

# Estimation of the Continuous and Discontinuous Leverage Effects

Yacine Aït-Sahalia\*

Department of Economics, Princeton University

Jianqing Fan<sup>†</sup>

Department of Operations Research and Financial Engineering, Princeton University

Roger J. A. Laeven<sup>‡</sup>

Department of Economics, University of Amsterdam

Christina Dan Wang<sup>§</sup>

Department of Statistics, Columbia University

and

Xiye Yang<sup>¶</sup>

Department of Economics, Rutgers University

August 23, 2016

## Abstract

This paper examines the leverage effect, or the generally negative covariation between asset returns and their changes in volatility, under a general setup that allows the log-price and volatility processes to be Itô semimartingales. We decompose the leverage effect into continuous and discontinuous parts and develop statistical methods to estimate them. We establish the asymptotic properties of these estimators. We also extend our methods and results (for the continuous leverage) to the situation where there is market microstructure noise in the observed returns. We show in Monte Carlo simulations that our estimators have good finite sample performance. When applying our methods to real data, our empirical results provide convincing evidence of the presence of the two leverage effects, especially the discontinuous one.

*Keywords:* Spot volatility, integrated volatility, jumps, co-jumps, high-frequency data, market microstructure noise

---

\*Supported in part by the NSF under grant SES-0850533.

†Supported in part by the NSF under grants DMS-1206464 and DMS-1406266.

‡Supported in part by the NWO under grant VIDI-2009.

§Supported in part by the NIH under grant R01-GM100474 and by a grant from the Bendheim Center for Finance.

¶Corresponding author. Email: [xiyeyang@econ.rutgers.edu](mailto:xiyeyang@econ.rutgers.edu)

# 1 Introduction

The “leverage effect” in financial markets refers to the phenomenon that stock returns and their volatility changes generally exhibit an inverse relation: when stock prices rise, volatilities tend to drop, and vice versa (see [Black \(1976\)](#)). Using absolute returns sampled at the five-minutes frequency as a simple volatility proxy, [Bollerslev et al. \(2006\)](#) found prolonged negative correlation between volatility and current and lagged stock returns.

In this paper, we develop estimators to quantify the leverage effect in a general continuous-time framework, allowing for general Itô semimartingale dynamics of the log-price and volatility processes, with a drift term, a continuous Brownian component and a discontinuous jump component. This represents an essential step towards fully discriminating the relation between stock returns and changes in volatility using high-frequency data: in our general framework, the leverage effect can take more sophisticated forms and exhibit richer dynamic properties than in more restricted models, such as continuous diffusion models with stochastic volatility.

The first question we address in this paper is how to statistically define the leverage effect in a continuous-time semimartingale model. We categorize the leverage effect according to the correlated stochastic components and define: (1) the continuous leverage effect (CLE) as the quadratic covariation between the continuous parts of the log-price and volatility processes; (2) the discontinuous leverage effect (DLE) as the quadratic covariation between their discontinuous parts; and (3) the total leverage effect (TLE) as the sum of the two. Next, we propose estimators for the CLE and DLE and establish their asymptotic properties by deriving the associated consistency results and central limit theorems. Due to the latency of the volatility estimators, the mathematical details required to obtain these results are involved and new to the literature.

Technically, our proposed estimators involve the asymptotic properties of (globally) odd functionals<sup>1</sup> of Itô semimartingales’ increments, required to deal with the inherent latency of volatility

---

<sup>1</sup>To be specific, we use functions such as  $f(x_1, x_2) = x_1 x_2^2$  and their truncated versions. Observe that this

and complementary to existing results in the literature. While at first sight, the study of odd functionals may not seem useful—for example, the expectation of odd powers of Brownian increments is zero, implying that odd functionals applied to it may not converge to a non-trivial limit—we show that after appropriate arrangement, the sum of certain odd functionals of the observed return process does converge to something very meaningful, namely, the continuous (discontinuous or total) leverage effect. Moreover, it is worth pointing out that, in our estimators, the number of successive increments in each summand of the functionals goes to infinity, as the time step goes to 0. In view of this feature, this paper fits into the literature of “pre-averaging” method (e.g., [Jacod et al. \(2009\)](#), [Jacod et al. \(2010\)](#) and Chapters 12 and 16 in [Jacod and Protter \(2011\)](#)), and those papers using spot volatility estimates as intermediate variables (e.g. [Jacod and Todorov \(2010\)](#), [Jacod and Rosenbaum \(2013\)](#), [Li et al. \(2016\)](#), [Vetter \(2015\)](#) and [Wang and Mykland \(2014\)](#), among others).

In most of the previous literature<sup>2</sup>, as a direct extension to the discrete-time leverage effect, the CLE takes the form of a constant, namely the *leverage parameter*. In recent years, the number of papers considering time-varying leverage effects is growing. To name a few, [Bandi and Renò \(2012\)](#) allow the leverage effect to be a function of the (stochastic) state of the firm (thus is time-varying), which is then summarized by the spot variance or spot volatility. In this paper, we impose no such structural assumption and allow the CLE to be a very general stochastic process, which may feature additional source(s) of randomness other than those given by log-prices and volatilities. Another example is [Wang and Mykland \(2014\)](#), who define the leverage effect in agreement with the 

---

function is odd with respect to  $x_1$ , but even in the second argument  $x_2$ . However, if we treat  $x = (x_1, x_2)$  as a whole, we have  $f(-x) = -f(x)$ . The use of odd functions here is due to the fact that the volatility process is latent, hence needs to be estimated from returns. Otherwise, if both the return and volatility series were observable, then our problem would not be essentially different from estimating quadratic covariation between two (or more) asset prices.

<sup>2</sup>For example, extensive literatures in finance and econometrics are based on the model proposed by [Heston \(1993\)](#), including [Aït-Sahalia et al. \(2013\)](#), [Vetter \(2012\)](#) and [Yu \(2005\)](#).

CLE and provide non-parametric estimators, with a particular interest in eliminating the impact from market microstructure noise. However, they assume both return and volatility processes to be continuous, whereas we allow both of them to have discontinuous parts, thus introducing a second potential source for the leverage effect, the DLE, which turns out to be particularly relevant empirically. Furthermore, [Kalnina and Xiu \(2016\)](#) define the leverage effect similarly, yet essentially differently (see [Remark 1](#)), through a quadratic covariation, and replace the volatility process by its functional (see [\(2.6\)](#)). They compare one estimator using only price data with another one that also includes VIX data, based on the assumption that VIX can be viewed as a functional of the volatility process. With this assumption, both of the processes of interest are directly observable, hence the estimation becomes relatively easy. While this is an appealing idea, we do not think VIX is a perfect approximation of a functional of the volatility process. It is contaminated by various noises, including artificial errors (see [Andersen et al. \(2015\)](#)). Instead of using VIX, we estimate the spot volatility process using price data.<sup>3</sup> In view of the differences in definitions, the econometric and asymptotic properties of our CLE estimator are very different from those of the two CLE estimators in [Kalnina and Xiu \(2016\)](#). Finally, among others, [Veraart and Veraart \(2012\)](#) extend the Heston model to allow for a time-varying feature, and [Curato \(2015\)](#) and [Curato and Sanfelici \(2015\)](#) employ a Fourier transform based method to estimate a stochastic continuous leverage effect.

The DLE stems from the potentially important co-jumping of log-prices and volatilities. Parametric models which allow for co-jumps have been proposed (see, e.g., [Eraker et al. \(2003\)](#)), although without focus on or an attempt at measuring the DLE. Also, their jump process is assumed

---

<sup>3</sup>This problem has also been studied by, for example, [Fan and Wang \(2008\)](#) and [Alvarez et al. \(2012\)](#). Moreover, [Jacod and Todorov \(2010\)](#) and [Li and Xiu \(2016\)](#) included the estimation of spot volatility as an intermediate step to their main goals. Just like the last two cited papers, the main target of this paper is not the estimation of the spot volatility process either, and nontrivial results follow right after we obtain the estimator of the spot volatility process.

to have finite activity and by nature its parametric form is restrictive. [Bandi and Renò \(2012\)](#) do mention the possibility of “co-jump leverage”, which is the conditional expectation of our DLE in a simpler model than the one considered here. However they do not attempt to provide an estimator of the co-jump leverage, which is among the main contributions of the present paper. The most closely related work in this respect is [Jacod and Todorov \(2010\)](#), but differences exist. These authors propose general procedures to test whether price and volatility jump together. Their initial definition of the general testing functional of the price and volatility processes includes DLE as a special case. However, certain technical conditions in the central limit theorems in the cited paper exclude the functional form we used to define DLE (see the discussion after [Theorem 4](#)). We have bridged this gap. The resulting testing procedure is different from those in [Jacod and Todorov \(2010\)](#).

Some of the statistical challenges arising in leverage effect estimation with high-frequency data are demonstrated in a specific example in [Aït-Sahalia et al. \(2013\)](#), who analyze the ability to measure the (constant) CLE in high-frequency data using methods that replace the unobservable volatility by a simple proxy. Under the well-known Heston model, they show that, at high-frequency and over short horizons, the estimated leverage parameter is close to zero instead of an expected strongly negative value. At longer horizons, the leverage effect is present, especially when using option-implied volatilities. They document that the source of the puzzle—why one may be unable to detect the leverage effect at high-frequency—can be traced back to the use of a simple realized volatility plug-in: it shows that one needs to be careful about the proxy employed in place of the unobservable volatility path. By contrast, our estimators for both the CLE and DLE are asymptotically unbiased, since our definitions are based on quadratic covariation only, the estimators do not involve squared estimated volatility changes.

In Monte Carlo simulations, we analyze the finite sample performance of our estimators and find that our limit theorems already provide good approximations at the pre-limiting level, confirming

the practical applicability of our estimators. As a by-product of potential interest in its own right, we also study the performance of leverage parameter estimators constructed not from variance and covariance (the usual definition) but from quadratic variation and covariation. We find that the bias in the latter<sup>4</sup> is much smaller than the bias in the former. This suggests that it is preferable to use quadratic variation and covariation in this setting.

We also extend our methods and theoretical as well as simulation results (for the CLE only, since it is far more challenging in the case of DLE; see the discussion at the end of Section 6) to the situation in which there is market microstructure noise in the observed returns. Upon applying our methods to real high-frequency equity index data, our empirical results provide convincing evidence of the presence of the two leverage effects, CLE and DLE, with the DLE being more pronounced than the CLE. To be more specific, we further decompose the DLE into two parts: one associated with positive price jumps, the other associated with negative price jumps. Their averaged absolute values both appear to be larger than that of the CLE. We also find that volatility jumps are more likely to coincide with negative price jumps than positive ones.

The remainder of the paper is organized as follows. In Section 2, we introduce the model setup, the assumptions and the definitions of the two leverage effects. In Section 3, we introduce the adopted blocking scheme and the estimators we propose. In Section 4, we present the Law of Large Numbers (LLN) results for our estimators and in Section 5 we provide the corresponding Central Limit Theorems (CLTs). In Section 6, we analyze the impact of market microstructure noise. We demonstrate the finite sample performance of our estimators by Monte Carlo simulations in Section 7, where we also compare two different estimators of the leverage parameter. Section 8 applies our estimators to real financial data. All proofs are deferred to the Appendix.

---

<sup>4</sup>Different from CLE and DLE estimation, without correction, a bias is still present in the case of leverage parameter estimation, because the quadratic variation of estimated spot volatility involves squared estimated volatility changes.

