HIGH FREQUENCY DATA ECONOMETRICS

Kim Christensen

PhD course, Aarhus University, 2016

Introduction

In the previous lecture, the concept of realized variance was introduced in so-called stochastic volatility models, where the efficient price process (observed with or without noise) has continuous sample paths.

Today, we extend the concepts to models with jumps, which makes it possible for the price process to exhibit discontinuous changes.

We review the associated theory of high-frequency estimation in this context, including estimation and inference about the quadratic variation and integrated variance. To the extent possible, we illustrate the concepts on either simulated or real data.

Motivation:

Financial asset prices often move in a way that is hard to reconcile with Brownian motion. If unexpected news arrives in the market, for instance, prices can change a lot in a few seconds. Indeed, the Efficient Market Hypothesis suggests that any new information should be absorbed in the price instantaneously. Jumps are thus natural and appealing ingredients to model such discrete events.

Moreover, if the log-price X_t is assumed to be a Lévy process, then the only such process with continuous paths is the Brownian motion (possibly with a drift). Thus, at least according to this notion, we can argue that models without jumps are “exceptions” to the rule that asset prices jump.

How often and by how much is debatable.

Figure 1: Jump in Kraft Foods (KFT) at 5-minute frequency.

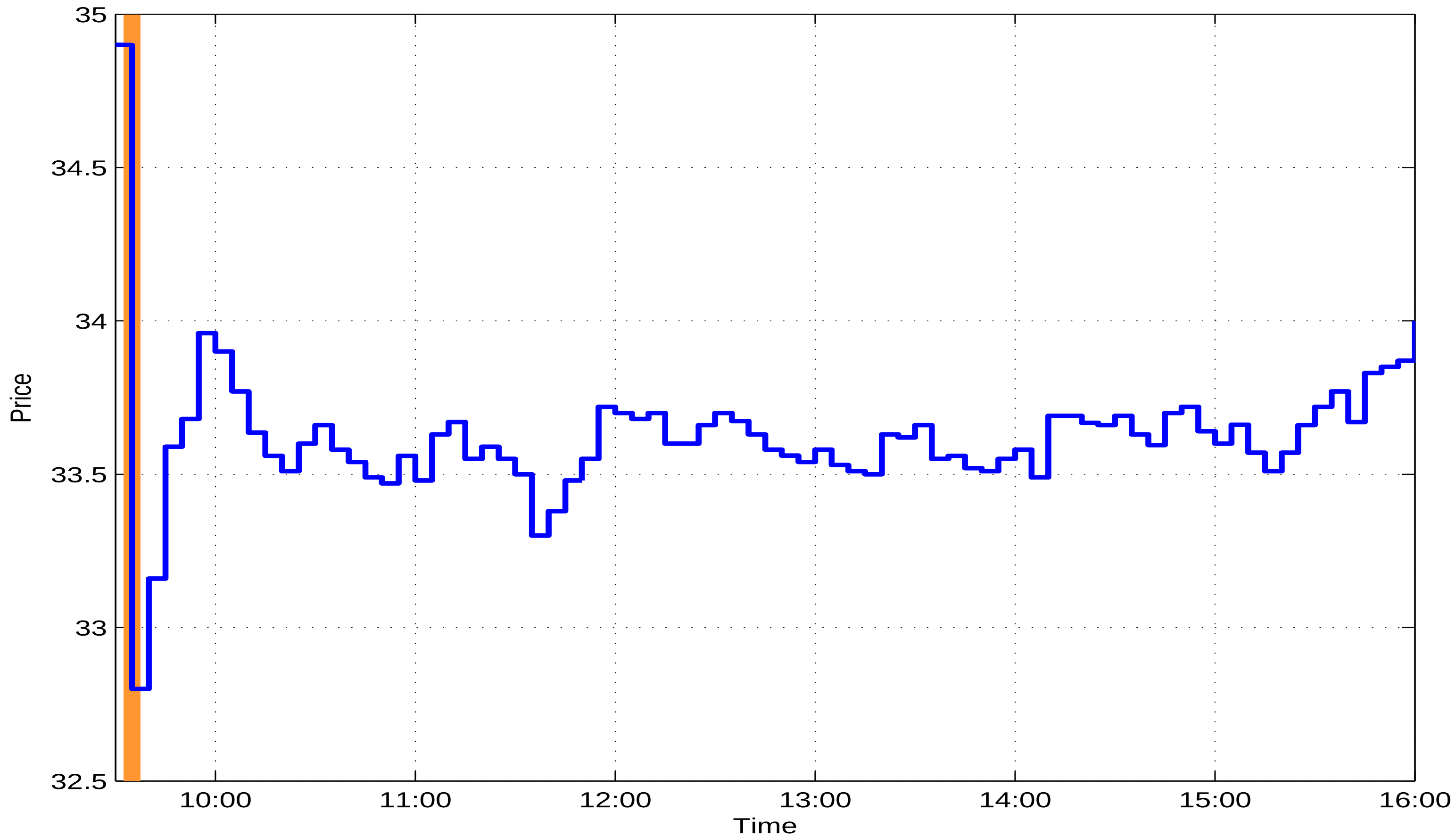
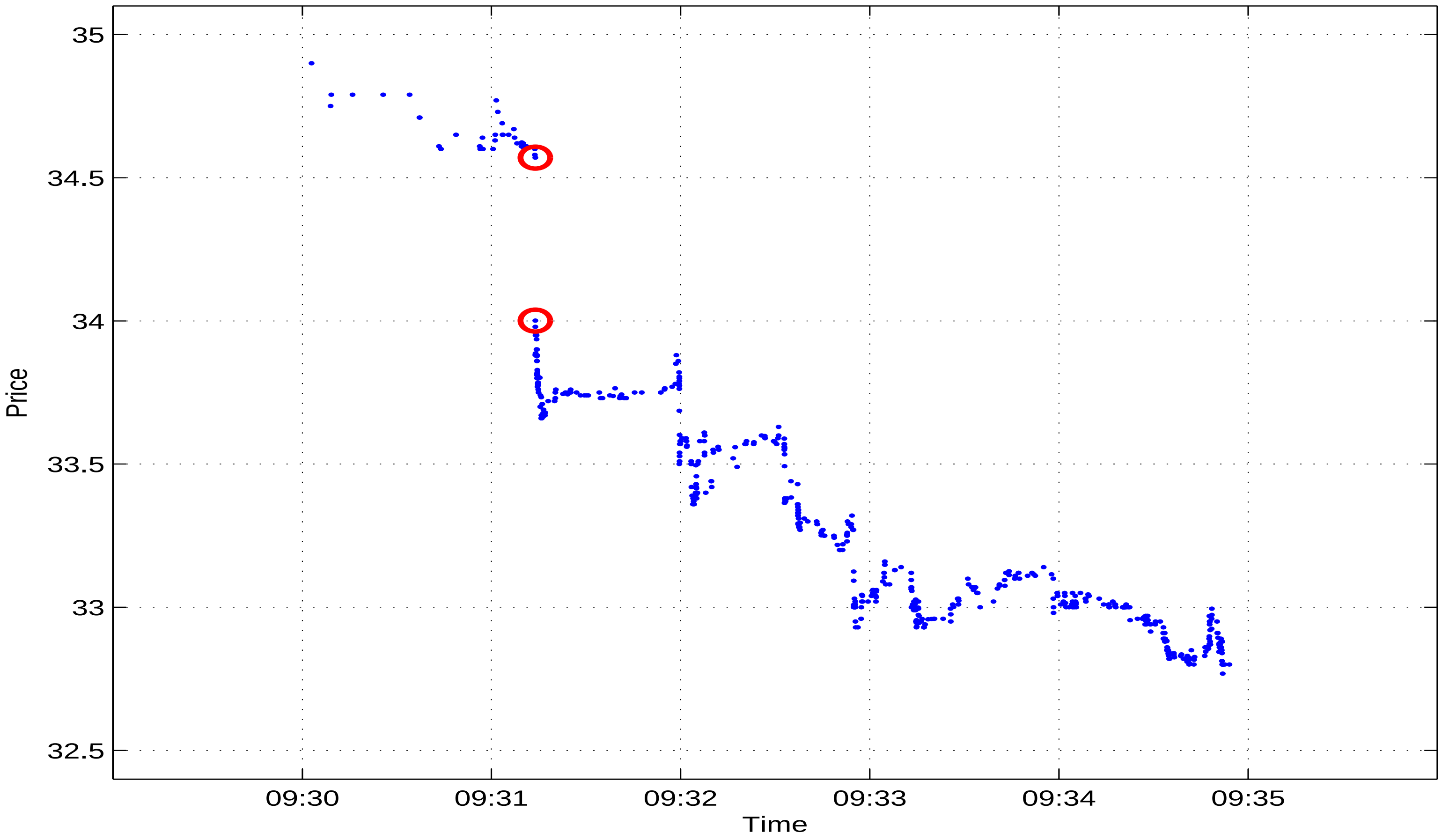


