NONPARAMETRIC LEVERAGE EFFECTS*

Federico M. Bandi† and Roberto Renò‡

October 16, 2008

Abstract

Vast empirical evidence points to the existence of a negative correlation, named "leverage effect," between shocks in volatility and shocks in returns. We provide a functional theory of leverage estimation in the context of a nonparametric, continuous-time, stochastic volatility model with jumps in returns, jumps in volatility, or both. Leverage is defined as a flexible function of the Markov state (spot variance). We show that its point-wise functional estimates have asymptotic properties (in terms of rates of convergence, limiting biases, and limiting variances) which crucially depend on the likelihood of the individual jumps and co-jumps as well as on the features of the jump size distributions. Empirically, we find economically important time-variation in stock market leverage with highly negative values associated with low and high spot volatility levels.

---

*We thank Robert Taylor and the participants at the Festschrift in honor of Peter Phillips (Singapore, July 14-15, 2008) for comments and discussions.

† *Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA. E-mail: federico.bandi@chicagogsb.edu*

‡ *Dipartimento di Economia Politica, Università di Siena, Piazza San Francesco 7, 53100, Siena, Italy. E-mail: reno@unisi.it.*

1 Introduction

Shocks to returns are negatively correlated with shocks to volatility. Due to its traditional economic justification in the case of stocks, this stylized fact has been termed "leverage effect": negative news decrease prices, increase financial leverage (i.e., the debt-to-equity ratio) and the overall riskiness of the company, thereby leading to higher stock volatility (Black, 1976). While the economics of leverage is still controversial in general, its empirical significance has been broadly established.

The time-varying volatility literature has emphasized the relevance of feedback effects between (innovations in) returns and volatility in a variety of parametric settings. Fundamental contributions in terms of modelling and pricing have been provided both in continuous time (Andersen et al., 2002, Bakshi et al., 1997, and Eraker et al., 2003, *inter alia*) and in discrete time (Engle and Ng, 1993, Glosten et al., 1993, Harvey and Shephard, 1996, and Jacquier et al., 2004, among others). Yu (2005) offers a review and an insightful assessment of the extant literature in discrete time.

This paper provides a nonparametric treatment of leverage estimation in the context of a stochastic volatility, jump-diffusion, model with discontinuities in returns, volatility, or both. The model is largely nonparametrically specified. Parametric assumptions are solely imposed on the distributions of the jump sizes for identification. Importantly, we allow leverage to be a general function of the Markov state (spot variance), and hence time-varying. We show that the limiting features of our studied kernel leverage estimator crucially depend on the continuity properties of the price and volatility processes.

Several cases are considered: absence of jumps in either process, jumps in returns, jumps in volatility, independent jumps in returns and volatility, contemporaneous jumps (or co-jumps) in returns and volatility. We show that the fastest convergence rate to a (mixed) normal distribution arises in the absence of jumps in both returns and volatility. The presence of jumps in returns (without jumps in volatility) does not affect the rate of convergence of the estimator as compared to the case with no jumps. However, it does affect asymptotic efficiency negatively by adding an additional term to the leverage estimator's limiting variance. The case of jumps in volatility (without jumps in returns) is quite different in that consistent estimation of the volatility process' diffusion function can only be conducted at a slower rate. This slower rate reduces the speed of convergence of the kernel leverage estimator. In particular, its limiting distribution is now driven by the asymptotic features of the spot variance's diffusion function estimator. Interestingly, the addition of jumps in the return process (in a model now with independent jumps in returns and volatility) does not modify this result. Since, in the presence of jumps in volatility, the asymptotic variance of the leverage estimator is already completely induced by that of the spot variance's diffusion estimator, the addition of independent jumps in returns has now (contrary to the continuous volatility case with jumps in returns) no effect even in terms of decreased asymptotic efficiency. Finally, allowing for co-jumps in returns and volatility may yield inconsistency of the leverage estimator unless the jump sizes are independent and the jumps in returns are mean zero. In this case, we show that the limiting distribution is completely driven by the features of the price/variance discontinuities and discuss ways to re-establish consistency of the leverage estimator by virtue of appropriate asymptotic bias corrections.

Naturally, kernel estimation of leverage effects requires suitable filtering of the spot variance series. We do so by considering kernel estimates of spot variance obtained by virtue of high-frequency asset price data and, when possible, allow for market microstructure noise as in Bandi and Renò (2008), BR henceforth. In particular, we show how the estimation error induced by spot variance estimation can be made asymptotically negligible for the purpose of leverage estimation. Our limit theory hinges on weaker conditions than stationarity. We solely assume recurrence[1] of the spot variance process, thereby allowing for a considerable amount of variance persistence in any given sample.

Applications of our methods and limiting results to S&P500 future data and bond future data suggest economically important nonlinearities in the leverage function and, hence, time variation in the correlation of return and volatility shocks. We find stock market leverage values around -0.3 for average volatility levels and drastically more negative leverage values for volatility levels that are either higher or lower than the average level. Bond leverage is also negative, statistically significant, and time-varying but generally smaller in absolute value (about -0.1). We now turn to the model.

2 The setting

Assume a complete probability space $(\Omega, \Im, P, \{\Im_t\}_{t\geq 0})$. Write continuously-compounded returns as $r_{t,t+1} = \log(p_{t+1}) - \log(p_t)$. We consider the system

$$
\begin{aligned}
r_{t,t+dt} &= d\log(p_t) = \mu_t dt + \sigma_t dW_t^r + dJ_t^r, & (1)\\
d\xi(\sigma_t^2) &= m_t dt + \Lambda(\sigma_t^2) dW_t^\sigma + dJ_t^\sigma, & (2)
\end{aligned}
$$

where $\{J_t^r, J_t^\sigma\}$ is a bi-dimensional compound Poisson process and, given the bi-variate standard Brownian motion $\{W_t^1, W_t^2\}$,

$$
\{dW_t^r, dW_t^\sigma\} = \{\rho(\sigma_t^2)dW_t^1 + \sqrt{1-\rho^2(\sigma_t^2)}dW_t^2, dW_t^1\}
$$

denotes increments of a bi-dimensional drift-less diffusion or, more explicitly in our framework, contemporaneous (continuous) shocks to returns and shocks to monotonic transformations of spot variance. Clearly, $< dW_t^r, dW_t^r > = < dW_t^s, dW_t^s > = dt$ and

$$
< dW_t^r, dW_t^s > \; = \rho(\sigma_t^2)dt,
$$

thereby implying that the function $\rho(\cdot)$ is a well-defined infinitesimal correlation between continuous shocks to returns and continuous shocks to spot variance if bounded between -1 and 1.

In order to specify the vector $\{J_t^r, J_t^\sigma\}$, we define three intensity functions: $\lambda_\sigma(\sigma_t^2)$, the intensity of jumps in volatility, $\lambda_r(\sigma_t^2)$, the intensity of jumps in returns, and $\lambda_{r,\sigma}(\sigma_t^2)$, the intensity of the co-jumps (see Remark 8 for more details). The jump sizes of $\xi(\sigma_t^2)$ and $\log p_t$ are determined by the random variables c_r and c_σ, respectively. We allow for correlation in both jump times and jump sizes, but not between times and sizes. We also assume independence between the jumps and the standard Brownian shocks W^1,W^2. The monotonic function $\xi(\cdot)$ in the variance process is introduced for generality. It is meant to allow for alternative specifications including the logarithmic model in, e.g., Jacquier et al. (1994) and the linear (in variance) models proposed by, e.g., Duffie et al. (2000) and Eraker et al. (2003). The object of econometric interest is $\rho(\cdot)$.

Assumption 1. *The drifts μ_t and m_t are adapted stochastic processes. The functions $\Lambda(\cdot)$, $\lambda_r(\cdot)$, $\lambda_\sigma(\cdot)$, $\lambda_{r,\sigma}(\cdot)$, and $\rho(\cdot)$ are at least twice continuously-differentiable Borel measurable functions of the*

[1] For a review of nonparametric methods for continuous-time models under Harris recurrence we refer the interested reader to Bandi and Phillips (2009).

Markov state. The random variables c_r and c_σ are bounded with probability one. All objects are such that a unique and recurrent strong solution of (1)-(2) *exists. We refer the reader to BR (2008) for details.*

We begin by assuming availability of n observations on both r_t and σ_t^2 in the time interval $[0, T]$. We denote by $\Delta_{n,T} = T/n$ the time distance between adjacent discretely-sampled observations. Our asymptotic design lets $\Delta_{n,T} \to 0$ with $T \to \infty$. The case of observability of σ_t^2 is, of course, unrealistic in practise. However, it is important in that it allows us to lay out the main ideas while avoiding the complications induced by spot variance estimation. This said, Section 8 discusses the case of spot variance estimation by virtue of kernel methods applied to high-frequency price data. As pointed out in the Introduction, this section presents conditions which guarantee that the estimation error associated with the spot variance estimates is asymptotically negligible. Importantly, these conditions take the nature of realistic intra-daily price formation mechanisms seriously and allow for market microstructure noise in spot variance estimation, when possible.

Define the following infinitesimal moments:

$$\begin{aligned}
\vartheta_{1,1}(\sigma^2) &= \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbf{E}\left[(\log p_{t+\Delta} - \log p_t) \left(\xi(\sigma_{t+\Delta}^2) - \xi(\sigma_t^2) \right) \middle| \sigma_t^2 = \sigma^2 \right], \\
\vartheta_k(\sigma^2) &= \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbf{E}\left[\left(\xi(\sigma_{t+\Delta}^2) - \xi(\sigma_t^2) \right)^k \middle| \sigma_t^2 = \sigma^2 \right], \qquad k = 1, 2, ...
\end{aligned}$$

and the kernel estimators

$$\widehat{\vartheta}_{1,1}(\sigma^2) = \frac{\sum_{i=1}^{n-1} \mathbf{K}\left(\frac{\sigma_{iT/n}^2 - \sigma^2}{h_{n,T}} \right) (\log p_{(i+1)T/n} - \log p_{iT/n}) \left(\xi(\sigma_{(i+1)T/n}^2) - \xi(\sigma_{iT/n}^2) \right)}{\Delta_{n,T} \sum_{i=1}^{n} \mathbf{K}\left(\frac{\sigma_{iT/n}^2 - \sigma^2}{h_{n,T}} \right)}, \tag{3}$$

$$\widehat{\vartheta}_k(\sigma^2) = \frac{\sum_{i=1}^{n-1} \mathbf{K}\left(\frac{\sigma_{iT/n}^2 - \sigma^2}{h_{n,T}} \right) \left(\xi(\sigma_{(i+1)T/n}^2) - \xi(\sigma_{iT/n}^2) \right)^k}{\Delta_{n,T} \sum_{i=1}^{n} \mathbf{K}\left(\frac{\sigma_{iT/n}^2 - \sigma^2}{h_{n,T}} \right)} \qquad k = 1, 2, ... \tag{4}$$

The function $\mathbf{K}(\cdot)$ satisfies the following conditions.