## 2 Model Setup and Definitions

### 2.1 Setting and Assumptions

We start with a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  on which various processes are defined, including a log-price and volatility process,  $(X_t)_{t \geq 0}$  and  $(\sigma_t)_{t \geq 0}$ , satisfying:

ASSUMPTION (H): The log-price and volatility processes are both Itô semimartingales:

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_{s-} dW_s + (\delta 1_{\{|\delta| \leq \kappa\}}) \star (\mu - \nu)_t + (\delta 1_{\{|\delta| > \kappa\}}) \star \mu_t, \quad (2.1)$$

$$\sigma_t = \sigma_0 + \int_0^t \tilde{a}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{b}_s dB_s + (\tilde{\delta} 1_{\{|\tilde{\delta}| \leq \kappa\}}) \star (\tilde{\mu} - \tilde{\nu})_t + (\tilde{\delta} 1_{\{|\tilde{\delta}| > \kappa\}}) \star \tilde{\mu}_t. \quad (2.2)$$

where  $W$  and  $B$  are independent standard Brownian motions;  $\mu$  is a Poisson random measure independent of the Brownian measure of  $W$  on  $(0, \infty) \times E$  with intensity measure  $\nu(dt, dx) = dt \otimes \lambda(dx)$ ,  $\lambda$  is a  $\sigma$ -finite measure without atoms on an auxiliary measurable set  $(E, \mathcal{E})$ ;  $\delta(\omega, t, x)$  is a predictable function on  $\Omega \times \mathbb{R}_+ \times E$ ;  $\tilde{\mu}, \tilde{\nu}$  and  $\tilde{\delta}$  satisfy the same assumption as their counterparts, respectively;  $\kappa$  is a positive constant; finally, the symbol  $\star$  denotes the (possibly stochastic) integral w.r.t. a random measure.<sup>5</sup> Additionally,

- (a) The processes  $a_t(\omega)$ ,  $\tilde{a}_t(\omega)$ ,  $\tilde{\sigma}_t(\omega)$ ,  $\tilde{b}_t(\omega)$ ,  $\sup_{x \in E} \frac{|\delta(\omega, t, x)|}{\gamma(x)}$  and  $\sup_{x \in E} \frac{|\tilde{\delta}(\omega, t, x)|}{\tilde{\gamma}(x)}$  are locally bounded, where  $\gamma(x)$  and  $\tilde{\gamma}(x)$  are nonnegative functions satisfying  $\int_E (\gamma(x)^2 \wedge 1) \lambda(dx) < \infty$  and  $\int_E (\tilde{\gamma}(x)^2 \wedge 1) \tilde{\lambda}(dx) < \infty$ , respectively;
- (b) All paths  $t \mapsto a_t(\omega)$ ,  $t \mapsto \tilde{a}_t(\omega)$ ,  $t \mapsto \tilde{\sigma}_t(\omega)$ ,  $t \mapsto \tilde{b}_t(\omega)$  are càglàd (left-continuous with right limits), and  $\tilde{\sigma}_+$  (right limit of  $\tilde{\sigma}$ ) is also an Itô semimartingale with a Grigelionis decomposition similar to (2.2);

---

<sup>5</sup>More specifically, we have  $(\delta 1_{\{|\delta| \leq \kappa\}}) \star (\mu - \nu)_t = \int_0^t \int_{\mathbb{R}} (\delta(\omega, s, x) 1_{\{|\delta(\omega, s, x)| \leq \kappa\}}) (\mu - \nu)(ds, dx)$  and  $(\delta 1_{\{|\delta| > \kappa\}}) \star \mu_t = \int_0^t \int_{\mathbb{R}} (\delta(\omega, s, x) 1_{\{|\delta(\omega, s, x)| > \kappa\}}) \mu(ds, dx)$ .



(c) The processes  $\sigma^2$  and  $\sigma_-^2$  (left limit of  $\sigma^2$ ) are bounded away from zero.

We may rewrite the volatility jump measure  $\tilde{\mu}$  as a linear combination of another two Poisson random measures  $\mu_1$  and  $\mu_2$ , where  $\mu_1$  has the same jump times as  $\mu$  (but the size part can be different), and  $\mu_2$  is independent of  $\mu$ . In other words,  $\mu_1$  represents the co-jump part and  $\mu_2$  represents the disjoint jump part.

## 2.2 Definitions of the Components of the Leverage Effect

A natural measure of the co-movement of two stochastic processes is their quadratic covariation. Let  $\Delta X_t = X_t - X_{t-}$  denote the jump size of  $X$  at time  $t$  and let  $\Delta\sigma_t^2$  be defined similarly for the process  $\sigma^2$ . The quadratic covariation of  $X$  and  $\sigma^2$  can be decomposed into two parts, as follows:

$$[X, \sigma^2]_T := [X, \sigma^2]_T^C + [X, \sigma^2]_T^D, \quad (2.3)$$

$$[X, \sigma^2]_T^C := 2 \int_0^T \sigma_{t-}^2 \tilde{\sigma}_t dt, \quad (2.4)$$

$$[X, \sigma^2]_T^D := \sum_{t \leq T} \Delta X_t \Delta \sigma_t^2. \quad (2.5)$$

We will call  $[X, \sigma^2]_T^C$  the *continuous leverage effect*,  $[X, \sigma^2]_T^D$  the *discontinuous leverage effect*, and  $[X, \sigma^2]_T$  the *total leverage effect*. We adopt the acronyms CLE, DLE and TLE.

**Remark 1.** We note that many related papers (e.g. [Kalnina and Xiu \(2016\)](#), [Curato \(2015\)](#), [Curato and Sanfelici \(2015\)](#), [Veraart and Veraart \(2012\)](#)) define CLE as the quadratic covariation between the Brownian motion parts of  $X$  and  $\sigma$ , i.e.  $\rho_t dt = d[W^X, W^\sigma]_t$ . The leverage effect defined in this way has the nice mathematical property that  $\rho_t \in [-1, 1]$  for any  $t$ . Yet we argue that the CLE defined in our way measures the co-movements between  $X$  and  $\sigma^2$ , instead of  $W^X$  and  $W^\sigma$ , hence is a more relevant quantity to finance and is closer to the original meaning of leverage. Moreover, the estimation of  $\rho_t$  involves estimating the volatility of volatility, i.e.  $[\sigma^2, \sigma^2]_T^C$ . This

is a very challenging task even in the simple case where neither  $X$  nor  $\sigma^2$  jumps. If the process  $\sigma^2$  jumps, it is still an open question whether there is a consistent estimator of  $[\sigma^2, \sigma^2]_T^C$  (it is not known whether all volatility jumps can be truncated out asymptotically).

In some applications, one may be interested in some functional of  $\sigma^2$ , say  $F(\sigma^2)$ , instead of  $\sigma^2$  itself. Then we define the continuous leverage effect as

$$[X, F(\sigma^2)]_T^C = 2 \int_0^T \nabla F(\sigma_{t-}^2) \sigma_{t-}^2 \tilde{\sigma}_t dt. \quad (2.6)$$

An example is the geometric Ornstein-Uhlenbeck model, with  $F(x^2) = \log x^2$ ,  $\tilde{\mu} \equiv 0$  and

$$d \log \sigma_t^2 = -\kappa \log \sigma_t^2 dt + \rho dW_t + \sqrt{1 - \rho^2} dB_t,$$

with parameters  $\kappa, \rho$  with  $|\rho| \leq 1$ . But for notational simplicity, we mainly focus on the  $\sigma^2$  case. The results in subsequent sections can be extended to the  $F(\sigma^2)$  case if the function  $F \in \mathbb{C}^2$ . [Kalnina and Xiu \(2016\)](#) define their continuous leverage effect in this way.

Additionally, we also consider a truncated version of the discontinuous leverage effect, where only co-jumps induced by large price jumps are selected:

$$[X, \sigma^2]_T^D(\epsilon) := \sum_{t \leq T} \Delta X_t \Delta \sigma_t^2 1_{|\Delta X_t| > \epsilon}, \text{ for a given } \epsilon > 0. \quad (2.7)$$

We call this the *tail discontinuous leverage effect*. The randomness of jumps stems from two sources: jump time and jump size. The jump time can be characterized by the jump intensity, which may or may not depend on the state of the underlying stochastic process(es), and the jump size is represented by the jump size distribution.<sup>6</sup> In our definition, we preserve both sources of

---

<sup>6</sup>In the definition of the co-jump leverage in [Bandi and Renò \(2012\)](#), the jump intensity is further related to the level of the volatility (or variance), and the random jump sizes (in return and volatility series) are replaced by their correlation. As a result, the first source of randomness is restricted, although such an assumption embeds an interesting driving force in the dynamics of the jump intensity process, while the second source of randomness is eliminated.

randomness without integrating any of them out, as a result of using the quadratic covariation. Consequently, our setting allows much flexibility. For example, there can be alternative source(s) of randomness for the jump intensity beyond, e.g., contemporaneous volatility. And the jump size distribution may depend on time hence may have time-varying mean, variance, etc. Similar to the case of integrated volatility, the estimation here will not be a statistical problem in the usual sense, since the objects to be estimated are random variables, instead of parameters. Similar examples in statistics include estimating residuals in regression, random effects in mixed effects models, and false discovery proportion (Fan et al., 2012).

### 3 Construction of the Estimators

Suppose the data are observed every  $\Delta_n = T/n$  units of time without any measurement error. The full grid containing all of the observation points is given by:

$$\mathcal{U} = \{0 = t_0^n < t_1^n < t_2^n < \dots < t_n^n = \lfloor T/\Delta_n \rfloor \Delta_n\}, \quad (3.1)$$

where  $t_j^n = j\Delta_n$  for each  $j$ . Then the increment (log-return) of log-prices over the  $j$ -th interval is  $\Delta_j^n X := X_{t_j^n} - X_{t_{j-1}^n}$ .

#### 3.1 The Continuous Leverage Effect (CLE)

We take  $\alpha_n = \alpha\Delta_n^\varpi$ , for some  $\alpha > 0$  and  $\varpi \in (0, \frac{1}{2})$  for detecting presence of jumps, and a local window of  $k_n$  time intervals for estimating the spot volatility, which is an integer that satisfies the condition

$$\frac{1}{K} \leq k_n \Delta_n^b \leq K, \quad \text{with } 0 < b < 1, \quad (3.2)$$

for some positive constant  $K$ . Let  $I_n^-(i) = \{i - k_n, \dots, i - 1\}$  (if  $i > k_n$ ) and  $I_n^+(i) = \{i + 1, \dots, i + k_n\}$  define two local windows in time of length  $k_n \Delta_n$ , just before and after time  $i\Delta_n$ . Denote the

downward truncated increment of  $X$  by  $\Delta_i^n X_{\alpha_n} = \Delta_i^n X \cdot 1_{\{|\Delta_i^n X| \leq \alpha_n\}}$ . Then we can define

$$\begin{aligned} \widehat{\langle X, \sigma^2 \rangle}_T^C &= \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \beta_i^n(X), \quad \text{with} \quad \beta_i^n(X) = \Delta_i^n X_{\alpha_n} (\hat{\sigma}_{i+}^2 - \hat{\sigma}_{i-}^2) \\ \hat{\sigma}_{i+}^2 &= \frac{1}{k_n \Delta_n} \sum_{j \in I_n^+(i)} (\Delta_j^n X_{\alpha_n})^2, \quad \hat{\sigma}_{i-}^2 = \frac{1}{k_n \Delta_n} \sum_{j \in I_n^-(i)} (\Delta_j^n X_{\alpha_n})^2. \end{aligned} \quad (3.3)$$

The estimator proposed in [Wang and Mykland \(2014\)](#) is given by

$$\widehat{\langle X, \sigma^2 \rangle}_T = 2 \sum_{i=0}^{K_n-2} (X_{\tau_{n,i+1}} - X_{\tau_{n,i}}) (\hat{\sigma}_{\tau_{n,i+1}}^2 - \hat{\sigma}_{\tau_{n,i}}^2), \quad \hat{\sigma}_{\tau_{n,i+1}}^2 = \frac{1}{M_n \Delta_n} \sum_{t_{n,i} \in (\tau_{n,i}, \tau_{n,i+1}]} (X_{t_{n,j+1}} - X_{t_{n,j}})^2,$$

where  $M_n$  is equivalent to  $k_n$  here and  $K_n = \lfloor n/M_n \rfloor$ . Note that the interval  $(\tau_{n,i}, \tau_{n,i+1}]$  contains  $M_n$  number of observation intervals like  $(t_{n,i}, t_{n,i+1}]$ . And the estimator in the cited paper is constructed by summing up (the product of) the increments of  $X$  and  $\hat{\sigma}^2$  over such large blocks, the size of which,  $M_n$ , goes to infinity. Hence, each summand only overlaps with its two adjacent ones. This simplifies the asymptotic analysis, but may not be efficient since such data aggregation leads to a coarse approximation of the integral (2.4). By contrast in (3.3), the summation is taken over each observation interval  $(t_{n,i}, t_{n,i+1}]$ . This provides a finer partition of the data in the time dimension, but complicates the asymptotic analysis in the current paper because each summand overlaps with about  $2k_n$  number of other ones within a local window.