Figure 2: Jump in Kraft Foods (KFT) at tick frequency.



The X process:

We assume that

$$X_t = X_0 + \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + \sum_{s=1}^{N_t} \Delta X_s, \qquad t \geq 0, \tag{1}$$

where

X_t: log-price at time t,

a_t: drift term,

σ_t: volatility process ($\sigma_t > 0$),

W_t: standard Brownian motion,

N_t: counting process,

ΔX_s: jump size.

So, N_t represents the total number of jumps in X that have occurred up to time t and ΔX_s denotes the corresponding, individual jump sizes.

We assume N_t is a finite-activity process (i.e., there is a finite number of jumps on finite time intervals, almost surely)...

... but most of what we talk about can be extended directly to infinite-activity, but finite-variation, jump processes.

Figure 3: Sample path of efficient log-price X_t.

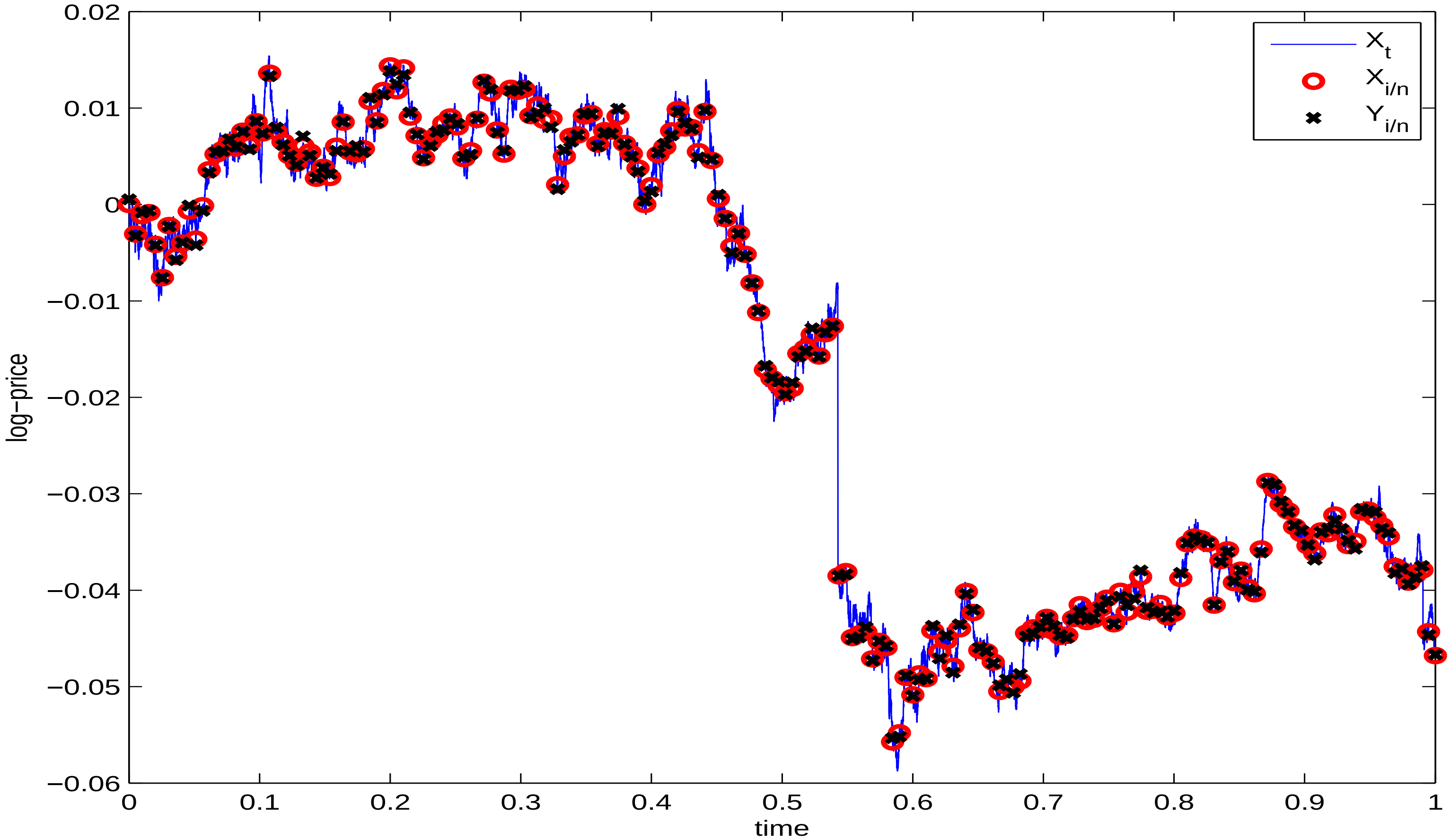


Figure 4: Sample path of diffusive variance.