Assumption 2. $\mathbf{K}(\cdot)$ *is a bounded, continuously-differentiable, symmetric, and nonnegative function whose derivative $\mathbf{K}'(\cdot)$ is absolutely integrable and bounded, and for which $\int \mathbf{K}(s)ds = 1$, $\mathbf{K}_1 = \int s^2 \mathbf{K}(s) ds < \infty$, and $\mathbf{K}_2 = \int \mathbf{K}^2(s) ds < \infty$.*

In what follows, the kernel estimators $\widehat{\vartheta}_{1,1}(\sigma^2)$ and $\widehat{\vartheta}_k(\sigma^2)$ (with $k = 1, \ldots$) will be employed to provide point-wise estimates of $\rho(\sigma^2)$ in various scenarios allowing for jumps in prices, jumps in volatility, or both. We will use the classical notation $\xrightarrow{p}$, $\Rightarrow$ to denote convergence in probability and weak convergence. The symbol $\Gamma_z(x)$ will be employed to define $h_{n,T}^2 \mathbf{K}_1 \left[z^{'}(x) \frac{s^{'}(x)}{s(x)} + \frac{1}{2} z^{''}(x) \right]$, where $s(dx)$ is the invariant measure (the speed measure, in the absence of variance jumps) of the spot variance process. Finally, the notation $\widehat{\overline{L}}_{\sigma^2}(T, \sigma^2) = \frac{\Delta_{n,T}}{h_{n,T}} \sum_{i=1}^{n} \mathbf{K}\left(\frac{\sigma_{iT/n}^2 - \sigma^2}{h_{n,T}} \right)$ will denote kernel estimates of the local time (at T and σ^2) of the underlying spot variance process. This quantity is known to play a role in characterizing the rate of convergence of nonparametric kernel estimates of recurrent Markov process (see, e.g., the review paper by Bandi and Phillips, 2009) and will be important in what follows. We begin with a relevant benchmark case.

3 The continuous case: $J^r = 0,\ J^\sigma = 0$

Write the kernel leverage estimator as

$$\widehat{\rho}(\sigma^2) = \frac{\widehat{\vartheta}_{1,1}(\sigma^2)}{\sigma\sqrt{\widehat{\vartheta}_2(\sigma^2)}}. \tag{5}$$

In the absence of discontinuities, we note that $\vartheta_{1,1}(\sigma^2) = \sigma\Lambda(\sigma^2)\rho(\sigma^2)$ and $\vartheta_2(\sigma^2) = \Lambda^2(\sigma^2)$. Hence, $\frac{\vartheta_{1,1}(\sigma^2)}{\sigma\sqrt{\vartheta_2(\sigma^2)}} = \rho(\sigma^2)$ and $\widehat{\rho}(\sigma^2)$ is expected to consistently estimate our object of interest. This result is shown in Theorem 1 below.

Theorem 1. *Assume* $J^r = J^\sigma = 0$. *Let* $n, T \to \infty$ *and* $h_{n,T} \to 0$ *so that*

$$\frac{v(T)}{h_{n,T}}\left(\Delta_{n,T}\log\frac{1}{\Delta_{n,T}}\right)^{1/2} \to 0,$$

where $\overline{L}_{\sigma^2}(T,\sigma^2) \sim v(T)$ *and* $v(T)$ *is a regularly-varying function at infinity, then* $\widehat{\rho}(\sigma^2) \xrightarrow{p} \rho(\sigma^2)$. *Furthermore, if* $\frac{h^5_{n,T}v(T)}{\Delta_{n,T}} \to C$, *where* C *is a suitable constant, then*

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left\{\widehat{\rho}(\sigma^2) - \rho(\sigma^2) - \widetilde{\Gamma}_\rho(\sigma^2)\right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2\left[1 - \frac{1}{2}\rho^2(\sigma^2)\right]\right), \tag{6}$$

where

$$\widetilde{\Gamma}_\rho(\sigma^2) = \frac{1}{\sigma\sqrt{\vartheta_2(\sigma^2)}}\Gamma_{\vartheta_{1,1}}(\sigma^2) - \frac{\vartheta_{1,1}(\sigma^2)}{2\sigma\sqrt{\vartheta_2^3(\sigma^2)}}\Gamma_{\vartheta_2}(\sigma^2). \tag{7}$$

Proof. *See Appendix A.*

Remark 1. In the absence of jumps in either volatility or returns, the leverage estimator converges at speed $\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}} \sim \sqrt{\frac{h_{n,T}v(T)}{\Delta_{n,T}}}$. In particular, both the numerator, $\widehat{\vartheta}_{1,1}(\sigma^2)$, and the denominator, $\sigma\sqrt{\widehat{\vartheta}_2(\sigma^2)}$, converge at this same velocity. The asymptotic distribution of $\widehat{\rho}(\sigma^2)$ is therefore a linear combination (with weights $\frac{1}{\sigma\sqrt{\vartheta_2(\sigma^2)}}$ and $-\frac{\vartheta_{1,1}(\sigma^2)}{2\sigma\sqrt{\vartheta_2^3(\sigma^2)}}$) of the limiting distributions of its components as evidenced by the resulting limiting bias ($\widetilde{\Gamma}_\rho(\sigma^2)$).

Remark 2. The asymptotic variance of the estimator is maximal (and equal to $\mathbf{K}_2$) for $\rho(\cdot) = 0$. It tends to $\frac{1}{2}\mathbf{K}_2$ as either $\rho(\cdot) \to 1$ or $\rho(\cdot) \to -1$.

4 The discontinuous case: $J^r \neq 0,\ J^\sigma = 0$

Consider the same estimator as in Eq. (5) above. The case with jumps in returns is presented in Theorem 2.

Theorem 2. *Assume* $J^r \neq 0$ *and* $J^\sigma = 0$. *Under the same conditions as in Theorem 1:*

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left\{\widehat{\rho}(\sigma^2) - \rho(\sigma^2) - \widetilde{\Gamma}_\rho(\sigma^2)\right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2\left[\left(1 - \frac{1}{2}\rho^2(\sigma^2)\right) + \frac{1}{2}\frac{\lambda_r(\sigma^2)\mathbf{E}[c_r^2]}{\sigma^2}\right]\right).$$

Proof. *See Appendix A.*

Remark 3. Allowing for jumps in returns only affects the (limiting) precision of the estimator. The asymptotic variance now contains an extra term ($\frac{1}{2}\frac{\lambda_r(\sigma^2)\mathbf{E}[c_r^2]}{\sigma^2}$) which, of course, depends on the frequency of the return jumps ($\lambda_r(\sigma^2)$) as well as on their size ($\mathbf{E}[c_r^2]$).

5 The discontinuous case: $J^r = 0,\ J^\sigma \neq 0$

When allowing for jumps in the variance process, $\widehat{\vartheta}_2(\sigma^2)$ estimates $\Lambda^2(\sigma^2)$ plus the second conditional moment of the jump part (i.e., $\lambda_\sigma(\sigma^2)\mathbf{E}(c_\sigma^2)$). In what follows, we show that $\lambda_\sigma(\sigma^2)\mathbf{E}(c_\sigma^2)$ can be identified, under parametric assumptions on the jump sizes, by virtue of the higher-order conditional moments (namely, $\widehat{\vartheta}_k(\sigma^2)$ with $k = 3, 4$) as proposed by BR (2008) in the case of nonparametric stochastic volatility modelling (see, also, Bandi and Nguyen, 2003, and Johannes, 2004). Hence, the form of the kernel leverage estimator in this section will be

$$\widetilde{\rho}(\sigma^2) = \frac{\widehat{\vartheta}_{1,1}(\sigma^2)}{\sigma f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\sigma^2)}, \tag{8}$$

where $f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$ is a specific function of the infinitesimal moments to which we now turn.

Assume the volatility jumps are exponentially distributed, i.e., $c_\sigma \sim \exp(\mu_\sigma)$, and $\xi(\sigma^2) = \sigma^2$. This specification is widely used in the parametric literature on volatility estimation (see, e.g., Eraker et al., 2003) and has been shown to perform very satisfactorily in a (semi-)nonparametric context (BR, 2008). The adopted model implies

$$\begin{aligned} \vartheta_2(\sigma^2) &= \Lambda^2(\sigma^2) + 2\mu_\sigma^2 \lambda_\sigma(\sigma^2), \\ \vartheta_3(\sigma^2) &= 6\mu_\sigma^3 \lambda_\sigma(\sigma^2), \\ \vartheta_4(\sigma^2) &= 24\mu_\sigma^4 \lambda_\sigma(\sigma^2). \end{aligned}$$

Hence,

$$\begin{aligned} \widehat{\mu}_\sigma &= \frac{1}{\overline{n}} \sum_{i=1}^{\overline{n}} \frac{\widehat{\vartheta}_4(\sigma^2_{i\overline{T}/\overline{n}})}{4\widehat{\vartheta}_3(\sigma^2_{i\overline{T}/\overline{n}})}, \\ \widehat{\lambda}_\sigma(\sigma^2) &= \frac{\widehat{\vartheta}_4(\sigma^2)}{24\widehat{\mu}_\sigma^4}, \end{aligned}$$

and

$$\widehat{\Lambda}^2(\sigma^2) = \widehat{\vartheta}_2(\sigma^2) - 2\widehat{\mu}_\sigma^2 \widehat{\lambda}_{\sigma^2}(\sigma^2),$$

are kernel estimators of μ_σ, $\lambda_\sigma(\sigma^2)$, and $\Lambda^2(\sigma^2)$, respectively.[2] These estimators have been shown to be consistent (with probability one) and (mixed) normally distributed (BR, 2008). Thus,

$$f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\sigma^2) = \widehat{\Lambda}^2(\sigma^2) = \widehat{\vartheta}_2(\sigma^2) - \frac{\widehat{\vartheta}_4(\sigma^2)}{12\widehat{\mu}_\sigma^2} = \widehat{\vartheta}_2(\sigma^2) - \frac{\widehat{\vartheta}_4(\sigma^2)}{12\left(\frac{1}{\overline{n}} \sum_{i=1}^{\overline{n}} \frac{\widehat{\vartheta}_4(\sigma^2_{i\overline{T}/\overline{n}})}{4\widehat{\vartheta}_3(\sigma^2_{i\overline{T}/\overline{n}})}\right)^2}.$$

Importantly, alternative estimation schemes may be adopted by imposing, for example, different distributional assumptions on c_σ. In these cases, our methods can of course be adapted accordingly.

[2] For technical reasons (discussed in BR, 2008, for instance), $\widehat{\mu}_\sigma$ is defined over a number of observations $\overline{n}$ growing to infinity over a fixed time span $\overline{T}$. The kernel estimators $\widehat{\vartheta}_3(\cdot)$ and $\widehat{\vartheta}_4(\cdot)$ continue to be defined over an enlarging time span ($T \to \infty$) for consistency.

In what follows, we assume using a slightly smaller bandwidth sequence to identify $\widehat{\vartheta}_4$ and $\widehat{\vartheta}_3$ for the purpose of $\widehat{\mu}_\sigma$ estimation. This choice will slightly simplify the look of the limiting biases of $\widehat{\rho}$ by preventing the insurgence of the asymptotic biases of $\widehat{\mu}_\sigma$.