It is interesting to compare the estimator (3.3) to those of the integrated volatility. In the latter case, estimators like bipower variation and multipower variation are the sum of even functions of the increments of the log-price process  $X$ . Here in our case with  $k_n = 1$ , observe that

$$\beta_i^n(X) \Delta_n = \sum_{j \in I_n^+(i)} \Delta_i^n X_{\alpha_n} (\Delta_j^n X_{\alpha_n})^2 - \sum_{j \in I_n^-(i)} \Delta_i^n X_{\alpha_n} (\Delta_j^n X_{\alpha_n})^2.$$

Clearly, the representative element of  $\beta_i^n(X)$  takes the function form  $f(x_1, x_2) = x_1 x_2^2$ , which is odd instead of even. Additionally,  $\beta_i^n(X)$  is an odd function of an increasing number of increments ( $k_n \rightarrow \infty$ ), the asymptotic properties of which have not been studied in the literature.

### 3.2 The Discontinuous Leverage Effect (DLE)

When estimating the DLE, we need to preserve jumps, instead of eliminating them as in the case of estimating CLE. Accordingly, we introduce the upward truncated increment  $\Delta_i^n X^{\alpha_n} = \Delta_i^n X \cdot 1_{\{|\Delta_i^n X| > \alpha_n\}}$ . Similarly, the estimator<sup>7</sup> of DLE is given by

$$[\widehat{X}, \widehat{\sigma^2}]_T^D = \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \Delta_i^n X^{\alpha_n} (\widehat{\sigma}_{i+}^2 - \widehat{\sigma}_{i-}^2). \quad (3.6)$$

We also consider the tail discontinuous leverage effect and accordingly use

$$[\widehat{X}, \widehat{\sigma^2}]_T^D(\epsilon) = \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \Delta_i^n X^{\alpha_n \vee \epsilon} (\widehat{\sigma}_{i+}^2 - \widehat{\sigma}_{i-}^2) \quad (3.7)$$

to estimate  $[X, \sigma^2]_T^D(\epsilon)$  in (2.7).

---

<sup>7</sup> In [Jacod and Todorov \(2010\)](#), the authors studied a very general limit functional  $U(F)_T$  (with minor changes in notations in what follows) given by

$$U(F)_T = \sum_{s \leq T} F(\Delta X_s, \sigma_{s-}^2, \sigma_s^2) 1_{\{\Delta X_s \neq 0\}}, \quad (3.4)$$

together with its estimator

$$U(F, k_n)_T = \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} F(\Delta_i^n X, \widehat{\sigma}_{i-}^2, \widehat{\sigma}_{i+}^2) 1_{\{|\Delta_i^n X| > \alpha_n\}}. \quad (3.5)$$

Here  $F$  is a function on  $\mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^*$  where  $\mathbb{R}_+^* = (0, \infty)$ . Observe that with the specification  $F(x, y, z) = x(z - y)$ , (3.4) and (3.5) would be equivalent to (2.5) and (3.6) (the discontinuous leverage effect and its estimator), respectively. As we will see, however, the specific functional form of  $F$  will allow us to derive central limit results that apply under much broader conditions regarding jump stochasticity than those applying to a generic  $F$ .

## 4 Convergence in Probability

### 4.1 Consistency of the CLE Estimator

**Theorem 1.** *Assume that either one of the following assumptions holds:*

- (a) *X is continuous and the jump part of  $\sigma^2$  has finite total variation;*
- (b) *X is discontinuous with  $\int(\gamma(x)^r \wedge 1)\lambda(dx) < \infty$ , for some  $r \in [0, 1)$ , and  $\alpha_n = \alpha\Delta_n^\varpi$ , for some  $\alpha > 0$  and  $\varpi \in [\frac{1}{2(2-r)}, \frac{1}{2})$ , and the disjoint jump part of  $\sigma^2$  has finite total variation.*

*Then, under Assumption (H) in Section 2.1 and (3.2), we have<sup>8</sup>*

$$[\widehat{X, \sigma^2}]_T^C \xrightarrow{u.c.p.} [X, \sigma^2]_T^C = 2 \int_0^T \sigma_{t-}^2 \tilde{\sigma}_t dt. \quad (4.1)$$

**Remark 2.** *To compare CLE estimation to the estimation of integrated volatility, observe that in Theorem 6.3 of Jacod (2012), the assumption on the jump activity index, namely  $r \in [0, 2)$ , is somewhat less restrictive than in Theorem 1, part (b). The reason is that price jumps affect the measurement of continuous leverage not only through volatility estimation, but also through the cross product with spot volatility, which will generate a bias term even without co-jumping.<sup>9</sup> Thus, price jumps cause more challenges for continuous leverage estimation than they do for integrated volatility estimation. Therefore, we can only allow for somewhat less active price jumps. Otherwise, their effect cannot be effectively eliminated.*

---

<sup>8</sup>Here,  $Z_T^n \xrightarrow{u.c.p.} Z_T$  means that the sequence of stochastic processes  $Z_T^n$  converges in probability, locally uniformly in time, to a limit  $Z_T$ , that is,  $\sup_{s \leq T} |Z_s^n - Z_s| \xrightarrow{\mathbb{P}} 0$  for all finite  $T$ .

<sup>9</sup>Refer to the decomposition of  $X$  in equation (A.1) and note that the conditional expectation of  $(\Delta_j^n X'')\sigma_{i_j^n}^2 \Delta_n$  is of the order  $\Delta_n^2$ , the same as  $(\Delta_j^n X')\sigma_{i_j^n}^2 \Delta_n$ .

## 4.2 Consistency of the DLE Estimator

Recall the jump activity index  $r$  introduced in Theorem 1, part (b).

**Theorem 2.** *Suppose that Assumption (H) holds and that condition (3.2) is satisfied. In addition, assume that  $\tilde{\mu}$  can be decomposed into two parts: one part that always jumps at the same time with  $\mu$  and satisfies (H) with the same  $r$ ; and another part that never jumps at the same time with  $\mu$  and has finite variation.*

(i) *If  $r \in [0, 1]$ , then  $[\widehat{X}, \widehat{\sigma^2}]_T^D$  converges in probability, in the Skorokhod topology, to  $[X, \sigma^2]_T^D$ .*

(ii) *Given some  $\epsilon > 0$ , if  $r \in [0, 2)$ , then  $[\widehat{X}, \widehat{\sigma^2}]_T^D(\epsilon)$  converges in probability, in the Skorokhod topology, to  $[X, \sigma^2]_T^D(\epsilon)$ .*

The results of Theorem 2 are direct consequences from Theorem 3.1 in Jacod and Todorov (2010). In case (i), we require  $r \in [0, 1]$  to satisfy condition (c) of their theorem with our particular choice of  $F$ . Furthermore, case (ii) satisfies condition (a) of their theorem.

It is interesting to note the tradeoff between jump activity and jump size in the assumptions: if all co-jumps are considered, then we must put a stronger restriction on the jump activity, as in part (i) of Theorem 2; but if we only consider co-jumps for which price jumps are large (larger than  $\epsilon$ ), then the restriction on the jump activity can be relaxed, as in part (ii) of Theorem 2. Intuitively speaking, we need to put certain restrictions on summability of the process to be estimated. Since the probability of having large price jumps is relatively low, by considering only co-jumps induced by large price jumps, we implicitly set such restriction.

## 4.3 Consistency of the TLE Estimator

Obviously, if  $X_n \xrightarrow{\mathbb{P}} X$  and  $Y_n \xrightarrow{\mathbb{P}} Y$ , then  $X_n + Y_n \xrightarrow{\mathbb{P}} X + Y$ . Therefore,  $[\widehat{X}, \widehat{\sigma^2}]_T^C + [\widehat{X}, \widehat{\sigma^2}]_T^D$  is a consistent estimator of  $[X, \sigma^2]_T$ .

## 5 The Central Limit Theorems

### 5.1 CLT for the CLE Estimator

**Theorem 3.** *For some finite and positive constant  $c$ , let  $k_n = \lfloor cn^b \rfloor$  with  $0 < b < 1$ . Assume that either one of the following assumptions holds:*

- (a)  *$X$  is continuous and the jump part of  $\sigma^2$  has finite total variation;*
- (b)  *$X$  is discontinuous with  $\int (\gamma(x)^r \wedge 1) \lambda(dx) < \infty$ , for some  $r \in [0, 1/2)$ , and  $\alpha_n = \alpha \Delta_n^\varpi$ , for some  $\alpha > 0$  and  $\varpi \in [\frac{3}{4(2-r)}, \frac{1}{2})$ , and the disjoint jump part of  $\sigma^2$  has finite total variation.*

Then, under Assumption (H),  $\sqrt{n}^{b\wedge(1-b)}([\widehat{X, \sigma^2}]_T^C - [X, \sigma^2]_T^C)$  converges stably in law to a limiting random variable defined on an extension of the original probability space. That is,

$$\sqrt{n}^{b\wedge(1-b)}([\widehat{X, \sigma^2}]_T^C - [X, \sigma^2]_T^C) \xrightarrow{\mathcal{L}_{st}} \int_0^T \eta_t dB_t, \quad (5.1)$$

where  $B$  is a standard Wiener process independent of  $\mathcal{F}$ , and  $\eta_s$  satisfies:

$$\int_0^T \eta_t^2 dt = \begin{cases} \frac{4}{c} \int_0^T \sigma_{t-}^6 dt, & \text{if } b < 1/2; \\ \frac{4}{c} \int_0^T \sigma_{t-}^6 dt + \frac{2cT}{3} \int_0^T \sigma_{t-}^2 d\langle \sigma^2, \sigma^2 \rangle_t, & \text{if } b = 1/2; \\ \frac{2cT}{3} \int_0^T \sigma_{t-}^2 d\langle \sigma^2, \sigma^2 \rangle_t, & \text{if } 1/2 < b < 1. \end{cases} \quad (5.2)$$

Note that the convergence rate is given by  $\sqrt{n}^{b\wedge(1-b)}$ . Hence, when  $b = 1/2$ , one gets the optimal convergence rate  $n^{1/4}$ . The feasible version is

$$\frac{\sqrt{n}^{b\wedge(1-b)}}{\sqrt{\widehat{V}_T^n(X, \alpha_n)}}([\widehat{X, \sigma^2}]_T^C - [X, \sigma^2]_T^C) \xrightarrow{\mathcal{L}_{st}} \mathcal{N}(0, 1), \quad (5.3)$$



where the standard normal random variable  $\mathcal{N}(0, 1)$  is independent of  $\mathcal{F}$ , and

$$\begin{aligned} \widehat{V}_T^n(X, \alpha_n) &= \frac{4}{15c\Delta_n^2} \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} (\Delta_i^n X_{\alpha_n})^6 \cdot 1_{\{0 < b \leq 1/2\}} + \frac{cT}{k_n} \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} (\Delta_i^n X_{\alpha_n})^2 \\ &\times \left( (\widehat{\sigma}_{i+}^2 - \widehat{\sigma}_{i-}^2)^2 - \frac{2}{3(k_n\Delta_n)^2} \sum_{j \in I_n^\pm(i)} (\Delta_j^n X_{\alpha_n})^4 \right) \cdot 1_{\{1/2 \leq b < 1\}}. \end{aligned} \quad (5.4)$$

Not surprisingly, the convergence rate is determined by the spot volatility estimates. There are two sources of errors in the estimation of spot volatility, as shown in the proof of Lemma 1 in the Appendix. We call the first one the *price variation error*, i.e.  $\xi_{\pm}^n(1)$  in the Appendix, since it only involves the log-price process and its limiting standard deviation is proportional to  $\sigma_{i\pm}^2$ , the spot value of the quadratic variation of  $X$ . And the second one could be termed the *volatility variation error*, i.e.  $\xi_{\pm}^n(2)$ , which is defined as the partial sum of integrals of volatility increments and has a limiting standard deviation related to the spot value of the quadratic variation of the  $\sigma^2$  process. Their convergence rates are given by  $1/\sqrt{k_n}$  and  $\sqrt{k_n\Delta_n}$  (this order hinges on the smoothness of the volatility process implied by equation (2.2)), respectively. When  $b < 1/2$ , in other words when the length of the local window  $k_n\Delta_n$  is relatively small, one can pretend volatility were constant, hence the price variation error would dominate. By contrast, when  $b > 1/2$ , such simplification does not hold anymore as the volatility variation error becomes dominant. Lastly, these two errors are of the same order when  $b = 1/2$ . In this case, one could estimate their corresponding limiting variances separately (see the two terms in (5.4)) and then choose an optimal value of  $c$  to minimize the total limiting variance (see a similar discussion in Wang and Mykland (2014)).