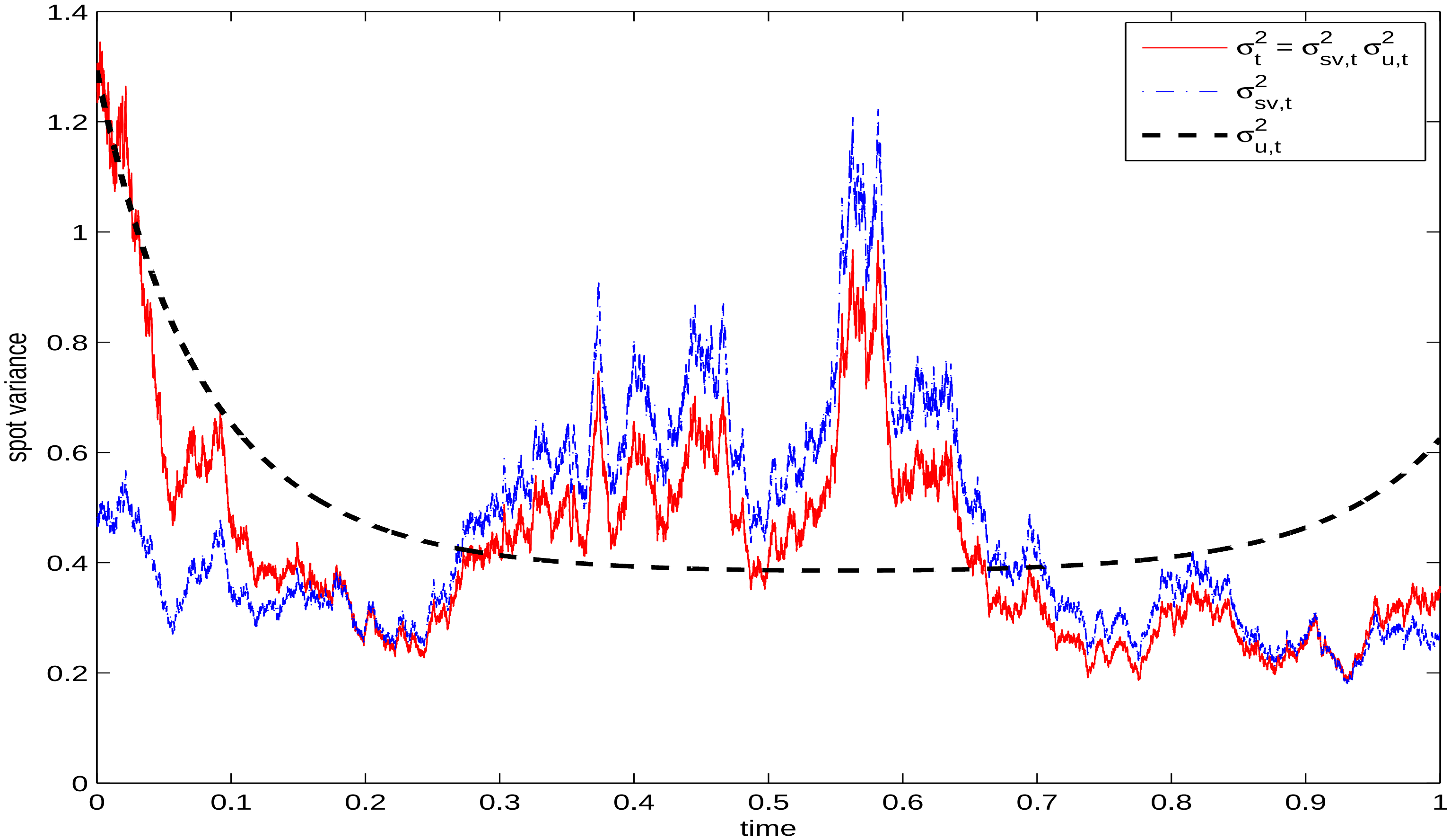
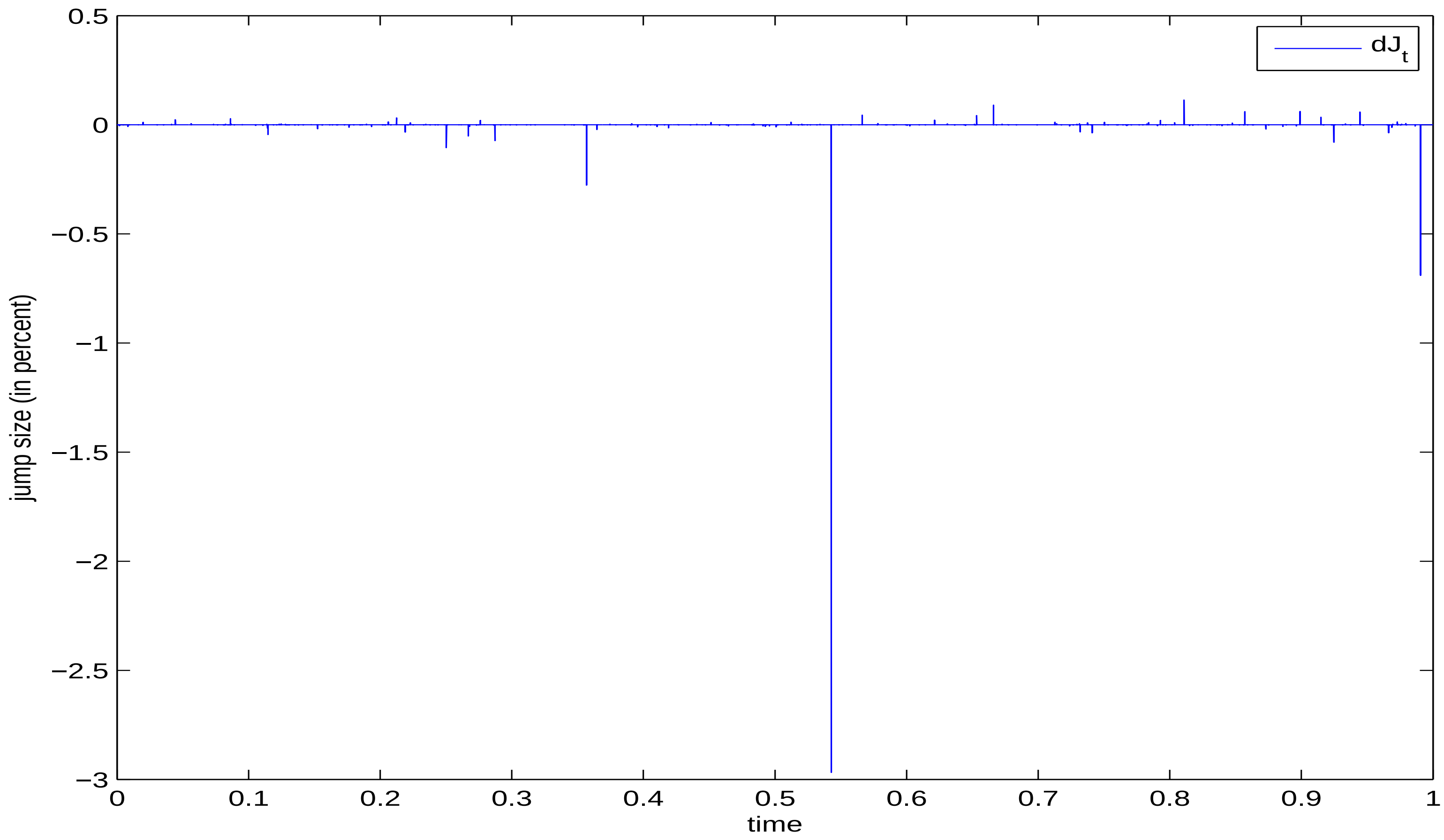


Figure 5: Sample path of jump process.



Quadratic variation

We define the quadratic variation (QV):

$$[X]_t = \underset{N \to \infty}{\text{p-lim}} \sum_{i=1}^{N} (X_{t_i} - X_{t_{i-1}})^2, \tag{2}$$

for any sequence of partitions $0 = t_0 < t_1 < \ldots < t_N = t$ with $\max(t_i - t_{i-1}) \to 0$ as $N \to \infty$ (e.g., Protter, 2004).

It can be shown that the QV process is always well-defined, if X is a semimartingale (which any arbitrage-free price process must be, e.g., Back (1991); Delbaen and Schachermayer (1994)).

It is this fundamental result from stochastic calculus that has motivated the increasing use of high-freuquency data to estimate financial volatility.

If X has the form (1):

$$[X]_t = \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{0 \leq s \leq t} |\Delta X_s|^2, \tag{3}$$

where $\Delta X_s = X_s - X_{s-}$.

The QV is thus equal to the integrated variance (IV) plus the sum of the squared jumps (JV). The model thus introduces an extra layer of risk via the jump component—compared to the previous lecture.

In fact, (3) is true for any semimartingale, irrespective of how crazy the jump process is (and it can be very wild indeed).