Theorem 3. *Assume $J^r = 0$ and $J^\sigma \neq 0$. If $\Delta_{n,T} \to 0$ (with $n, T \to \infty$), $h_{n,T} \to 0$,*

$$h_{n,T} v(T) \to \infty,$$

and

$$\frac{v(T)}{h_{n,T}} \left(\Delta_{n,T} \log \frac{1}{\Delta_{n,T}} \right)^{1/2} \to 0,$$

where $\overline{L}_{\sigma^2}(T, \sigma^2) \sim v(T)$ and $v(T)$ is a regularly-varying function at infinity, then $\widetilde{\rho}(\sigma^2) \xrightarrow{p} \rho(\sigma^2)$. Further, if $\rho(\sigma^2) \neq 0$ and $h^5_{n,T} v(T) \to C$, where C is a suitable constant, then

$$\sqrt{h_{n,T} \widehat{\overline{L}}_{\sigma^2}(T, \sigma^2)} \left\{ \widetilde{\rho}(\sigma^2) - \rho(\sigma^2) - \widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) \right\} \Rightarrow \mathbf{N} \left(0, \mathbf{K}_2 \frac{\rho^2(\sigma^2)}{4\Lambda^4(\sigma^2)} \lambda_\sigma(\sigma^2) \mathbf{E} \left(\left(c_\sigma^2 - \frac{1}{12\mu_\sigma^2} c_\sigma^4 \right)^2 \right) \right) \quad (9)$$

with

$$\widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) = \frac{1}{\sigma \Lambda(\sigma^2)} \Gamma_{\vartheta_{1,1}}(\sigma^2) - \frac{\vartheta_{1,1}(\sigma^2)}{2\sigma \Lambda^3(\sigma^2)} \left(\Gamma_{\vartheta_2}(\sigma^2) - \frac{1}{12\mu_\sigma^2} \Gamma_{\vartheta_4}(\sigma^2) \right).$$

Proof. *See Appendix A.*

Remark 4. The estimator $f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$ now converges to $\Lambda^2(\cdot)$ at a slower speed than $\widehat{\vartheta}_2(\cdot)$ for the case of no jumps ($\sqrt{h_{n,T} \widehat{\overline{L}}_{\sigma^2}(T, \sigma^2)}$ versus $\sqrt{\frac{h_{n,T} \widehat{\overline{L}}_{\sigma^2}(T, \sigma^2)}{\Delta_{n,T}}}$). Since $\widehat{\vartheta}_{1,1}(\cdot)$ continues to converge at speed $\sqrt{\frac{h_{n,T} \widehat{\overline{L}}_{\sigma^2}(T, \sigma^2)}{\Delta_{n,T}}}$, not only is the slower speed of convergence of $f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$ driving the rate of convergence of $\widehat{\rho}(\cdot)$ but, also, of course, the asymptotic variance of the leverage estimator is fully determined by the asymptotic variance of $f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$ (times a term $\frac{\vartheta^2_{1,1}(\sigma^2)}{4\sigma^2 \Lambda^6(\sigma^2)} = \frac{\rho^2(\sigma^2)}{4\Lambda^4(\sigma^2)}$ which readily derives from the delta method - see, e.g., Remark 1 above).

Remark 5. Under the assumed exponential jumps, the asymptotic variance in Eq. (9) can be more explicitly expressed as $46\mathbf{K}_2 \frac{\rho^2(\sigma^2)\lambda_\sigma(\sigma^2)}{\Lambda^4(\sigma^2)} \mu_\sigma^4$.

6 The discontinuous case: $J^r \neq 0$, $J^\sigma \neq 0$, with independent jumps

Consider the same estimator as in Eq. (8) above.

Theorem 4. *Assume $J^r \neq 0$, $J^\sigma \neq 0$, and $J^r \perp J^\sigma$. Under the same conditions as in Theorem 3:*

$$\sqrt{h_{n,T} \widehat{\overline{L}}_{\sigma^2}(T, \sigma^2)} \left\{ \widetilde{\rho}(\sigma^2) - \rho(\sigma^2) - \widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) \right\} \Rightarrow \mathbf{N} \left(0, \mathbf{K}_2 \frac{\rho^2(\sigma^2)}{4\Lambda^4(\sigma^2)} \lambda_\sigma(\sigma^2) \mathbf{E} \left(\left(c_\sigma^2 - \frac{1}{12\mu_\sigma^2} c_\sigma^4 \right)^2 \right) \right) \quad (10)$$

with

$$\widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) = \frac{1}{\sigma \Lambda(\sigma^2)} \Gamma_{\vartheta_{1,1}}(\sigma^2) - \frac{\vartheta_{1,1}(\sigma^2)}{2\sigma \Lambda^3(\sigma^2)} \left(\Gamma_{\vartheta_2}(\sigma^2) - \frac{1}{12\mu_\sigma^2} \Gamma_{\vartheta_4}(\sigma^2) \right).$$

Proof. *See Appendix A.*

Remark 6. Adding independent jumps in returns to the case of jumps in volatility does not modify the limiting distribution of $\widetilde{\rho}(\cdot)$ ((10) is the same as (9)). This is, of course, in contrast to the case where jumps in returns are added to the case of no jumps. Here the addition of independent return jumps does not translate into efficiency losses, as implied by a higher asymptotic variance, since the limiting variance of the leverage estimator is again only driven by the denominator, $\sigma f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$.

7 The discontinuous case: $J^r \neq 0,\ J^\sigma \neq 0$, with correlated jumps

Finally, we allow for correlated jumps and, again, evaluate the estimator in Eq. (8).

Theorem 5. *Assume $J_r = J_r^* \neq 0$, $J_\sigma = J_\sigma^* \neq 0$, and the intensity of common shocks $\lambda_r(\sigma^2) = \lambda_\sigma(\sigma^2) = \lambda_{r,\sigma}(\sigma^2) \neq 0$. If $\Delta_{n,T} \to 0$ (with $n, T \to \infty$), $h_{n,T} \to 0$,*

$$h_{n,T} v(T) \to \infty,$$

and

$$\frac{v(T)}{h_{n,T}} \left(\Delta_{n,T} \log \frac{1}{\Delta_{n,T}} \right)^{1/2} \to 0,$$

where $\overline{L}_{\sigma^2}(T, \sigma^2) \sim v(T)$ and $v(T)$ is a regularly-varying function at infinity, then

$$\widetilde{\rho}(\sigma^2) \xrightarrow{p} \Xi(\sigma^2) = \rho(\sigma^2) + \frac{\lambda_{r,\sigma}(\sigma^2)}{\sigma \Lambda(\sigma^2)} \mathbf{E}[c_r c_\sigma].$$

Further, if $h_{n,T}^5 v(T) \to C$, where C is a suitable constant, then

$$\sqrt{h_{n,T} \widehat{\overline{L}}_{\sigma^2}(T, \sigma^2)} \left\{ \widetilde{\rho}(\sigma^2) - \Xi(\sigma^2) - \widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) \right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2 V_\Xi\right) \tag{11}$$

with

$$V_\Xi = \frac{\lambda_{r,\sigma}(\sigma^2)}{\sigma^2 \Lambda^2(\sigma^2)} \mathbf{E}\left[\left(c_r c_\sigma - \frac{\vartheta_{1,1}(\sigma^2)}{2\Lambda^2(\sigma^2)} \left(c_\sigma^2 - \frac{c_\sigma^4}{12\mu_\sigma^2} \right) \right)^2 \right],$$

and

$$\widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) = \frac{1}{\sigma \Lambda(\sigma^2)} \Gamma_{\vartheta_{1,1}}(\sigma^2) - \frac{\vartheta_{1,1}(\sigma^2)}{2\sigma \Lambda^3(\sigma^2)} \left(\Gamma_{\vartheta_2}(\sigma^2) - \frac{1}{12\mu_\sigma^2} \Gamma_{\vartheta_4}(\sigma^2) \right),$$

with

$$\vartheta_{1,1}(\sigma^2) = \sqrt{\sigma^2} \Lambda(\sigma^2) \rho(\sigma^2) + \lambda_{r,\sigma}(\sigma^2) \mathbf{E}[c_r c_\sigma].$$

Proof. *See Appendix A.*

Remark 7. In this case, the kernel leverage estimator is inconsistent. Since both the numerator, $\widehat{\vartheta}_{1,1}(\cdot)$, and the denominator, $f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$, converge at the same rate, the limiting distribution of $\widetilde{\rho}(\cdot)$ is that of a linear combination of $\widehat{\vartheta}_{1,1}(\cdot)$ and $f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$.

Corollary to Theorem 5 (A relevant example: Independent jump sizes with mean zero return jumps). Under the assumptions of Theorem 5, if $c_r \perp c_\sigma$, $\mathbf{E}[c_r] = 0$, $\mathbf{E}[c_r^2] = \sigma_r^2$, and $c_\sigma \sim \exp(\mu_\sigma)$, then $\widetilde{\rho}(\sigma^2) \xrightarrow{p} \rho(\sigma^2)$ and consistency is preserved. In addition:

$$\sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\left\{\widetilde{\rho}(\sigma^2) - \rho(\sigma^2) - \widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2)\right\} \Rightarrow \mathbf{N}\left(0, \mathbf{K}_2 \frac{\lambda_{r,\sigma}(\sigma^2)}{\sigma^2\Lambda^2(\sigma^2)}\left[2\sigma_r^2\mu_\sigma^2 + 46\frac{\sigma^2\rho^2(\sigma^2)}{\Lambda^2(\sigma^2)}\mu_\sigma^4\right]\right), \quad (12)$$

with

$$\widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2) = \frac{1}{\sigma\Lambda(\sigma^2)}\Gamma_{\vartheta_{1,1}}(\sigma^2) - \frac{\vartheta_{1,1}(\sigma^2)}{2\sigma\Lambda^3(\sigma^2)}\left(\Gamma_{\vartheta_2}(\sigma^2) - \frac{1}{12\mu_\sigma^2}\Gamma_{\vartheta_4}(\sigma^2)\right),$$

and

$$\vartheta_{1,1}(\sigma^2) = \sqrt{\sigma^2}\Lambda(\sigma^2)\rho(\sigma^2).$$

Remark 8. (Contemporaneous and non-contemporaneous jumps) The theorem solely assumes contemporaneous jumps with an infinitesimal probability of co-jumps equal to $\lambda_{r,\sigma}(\sigma^2)dt$. This is a classical case of dependence in the parametric literature. It is considered, for example, in model SVCJ in Eraker et al. (2003).[3] More generally, we could assume $J_r = J_r^* + J_r^{||}$ and $J_\sigma = J_\sigma^* + J_\sigma^{||}$, with $J_r^* \perp J_\sigma^*$, $J_r^* \perp J_r^{||}$, $J_r^* \perp J_\sigma^{||}$ and $J_\sigma^* \perp J_\sigma^{||}$, $J_\sigma^* \perp J_r^{||}$. More explicitly, we could assume that both processes comprise two components, $J_{r,\sigma}^*$ and $J_{r,\sigma}^{||}$, that are independent of each other and all others with the exception of $J_r^{||}$ and $J_\sigma^{||}$, which are dependent. Denote now by $c_{r,\sigma}^*$ and $\lambda_{r,\sigma}^*$ the jump sizes and intensities of the jumps of the independent components $J_{r,\sigma}^*$. Similarly, denote by $c_{r,\sigma}^{||}$ and $\lambda_{r,\sigma}^{||} = \lambda_r^{||} = \lambda_\sigma^{||}$ the jump sizes of the dependent components and the (common) intensity of the common shocks. The result in Eq. (11) continues to hold and may be re-written as follows:

$$\sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\left\{\widetilde{\rho}(\sigma^2) - \Xi(\sigma^2) - \widetilde{\Gamma}_{\widetilde{\rho}}(\sigma^2)\right\}$$
$$\Rightarrow \mathbf{N}\left(0, \mathbf{K}_2\left[\begin{array}{c} \frac{\lambda_{r,\sigma}^{||}(\sigma^2)}{\sigma^2\Lambda^2(\sigma^2)}\mathbf{E}\left[\left(c_r^{||}c_\sigma^{||} - \frac{\vartheta_{1,1}(\sigma^2)}{2\Lambda^2(\sigma^2)}\left((c_\sigma^{||})^2 - \frac{(c_\sigma^{||})^4}{12\mu_\sigma^2}\right)\right)^2\right] \\ +\frac{\vartheta_{1,1}^2(\sigma^2)}{4\sigma^2\Lambda^6(\sigma^2)}\lambda_\sigma^*\mathbf{E}\left[\left((c_\sigma^*)^2 - \frac{(c_\sigma^*)^4}{12\mu_\sigma^4}\right)^2\right]\end{array}\right]\right),$$

where

$$\Xi(\sigma^2) = \rho(\sigma^2) + \frac{\lambda_{r,\sigma}^{||}(\sigma^2)}{\sigma\Lambda(\sigma^2)}\mathbf{E}[c_r^{||}c_\sigma^{||}].$$

In this case, the limiting value of $\widetilde{\rho}(\sigma^2)$ is the same as that in Theorem 5. The same is true for the convergence rate. However, the limiting variance depends now explicitly on the sizes and common intensity of the dependent jumps ($c_{r,\sigma}^{||}$ and $\lambda_{r,\sigma}^{||}$) as well as on the size and intensity of the independent volatility jumps (c_σ^* and λ_σ^*).

