_0^\infty \frac{\sqrt{u}\sin(R\ln(u))}{u^2+1}\sin(xu)du| \le KR^2(x^{1/2-\iota}\wedge x^{-3/2+\iota})$ (for the first part of this bound, we do not need the above decomposition but only the inequality $|\sin(x)/x| \le 1$) for any $\iota > 0$. Similar results hold for $\int_0^\infty \frac{\sqrt{u}\cos(R\ln(u))}{u^2+1}\sin(xu)du$ as well, and we thus have $|\Pi(R,z)| \le \exp(\pi R)R^2\widehat{\Pi}(z)$ for some $\widehat{\Pi}(z)$, with $\widehat{\Pi}(z) = o(z^{1/2-\iota})$

for $z \to 0$ and $\widehat{\Pi}(z) = o(z^{-3/2+\iota})$ for $z \to \infty$ and arbitrary small $\iota > 0$. From here, an application of Fubini's theorem, together with the results in (A.19), (A.20), (A.21), (A.22), (A.23), and (A.24), yields

$$\mathbb{E}\left\{\int_0^\infty\int_0^\infty\int_0^\infty w(x)\Pi(R,xu)\Pi(R,xv)|\widehat{\mathcal{L}}(u) - \mathcal{L}(u)||\widehat{\mathcal{L}}(v)- \mathcal{L}(v)|dudvdx\right\}$$
$$\leq K\exp(\pi R)R^4\left(T^{-1}+n^{-1+\beta/2+\iota}\right)\int_0^\infty\frac{w(x)(1+x)}{x^{3-\iota}}dx$$
$$\leq K\exp(\pi R)R^4\left(T^{-1}+n^{-1+\beta/2+\iota}\right),$$
$$\mathbb{E}\left\{\int_0^\infty\int_0^\infty\int_0^\infty \chi^2(R^2x)\Pi(R,xu)\Pi(R,xv)|\widehat{\mathcal{L}}(u) - \mathcal{L}(u)||\widehat{\mathcal{L}}(v)- \mathcal{L}(v)|dudvdx\right\}$$
$$\leq K\exp(\pi R)R^4\left(T^{-1}+n^{-1+\beta/2+\iota}\right)\int_0^\infty\frac{\chi^2(R^2x)(1+x)}{x^{3-\iota}}dx$$
$$\leq K\exp(\pi R)R^8\left(T^{-1}+n^{-1+\beta/2+\iota}\right).$$

From here, the rate in (11) and its analogues for the integrated squared estimation error in (12) and (14) follow. □

[Received September 2011. Revised March 2012.]

REFERENCES

Ait-Sahalia, Y., and Jacod, J. (2009), "Testing for Jumps in a Discretely Observed Process," *The Annals of Statistics*, 37, 184–222. [622]

Andersen, T. G., and Bollerslev, T. (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158. [629]

Andersen, T. G., Bollerslev, T., Diebold, F., and Labys, P. (2001), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55. [623]

Barndorff-Nielsen, O., and Shephard, N. (2001), "Non-Gaussian Ornstein–Uhlenbeck-Based Models and Some of Their Uses in Financial Fconomics," *Journal of the Royal Statistical Society,* Series B, 63, 167–241. [627,628,631,632]

——— (2004), "Power and Bipower Variation With Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37. [622,629]

——— (2006), "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1–30. [622,629]

Black, F. (1976), "Studies of Stock Price Volatility Changes," in *Proceedings of the Business and Economics Section of the American Statistical Association*, pp. 177–181. [622]

Bollerslev, T., and Todorov, V. (2011), "Estimation of Jump Tails," *Econometrica*, 79, 1727–1783. [627]

Comte, F., and Genon-Catalot, V. (2006), "Penalized Projection Estimator for Volatility Density," *Scandinavian Journal of Statistics*, 33, 875–893. [623,626]

Fan, J. (1991), "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *The Annals of Statistics*, 19, 1257–1272. [626]

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications* (Vol. 2), New York: Wiley. [624]

Jacod, J. (2008), "Asymptotic Properties of Power Variations and Associated Functionals of Semimartingales," *Stochastic Processes and Their Applications*, 118, 517–559. [630]

Jacod, J., and Shiryaev, A. N. (2003), *Limit Theorems For Stochastic Processes* (2nd ed.), Berlin: Springer-Verlag. [634]

Klüppelberg, C., Lindner, A., and Maller, R. (2004), "A Continuous Time GARCH Process Driven by a Lévy Process: Stationarity and Second Order Behavior," *Journal of Applied Probability*, 41, 601–622. [622]

Koenker, R., and Mizera, I. (2010), "Quasi-Concave Density Estimation," *The Annals of Statistics*, 38, 2998–3027. [629]

Kryzhniy, V. (2003a), "Direct Regularization of the Inversion of Real-Valued Laplace Transforms," *Inverse Problems*, 19, 573–583. [623,624,625,632]

——— (2003b), "Regularized Inversion of Integral Transformations of Mellin Convolution Type," *Inverse Problems*, 19, 1227–1240. [623,624]

Mancini, C. (2009), "Non-Parametric Threshold Estimation for Models With Stochastic Diffusion Coefficient and Jumps," *Scandinavian Journal of Statistics*, 36, 270–296. [630]

Mykland, P., and Zhang, L. (2009), "Inference for Continuous Semimartingales Observed at High Frequency," *Econometrica*, 77, 1403–1445. [622]

Prakasa Rao, B. (1988), "Statistical Inference From Sampled Data for Stochastic Processes," in *Contemporary Mathematics* (Vol. 80), *Statistical Inference from Stochastic Processes*, ed. N. U. Prabhu, Providence, RI: American Mathematical Society, pp. 249–284. [626]

Sato, K. (1999), *Lévy Processes and Infinitely Divisible Distributions*, Cambridge: Cambridge University Press. [628]

Tikhonov, A., and Arsenin, V. (1977), *Solutions of Ill-Posed Problems*, Washington, DC: Winston. [624,625]

Todorov, V., and Tauchen, G. (2012), "The Realized Laplace Transform of Volatility," *Econometrica*, 80, 1105–1127. [622,623,624,625,630]

Todorov, V., Tauchen, G., and Grynkiv, I. (2011), "Realized Laplace Transforms for Estimation of Jump-Diffusive Volatility Models," *Journal of Econometrics*, 164, 367–381. [628]

Van Es, B., Spreij, P., and Van Zanten, H. (2003), "Nonparametric Volatility Density Estimation," *Bernoulli*, 9, 451–465. [623,626]