The data

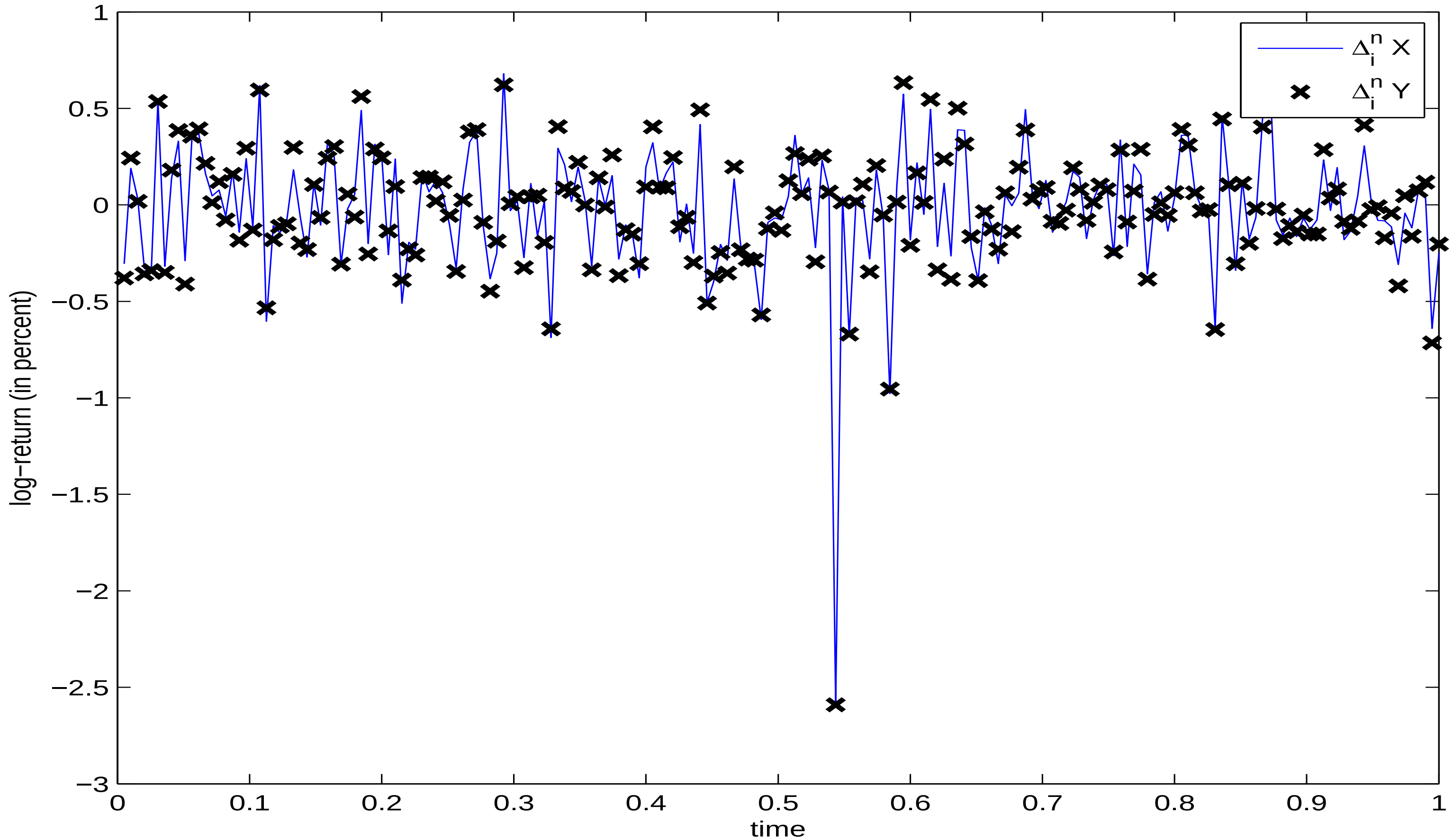
Throughout the rest of this lecture, we restrict attention to the unit interval by setting $t = 1$. We then assume that the process is observed at equidistant time points $t_i = i/N$, for $i = 0, 1, \ldots, N$.

We compute the increments of X:

$$\Delta_i^N X = X_{i/N} - X_{(i-1)/N}, \quad \text{for } i = 1, \ldots, N. \tag{4}$$

Note that as X_t is a log-price, $\Delta_i^N X$ can be interpreted as a continuously compounded return between time $t_i = i/N$ and $t_{i-1} = (i-1)/N$.

Figure 6: Efficient log-return $\Delta_i^N X$.



Realized variance

You already encountered the realized variance (RV) in the previous lecture. Recall that this estimator is defined as:

$$RV^N = \sum_{i=1}^{N} |\Delta_i^N X|^2. \tag{5}$$

RV^N is the sum of the squared increments of X. Note that from (2) – (3), it follows that RV^N is a consistent estimator of $[X]_1$:

$$RV^N \xrightarrow{p} [X]_1 \text{ as } N \to \infty. \tag{6}$$

Thus, RV^N no longer estimates the IV, but the combined return variation induced by the diffusive volatility and the jump part.

If X was a stochastic volatility model (i.e., without the jump part) $\mathrm{d}X_t = a_t \mathrm{d}t + \sigma_t \mathrm{d}W_t$, then RV^N admits a CLT:

$$\sqrt{N}\left(RV^N - \int_0^1 \sigma_s^2 \mathrm{d}s\right) \xrightarrow{d_s} MN\left(0, 2\int_0^1 \sigma_s^4 \mathrm{d}s\right) \tag{7}$$

$\int_0^1 \sigma_s^4 \mathrm{d}s$: Integrated quarciticy (IQ)

When X has a jump component, and assuming that there are no common jumps in σ and J, this results changes to (e.g., Veraart, 2011):

$$\sqrt{N}\left(RV^N - [X]_1\right) \xrightarrow{d_s} MN\left(0, 2\int_0^1 \sigma_s^4 \mathrm{d}s + 4\sum_{0\leq s\leq 1} \sigma_s^2 |\Delta X_s|^2\right), \tag{8}$$

Note that the estimation error increases—i.e., it is more difficult to estimate the QV—with high volatility and large jumps.

Jump-robust measurement of IV

This implies that—in the presence of jumps—RV is not informative about the IV itself. Nonetheless, it may still be interesting to estimate that piece of the QV. This can be useful, for example, if we are looking to:

$\rightarrow$ Split the IV from the JV in order to measure their relative contribution to QV,

$\rightarrow$ Devise a statistical test for the presence of a jump component,

$\rightarrow$ Build a time series model for the IV.

There are two main approaches, which can accomplish this goal:

1. Bi- or multi-power variation

2. Truncation

We cover each approach in turn. Alternative ways of estimating the IV in the presence of jumps not covered in this class, include the quantile-based estimators of Christensen, Oomen, and Podolskij (2010) and Andersen, Dobrev, and Schaumburg (2012).

Recent papers that show how to improve the finite sample jump robustness of the BV or make it more efficient are also beyond our scope; see Corsi, Pirino, and Renò (2010) and Mykland, Shephard, and Sheppard (2012).

Bipower variation

Barndorff-Nielsen and Shephard (2004) proposed the bipower variation (BV) estimator, which is an intuitive way to separate the diffusive- and jump-variation components of $[X]_1$. The (1,1)-bipower variation (see below for a more general version) is defined as:

$$BV^N = \frac{N}{N-1}\frac{\pi}{2}\sum_{i=2}^{N} |\Delta_{i-1}^N X||\Delta_i^N X|. \tag{9}$$

In words, BV replaces $|\Delta_i^N X|^2$ of RV by multiplying together contiguous high-frequency returns... suitably normalized by $\pi/2$.

The factor $N/(N-1)$ is applied to improve the finite sample properties of BV^N. It corrects for the "loss" of a summand.

It holds that, if X follows (1), as $N \to \infty$:

$$BV^N \xrightarrow{p} \int_0^1 \sigma_s^2 \mathrm{d}s. \tag{10}$$

So BV^N is a consistent and jump-robust measure of IV, irrespective of any jumps in X.