Remark 9. (Re-establishing consistency) Methods have been put forward to identify co-jumps. Gobbi and Mancini (2008), for example, suggest identifying the contemporaneous discontinuities of two generic jump-diffusion processes X_1 and X_2 by virtue of products of the type

$$\Delta X_1 \mathbf{1}_{\left((\Delta X_1)^2 \geq r(\Delta_{n,T})\right)}\Delta X_2 \mathbf{1}_{\left((\Delta X_2)^2 \geq r(\Delta_{n,T})\right)},$$

[3] Their model SVIJ assumes independence of the jumps as in Section 6 above.

where $r(\delta)$ is a function such that $\frac{\delta \log\left(\frac{1}{\delta}\right)}{r(\delta)} \to 0$ when $\delta \to 0$. Asymptotically (for $\Delta_{n,T} \to 0$), the indicators effectively eliminate variations which are smaller than a threshold. Since the threshold is modelled based on the modulus of continuity of Brownian notion, the variations being eliminated are of the Brownian type, thereby leading to identification of the contemporaneous Poisson jumps.

Once a time-series of co-jumps is formed, the corresponding intensity ($\lambda^{\|}_{1,2}$) may be evaluated, possibly under an assumption of constancy, by computing the in-sample frequency of co-jumps. Similarly, the expected first cross-moment ($\mathbf{E}[c^{\|}_1 c^{\|}_2]$) can be consistently identified by virtue of sample averages of the co-jumps (under, of course, stationarity of the jump distributions). Along with $\widehat{\Lambda}^2(\cdot) = f(\widehat{\vartheta}_2, \widehat{\vartheta}_3, \widehat{\vartheta}_4)(\cdot)$, both estimated quantities may then be employed to re-establish consistency by virtue of $\widetilde{\rho}(\sigma^2) - \frac{\widehat{\lambda}^{\|}_{r,\sigma}(\sigma^2)}{\sigma\sqrt{f(\widehat{\vartheta}_2,\widehat{\vartheta}_3,\widehat{\vartheta}_4)(\sigma^2)}}\widehat{\mathbf{E}}[c^{\|}_r c^{\|}_\sigma]$.

8 Allowing for spot variance estimation

When implementing $\widehat{\vartheta}_{1,1}(\sigma^2)$ and $\widehat{\vartheta}_k(\sigma^2)$ with $k = 1, 2, 3, 4$, one must replace $\sigma^2_{iT/n}$ with spot variance estimates $\widehat{\sigma}^2_{iT/n}$. To this extent, assume availability of k (possibly not equi-spaced) high-frequency observations over each time interval $[i\Delta_{n,T}, i\Delta_{n,T} + \phi_{n,T}]$. Assume these k observations are employed to estimate the *integrated variance* of the logarithmic price process over each interval (i.e., $\int_{i\Delta_{n,T}}^{i\Delta_{n,T}+\phi_{n,T}} \sigma^2_s ds$) by virtue of a generic estimator $\widehat{V}_{iT/n}$. We make the following assumption about $\widehat{V}_{iT/n}$.

Assumption 3. $\widehat{V}_{iT/n}$ *is such that*

$$\mathbf{E}_{\sigma^2}\left(\phi^{\beta}_{n,T} k^{\alpha}\left(\frac{\widehat{V}_{iT/n}}{\phi_{n,T}} - \frac{\int_{iT/n}^{iT/n+\phi_{n,T}} \sigma^2_s ds}{\phi_{n,T}}\right)\right) \overset{a}{\approx} 0$$

and

$$\mathbf{V}_{\sigma^2}\left(\phi^{\beta}_{n,T} k^{\alpha}\left(\frac{\widehat{V}_{iT/n}}{\phi_{n,T}} - \frac{\int_{iT/n}^{iT/n+\phi_{n,T}} \sigma^2_s ds}{\phi_{n,T}}\right)\right) \overset{a}{\approx} \left(a\left(\sigma^4_{iT/n}\right)^{\eta} + b\right)$$

with $\alpha \in (0, \frac{1}{2}]$ and $\beta \in [0, 1]$ *given* T *and* n. $\mathbf{E}_{\sigma^2}$ *and* $\mathbf{V}_{\sigma^2}$ *denote expectation and variance conditional on the spot volatility path.* a, b, *and* η *are numbers. The symbol* $\overset{a}{\approx}$ *denotes asymptotic equivalence for a large* k *and a small* $\phi_{n,T}$.

Now define $\widehat{\sigma}^2_{iT/n} = \frac{\widehat{V}_{iT/n}}{\phi_{n,T}}$ and let $\phi_{n,T} \to 0$ as $k \to \infty$. Under Assumption 3, BR (2008) show that a theory of consistent *spot variance* estimation based on $\widehat{\sigma}^2_{iT/n}$ can be derived by controlling the rate at which $\phi_{n,T}$ vanishes as k goes off to infinity. Importantly, they show that Assumption 3 is satisfied by virtually all know estimators $\widehat{V}_{iT/n}$ of integrated variance. For example, for appropriate choices of α, β, a, b, and η, it is satisfied by the classical realized variance estimator (Andersen et al., 2003, and Barndorff-Nielsen and Shephard, 2002). For alternative, appropriate choices of the same parameters, and appropriate choices of the number of subsamples/autocovariances, it is satisfied by the two-scale estimator of Zhang et al. (2005) as well as by the family of flat-top symmetric kernels suggested by Barndorff-Nielsen et al. (2009). Both are, among other estimators, robust to market microstructure noise. Specifically, BR (2008) show that, for a general class of integrated variance estimators $\widehat{V}_{iT/n}$,

$$\phi^{\beta}_{n,T} k^{\alpha}\left(\widehat{\sigma}^2_{iT/n} = \frac{\widehat{V}_{iT/n}}{\phi_{n,T}} - \sigma^2_{iT/n}\right) \Rightarrow \mathbf{MN}\left(0, a\left(\sigma^4_{iT/n}\right)^{\eta} + b\right)$$

if $\phi_{n,T}^{\beta} k^{\alpha} \to \infty$ and $\phi_{n,T}^{\beta} k^{\alpha} \left(\phi_{n,T} \log\left(\frac{1}{\phi_{n,T}}\right)\right)^{1/2} \to 0$ (under additional, more specific, conditions). Appendix A in BR (2008) details the values of α, β, a, b, and η for a variety of spot variance estimators $\widehat{\sigma}_{iT/n}^2$ with different robustness properties with respect to jumps and market microstructure noise. For interesting, alternative approaches to spot variance estimation using realized variance and Fourier estimates, we refer the reader to the recent work of Malliavin and Mancino (2008) and Kristensen (2008).

If feasibility is restored by employing $\widehat{\sigma}_{iT/n}^2$ in place of the unobservable $\sigma_{iT/n}^2$, $\forall i = 1, ..., n$, the induced estimation error must be controlled by relating the limiting properties of n, T, and $\Delta_{n,T}$ to those of $\phi_{n,T}$ and k. The following theorem does so.

Theorem 6 *Under Assumption 3, $\phi_{n,T} \to 0$, $k \to \infty$, and*

$$\frac{Tv(T)^{-1}}{\Delta_{n,T} h_{n,T} k^{\alpha} \phi_{n,T}^{\beta}} + \frac{Tv(T)^{-1}}{\Delta_{n,T} h_{n,T}} \left(\phi_{n,T} \log\left(\frac{1}{\phi_{n,T}}\right)\right)^{1/2} \to 0$$

the consistency results in Theorems 1-5 hold when replacing $\sigma_{iT/n}^2$ with $\widehat{\sigma}_{iT/n}^2$.

Under Assumption 3, $\phi_{n,T} \to 0$, $k \to \infty$, and

$$\frac{Tv(T)^{-1/2}}{\Delta_{n,T}^{3/2} h_{n,T}^{1/2} k^{\alpha} \phi_{n,T}^{\beta}} + \frac{Tv(T)^{-1/2}}{\Delta_{n,T}^{3/2} h_{n,T}^{1/2}} \left(\phi_{n,T} \log\left(\frac{1}{\phi_{n,T}}\right)\right)^{1/2} \to 0$$

the weak convergence results in Theorems 1 and 2 hold when replacing $\sigma_{iT/n}^2$ with $\widehat{\sigma}_{iT/n}^2$.

Under Assumption 3, $\phi_{n,T} \to 0$, $k \to \infty$, and

$$\frac{Tv(T)^{-1/2}}{\Delta_{n,T} h_{n,T}^{1/2} k^{\alpha} \phi_{n,T}^{\beta}} + \frac{Tv(T)^{-1/2}}{\Delta_{n,T} h_{n,T}^{1/2}} \left(\phi_{n,T} \log\left(\frac{1}{\phi_{n,T}}\right)\right)^{1/2} \to 0$$

the weak convergence results in Theorems 3 through 5 hold when replacing $\sigma_{iT/n}^2$ with $\widehat{\sigma}_{iT/n}^2$.

We now turn to the empirical work.

9 Leverage estimates: stocks and bonds

We apply the leverage estimator to index futures and bond future data. We employ high-frequency S&P500 future prices from January 1990 to December 2007 (4,344 days). The bond data are high-frequency 30-year US Treasury bond future prices from January 1990 to October 2003 for a total of 3,231 daily observations.

We express the quantities of interest as a function of spot volatility σ. Each day, we use intra-daily observations (interpolated on a 5-minute grid, 84 intervals per day) to construct (daily) spot variance estimates (hence, $k = 84$). To do so, we use the combination of the threshold technique advocated by Mancini (2007) and bipower variation, as proposed by Corsi et al. (2008) - $TBPV$. We refer the interested reader to Corsi et al. (2008) for details on construction. In particular, Theorem 2.3 in Corsi et al. (2008) implies that $TBPV$ satisfies Assumption 3 with $\alpha = 1/2$, $\beta = 0$, $a \simeq 2.6$, $b = 0$, and $\eta = 1$

(BR, 2008).[4] Thus, our spot volatility estimator is:

$$\widehat{\sigma}_{iT/n} = \sqrt{\frac{TBPV}{\phi_{n,T}}}.$$

We consider our more general case of jumps in both returns and variance. The nonparametric estimates $\widehat{\vartheta}_{1,1}(\sigma)$ and $\widehat{\vartheta}_k(\sigma)$ with $k = 1, 2, 3, 4$ are implemented with $h_{n,T} = h_s \hat{s} n^{-\frac{1}{5}}$, where $\hat{s}$ is the standard deviation of the time-series of spot volatilities. Based on preliminary investigation, we set $h_s = 2$ for $\widehat{\vartheta}_{1,1}(\sigma)$ and $\widehat{\vartheta}_2(\sigma)$ and use $h_s = 4$ for $\widehat{\vartheta}_1(\sigma)$, $\widehat{\vartheta}_3(\sigma)$, and $\widehat{\vartheta}_4(\sigma)$. Nonparametric identification is conducted using a first-order correction in $\Delta_{n,T}$. This correction is immaterial asymptotically but has the potential to improve finite-sample performance, particularly when evaluating the intensity of volatility jumps (which, of course, plays a role in the denominator of $\widetilde{\rho}(\sigma)$). Specifically, assuming exponential jumps in volatility with parameter μ_σ as done earlier, we have

$$\begin{aligned}
\vartheta_2(\sigma) &\approx \Lambda^2(\sigma) + 2\mu_\sigma^2 \lambda_\sigma(\sigma), \\
\vartheta_3(\sigma) &\approx 6\mu_\sigma^3 \lambda_\sigma(\sigma) + 3\vartheta_1(\sigma)\vartheta_2(\sigma)\Delta, \\
\vartheta_4(\sigma) &\approx 24\mu_\sigma^4 \lambda_\sigma(\sigma) + \left[3\left(\vartheta_2(\sigma)\right)^2 + 4\vartheta_1(\sigma)\vartheta_3(\sigma)\right]\Delta,
\end{aligned}$$