Here, when there are no price jumps but only volatility jumps (with finite total variation), the convergence rate of the CLE estimator does not change at all.<sup>10</sup> The reason is that our definition of CLE is solely based on time domain properties of the underlying stochastic processes and is not a function of spot volatility. Therefore, volatility jumps do not change the convergence rate of the

---

<sup>10</sup>But in Bandi and Renò (2012), volatility jumps will slow down the convergence rate.

estimators.

## 5.2 CLT for the DLE Estimator

**Theorem 4.** *Suppose that Assumption (H) holds and that condition (3.2) is satisfied. In addition, assume that  $\tilde{\mu}$  can be decomposed into two parts: one part that always jumps at the same time with  $\mu$  and satisfies (H) with the same  $r$ ; and another part that never jumps at the same time with  $\mu$  and has finite variation. Let  $\alpha_n = \alpha \Delta_n^\varpi$  with  $\varpi \in (0, 1/2)$  and  $u_n = n^{b \wedge (1-b)}$ .*

(i) *If  $r \in [0, 1)$ , then, for any  $\varpi \in [\frac{1}{2(2-r)}, \frac{1}{2})$  and any finite  $T$ ,  $\sqrt{u_n}([\widehat{X}, \widehat{\sigma^2}]_T^D - [X, \sigma^2]_T^D)$  converges stably in law to the random variable*

$$\mathcal{D}_T = \sum_{p \geq 1} \Delta X_{T_p} (V_p^+ - V_p^-) 1_{\{T_p \leq T\}}, \quad (5.5)$$

where  $V_p^+$  and  $V_p^-$  are independent normal variables with  $\mathcal{F}$ -conditional variance given in (A.8) and  $\{T_p\}$  is the set of jump times.

(ii) *Given some  $\epsilon > 0$ , if  $r \in [0, 3/2)$ , then, for any  $\varpi \in [\frac{1}{4(2-r)}, \frac{1}{2})$ ,*

$$\sqrt{u_n}([\widehat{X}, \widehat{\sigma^2}]_T^D(\epsilon) - [X, \sigma^2]_T^D(\epsilon))$$

*converges stably in law to the process*

$$\mathcal{D}_T^\epsilon = \sum_{p \geq 1} \Delta X_{T_p} 1_{\{|\Delta X_{T_p}| > \epsilon\}} (V_p^+ - V_p^-) 1_{\{T_p \leq T\}}.$$

*Since the estimation errors in these cases come from estimating pre- and post-jump volatilities, the  $\mathcal{F}$ -conditional variances of  $\mathcal{D}_T$  and  $\mathcal{D}_T^\epsilon$  can be estimated by adapting (5.4) accordingly. For*

instance,

$$\begin{aligned} \text{Var}(\widehat{\mathcal{D}_T | \mathcal{F}})(X, \alpha_n) &= \frac{2}{3ck_n \Delta_n^2} \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} (\Delta_i^n X^{\alpha_n})^2 \sum_{j \in I_n^\pm(i)} (\Delta_j^n X_{\alpha_n})^4 \cdot 1_{\{0 < b \leq 1/2\}} + \frac{c}{k_n} \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} (\Delta_i^n X^{\alpha_n})^2 \\ &\quad \times \left( (\widehat{\sigma}_{i+}^2 - \widehat{\sigma}_{i-}^2)^2 - \frac{2}{3(k_n \Delta_n)^2} \sum_{j \in I_n^\pm(i)} (\Delta_j^n X_{\alpha_n})^4 \right) \cdot 1_{\{1/2 \leq b < 1\}}. \end{aligned}$$

As in the CLE case, the optimal convergence rates in both (i) and (ii) are  $n^{1/4}$ , corresponding to  $b = 1/2$ .

In connection to Theorem 3.2 in [Jacod and Todorov \(2010\)](#), observe that when  $r > 0$ ,  $F(x, y, z) = x(z - y)$  fails to satisfy the conditions in that theorem. In the general case where  $\epsilon = 0$ , i.e. part (i), Theorem 4 provides a CLT not just for the case of finite jump activity, but also for the infinite jump activity case with jump index  $r < 1$ .

Note also that we encounter a similar tradeoff between jump activity and jump size as in Theorem 2.

### 5.3 CLT for the TLE Estimator

For the purpose of estimating the TLE, we no longer need to distinguish between CLE and DLE and therefore can impose a less stringent restriction on the jump activity index:

**Theorem 5.** *Suppose that Assumption (H) holds and that condition (3.2) is satisfied. Let  $\alpha_n$  and  $u_n$  be as before and let*

$$[\widehat{X, \sigma^2}]_T := \sum_{i=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} (\Delta_i^n X) (\widehat{\sigma}_{i+}^2 - \widehat{\sigma}_{i-}^2).$$

If  $r \in [0, 3/2)$ , then, for any finite  $T$ , we have

$$\sqrt{u_n}([\widehat{X, \sigma^2}]_T - [X, \sigma^2]_T) \xrightarrow{\mathcal{L}_{st}} \int_0^T \eta_s dB_s + \mathcal{D}_T.$$

Again, the optimal convergence rate  $n^{1/4}$  is achieved when  $b = 1/2$ .

## 6 Market Microstructure Noise

In high-frequency financial applications, the presence of market microstructure noise in asset prices can be non-negligible. To deal with market microstructure noise, we employ pre-averaging. The contaminated log return process  $Y_t$  is observed every  $\Delta t_{n,i} = T/n$  units of time, at times  $0 = t_{n,0} < t_{n,1} < t_{n,2} < \dots < t_{n,n} = T$ . The noise term has the following structure:

**Assumption 1.**

$$Y_t = X_t + \epsilon_t, \text{ where the } \epsilon_t \text{'s are i.i.d. } N(0, a^2) \text{ and } \epsilon_t \perp\!\!\!\perp W_t \text{ and } B_t, \text{ for all } t \geq 0. \quad (6.1)$$

Blocks are defined on a much less dense grid of  $\tau_{n,i}$ 's, also spanning  $[0, T]$ , so that

$$\text{block } i = \{t_{n,j} : \tau_{n,i} \leq t_{n,j} < \tau_{n,i+1}\} \quad (6.2)$$

(the last block, however, includes  $T$ ). We define the block size,  $M_{n,i}$ , by

$$M_{n,i} = \#\{j : \tau_{n,i} \leq t_{n,j} < \tau_{n,i+1}\}. \quad (6.3)$$

In principle, the block size can vary across the trading period  $[0, T]$ , but for this development we take  $M_{n,i} = M$ : it depends on the sample size  $n$ , but not on the block index  $i$ . We then use as an estimated value of the efficient price in the time period  $[\tau_{n,i}, \tau_{n,i+1})$ :

$$\hat{X}_{\tau_{n,i}} = \frac{1}{M} \sum_{t_{n,j} \in [\tau_{n,i}, \tau_{n,i+1})} Y_{t_{n,j}}.$$

Let  $J_n^-(i) = \{i - k_n M, \dots, i - M\}$  (if  $i > k_n M$ ) and  $J_n^+(i) = \{i + M, \dots, i + k_n M\}$  define two local windows in time of length  $k_n M \Delta_n$  just before and after the interval  $[i \Delta_n, (i + 1) \Delta_n)$ . Denote the truncated increment of  $\hat{X}$  over  $M \Delta_n$  by  $\Delta_j^n \hat{X}_{\alpha_n} = \Delta_j^n \hat{X} \cdot 1_{\{|\Delta_j^n \hat{X}| \leq \alpha_n\}}$ . Then we can define

$$\begin{aligned} [\widehat{X}, \widehat{\sigma^2}]_T^C &= \sum_{i=k_n M+1}^{n-k_n M} \Delta_i^n \hat{X}_{\alpha_n} (\hat{\sigma}_{i+}^2 - \hat{\sigma}_{i-}^2) \\ \hat{\sigma}_{i+}^2 &= \frac{3}{2k_n M \Delta_n} \sum_{j \in J_n^+(i)} \Delta_j^n \hat{X}_{\alpha_n}^2, \quad \hat{\sigma}_{i-}^2 = \frac{3}{2k_n M \Delta_n} \sum_{j \in J_n^-(i)} \Delta_j^n \hat{X}_{\alpha_n}^2. \end{aligned} \quad (6.4)$$

**Theorem 6.** Under the same conditions as in Theorem 1 except that  $M = \lfloor c_1 \sqrt{n} \rfloor$  and  $k_n = \lfloor \frac{c}{\sqrt{c_1}} n^{b/2} \rfloor$  for some finite and positive constant  $c_1$ ,

$$[\widehat{X}, \widehat{\sigma^2}]_T^C \xrightarrow{u.c.p.} [X, \sigma^2]_T^C = \int_0^T 2 \sigma_{t-}^2 \tilde{\sigma}_t dt. \quad (6.5)$$

**Theorem 7.** Under the same conditions as in Theorem 3, except that  $M = \lfloor c_1 \sqrt{n} \rfloor$  and  $k_n = \lfloor \frac{c}{\sqrt{c_1}} n^{b/2} \rfloor$  for some finite and positive constant  $c_1$ ,  $n^{\frac{b \wedge (1-b)}{4}} ([\widehat{X}, \widehat{\sigma^2}]_T^C - [X, \sigma^2]_T^C)$  converges stably in law to a limiting variable defined on an extension of the original probability space. That is,

$$n^{\frac{b \wedge (1-b)}{4}} \left( [\widehat{X}, \widehat{\sigma^2}]_T^C - [X, \sigma^2]_T^C \right) \xrightarrow{\mathcal{L}_{st}} \int_0^T \tilde{\eta}_s dB_s, \quad (6.6)$$

where  $B$  is a standard Wiener process independent of  $\mathcal{F}$ , and  $\tilde{\eta}_s$  satisfies:

$$\begin{aligned} \int_0^T \tilde{\eta}_s^2 ds &= \left( \frac{4\sqrt{c_1}}{c} \int_0^T \sigma_{t-}^6 dt + \frac{24a^2}{cc_1^{3/2}T} \int_0^T \sigma_{t-}^4 dt + \frac{54a^4}{cc_1^{7/2}T^2} \int_0^T \sigma_{t-}^2 dt \right) 1_{\{0 < b \leq 1/2\}} \\ &+ \left( \frac{2c\sqrt{c_1}T}{3} \int_0^T \sigma_{t-}^2 d\langle \sigma^2, \sigma^2 \rangle_t \right) 1_{\{1/2 \leq b < 1\}}. \end{aligned}$$

Similar to the case without market microstructure noise, we shall assess the asymptotic normality of the standardized statistics

$$\frac{n^{\frac{b \wedge (1-b)}{4}} \left( [\widehat{X}, \widehat{\sigma^2}]_T^C - [X, \sigma^2]_T^C \right)}{\sqrt{\int_0^T \tilde{\eta}_s^2 ds}}$$

in the simulation study.