Intuition:

Assume there is a jump in $\Delta_i^N X$. Then, with a probability going to one, there is none in $\Delta_{i-1}^N X$. That is, each jump term is—sooner or later—paired with a "continuous" return. The jump term is $O_p(1)$, while the other is $O_p(N^{-1/2})$ (and so is the product).

As there are only finitely many terms with jumps (in finite time intervals), all jumps terms are knocked out of the plim.

It therefore suffices to look at "no jump" terms. Assume σ is constant and no drift, so that $|\Delta_{i-1}^N X| = \sigma|\Delta_{i-1}^N W|$ and $|\Delta_i^N X| = \sigma|\Delta_i^N W|$. Then, $|\Delta_{i-1}^N X||\Delta_i^N X| = \sigma^2|\Delta_{i-1}^N W||\Delta_i^N W|$.

$E(|\underbrace{\sqrt{N}\Delta_i^N W}_{\stackrel{d}{=}N(0,1)}|) = \sqrt{2/\pi}$ explains the normalization.

Taken together:

$$RV^N - BV^N \xrightarrow{p} \sum_{0\leq s\leq 1} |\Delta X_s|^2 \tag{11}$$

i.e., the difference between RV and BV is a consistent estimator of the sum of the squared jumps

Figure 7: 5-minute RV – annualized std.

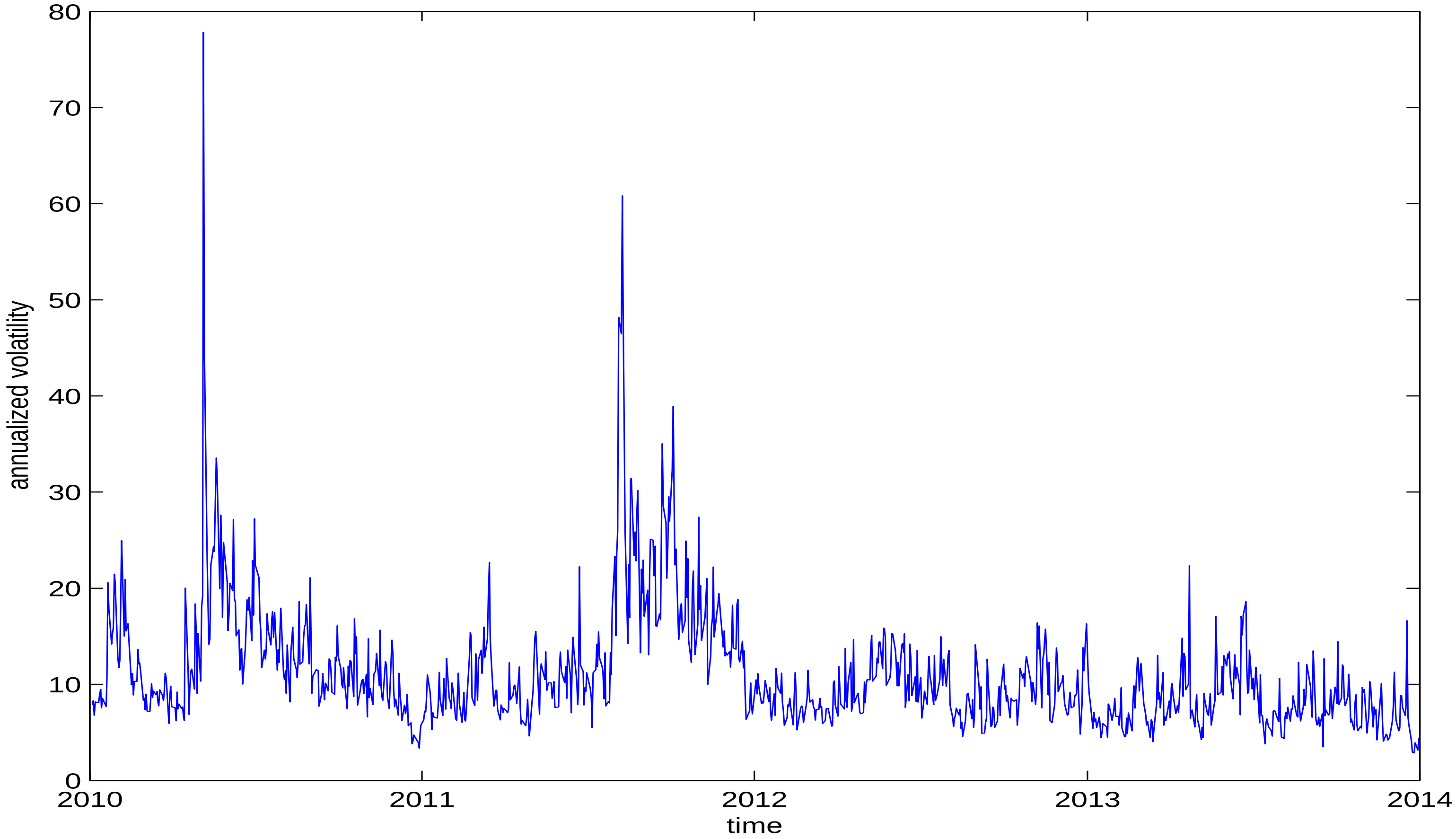
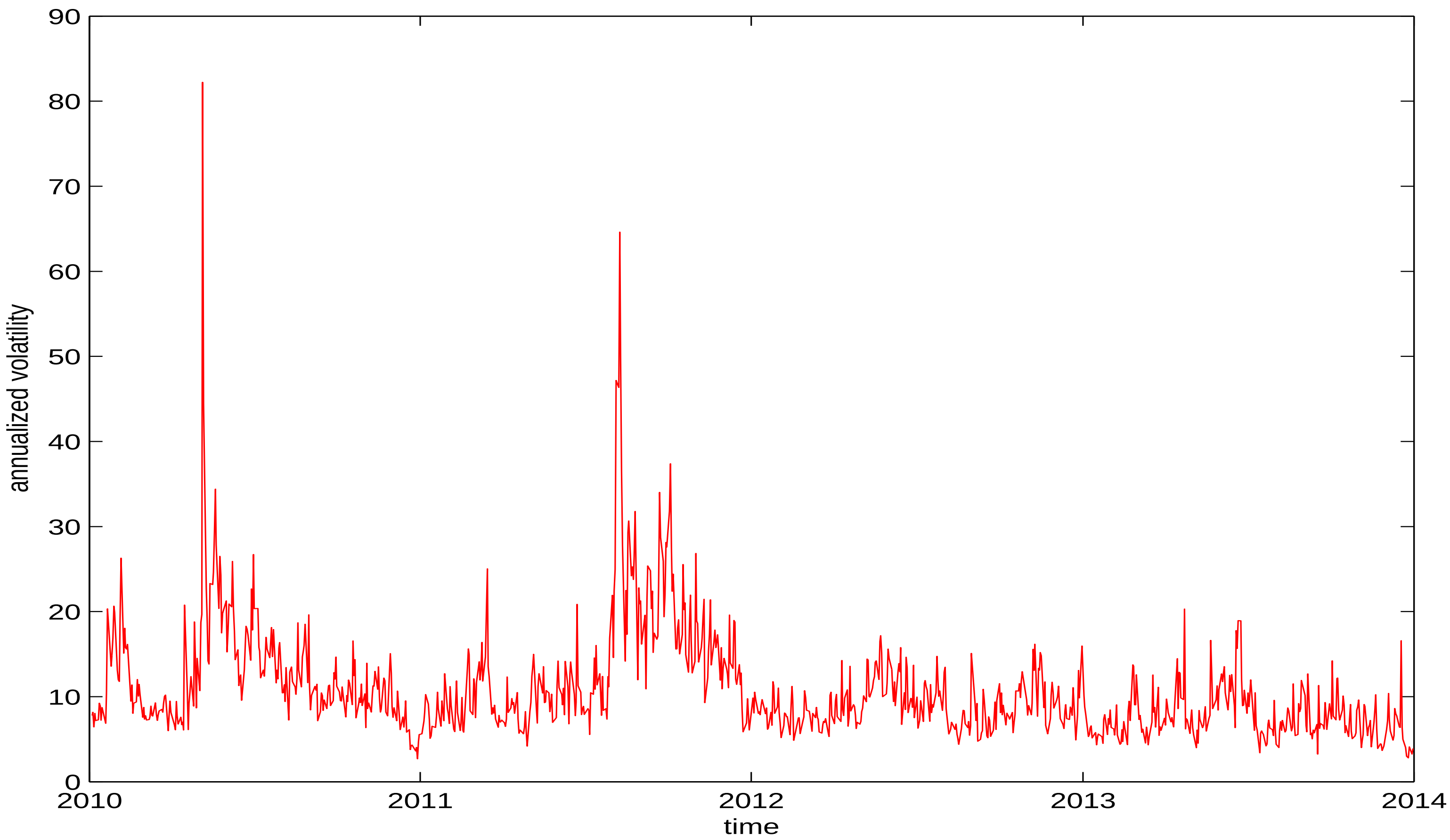