Hence, we identify the system through:

$$\begin{aligned}
\widetilde{\mu}_\sigma &= \frac{1}{4}\sum_{i=1}^{\overline{n}} \left(\frac{\widehat{\vartheta}_4(\widehat{\sigma}_{i\overline{T}/\overline{n}}) - 3\Delta_{n,T}\left[\left(\widehat{\vartheta}_2(\widehat{\sigma}_{i\overline{T}/\overline{n}})\right)^2 + 4\widehat{\vartheta}_1(\widehat{\sigma}_{i\overline{T}/\overline{n}})\widehat{\vartheta}_3(\widehat{\sigma}_{i\overline{T}/\overline{n}})\right]}{\widehat{\vartheta}_3(\widehat{\sigma}_{i\overline{T}/\overline{n}}) - 3\Delta_{n,T}\widehat{\vartheta}_1(\widehat{\sigma}_{i\overline{T}/\overline{n}})\widehat{\vartheta}_2(\widehat{\sigma}_{i\overline{T}/\overline{n}})} \right), \\
\widetilde{\lambda}_\sigma(\sigma) &= \frac{\widehat{\vartheta}_4(\sigma) - \Delta_{n,T}\left[3\left(\widehat{\vartheta}_2(\sigma)\right)^2 + 4\widehat{\vartheta}_1(\sigma)\widehat{\vartheta}_3(\sigma)\right]}{24\widehat{\mu}_\sigma^4}, \\
\widetilde{\Lambda}^2(\sigma) &= \widehat{\vartheta}_2(\sigma) - 2\widehat{\mu}_\sigma^2\widehat{\lambda}_\sigma(\sigma),
\end{aligned}$$

and, of course,

$$\widetilde{\widetilde{\rho}}(\sigma) = \frac{\widehat{\vartheta}_{1,1}(\sigma)}{\sqrt{\sigma\widetilde{\Lambda}^2(\sigma)}},$$

with $\widehat{\vartheta}_{1,1}(\cdot)$ as defined in Eq. (3). We apply a similar first-order correction to evaluate the confidence bands. These are obtained by using the limiting results in Section 6 for the case with independent return/volatility jumps. Finally, when estimating μ_σ we weigh the addend by virtue of the estimated local time at $\widehat{\sigma}_{iT/n}$, $\forall i = 1, ..., n$. [5]

Empirical findings are presented in Fig. 1 and Fig. 2. The left column of Fig. 1 contains the S&P500 future volatility's diffusion function, the intensity of the jumps in volatility (expressed in terms of number of yearly jumps), and the leverage estimates. The right column provides corresponding results for the 30-year Treasury bond future data. In all cases, spot volatility is expressed (on the horizontal axis) in daily, percentage terms.

[4] Contrary to threshold bipower variation, for the classical bipower variation estimator of Barndorff-Nielsen and Shephard (2004), Assumption 3 is not satisfied. However, the conditions in Theorem 6 are still valid with $\alpha = \beta = 1/2$ (see BR, 2008).

[5] In the case of bonds, we also compute the average in the definition of $\widehat{\mu}_\sigma$ after trimming the (noisy) 10% upper and lower addends.

In the case of the S&P500 index the leverage estimates are reversed hump-shaped about roughly $(-0.5, -0.3)$ with the maximum value at about $\sigma = 0.5\%$. The nonlinearity is significant. It is economically indicative of time-varying (with volatility) correlations between shocks to returns and shocks to volatility, with the strongest negative correlation associated with either low ($\sigma = 0.1\%$) or high ($\sigma = 1.7\%$) volatility levels (see Fig. 2). Turning to bonds, the leverage dynamics are less striking. The overall variation is now only between 0 and -0.2. In this case, the leverage values tend to hover around -0.1.[6]

We of course emphasize that, in light of Theorem 5, the reported leverage estimates are consistent only in the absence of co-jumps (or if, in the presence of co-jumps, the jump sizes are independent and the mean of the return jumps is equal to zero, as sometimes assumed in the literature - see the Corollary to Theorem 5). In the presence of unrestricted co-jumps, the estimates contain a limiting bias whose magnitude may be negligible should the probability of co-jumps (or $\mathbf{E}[c_r c_\sigma]$) be small. We view our current results as being suggestive, but of course far from conclusive, and leave a more thorough empirical investigation, inclusive of potential bias corrections (as in Remark 9), when needed, for future work.

10 Conclusions

We adopt a flexible nonparametric specification in the family of discontinuous stochastic volatility models in order to provide a framework to better understand the nature of the correlation between return and volatility shocks. We show that kernel estimates of leverage effects have asymptotic sampling distributions which crucially depend on the features of objects that are fundamentally hard to pin down, namely the probability and size distribution of the individual and joint discontinuities in the return and volatility sample paths. We discuss the nature of this dependence and its implications, while providing tools for feasible identification of (potentially time-varying) leverage effects under mild parametric structures and weak recurrence assumptions. Our empirical work shows that, for stock index futures, stronger leverage effects are associated with low and high volatility regimes. This novel finding points to the importance of time-varying dynamics in the relation between shocks to stock returns and shocks to volatility and warrants further empirical analysis.

11 Proofs

We consider $\xi(x) = x$ for brevity and, when not differently indicated, $c_\sigma \sim \exp(\mu_\sigma)$. Alternative specifications for $\xi(\cdot)$ and c_σ may be treated similarly. The notation $\widetilde{\mathbf{K}}(A_i)$ denotes

$$\widetilde{\mathbf{K}}(A_i) = \frac{\sum_{i=1}^{n-1} \mathbf{K}\left(\frac{\sigma^2_{iT/n} - \sigma^2}{h_{n,T}}\right) A_i}{\Delta_{n,T} \sum_{i=1}^{n} \mathbf{K}\left(\frac{\sigma^2_{iT/n} - \sigma^2}{h_{n,T}}\right)}.$$

Lemma A.1. *Given a Borel measurable bounded function $g(\cdot)$, consider the quantity*

$$\Psi(\sigma^2)_k = \widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n} \left(\log p_{s-} - \log p_{iT/n}\right)^k \int_Z g(z) \bar{\nu}_\sigma(ds, dz)\right),$$

where $\bar{\nu}_\sigma$ is the compensated measure of J^σ. If

[6] The point estimate of $\widetilde{\mu}_\sigma$ is 0.2045 for stocks and 0.0688 for bonds.

$$\frac{v(T)}{h_{n,T}}\left(\Delta_{n,T}\log\frac{1}{\Delta_{n,T}}\right)^{1/2}\to 0,$$

we have

$$\Psi(\sigma^2)_0 = O_p\left(\sqrt{\frac{1}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right)$$

and

$$\Psi(\sigma^2)_1 = O_p\left(\sqrt{\frac{\Delta_{n,T}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right).$$

Moreover:

$$\sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\Psi(\sigma^2)_0 \Rightarrow \mathbf{N}\left(0,\mathbf{K}_2\lambda_\sigma(\sigma^2)\mathbf{E}[g^2]\right),$$

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\Psi(\sigma^2)_1 \Rightarrow \mathbf{N}\left(0,\frac{1}{2}\mathbf{K}_2\left[\lambda_r(\sigma^2)\mathbf{E}[c_r^2]+\sigma^2\right]\lambda_\sigma(\sigma^2)\mathbf{E}[g^2]\right).$$

Remark to Lemma A.1. Similar results hold if we replace $\bar{\nu}_\sigma$ with $\bar{\nu}_r$ (the compensated measure of J^r) or $\bar{\nu}_{r,\sigma}$ (the compensated measure of the contemporaneous jumps between r and σ^2) and, of course, if we replace $\log p_{s-} - \log p_{iT/n}$ with $\sigma^2_{s-} - \sigma^2_{iT/n}$.

Lemma A.2. *Consider the quantity*

$$\Phi(\sigma^2)_k = \widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\log p_{s-} - \log p_{iT/n}\right)^k \Lambda(\sigma^2_{s-})dW^\sigma_s\right).$$

If

$$\frac{v(T)}{h_{n,T}}\left(\Delta_{n,T}\log\frac{1}{\Delta_{n,T}}\right)^{1/2}\to 0$$

we have

$$\Phi(\sigma^2)_0 = O_p\left(\sqrt{\frac{1}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right)$$

and

$$\Phi(\sigma^2)_1 = O_p\left(\sqrt{\frac{\Delta_{n,T}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right).$$

Moreover:

$$\sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\Phi(\sigma^2)_0 \Rightarrow \mathbf{N}\left(0,\mathbf{K}_2\Lambda^2(\sigma^2)\right),$$

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\Phi(\sigma^2)_1 \Rightarrow \mathbf{N}\left(0,\frac{1}{2}\mathbf{K}_2\left[\lambda_r(\sigma^2)\mathbf{E}[c_r^2]+\sigma^2\right]\Lambda^2(\sigma^2)\right).$$

Remark to Lemma A.2. Similar results hold if we replace W^σ with W^r and, of course, if we replace $\log p_{s-} - \log p_{iT/n}$ with $\sigma^2_{s-} - \sigma^2_{iT/n}$.

Proof of Lemma A.1. We prove the lemma for $\Psi(\sigma^2)_1$. The case $\Psi(\sigma^2)_0$ follows analogously. Let T be fixed and define:

$$\begin{aligned}\sqrt{\frac{h_{n,T}}{\Delta_{n,T}}}\Psi^{num} &:= \frac{1}{\sqrt{h_{n,T}\Delta_{n,T}}}\sum_{i=1}^{n-1}\mathbf{K}\left(\frac{\sigma^2_{iT/n}-\sigma^2}{h_{n,T}}\right)\int_{iT/n}^{(i+1)T/n}\left(\log p_{s-} - \log p_{iT/n}\right)\int_Z g(z)\bar{\nu}_\sigma(ds,dz)\\ &:= \sum_{i=1}^{n-1}u_{iT/n,(i+1)T/n},\end{aligned}$$

where the $u_{iT/n,(i+1)T/n}s$ are square-integrable martingale difference sequences. We immediately have

$$\sum_{i=1}^{n-1}\mathbf{E}[u_{iT/n,(i+1)T/n}|\Im_{iT/n}] = 0$$

and, by virtue of Ito's Lemma on $\left(\log p_{s-}-\log p_{iT/n}\right)^2$:

$$\begin{aligned} & \sum_{i=1}^{n-1}\mathbf{E}[u^2_{iT/n,(i+1)T/n}|\Im_{iT/n}] \\ & \underset{\Delta_{n,T}\to 0}{\overset{p}{\to}} \frac{1}{2}\frac{1}{h_{n,T}}\int_0^T \mathbf{K}^2\left(\frac{\sigma^2_{s-}-\sigma^2}{h_{n,T}}\right)\left(\sigma^2_{s-}+\lambda_r(\sigma^2_{s-})\mathbf{E}[c_r^2]\right)\lambda_\sigma(\sigma^2_{s-})\mathbf{E}[g^2]ds \\ = & \widetilde{\mathbf{V}}_T. \end{aligned}$$