Although the effects of microstructure noise have been well understood in the context of estimating diffusion related quantities (such as integrated and spot volatility), it seems that, for jump related quantities, all existing methods dealing with microstructure noise become powerless. For example, [Bücher and Vetter \(2013\)](#) explain why the pre-averaging method fails when estimating the Lévy measure of log-price jumps (see Section 5.4 of the cited paper). For the same reason, a

direct application of the pre-averaging method does not yield a consistent estimator of the discontinuous leverage effect. It is still an open question whether a consistent estimator exists or not. We leave this for further research.

## 7 Monte Carlo Simulations

Throughout this section, the time unit is measured in years, hence the time span of one trading day is  $T = 1/252$ .

### 7.1 CLE Without Jumps

We first examine the finite sample performance of the CLE estimator. We use the Heston model to generate by simulations the log-price process  $X_t$  and the volatility process  $\sigma_t^2$ :

$$\begin{cases} dX_t = (\mu - \sigma_t^2/2)dt + \sigma_t dW_t \\ d\sigma_t^2 = \kappa(\theta - \sigma_t^2)dt + \eta\sigma_t(\rho dW_t + \sqrt{1 - \rho^2}dV_t), \end{cases} \quad (7.1)$$

where  $W$  and  $V$  are independent standard Brownian motions. We take the following parameter values:  $\theta = 0.1, \eta = 0.5, \kappa = 5, \rho = -0.8$  and  $\mu = 0.05$ .

We simulate data using three total time span specifications: one week (5 trading days), one month (21 trading days) and one year (252 trading days). Within each trading day, the number of observations is 4680, corresponding to sampling every 5 seconds for a 6.5 hours trading day (390 minutes), which mimics the empirical data we are going to analyze in Section 8. We repeat each simulation run 5000 times. We take  $k_n = \lfloor \sqrt{T/\Delta_n} \rfloor$ .

Figure 1 presents the simulation results. The first row gives the densities of the realized (red solid line) and estimated (blue dashed line) continuous leverage effects. In the weekly case, where the total time span  $T$  is relatively small, the variance of the estimates is much larger than that of the true (realized) values. As the time span increases, these two densities become more similar.

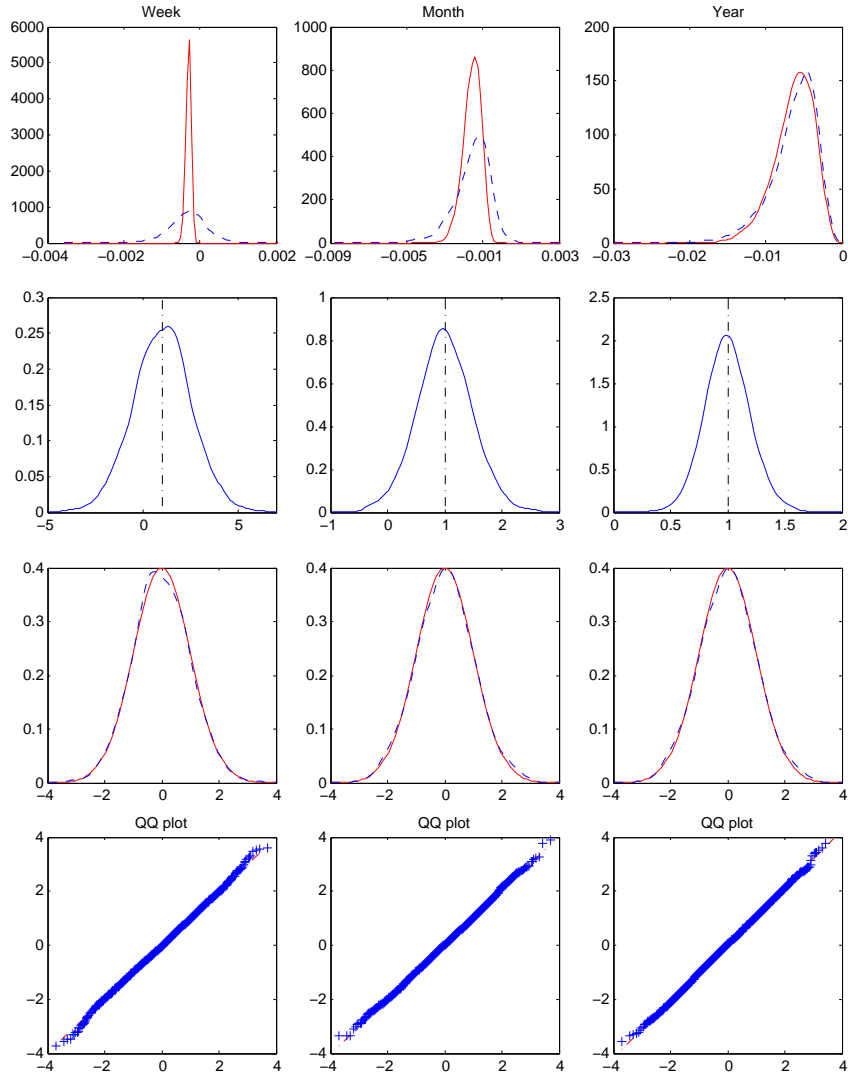


Figure 1: The first three rows are densities of: (1) the realized (red and solid) and estimated (blue and dashed) continuous leverage effects; (2) ratios of the estimated CLEs and the corresponding true values; (3) the standardized estimation errors (blue and dashed, the left hand side of (5.3)) and a standard normal random variable (red and solid). The last row provides quantile-quantile plots of the sample quantiles of the standardized estimation errors versus theoretical quantiles from a normal distribution. Each column corresponds to a different total time span.

The second row of graphs measures the estimation precision. We plot the densities of ratios of the estimated CLEs to the corresponding true (realized) ones. As expected, the precision increases with time span and sampling frequency. Although in the weekly case the estimation precision is arguably still somewhat low, the mean estimation errors (of order  $10^{-6}$ ) remain small compared to the sample mean of the true values (of order  $10^{-3}$ ). And the mean value of the ratios is around 1. Both rows confirm that the CLE estimator is indeed unbiased.

The graphs in the last two rows assess the asymptotic normality in finite samples: in the third row we plot the densities of the standardized estimation errors (blue dashed line), i.e. the LHS of (5.3), and the standard normal density (red solid line) as benchmark; the last row presents the corresponding QQ plots. Contrary to the first two rows, the results in the last two rows are not very different across time span. In all the three cases, both densities and quantiles are quite close to those of a standard normal random variable. These results suggest that the somewhat poor finite sample performance in the weekly case is due to large variance and not to bias.

In addition, we plot the rejection rates of the test for the presence of CLE against significant levels in Figure 2. The plot is similar to the ROC curve. The null hypothesis of  $CLE = 0$  is tested by using (5.3). Since we assume the presence of CLE in the simulated model, this illustrates the finite sample power of the test. In the weekly case, the rejection rates of both one-sided and two-sided tests are just above the corresponding significance levels, indicating low power of the tests. However, the tests become much more effective in rejecting the false null hypothesis in the monthly case. Even when the significance level is small, we obtain reasonably large rejection rates. This is due to the fact that, in the monthly case, we not only have more observations but also the magnitude of the latent true values becomes larger compared to the estimated standard errors. Finally, in the yearly case, as the significance level increases, the rejection rate goes to 1 very rapidly. The results presented here will serve as guidance to choose a suitable time span to test for the presence of CLE in the empirical study that follows.



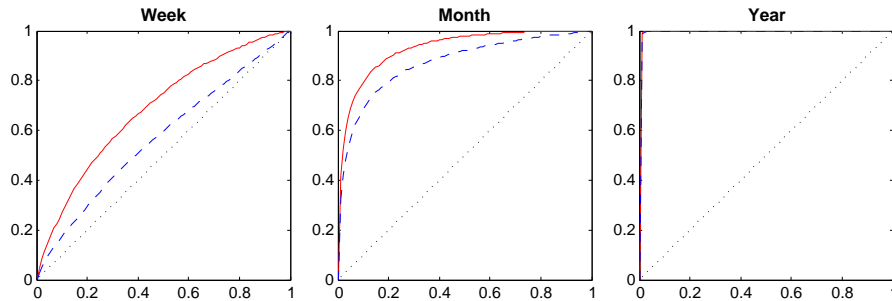


Figure 2: Testing powers ( $y$ -axis) against significant levels ( $x$ -axis) for different specifications of total time span. The red solid curve corresponds to the power of the one-sided test, and the blue dashed curve corresponds to that of the two-sided test.

## 7.2 CLE With Nuisance Jumps

Next, we consider the following model, which incorporates jumps:

$$\begin{cases} dX_t = (\mu - \sigma_t^2/2)dt + \sigma_t dW_t + J_t^X dN_t \\ d\sigma_t^2 = \kappa(\theta - \sigma_t^2)dt + \eta\sigma_t(\rho dW_t + \sqrt{1 - \rho^2}dV_t) + J_t^\sigma dN_t, \end{cases} \quad (7.2)$$

where  $N_t$  is a Poisson process with intensity  $\lambda$ , and where  $J_t^X$  and  $J_t^\sigma$  are the jump sizes of the log-price and volatility processes at time  $t$ , respectively. The parameters of the diffusion part are taken to be the same as in Section 7.1. The density of price jumps is taken as an asymmetric double exponential distribution

$$f_X(x) = \begin{cases} \frac{p}{\gamma_d} \exp\left(\frac{x}{\gamma_d}\right), & -\infty < x \leq 0; \\ \frac{1-p}{\gamma_u} \exp\left(-\frac{x}{\gamma_u}\right), & 0 < x < \infty; \end{cases}$$

where  $\gamma_u, \gamma_d > 0$  and  $0 \leq p \leq 1$ , while the density of volatility jumps is specified as the exponential distribution

$$f_\sigma(x) = \frac{1}{\gamma_\sigma} \exp\left(-\frac{x}{\gamma_\sigma}\right), \quad x \geq 0.$$

We set  $\gamma_u = 0.008$ ,  $\gamma_d = 0.018$ ,  $p = 0.6$  and  $\lambda = 300$ .

We vary the values of  $\gamma_\sigma$  to study the effects of volatility jumps on the performance of our estimator  $[\widehat{X}, \widehat{\sigma^2}]_T^C$  and the power of the test for the null hypothesis of  $\text{CLE} = 0$ . Recall the result in Theorem 3 that volatility jumps do not change the convergence rate but increase the asymptotic variance. Based on the results in the previous subsection, we choose  $T = 21/252$  (one trading month) so that it is easier to see the differences in the power of the test under various choices of volatility jump sizes. We use again 5-sec sampling yielding 4680 observations per trading day, replicated 5000 times. We take  $k_n = \lfloor \sqrt{T/\Delta_n} \rfloor$ ,  $\alpha = 5\sqrt{\text{BV}_T}$ , where  $\text{BV}_T$  stands for the bipower variation in the selected time span  $T$ , and  $\varpi = 0.49$ .

Figure 3 presents the simulation results. As in Section 7.1, the densities of the standardized estimation errors (LHS of (5.3)) again follow closely the standard normal density. As can be expected, large volatility jump sizes have an adverse effect on the power of the test: the larger  $\gamma_\sigma$ , the lower the power of the test.

### 7.3 CLE with Noise

This subsection aims to verify in finite samples the asymptotic normality of the standardized noise-robust statistics in Section 6. The number of simulations is 5000. We set  $T = 1/252$  (one trading day) and  $n = 10^6$ . In addition, we take  $c = 1$ ,  $c_1 = 1$ ,  $b = 1/2$  and, accordingly,  $k_n = \lfloor n^{1/4} \rfloor$ , and choose again  $\alpha = 5\sqrt{\text{BV}_T}$  and  $\varpi = 0.49$ . We first simulate the processes of Section 7.1 with market microstructure noise (but without jumps). The microstructure noise is simulated from  $N(0, 0.005^2)$ . The results are shown in Figure 4. Next, we simulate the processes of Section 7.2 with market microstructure noise (and jumps). The results are shown in Figure 5. We observe a close fit to the standard normal distribution in both cases.