Figure 8: 5-minute BV – annualized std.



Jump proportion

This allows to measure the contribution of jumps to $[X]_1$. In particular, the jump variation (JV) can be found as:

$$JV = 1 - \frac{\int_0^1 \sigma_s^2 \mathrm{d}s}{[X]_1}, \tag{12}$$

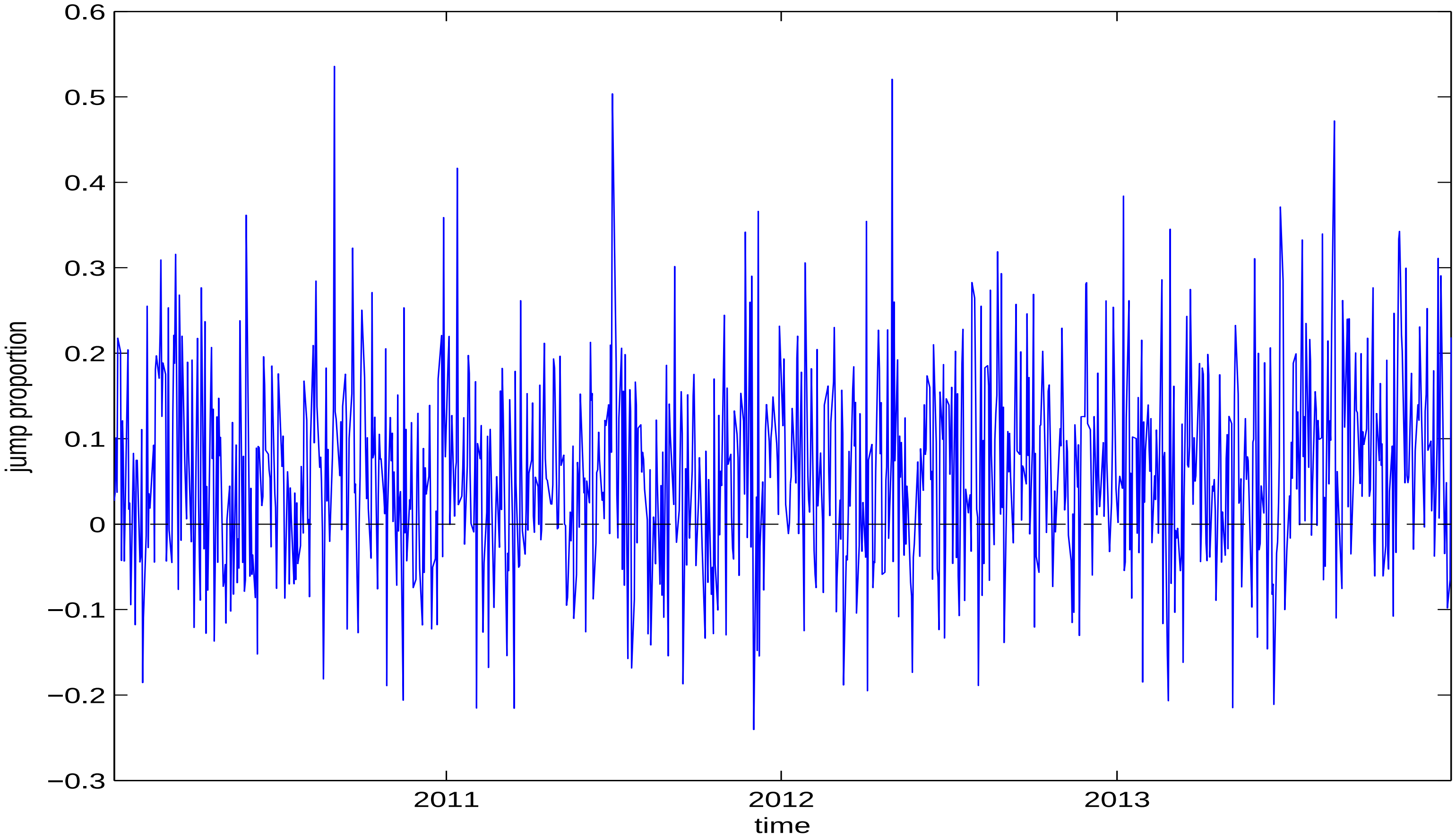
i.e., the proportion induced by discrete price changes.

The empirical counterpart:

$$\hat{JV} = 1 - \frac{BV^N}{RV^N}, \tag{13}$$

This result has been widely exploited in empirical work to study the role of the jump component (e.g., Table 1 in Christensen, Oomen, and Podolskij, 2014, for an overview of the literature).

Figure 9: Jump proportion from 5-minute data.



Jump testing

We can extend the above to construct a statistical test for the presence of jumps. In particular, under the null hypothesis that X is continuous, it follows from Barndorff-Nielsen and Shephard (2006) that:

$$\sqrt{N}\begin{pmatrix} RV^N - \int_0^1 \sigma_s^2 \mathrm{d}s \\ BV^N - \int_0^1 \sigma_s^2 \mathrm{d}s \end{pmatrix} \xrightarrow{d_s} MN\left(\mathbf{0}, \int_0^1 \sigma_s^4 \mathrm{d}s \times \begin{bmatrix} 2 & 2 \\ 2 & 2.6 \end{bmatrix}\right) \tag{14}$$

This implies that if there are no jumps in X, so that $[X]_1 = \int_0^1 \sigma_s^2 \mathrm{d}s$, BV^N is asymptotically less efficient than RV^N (RV^N is the MLE in the parametric version $X_t = \sigma W_t$ of the problem, so it is hard to beat).

i.e., robustness has a cost (was that expected?).

Applying the delta rule and the properties of stable convergence in law, we achieve the following t-statistic (under the null):

$$T^N = \frac{\sqrt{N}(RV^N - BV^N)}{\sqrt{0.6 \int_0^1 \sigma_s^4 \mathrm{d}s}} \xrightarrow{d} N(0,1). \tag{15}$$

Tables with critical values from the standard normal can be used to assess the statistical significance of the deviation from the null, against a right-tailed, one-sided alternative...

... because, conditional on the presence of jumps in X, $RV^N - BV^N = \sum_{0 \le s \le 1} |\Delta X_s|^2 > 0$, so $T^N \to \infty$. The test thus has unit power and is consistent, asymptotically.

Barndorff-Nielsen and Shephard (2006) propose other transformations of the bivariate CLT, which exhibit better finite sample properties.