Now write

$$\begin{aligned} & \sum_{i=1}^{n-1}\mathbf{E}\left[u^2_{iT/n,(i+1)T/n}\mathbf{1}_{(|u_{iT/n,(i+1)T/n}|>\epsilon)}|\Im_{iT/n}\right] \\ = & \sum_{i=1}^{n-1}\mathbf{E}\left[u^2_{iT/n,(i+1)T/n}|\Im_{iT/n}\right]-\sum_{i=1}^{n-1}\mathbf{E}\left[u^2_{iT/n,(i+1)T/n}\mathbf{1}_{(|u_{iT/n,(i+1)T/n}|\leq\epsilon)}|\Im_{iT/n}\right] \\ = & \widetilde{\mathbf{V}}_T-\sum_{i=1}^{n-1}\mathbf{E}\left[u^2_{iT/n,(i+1)T/n}\mathbf{1}_{\left(\left|\mathbf{K}\left(\frac{\sigma^2_{iT/n}-\sigma^2}{h_{n,T}}\right)\int_{iT/n}^{(i+1)T/n}\left(\log p_{s-}-\log p_{iT/n}\right)\int_Z g(z)\bar{\nu}_\sigma(ds,dz)\right|\leq\epsilon\sqrt{h_{n,T}\Delta_{n,T}}\right)}|\Im_{iT/n}\right], \end{aligned} \tag{13}$$

but $\mathbf{K}\left(\frac{\sigma^2_{iT/n}-\sigma^2}{h_{n,T}}\right)\int_{iT/n}^{(i+1)T/n}\left(\log p_{s-}-\log p_{iT/n}\right)\int_Z g(z)\bar{\nu}_\sigma(ds,dz)=O_p(\Delta_{n,T})$. Hence, the indicator converges in probability to 1 and, given boundedness of $\widetilde{\mathbf{V}}_T$, Eq. (13) converges in probability to 0 (as $\Delta_{n,T}\to 0$). This is a conditional Lindeberg condition. Using Theorem VIII.3.33 in Jacod and Shiryaev (2002), we conclude that, for each T, $\sqrt{\frac{h_{n,T}}{\Delta_{n,T}}}\Psi^{num}\Rightarrow W\left(\widetilde{\mathbf{V}}_T\right)$ and W is an independent Brownian motion. This implies that

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\Psi(\sigma^2)_1=\frac{\sqrt{\frac{h_{n,T}}{\Delta_{n,T}}}\Psi^{num}}{\sqrt{\frac{\Delta_{n,T}}{h_{n,T}}\sum_{i=1}^n\mathbf{K}\left(\frac{\sigma^2_{iT/n}-\sigma^2}{h_{n,T}}\right)}}\Rightarrow W\left(\frac{\widetilde{\mathbf{V}}_T}{\frac{1}{h_{n,T}}\int_0^T\mathbf{K}\left(\frac{\sigma^2_{s-}-\sigma^2}{h_{n,T}}\right)}\right),$$

uniformly in T. By the Quotient limit theorem (see, e.g., Revuz and Yor, 1998, Theorem 3.12) we now have that, as $T\to\infty$ (with $h_{n,T}\to 0$),

$$\frac{\widetilde{\mathbf{V}}_T}{\frac{1}{h_{n,T}}\int_0^T\mathbf{K}\left(\frac{\sigma^2_{s-}-\sigma^2}{h_{n,T}}\right)}\overset{p}{\longrightarrow}\frac{1}{2}\mathbf{K}_2\left(\sigma^2+\lambda_r(\sigma^2)\right)\lambda_\sigma(\sigma^2)\mathbf{E}[g^2]$$

which, using Skorohod embedding arguments as in Theorem 4.1 in Van Zanten (2000), for example, gives the desired result. ■

Proof of Lemma A.2. The proof follows the same lines as that of Lemma A.1.

Proof of Theorem 1. Under the assumptions of the theorem, Bandi and Phillips (2003) prove that $\widehat{\vartheta}_2(\sigma^2)\to\vartheta_2(\sigma^2)=\Lambda^2(\sigma^2)$ with probability one and

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left(\widehat{\vartheta}_2(\sigma^2)-\Lambda^2(\sigma^2)-\Gamma_{\vartheta_2}(\sigma^2)\right)\Rightarrow\mathbf{N}\left(0,2\mathbf{K}_2\Lambda^4(\sigma^2)\right). \tag{14}$$

Now consider $\widehat{\vartheta}_{1,1}$. Itô's lemma gives:

$$\begin{aligned} & \widehat{\vartheta}_{1,1} \\ = & \widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\sigma^2_{s-}-\sigma^2_{iT/n}\right)\mu_{s-}ds\right)+\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\sigma^2_{s-}-\sigma^2_{iT/n}\right)\sigma_{s-}dW_s^r\right) \\ & +\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\log(p_{s-})-\log\left(p_{iT/n}\right)\right)m_{s-}ds\right)+\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\log(p_{s-})-\log\left(p_{iT/n}\right)\right)\Lambda(\sigma^2_{s-})dW_s^\sigma\right) \\ & +\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\rho(\sigma^2_{s-})\sigma_{s-}\Lambda(\sigma^2_{s-})ds\right) \\ : & =\widehat{\vartheta}^c_{1,1}. \end{aligned} \tag{15}$$

Using Lemma A.2. we get:

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left(\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\sigma_{s-}^2-\sigma_{iT/n}^2\right)\sigma_{s-}dW_s^r\right)\right)\Rightarrow\mathbf{N}\left(0,\frac{1}{2}\mathbf{K}_2\Lambda^2(\sigma^2)\sigma^2\right)$$

and

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left(\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\log(p_{s-})-\log\left(p_{iT/n}\right)\right)\Lambda(\sigma_{s-}^2)dW_s^\sigma\right)\right)\Rightarrow\mathbf{N}\left(0,\frac{1}{2}\mathbf{K}_2\Lambda^2(\sigma^2)\sigma^2\right).$$

The asymptotic covariance between

$$\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\sigma_{s-}^2-\sigma_{iT/n}^2\right)\sigma_{s-}dW_s^r\right)$$

and

$$\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\log(p_{s-})-\log\left(p_{iT/n}\right)\right)\Lambda(\sigma_{s-}^2)dW_s^\sigma\right)$$

is equal to

$$\frac{1}{2}\mathbf{K}_2\Lambda^2(\sigma^2)\sigma^2\rho^2(\sigma^2)\left(\frac{\Delta_{n,T}}{\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)h_{n,T}}\right).$$

Hence,

$$\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left(\widehat{\vartheta}_{1,1}-\rho(\sigma^2)\sigma\Lambda(\sigma^2)-\Gamma_{\vartheta_{1,1}}(\sigma^2)\right)\Rightarrow\mathbf{N}\left(0,\mathbf{K}_2\Lambda^2(\sigma^2)\sigma^2(1+\rho^2(\sigma^2))\right)$$

since, by the Quotient limit theorem,

$$\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\rho(\sigma_{s-}^2)\sigma_{s-}\Lambda(\sigma_{s-}^2)ds\right)-\vartheta_{1,1}(\sigma^2)=\Gamma_{\vartheta_{1,1}}(\sigma^2)+o_p(h_{n,T}^2).$$

In the same way,

$$\begin{aligned}&\widehat{\vartheta}_2(\sigma^2)\\=\;&\widetilde{\mathbf{K}}\left(2\int_{iT/n}^{(i+1)T/n}\left(\sigma_{s-}^2-\sigma_{iT/n}^2\right)m_{s-}ds\right)+\widetilde{\mathbf{K}}\left(2\int_{iT/n}^{(i+1)T/n}\left(\sigma_{s-}^2-\sigma_{iT/n}^2\right)\Lambda(\sigma_{s-}^2)dW_s^\sigma\right)\\&+\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\Lambda^2(\sigma^2)ds\right).\end{aligned}$$

Hence, the asymptotic covariance between $\widehat{\vartheta}_{1,1}$ and $\widehat{\vartheta}_2$ is given by:

$$Asycov(\widehat{\vartheta}_{1,1},\widehat{\vartheta}_2)=\mathbf{K}_2\frac{\Delta_{n,T}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\left(2\Lambda^3(\sigma^2)\sigma\rho(\sigma^2)\right).$$

Finally, by the delta method:

$$\begin{aligned}&\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left\{\widehat{\rho}(\sigma^2)-\rho(\sigma^2)\right\}\\\overset{asy}{\approx}\;&\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\left\{\frac{1}{\sigma\Lambda(\sigma^2)}\left\{\widehat{\vartheta}_{1,1}(\sigma^2)-\vartheta_{1,1}(\sigma^2)\right\}-\frac{\vartheta_{1,1}}{2\sigma\Lambda^3(\sigma^2)}\left\{\widehat{\vartheta}_2(\sigma^2)-\vartheta_2(\sigma^2)\right\}\right\}.\end{aligned}\tag{16}$$

Hence,

$$
\begin{aligned}
& Asyvar\left(\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\widehat{\rho}(\sigma^2)\right) \\
= & \left[\frac{1}{\sigma^2\Lambda^2(\sigma)}Asyvar\left(\widehat{\vartheta}_{1,1}(\sigma^2)\right) - \frac{\vartheta_{1,1}}{\sigma^2\Lambda^4(\sigma^2)}Asycov\left(\widehat{\vartheta}_{1,1}(\sigma^2),\widehat{\vartheta}_2(\sigma^2)\right) + \frac{\vartheta_{1,1}^2}{4\sigma^2\Lambda^6(\sigma^2)}Asyvar\left(\widehat{\vartheta}_2(\sigma^2)\right)\right] \\
= & \frac{1}{\sigma^2\Lambda^2(\sigma)}\left[Asyvar\left(\widehat{\vartheta}_{1,1}(\sigma^2)\right) - \frac{\sigma\rho(\sigma^2)}{\Lambda(\sigma^2)}Asycov\left(\widehat{\vartheta}_{1,1}(\sigma^2),\widehat{\vartheta}_2(\sigma^2)\right) + \left(\frac{\sigma\rho(\sigma^2)}{2\Lambda(\sigma^2)}\right)^2 Asyvar\left(\widehat{\vartheta}_2(\sigma^2)\right)\right] \\
= & \frac{1}{\sigma^2\Lambda^2(\sigma)}\left[\Lambda^2(\sigma^2)\sigma^2(1+\rho^2(\sigma^2)) - \frac{\sigma\rho(\sigma^2)}{\Lambda(\sigma^2)}(2\Lambda^3(\sigma^2)\sigma\rho(\sigma^2)) + \left(\frac{\sigma\rho(\sigma^2)}{2\Lambda(\sigma^2)}\right)^2 2\Lambda^4(\sigma^2)\right] \\
= & \mathbf{K}_2\left(1-\frac{1}{2}\rho^2(\sigma^2)\right).
\end{aligned}
$$

As for the asymptotic bias, clearly

$$
\widetilde{\Gamma}_\rho(\sigma^2) = \frac{1}{\sigma\sqrt{\vartheta_2(\sigma^2)}}\Gamma_{\vartheta_{1,1}}(\sigma^2) - \frac{\vartheta_{1,1}(\sigma^2)}{2\sigma\sqrt{\vartheta_2^3(\sigma^2)}}\Gamma_{\vartheta_2}(\sigma^2).
$$

■

Proof of Theorem 2. Since $J^\sigma = 0$, the speed of convergence and asymptotic distribution of $\widehat{\vartheta}_2$ do not change. On the other hand, Ito's lemma now implies that

$$
\widehat{\vartheta}_{1,1} = \widehat{\vartheta}_{1,1}^c + \widetilde{\mathbf{K}}\left(\sum_{s\in[iT/n,(i+1)T/n[}\left[\Delta\log(p_s)(\sigma_{s-}^2 - \sigma_{iT/n}^2)\right]\right),
$$

where $\widehat{\vartheta}_{1,1}^c$ is defined in Eq. (15). The extra term does not affect consistency (nor does it contribute to the asymptotic bias) and does not change the speed of convergence. However, the limiting variance of $\widehat{\vartheta}_{1,1}$ (and $\widehat{\rho}(\sigma^2)$) does change. Notice, in fact, that by Lemma A.1.,

$$
\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\widetilde{\mathbf{K}}\left(\int_{iT/n}^{(i+1)T/n}\left(\sigma_{s-}^2 - \sigma_{iT/n}^2\right)\int_Z c_r\bar{\nu}_r(ds,dz)\right) \Rightarrow \mathbf{N}\left(0,\frac{1}{2}\mathbf{K}_2\Lambda^2(\sigma^2)\lambda_r(\sigma^2)\mathbf{E}[c_r^2]\right).
$$