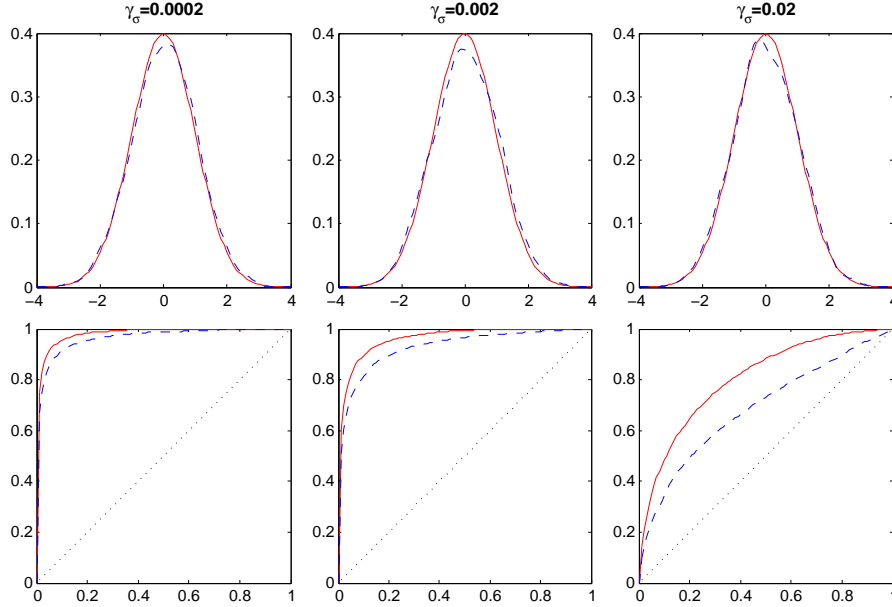


Figure 3: Standardized estimation error densities and testing powers. First row: the red solid line corresponds to the standard normal density, the blue dashed line to the standardized estimation error density. Second row: the red solid line corresponds to the one-sided test, the blue dashed line to the two-sided test.

## 7.4 DLE

In this subsection, we work with a stochastic volatility model given by

$$\begin{cases} dX_t = \sqrt{V_t^1 + V_t^2} dW_t + \int_{\mathbb{R}} x \mu(dt, dx, dy), \\ dV_t^1 = \kappa_1(\theta - V_t^1) dt + \eta \sqrt{V_t^1} dW'_t, \\ dV_t^2 = -\kappa_2 V_t^2 dt + \int_{\mathbb{R}} y \mu(dt, dx, dy), \end{cases}$$

where  $W$  and  $W'$  are independent standard Brownian motions, and the Poisson random measure  $\mu$  has compensator

$$\nu(dt, dx, dy) = \frac{\lambda}{(h-l)(u-d)} \mathbf{1}_{\{x \in [-h, -l]\}} \mathbf{1}_{\{y \in [d, u]\}},$$

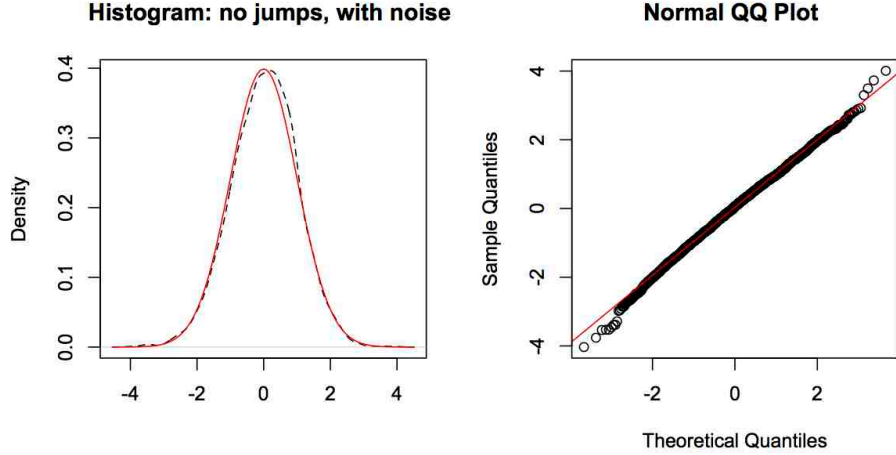


Figure 4: Standardized estimation error densities and QQ plots with market microstructure noise but without jumps, with  $T = 1/252$  (one trading day) and  $n = 10^6$ . The plots show a close fit to  $N(0, 1)$ .

for  $0 < l < h$  and  $0 < d < u$ . The parameter settings used in the simulations are given in Table 1.

Table 1: Parameter settings used in simulation

Case	Parameters								
	$\kappa_1$	$\theta$	$\eta$	$\kappa_2$	$\lambda$	$l$	$h$	$d$	$u$
I-j	5.04	0.4	0.2	126	126	0.1	1.0420	0.04	0.76
II-j	5.04	0.4	0.2	126	252	0.1	0.7197	0.04	0.36
III-j	5.04	0.4	0.2	126	1008	0.1	0.3275	0.04	0.06

We set  $T = 5/252$  (one trading week) and simulate 5000 replications. On each day, we consider sampling  $n = 390$  and  $n = 4680$  times. For the calculation of the local volatility estimators we use again a window of  $k_n = \lfloor \sqrt{T/\Delta_n} \rfloor$  and choose the truncation parameters again as  $\alpha = 5\sqrt{BV_T}$

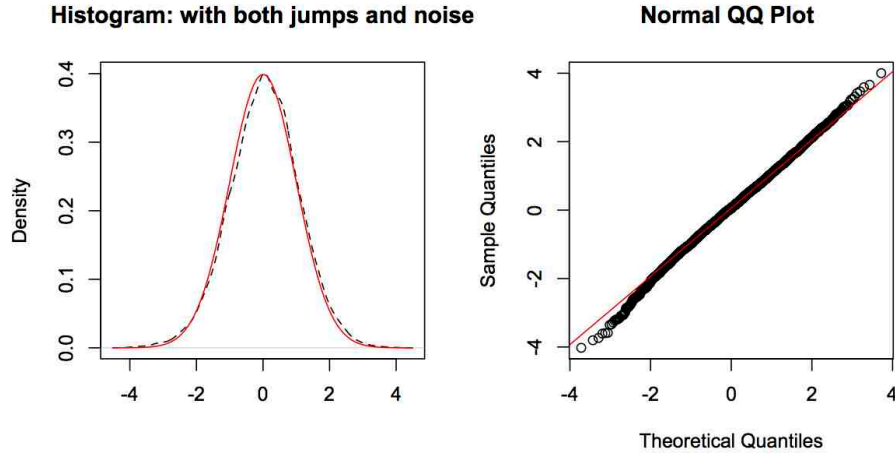


Figure 5: Standardized estimation error densities and QQ plots with both jumps and market microstructure noise, with  $T = 1/252$  (one trading day) and  $n = 10^6$ . Again, we observe a close fit to  $N(0, 1)$  in both plots.

and  $\varpi = 0.49$ .

The first row in Figure 6 shows the densities of the standardized estimation errors. In all three cases, they are very close to the standard normal density, demonstrating that the asymptotic normality holds quite well in finite samples. The second row displays the sample rejection rates for the false null hypothesis of  $DLE = 0$  (finite sample testing power). As one can see, the power increases with sampling frequency, and decreases as the number of jumps gets larger but the jump sizes become smaller.

## 7.5 The Estimation of the Leverage Parameter

In this subsection, we study the estimation of the leverage parameter,  $\rho$ , which is also the correlation between the two Wiener processes appearing in  $X_t$  and  $\sigma_t^2$ . We will consider again the Heston model,

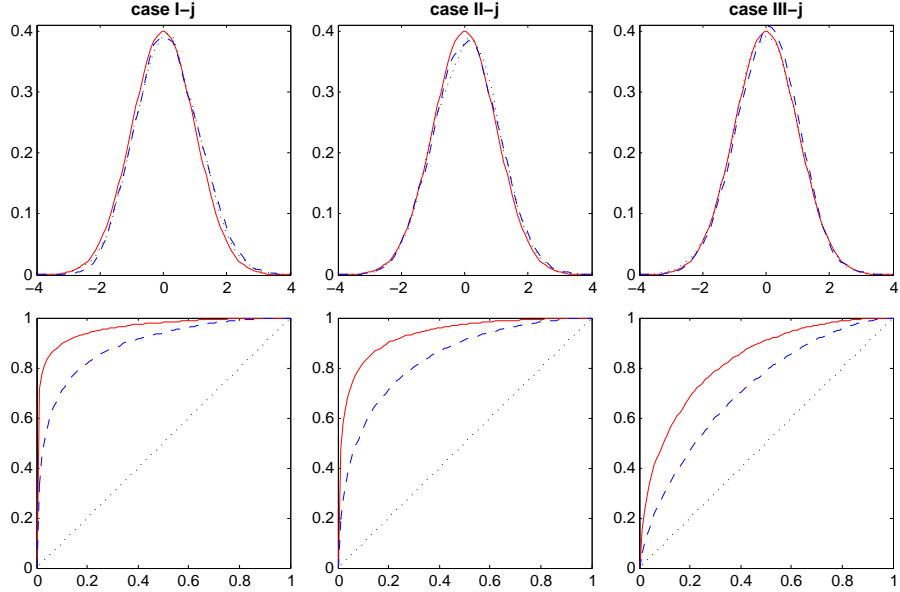


Figure 6: Standardized estimation error densities and testing powers. First row: the red solid line corresponds to the standard normal density, the blue dashed line to the standardized estimation error density in case  $n = 390$  (1-min), and the black dotted line to the standardized estimation error density in case  $n = 4680$  (5-sec). Second row: the  $y$ -axis represents the sample rejection rate while the  $x$ -axis gives the significant level; the red solid line corresponds to  $n = 4680$ , the blue dashed line corresponds to  $n = 390$ .

as in (7.1). Note that in the Heston model

$$\rho = \frac{\langle X, \sigma^2 \rangle_T}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_T}}.$$

We will compare estimation results for  $\rho$  using different ways to estimate the volatility of volatility:

1. Smoothed simple estimator and its unsmoothed counterpart:

$$\langle \sigma^2, \sigma^2 \rangle_{T, \text{naive}}^S = \frac{1}{k_n} \sum_{j=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} (\hat{\sigma}_{j+}^2 - \hat{\sigma}_{j-}^2)^2, \quad \langle \sigma^2, \sigma^2 \rangle_{T, \text{naive}}^U = \sum_{j=1}^{\lfloor T/(k_n \Delta_n) \rfloor - 1} (\hat{\sigma}_{j k_n+}^2 - \hat{\sigma}_{j k_n-}^2)^2.$$

2. Smoothed sophisticated estimator and its unsmoothed counterpart:

$$\langle \sigma^2, \sigma^2 \rangle_{T, \text{soph}}^S = \frac{1}{k_n} \sum_{j=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \left( \frac{3}{2} (\hat{\sigma}_{j+}^2 - \hat{\sigma}_{j-}^2)^2 - \frac{6}{k_n} \hat{\sigma}_j^4 \right),$$

$$\langle \sigma^2, \sigma^2 \rangle_{T, \text{soph}}^U = \frac{1}{k_n} \sum_{j=1}^{\lfloor T/(k_n \Delta_n) \rfloor - 1} \left( \frac{3}{2} (\hat{\sigma}_{j k_n+}^2 - \hat{\sigma}_{j k_n-}^2)^2 - \frac{6}{k_n} \hat{\sigma}_{j k_n}^4 \right), \quad \hat{\sigma}_j^4 = \frac{1}{6 k_n \Delta_n^2} \sum_{l \in I_n^\pm(j)} (\Delta_l^n X)^4.$$

Although this estimator can correct for the bias present in the simple estimator, the subtraction sometimes introduces negative estimated values for the positive volatility of volatility. To avoid this and maintain the bias-correction, one can use instead the ratio statistics below as an alternative estimator (for further details about the asymptotics, refer to [Vetter \(2015\)](#)).