Multi-power variation

To implement the jump test, the IQ has to be estimated in a jump-robust fashion (else the power is severely impaired). It appears that we can handle this problem with a jump-robust estimator, say, of the form $\sum_{i=2}^{N} |\Delta_{i-1}^{N} X|^2 |\Delta_i^N X|^2 \ldots$ but this turns out not to work.

To do it right, we adopt the class of multipower variation:

$$MPV^N(r) = \frac{N}{N-I} \frac{1}{\prod_i \mu_{r_i}} N^{\sum r_i/2-1} \sum_{i=I}^{N} \prod_{j=1}^{I} |\Delta_{i+j-1}^{N} X|^{r_i} \qquad (16)$$

where $r = (r_1, \ldots, r_I)$ with $r_i > 0$ is an index of powers, while $\mu_x = E(|N(0,1)|^x) = \pi^{-1/2} 2^{x/2} \Gamma((x+1)/2)$ and Γ is the Gamma function.

Thus, multipower variation generalizes the bipower estimator to arbitrary lag lengths and general powers. $N/(N-I)$ is a finite sample term.

Note that:

$\rightarrow$ The (1,1)-bipower variation in (9) has $I = 2$ and $r_i = 1$.

$\rightarrow$ An alternative estimator in this class, which is analyzed below, is the tripower variation (TV). It takes $I = 3$ and $r_i = 2/3$.

Set $r_+ = \max r_i$. Then, if $r_+ < 2$:

$$MPV^N(r) \xrightarrow{p} \int_0^1 \sigma_s^{\sum r_i} \mathrm{d}s. \tag{17}$$

On the other hand, if $r_+ = 2$, the estimator is not jump-robust and converges in probability to some complicated mixture of diffusive volatility and the jump part (e.g., Barndorff-Nielsen and Shephard, 2004)

This shows why the BV “adaption” above cannot estimate IQ robustly. To do that, we need to include at least one more lag.

Two examples often used in practice:

$\rightarrow$ Tripower quarticity (TQ): $I = 3$ and $r_i = 4/3$.

$\rightarrow$ Quadpower quarticity (QQ): $I = 4$ and $r_i = 1$.

Figure 10: Jump t-statistic 5-minute data (via quadpower quarticity).

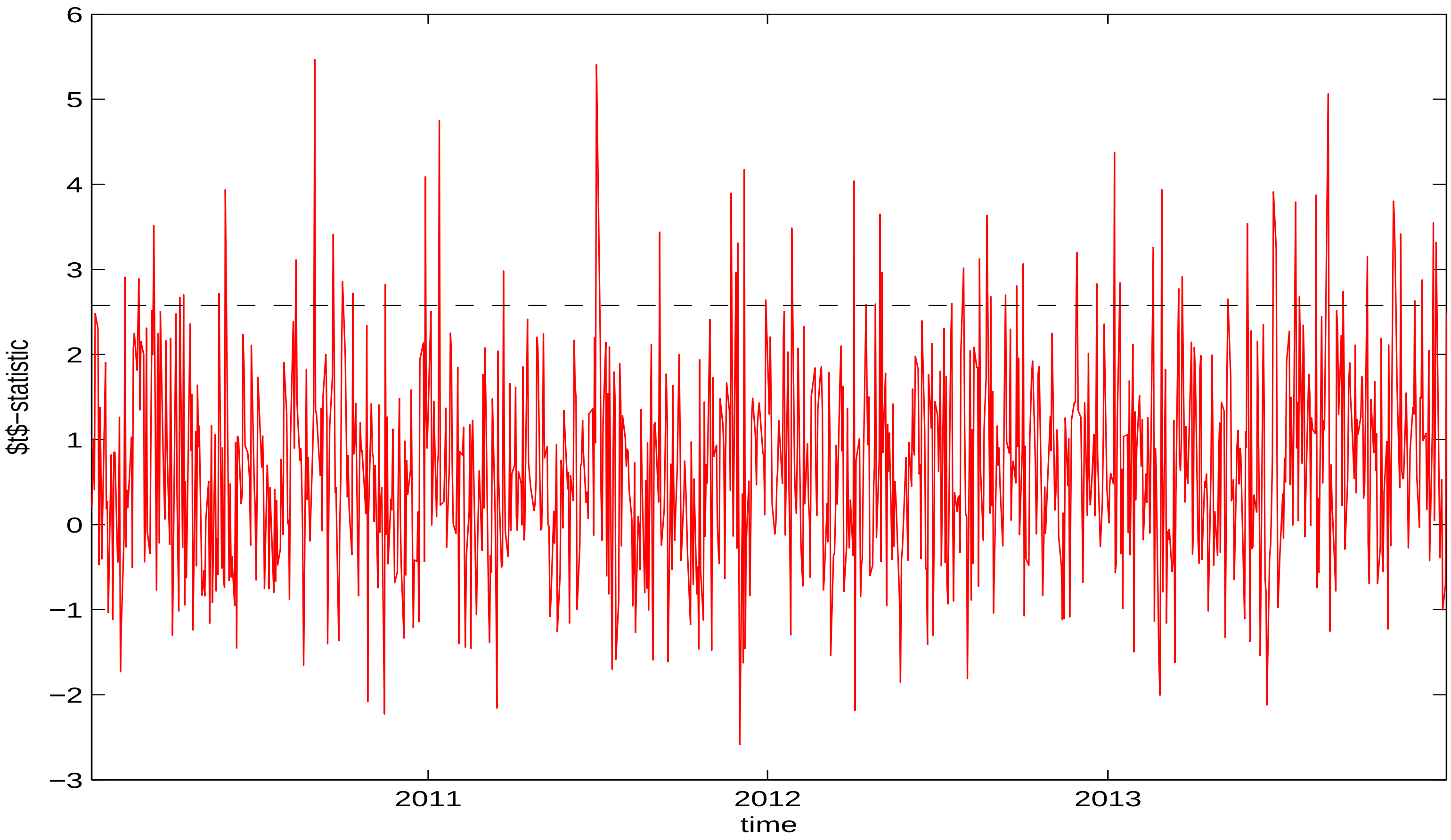
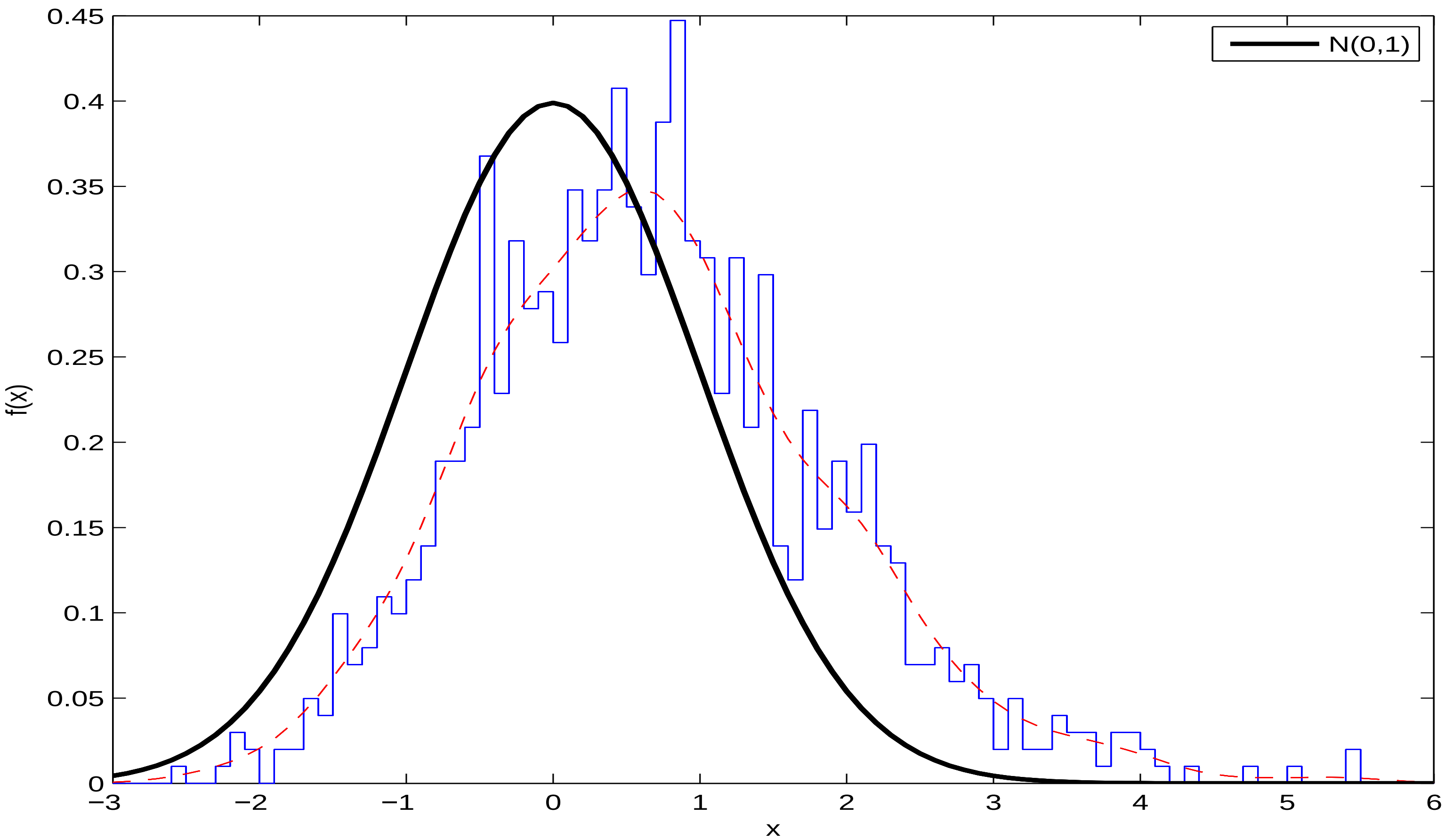