This implies that, by the delta method in Eq. (16):

$$
Asyvar\left(\sqrt{\frac{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}{\Delta_{n,T}}}\widehat{\rho}(\sigma^2)\right) = \mathbf{K}_2\left[\left(1-\frac{1}{2}\rho^2(\sigma^2)\right) + \frac{1}{2}\frac{\lambda_r(\sigma^2)\mathbf{E}[c_r^2]}{\sigma^2}\right].
$$

■

Proof of Theorem 3. BR (2008) show that $\widehat{\Lambda}^2(\sigma^2) \xrightarrow{p} \Lambda^2(\sigma^2)$ and

$$
\sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\left(\widehat{\Lambda}^2(\sigma^2) - \Lambda^2(\sigma^2)\right) \Rightarrow \mathbf{N}\left(0,\mathbf{K}_2\lambda_\sigma(\sigma^2)\mathbf{E}\left(\left((c_\sigma)^2 - \frac{1}{12\mu_\sigma^2}(c_\sigma)^4\right)^2\right)\right). \tag{17}
$$

Ito's lemma now implies

$$
\widehat{\vartheta}_{1,1} = \widehat{\vartheta}_{1,1}^c + \widetilde{\mathbf{K}}\left(\sum_{s\in[iT/n,(i+1)T/n[}\left[\left(\log(p_{s-}) - \log p_{iT/n}\right)\Delta\sigma_s^2\right]\right)
$$

which gives, as earlier, $\widehat{\vartheta}_{1,1} = O_p\left(\sqrt{\frac{\Delta_{n,T}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right)$. Hence, the rate convergence of $\widehat{\Lambda}^2(\sigma^2)$ is slower and therefore dominating. In other words, from Eq. (16), the limiting variance of $\widehat{\rho}(\sigma^2)$ solely depends on

$$
\sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\frac{\vartheta_{1,1}}{2\sigma\Lambda^3(\sigma^2)}\left\{\widehat{\Lambda}^2(\sigma^2) - \Lambda^2(\sigma^2)\right\} = \sqrt{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\frac{\rho(\sigma^2)}{2\Lambda^2(\sigma^2)}\left\{\widehat{\Lambda}^2(\sigma^2) - \Lambda^2(\sigma^2)\right\},
$$

thereby giving the stated result. ■

Proof of Theorem 4. See BR (2008), Theorem 7. ■

Proof of Theorem 5. The result in Eq. (17) still holds. Using Itô's lemma in the presence of contemporaneous jumps gives:

$$\widehat{\vartheta}_{1,1} = \widehat{\vartheta}_{1,1}^{c} + \widetilde{\mathbf{K}}\left(\sum_{s\in[iT/n,(i+1)T/n[} \left[\Delta\log(p_s)\Delta\sigma_s^2 + \left(\log(p_{s-}) - \log p_{iT/n}\right)\Delta\sigma_s^2 + \Delta\log(p_s)(\sigma_{s-}^2 - \sigma_{iT/n}^2)\right]\right).$$

From Lemma A.1. we now obtain that the contemporaneous jump part is $O_p\left(\sqrt{\frac{1}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right)$ while all other terms are $O_p\left(\sqrt{\frac{\Delta_{n,T}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}}\right)$. Thus, immediately,

$$Asyvar(\widehat{\vartheta}_{1,1}(\sigma^2)) = \mathbf{K}_2\left(\frac{1}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\right)\lambda_{r,\sigma}(\sigma^2)\mathbf{E}[c_r^2 c_\sigma^2].$$

Similarly, since the lower-order term in $\widehat{\vartheta}_2(\sigma^2)$ is

$$\widetilde{\mathbf{K}}\left(\sum_{s\in[iT/n,(i+1)T/n[}(\Delta\sigma_s^2)^2\right)$$

and in $\widehat{\vartheta}_4(\sigma^2)$ is

$$\widetilde{\mathbf{K}}\left(\sum_{s\in[iT/n,(i+1)T/n[}(\Delta\sigma_s^2)^4\right),$$

we have

$$\begin{aligned} Asycov(\widehat{\vartheta}_{1,1}(\sigma^2),\widehat{\Lambda}^2(\sigma^2)) &= Asycov\left(\widehat{\vartheta}_{1,1}(\sigma^2),\widehat{\vartheta}_2(\sigma^2) - \frac{\widehat{\vartheta}_4(\sigma^2)}{12\widehat{\mu}_\sigma^2}\right) \\ &= \mathbf{K}_2\left(\frac{1}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\right)\lambda_{r,\sigma}(\sigma^2)\left(\mathbf{E}[c_r c_\sigma^3] - \frac{\mathbf{E}[c_r c_\sigma^5]}{12\mu_\sigma^2}\right). \end{aligned}$$

Due to the presence of co-jumps, compensation of the object $\sum_{s\in[iT/n,(i+1)T/n[}\left[\Delta\log(p_s)\Delta\sigma_s^2\right]$ requires subtraction of $\int_{iT/n}^{(i+1)T/n}\lambda_{r,\sigma}(\sigma_{s-}^2)\mathbf{E}[c_r c_\sigma]ds$. This term contributes to the probability limit of $\widehat{\vartheta}_{1,1}(\sigma^2)$ which now is $\rho(\sigma^2)\sigma\Lambda(\sigma^2) + \lambda_{r,\sigma}(\sigma^2)\mathbf{E}[c_r c_\sigma]$. Naturally, the term also changes the probability limit of $\widehat{\rho}(\sigma^2)$ giving $\widehat{\rho}(\sigma^2) \xrightarrow{p} \rho(\sigma^2) + \frac{\lambda_{r,\sigma}(\sigma^2)\mathbf{E}[c_r c_\sigma]}{\sigma\Lambda(\sigma^2)}$. Finally, the delta method yields:

$$\begin{aligned} & Asycov(\widehat{\rho}(\sigma^2)) \\ = \; & \frac{\mathbf{K}_2\frac{\lambda_{r,\sigma}(\sigma^2)}{\sigma^2\Lambda^2(\sigma^2)}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\left(\mathbf{E}[c_r^2c_\sigma^2] - \frac{\vartheta_{1,1}(\sigma^2)}{\Lambda^2(\sigma^2)}2\mathbf{E}\left[c_r c_\sigma\left(c_\sigma^2 - \frac{c_\sigma^4}{12\mu_\sigma^2}\right)\right] + \left(\frac{\vartheta_{1,1}(\sigma^2)}{2\Lambda^2(\sigma^2)}\right)^2\mathbf{E}\left[\left(c_\sigma^2 - \frac{c_\sigma^4}{12\mu_\sigma^2}\right)^2\right]\right) \\ = \; & \frac{\mathbf{K}_2\frac{\lambda_{r,\sigma}(\sigma^2)}{\sigma^2\Lambda^2(\sigma^2)}}{h_{n,T}\widehat{\overline{L}}_{\sigma^2}(T,\sigma^2)}\mathbf{E}\left[\left(c_r c_\sigma - \frac{\vartheta_{1,1}(\sigma^2)}{2\Lambda^2(\sigma^2)}\left(c_\sigma^2 - \frac{c_\sigma^4}{12\mu_\sigma^2}\right)\right)^2\right], \end{aligned}$$

thereby leading to the stated result.■

Proof of Theorem 6. Denote by $\widetilde{\vartheta}_{1,1}(\sigma^2)$ and by $\widetilde{\vartheta}_k(\sigma^2)$ with $k = 2,3,4$, the moment estimators constructed using estimated spot variance in place of the true, unknown spot variance. Using the method of proof of Theorem 2 of BR (2008), we can show that

$$\widetilde{\vartheta}_{1,1}(\sigma^2) = \widehat{\vartheta}_{1,1}(\sigma^2) + O_p\left(\frac{Tv(T)^{-1}}{\Delta_{n,T}h_{n,T}k^\alpha\phi_{n,T}^\beta} + \frac{Tv(T)^{-1}}{\Delta_{n,T}h_{n,T}}\left(\phi_{n,T}\log\left(\frac{1}{\phi_{n,T}}\right)\right)^{1/2}\right)$$

and

$$\widetilde{\vartheta}_k(\sigma^2) = \widehat{\vartheta}_k(\sigma^2) + O_p\left(\frac{Tv(T)^{-1}}{\Delta_{n,T}h_{n,T}k^\alpha\phi_{n,T}^\beta} + \frac{Tv(T)^{-1}}{\Delta_{n,T}h_{n,T}}\left(\phi_{n,T}\log\left(\frac{1}{\phi_{n,T}}\right)\right)^{1/2}\right)$$

with $k = 2, 3, 4$. Therefore,

$$\frac{Tv(T)^{-1}}{\Delta_{n,T} h_{n,T} k^{\alpha} \phi_{n,T}^{\beta}} + \frac{Tv(T)^{-1}}{\Delta_{n,T} h_{n,T}} \left(\phi_{n,T} \log \left(\frac{1}{\phi_{n,T}} \right) \right)^{1/2} \to 0$$

is sufficient for consistency of the feasible leverage estimator. For weak convergence, this condition ought to be strenghtened. The relevant condition is

$$\frac{Tv(T)^{-1/2}}{\Delta_{n,T}^{3/2} h_{n,T}^{1/2} k^{\alpha} \phi_{n,T}^{\beta}} + \frac{Tv(T)^{-1/2}}{\Delta_{n,T}^{3/2} h_{n,T}^{1/2}} \left(\phi_{n,T} \log \left(\frac{1}{\phi_{n,T}} \right) \right)^{1/2} \to 0,$$

for Theorem 1 and 2 (i.e., when the convergence rate is $\sqrt{\frac{h_{n,T} v(T)}{\Delta_{n,T}}}$). It is

$$\frac{Tv(T)^{-1/2}}{\Delta_{n,T} h_{n,T}^{1/2} k^{\alpha} \phi_{n,T}^{\beta}} + \frac{Tv(T)^{-1/2}}{\Delta_{n,T} h_{n,T}^{1/2}} \left(\phi_{n,T} \log \left(\frac{1}{\phi_{n,T}} \right) \right)^{1/2} \to 0,$$

for Theorems 3, 4, and 5 (i.e., when the convergence rate is $\sqrt{h_{n,T} v(T)}$). ■

References

[1] Andersen, T.G., L. Benzoni, and J. Lund (2002). Towards an empirical foundation for continuous-time equity return models. *Journal of Finance* 57, 1047-1091.

[2] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica* 71, 579-625.

[3] Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance* 52, 2003-2049.

[4] Bandi, F.M. and T. Nguyen (2003). On the functional estimation of jump-diffusion models. *Journal of Econometrics* 116, 293-328.

[5] Bandi, F.M. and P.C.B. Phillips (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica* 71, 241-283.

[6] Bandi, F.M. and P.C.B. Phillips (2009). Nonstationary continuous-time models. In *Handbook of Financial Econometrics* (L.P. Hansen and Y. Ait-Sahalia, Eds.) Elsevier Science, New York, forthcoming.

[7] Bandi, F. M. and R. Reno' (2008). Nonparametric Stochastic Volatility. Available at SSRN: http://ssrn.com/abstract=1158438.

[8] Barndorff-Nielsen, O.E., P. Hansen, A. Lunde and N. Shephard (2009). Designing realized kernels to measure ex-post variation of equity prices in the presence of noise. *Econometrica*, forthcoming.

[9] Barndorff-Nielsen, O.E., and N. Shephard (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.

[10] Barndorff-Nielsen, O.E., and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2, 1-37.

[11] Black, F. (1976). Studies of stock market volatility changes. Proceedings of the American Statistical Association, Business and Economic Statistics section, 177-181.

[12] Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes.* Chapman & Hall.