3. Smoothed ratio estimator:

$$\langle \sigma^2, \sigma^2 \rangle_{T, \text{ratio}}^S = \frac{\frac{1}{k_n} \sum_{j=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \frac{3}{2} (\hat{\sigma}_{j+}^2 - \hat{\sigma}_{j-}^2)^2}{\sum_{i=0}^{\ell} r^i}, \quad r = \frac{\sum_{j=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \frac{6}{k_n} \hat{\sigma}_j^4}{\sum_{j=k_n+1}^{\lfloor T/\Delta_n \rfloor - k_n} \frac{3}{2} (\hat{\sigma}_{j+}^2 - \hat{\sigma}_{j-}^2)^2}, \quad (7.3)$$

where  $\ell$  is a given positive integer. Note that when  $\ell = \infty$  and  $r < 1$ , it reduces to case 2. The advantage is that the modified quantity is non-negative.

With these estimators of the volatility of volatility, we study the following 7 estimators of  $\rho$  in

simulations:

$$\left\{ \begin{array}{l} \rho_1 = \frac{\langle X, \sigma^2 \rangle_T}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_T}}, \quad \rho_2 = \frac{[\widehat{X, \sigma^2}]_T^C}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_T}}, \\ \rho_3 = \frac{[\widehat{X, \sigma^2}]_T^C}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_{T, \text{naive}}^U}}, \quad \rho_4 = \frac{[\widehat{X, \sigma^2}]_T^C}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_{T, \text{soph}}^U}}, \\ \rho_5 = \frac{[\widehat{X, \sigma^2}]_T^C}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_{T, \text{naive}}^S}}, \quad \rho_6 = \frac{[\widehat{X, \sigma^2}]_T^C}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_{T, \text{soph}}^S}}, \\ \rho_7 = \frac{[\widehat{X, \sigma^2}]_T^C}{\sqrt{\langle X, X \rangle_T \langle \sigma^2, \sigma^2 \rangle_{T, \text{ratio}}^S}}. \end{array} \right.$$

We will examine how the choice of the tuning parameter  $c$ , introduced in Theorem 3, impacts the MSE of the leverage parameter estimator. In the simulations, the parameterization of the Heston model is set as:  $\theta = 0.06, \eta = 1, \kappa = 10, \rho = -0.8$  and  $\mu = 0.05$ . Furthermore, the total time span is  $T = 21/252$  (monthly) and  $\Delta_n = (1/252)/4680$  (5-sec) or  $\Delta_n = (1/252)/23400$  (1-sec). Each simulation run is repeated 5000 times. We take  $b = 1/2$ .

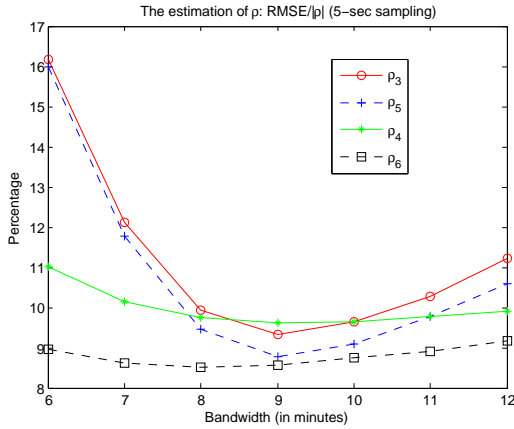


Figure 7: Change of root MSE against a change of the tuning parameter  $c$  (5-sec).

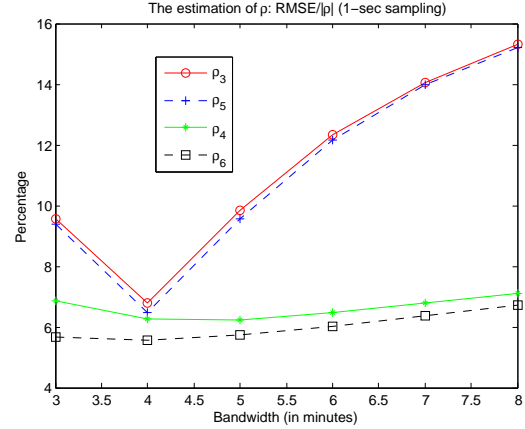


Figure 8: Change of root MSE against a change of the tuning parameter  $c$  (1-sec).



In Figures 7 and 8, we first consider the MSEs of the estimators  $\rho_3$  to  $\rho_6$ . The  $x$ -axis indicates how we vary the number of local observations (controlled by the tuning parameter  $c$  that first appears in Theorem 3) to estimate the spot volatility. If the value on the  $x$ -axis is  $x$ , the number of local observations used in calculating the spot volatility is  $x$  minutes  $\times$  12 (5-sec case) or  $x$  minutes  $\times$  60 (1-sec case). It is clearly apparent that the MSEs of  $\rho_4$  and  $\rho_6$  are more steady over the different specifications of the tuning parameter  $c$ , hence less sensitive to the choice of  $c$  compared to the MSEs of  $\rho_3$  and  $\rho_5$ . Furthermore, the MSEs of  $\rho_4$  and  $\rho_6$  are smaller than those of  $\rho_3$  and  $\rho_5$ , with  $\rho_6$  having the smallest MSE among the four estimators.

The quality of the estimation of  $\rho$  obviously depends on the quality of the estimation of the volatility of volatility. The method used to estimate the volatility of volatility in  $\rho_3$  and  $\rho_5$  introduces biases. It is visible in Figures 7 and 8 that, when the sampling frequency increases, the influence of the bias in  $\rho_3$  and  $\rho_5$  becomes more severe, as indicated by a much larger MSE in the second plot. Although the estimators of the volatility of volatility in  $\rho_4$  and  $\rho_6$  are not biased, they often produce negative values. This negative estimation for a positive quantity is obviously detrimental to the estimation of  $\rho$ . So we also study the adjusted version of the estimator of the volatility of volatility as shown in equation (7.3) and hence the correlation estimator  $\rho_7$  in Figure 9.

More specifically, in Figure 9, we study the behavior of  $\rho_3$ ,  $\rho_6$  and  $\rho_7$ . Here,  $\rho_3$  has the simplest form among the three. The drawback of  $\rho_3$  is that it does not incorporate bias correction. But the advantage of  $\rho_3$  is its simple form and the fact that it always yields positive estimates for the positive volatility of volatility. Compared to  $\rho_3$ ,  $\rho_6$  corrects for the bias in the estimation of the volatility of volatility and as such is a better estimator of  $\rho$ . However,  $\rho_6$  sometimes produces negative values for the estimation of a positive quantity which will be taken into the square root to estimate  $\rho$ . This drawback of  $\rho_6$  makes it impossible to estimate  $\rho$  occasionally. To fix this negativity problem,  $\rho_7$  seems to be the ideal estimator of  $\rho$ :  $\rho_7$  corrects for the bias in the estimation of the volatility of volatility and  $\rho_7$  will also make sure every estimated value of a necessarily positive quantity is

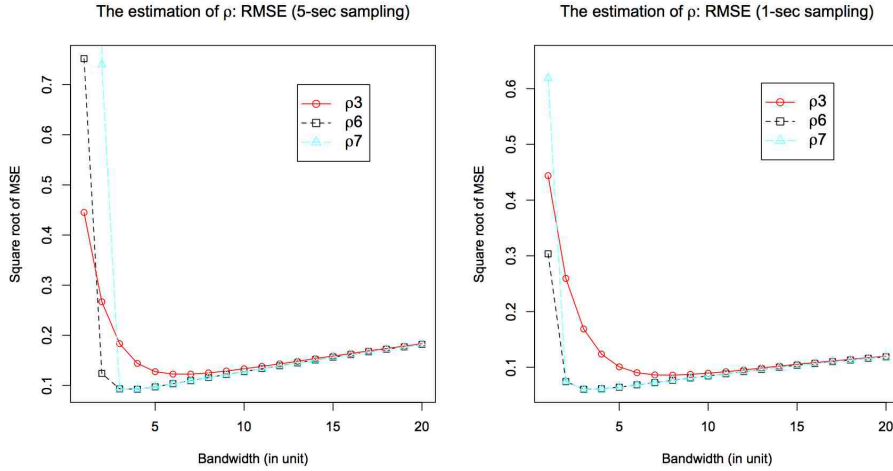


Figure 9: MSE comparison of the estimation of  $\rho$ . This figure plots the change of the root MSE against a change of the tuning parameter  $c$  ( $x$ -axis) for  $\rho_3, \rho_6$  and  $\rho_7$ . All curves exhibit first a decreasing and next an increasing trend in root MSE.

positive. In addition, the MSE of  $\rho_7$  is almost of the same size as that of  $\rho_6$ . This can be seen in Figure 9. Both  $\rho_7$  and  $\rho_6$  achieve the minimum MSE at the same value of the tuning parameter  $c$ . By contrast,  $\rho_3$  has a much larger MSE than the other two estimators (except when  $c = 1$ ). This larger MSE is due to the uncorrected bias.

We also analyze the degree of accuracy the estimators of  $\rho$  reach relative to the true value (taken to be  $-0.8$ ). The densities of  $\rho_1, \dots, \rho_6$  (in the case of their optimal choices of bandwidth shown above) are presented in Figure 10. In the figure, the adjective “true” refers to the idealized hypothetical situation in which the volatility process is actually observable. The density of  $\rho_7$  is almost identical to that of  $\rho_6$  and is therefore omitted.

The results of this subsection show that a non-naive estimator of the volatility of volatility should be applied in empirical studies. They also show that the tuning parameter should be sufficiently large to achieve a good MSE. Of course, the tuning parameter can be optimized by

minimizing the asymptotic variance. Moreover, this simulation study demonstrates how challenging it is to estimate the standardized measure  $\rho_t$ . Even in the case where the process  $\rho_t$  is constant and there are no volatility jumps, it is still rather difficult to recover the constant leverage parameter  $\rho$  in finite samples. Note, however, that this is a shortcoming of the standardized measure, not of the non-standardized measure, i.e., the definition we use for CLE.

## 8 Empirical Results

In our empirical application, we use a data set consisting of 5-seconds Dow Jones equity index data from July 1st, 2003, until February 8th, 2013, covering 2426 trading days (502 weeks). Since the volatility signature plot does not support the existence of microstructure noise in our sample<sup>11</sup>, we only use the non-noise-robust estimators. As in the Monte Carlo simulations, we take the truncation parameters to be  $\alpha = 5 \times \sqrt{BV_T}$  and  $\varpi = 0.49$ , and we take  $b = 1/2$ . The testing results are in Tables 2 and 3.

The whole sample period is subdivided into disjoint 1-week or 4-week time intervals. In each given time interval, the one-sided test for the null hypothesis of  $CLE = 0$  is conducted based on (5.3). The percents of rejections are reported in Table 2. From the table, one can see that, with

---

<sup>11</sup>A possible explanation is that the Dow Jones Index is constructed from very liquid stocks. In fact, the realized variation first stays roughly the same as we increase the sampling frequency from 30-minutes to about 3-minutes, and then decreases (rather than increases) as we further increase the frequency to 5-seconds. This is consistent with the finding that 5-seconds returns are positively correlated, which might be explained by self- and/or mutual excitation in jumps (see e.g. [Aït-Sahalia et al. \(2015\)](#) and [Boswijk et al. \(2015\)](#)). For instance, it could be the case that occurred jumps tend to excite other jumps with the same sign in the short run. If such scenario happens frequently, then the sample autocorrelations can be positive. Yet we don't expect such same-sign-jump-exciting effect to last long, perhaps only a few minutes.