[13] Corsi, F., D. Pirino, and R. Renò (2008). Volatility forecasting: The jumps do matter. *Working paper.*

[14] Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343-1376.

[15] Engle, R. and V. Ng (1993). Measuring and testing the impact of news in volatility. *Journal of Finance* 43, 1749-1778.

[16] Eraker, B., M. Johannes, and N. Polson (2003). The impact of jumps in volatility and returns. *Journal of Finance* 12, 69-1300.

[17] Glosten, L.R., R. Jagannathan, and D. Runkle (1993). Relationship between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779-1802.

[18] Gobbi, F. and C. Mancini (2008). Identifying the covariation between the diffusion parts and the co-jumps given discrete observations. *Working paper.*

[19] Harvey, A.C. and N. Shephard (1996). The estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics* 14, 429-434.

[20] Jacod, J. and A.N. Shiryaev (2002). Limit theorems for stochastic processes. Springer.

[21] Jacquier, E., N. Polson, and P. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* 12, 371-389.

[22] Jacquier, E., N. Polson, and P. Rossi (2004). Bayesian analysis of stochastic volatility models with fat tails and correlated errors. *Journal of Econometrics* 122, 185-212.

[23] Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance,* 227-260.

[24] Kristensen, D. (2008). Nonparametric filtering of the realized spot volatility: a kernel-based approach. *Econometric Theory,* forthcoming.

[25] Malliavin, P. and M.E. Mancino (2008). A Fourier transform method for nonparametric estimation of volatility. *Annals of Statistics,* forthcoming.

[26] Mancini, C. (2007). Nonparametric threshold estimation for models with stochastic diffusion coefficients and jumps. *Scandinavian Journal of Statistics*, forthcoming.

[27] Revuz, D. and M. Yor (1998). *Continuous Martingales and Brownian Motion.* Springer-Verlag.

[28] Van Zanten, J.H. (2000). The stable central limit theorem for local martingales with bounded jumps via Skorohod embedding. *Working paper.*

[29] Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics* 127, 165-178.

[30] Zhang, L., P. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394-1411.

S&P 500 **30y Treasury Bond**

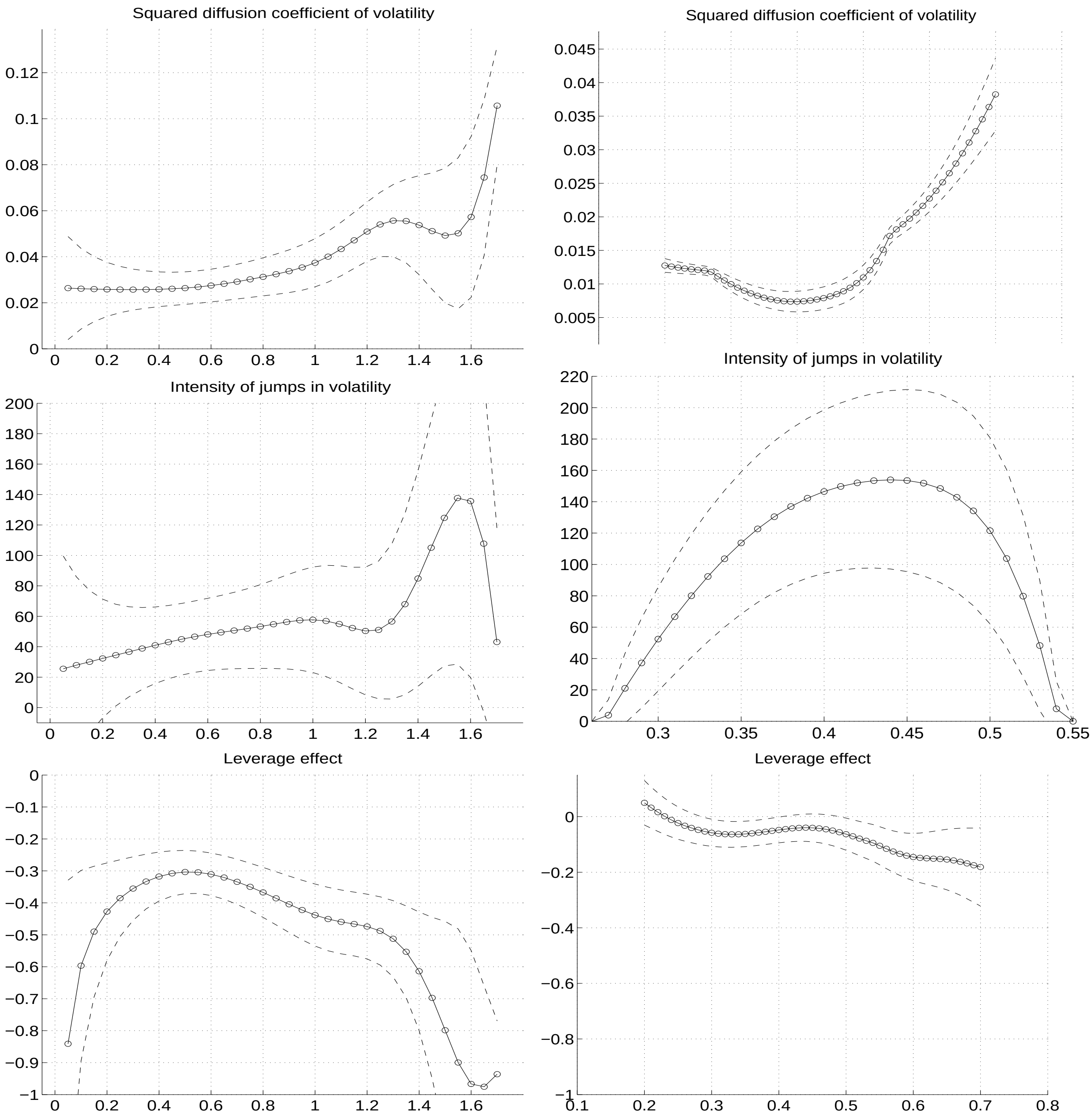
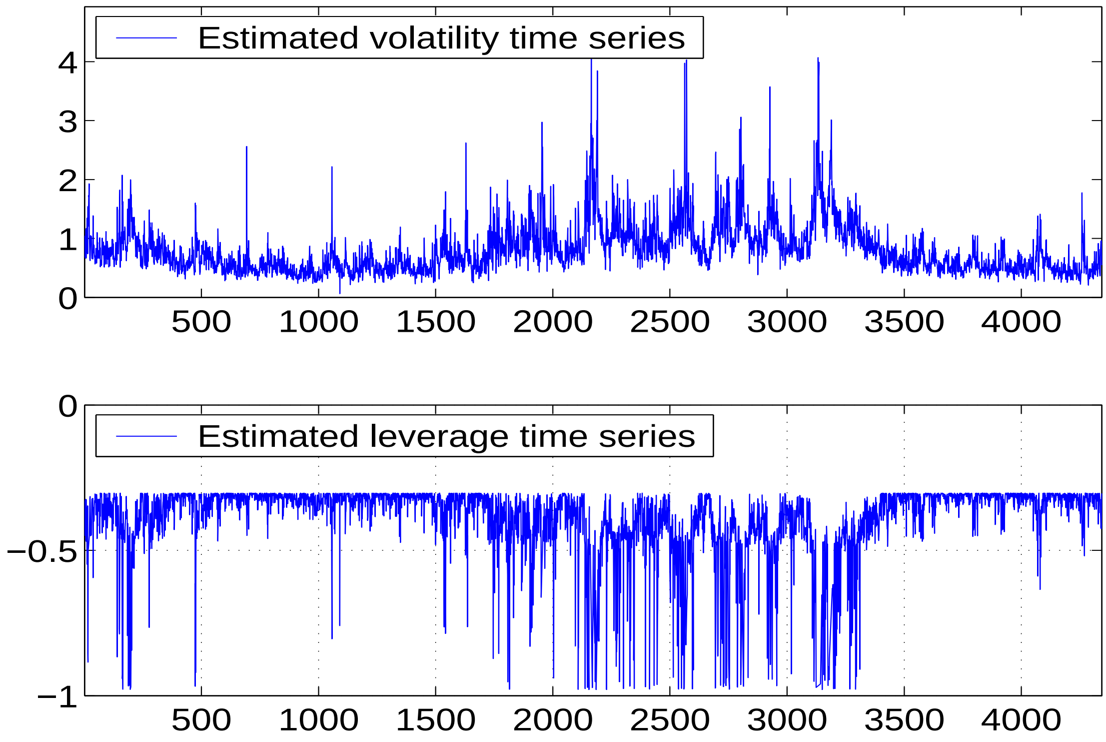


Figure 1: Reports estimates of the function $\lambda(\sigma)$ (top), $\Lambda^2(\sigma)$ (medium) and $\rho(\sigma)$ (bottom), on the S&P 500 time series (left column) and the 30 years Treasury Bond time series (right column), obtained with the estimators described in the paper. The quantities are expressed on daily basis in percentage form, except the intensity of jumps which is yearly.

S&P500



30y Treasury Bond

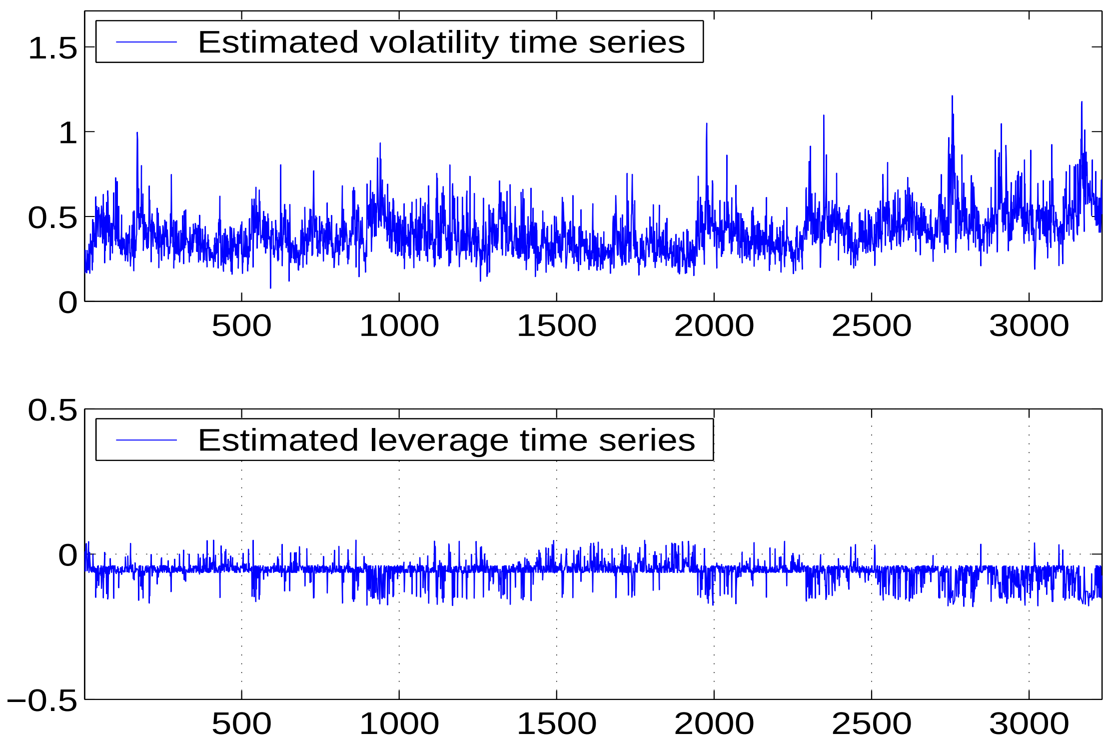


Figure 2: Reports estimated time series of spot volatility and of estimated leverage $\widehat{\rho}(\widehat{\sigma}_{iT/n})$ for the considered samples.