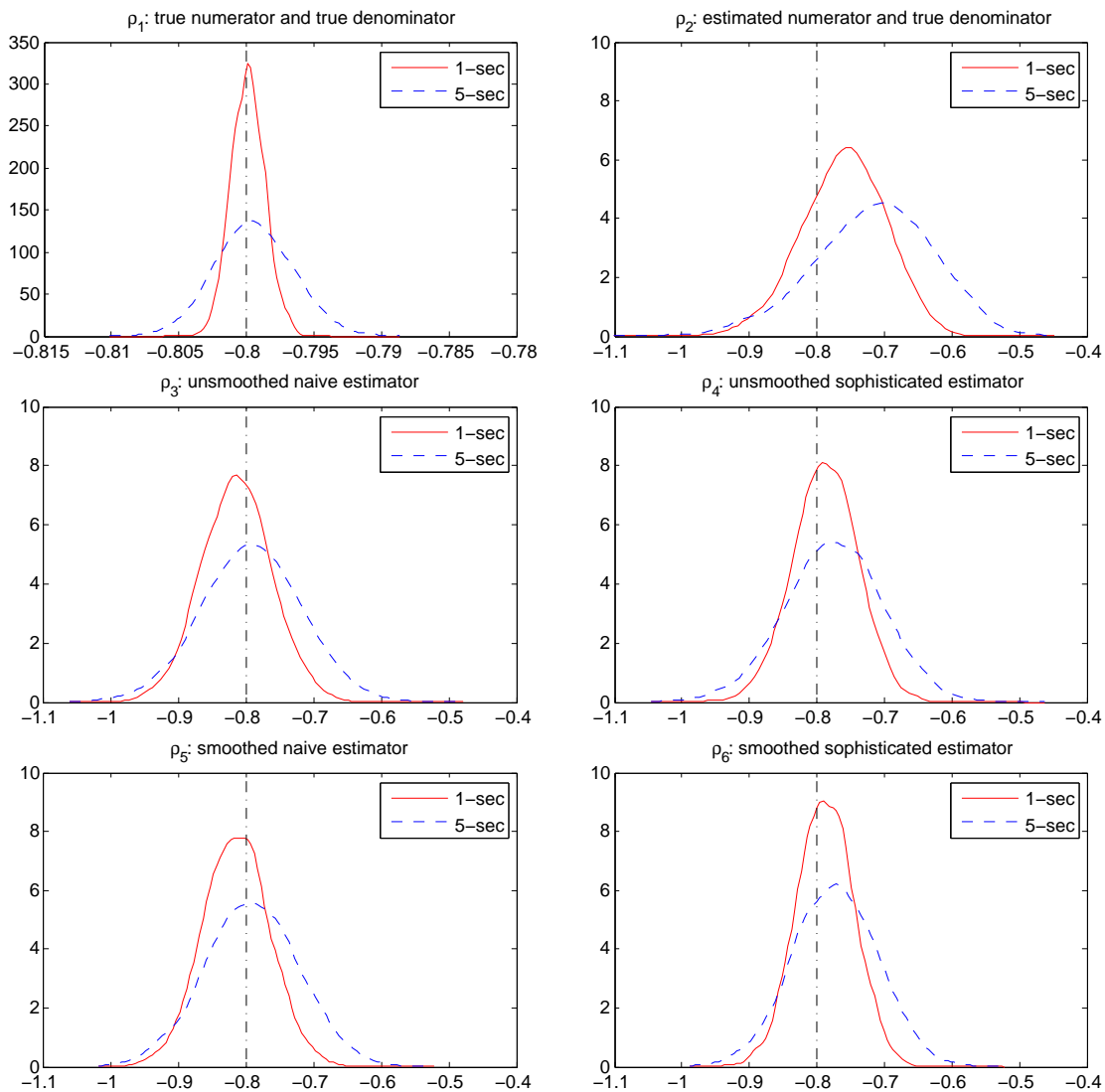


Figure 10: Densities of the leverage parameter estimators.

weekly data, the rejection rates are not very high, while with monthly data, the rejection rates are much higher. This is expected, as a larger sample size yields a higher power, and this finding is consistent with what we have found in the simulation study (cf. Section 7.1). Had there been

no continuous leverage effect within each week, then we would not find more supportive evidence for its presence with monthly data. Therefore, we conclude that there is positive evidence for the presence of the continuous leverage effect.

Table 2: Testing for the absence of CLE

		Rejection rates for a given significance level		
		1%	5%	10%
1-week	$c = 1$	22.91%	35.66%	42.23%
	$c = 2$	24.30%	37.25%	44.02%
4-weeks	$c = 1$	39.68%	60.32%	67.46%
	$c = 2$	36.51%	55.56%	65.08%

As for the DLE, we first employ the procedure introduced in [Aït-Sahalia and Jacod \(2009\)](#) to test whether the returns jump or not within each week.<sup>12</sup> Next, we apply our estimation and testing procedures for the DLE to those weeks containing return jumps. Table 3 displays the testing results for the absence of DLE. Here, we further classify the DLE according to the sign of the return jump. The terms “pos” and “neg” in the third column stand for the discontinuous leverage effects from positive and negative return jumps, respectively, while “all” means all return jumps together. More precisely, “all” refers to the estimator (3.6), whereas “pos” computes (3.6) with the further restriction to  $\Delta_j^n X^{\alpha_n} > 0$  and “neg” is defined similarly. Furthermore, we also consider the tail DLE with three thresholds, namely 0.2%, 0.3% and 0.4%. One readily sees that the evidence for the presence of the various types of DLE is very strong. Another very interesting finding is that, at any threshold level, negative return jumps are more likely to be accompanied by volatility jumps than positive ones.

<sup>12</sup>See also [Fan and Fan \(2011\)](#) for an improved version of the test.

Table 3: Testing for the absence of DLE

Jump size	# of weeks	DLE	Rejection rates for a given significance level					
			$c = 1$			$c = 2$		
			1%	5%	10%	1%	5%	10%
any size	495	all	62.83%	71.31%	74.95%	70.91%	77.58%	79.80%
		pos	82.22%	87.27%	88.48%	81.62%	86.06%	89.29%
		neg	87.68%	91.52%	92.93%	88.48%	90.71%	92.32%
> 0.20%	221	all	57.01%	61.99%	67.42%	59.28%	66.52%	73.30%
		pos	35.29%	41.18%	43.44%	37.10%	40.72%	45.25%
		neg	46.15%	49.77%	53.39%	45.70%	52.04%	56.11%
> 0.30%	112	all	51.79%	58.93%	62.50%	58.93%	65.18%	72.32%
		pos	24.11%	28.57%	30.36%	28.57%	30.36%	33.93%
		neg	39.29%	39.29%	42.86%	42.86%	44.64%	48.21%
> 0.40%	61	all	47.54%	52.46%	55.74%	54.10%	60.66%	62.30%
		pos	18.03%	22.95%	26.23%	26.23%	27.87%	31.15%
		neg	34.43%	39.34%	39.34%	36.07%	40.98%	40.98%

Furthermore, to get an idea about the relative magnitudes of the CLE and DLE, we compute the ratios of  $|\text{CLE}|$ ,  $|\text{DLE}_{\text{pos}}|$  and  $|\text{DLE}_{\text{neg}}|$  to their sum on the basis of 4-week data. We report the resulting sample means and standard errors of these ratios in Table 4. The results show that  $|\text{DLE}_{\text{pos}}|$  and  $|\text{DLE}_{\text{neg}}|$  are larger than  $|\text{CLE}|$  on average, suggesting that the co-jumps between the price and volatility processes are particularly relevant. Specifically,  $|\text{DLE}|$ , which is the sum of  $|\text{DLE}_{\text{pos}}|$  and  $|\text{DLE}_{\text{neg}}|$ , accounts for 73.84% of the leverage effect, whereas  $|\text{CLE}|$  accounts for only 26.15%.

Table 4: Relative magnitude

	CLE	DLE-pos	DLE-neg
mean	0.2615	0.3384	0.4000
std. err.	0.1657	0.1573	0.1337

## References

- Aït-Sahalia, Y., J. Cacho-Diaz, and R. J. Laeven (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics* 117, 585–606.
- Aït-Sahalia, Y., J. Fan, and Y. Li (2013). The leverage effect puzzle: Disentangling sources of bias at high frequency. *Journal of Financial Economics* 109, 224–249.
- Aït-Sahalia, Y. and J. Jacod (2009). Testing for jumps in a discretely observed process. *Annals of Statistics* 37, 184–222.
- Aït-Sahalia, Y. and J. Jacod (2014). *High-Frequency Financial Econometrics*. Princeton University Press.
- Alvarez, A., F. Panloup, M. Pontier, and N. Savy (2012). Estimation of the instantaneous volatility. *Statistical Inference for Stochastic Processes* 15, 27–50.
- Andersen, T., O. Bondarenko, and M. Gonzalez-Perez (2015). Exploring return dynamics via corridor implied volatility. *Review of Financial Studies* 28(10), 2902–2945.
- Bandi, F. M. and R. Renò (2012). Time-varying leverage effects. *Journal of Econometrics* 169, 94–113.

- Black, F. (1976). Studies of stock price volatility changes. In *Proceedings of the 1976 Meetings of the American Statistical Association*, pp. 171–181.
- Bollerslev, T., J. Litvinova, and G. Tauchen (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* 4, 353–384.
- Boswijk, P. H., R. J. Laeven, and X. Yang (2015). Testing for self-excitation in jumps. working paper.
- Bücher, A. and M. Vetter (2013). Nonparametric inference on Lévy measures and copulas. *The Annals of Statistics* 41(3), 1485–1515.
- Curato, I. V. (2015). Estimation of the stochastic leverage effect using the Fourier transform method. Available at SSRN: <http://ssrn.com/abstract=2615271> or <http://dx.doi.org/10.2139/ssrn.2615271>.
- Curato, I. V. and S. Sanfelici (2015). *Measuring the leverage effect in a high frequency trading framework*, Chapter 24, pp. 425–446. Elsevier.
- Eraker, B., M. S. Johannes, and N. Polson (2003). The impact of jumps in equity index volatility and returns. *The Journal of Finance* 58, 1269–1300.
- Fan, J., X. Han, and W. Gu (2012). Estimating false discovery proportion under arbitrary covariance dependence. *Journal of the American Statistical Association* 107(499), 1019–1035.
- Fan, J. and Y. Wang (2008). Spot volatility estimation for high-frequency data. *Statistics and Its Interface* 1, 279–288.
- Fan, Y. and J. Fan (2011). Testing and detecting jumps based on a discretely observed process. *Journal of Econometrics* 164, 331–344.



- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6, 327–343.
- Jacod, J. (2012). Statistics and high frequency data. In A. L. Matthieu Kessler and M. Sorensen (Eds.), *Statistical Methods for Stochastic Differential Equations*, pp. 191–310. Taylor and Francis.
- Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and Their Applications* 119, 2249–2276.
- Jacod, J., M. Podolskij, and M. Vetter (2010). Limit theorems for moving averages of discretized processes plus noise. *Annals of Statistics* 38, 1478–1545.
- Jacod, J. and P. Protter (2011). *Discretization of Processes*. Springer-Verlag.
- Jacod, J. and M. Rosenbaum (2013). Quarticity and other functionals of volatility: Efficient estimation. *Annals of Statistics* 41(3), 1462–1484.
- Jacod, J. and V. Todorov (2010). Do price and volatility jump together? *Annals of Applied Probability* 20, 1425–1469.
- Kalnina, I. and D. Xiu (2016). Nonparametric estimation of the leverage effect using information from derivatives markets. *Journal of the American Statistical Association*, forthcoming.
- Li, J., V. Todorov, and G. Tauchen (2016). Adaptive estimation of continuous-time regression models using high-frequency data. Working paper.
- Li, J. and D. Xiu (2016). Generalized method of integrated moments for high-frequency data. *Econometrica*, forthcoming.

- Veraart, A. E. D. and L. A. M. Veraart (2012). Stochastic volatility and stochastic leverage. *Annals of Finance* 8(2-3), 205–223.
- Vetter, M. (2012). Estimation of correlation for continuous semimartingales. *Scandinavian Journal of Statistics* 39(4), 757–771.
- Vetter, M. (2015). Estimation of integrated volatility of volatility with applications to goodness-of-fit testing. *Bernoulli* 21(4), 2393–2418.
- Wang, D. C. and P. A. Mykland (2014). The estimation of leverage effect with high-frequency data. *Journal of the American Statistical Association* 109, 197–215.
- Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics* 127(2), 165–178.