Figure 11: Kernel density estimate of t-statistic.



Jump-robust inference about IV

The above stochastic convergence of multipower variation can be extended to a central limit theorem, if $r_+ < 1$:

$$\sqrt{N}\left(MPV^N(r) - \int_0^1 \sigma_s^{\sum r_i} \mathrm{d}s\right) \xrightarrow{d_s} MN\left(0, c_{\mu_r} \int_0^1 \sigma_s^{2\sum r_i} \mathrm{d}s\right), \quad (18)$$

where c_{μ_r} is a known (but complicated) function of μ_{r_i}.

This implies that we cannot draw inference about the IV (e.g., construct a $(1-\alpha)\%$ confidence interval) with BV^N. Indeed, the limiting distribution of BV^N is not even mixed Gaussian (e.g., Vetter, 2010).

A substitute is tripower variation (TV):

$$TV^N = \frac{N}{N-2}\frac{1}{\mu_{2/3}^3}\sum_{i=3}^{N}|\Delta_{i-2}^N X|^{2/3}|\Delta_{i-1}^N X|^{2/3}|\Delta_i^N X|^{2/3}, \tag{19}$$

Then, in model (1),

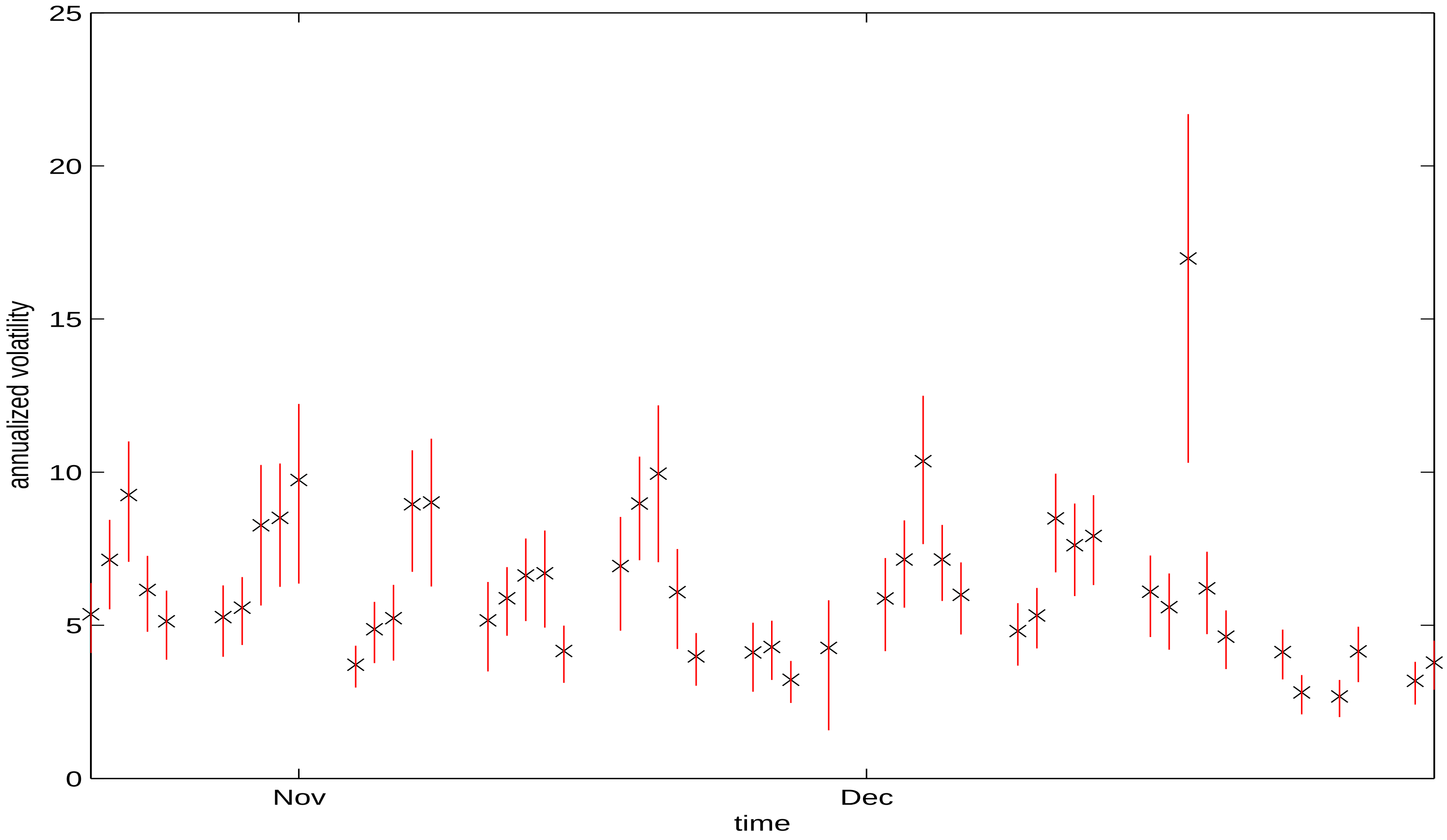
$$\sqrt{N}\left(TV^N - \int_0^1 \sigma_s^2 \mathrm{d}s\right) \xrightarrow{d_s} MN\left(0, 3.06\int_0^1 \sigma_s^4 \mathrm{d}s\right). \tag{20}$$

A two-sided $(1-\alpha)\%$ CI for the IV can then be computed by:

$$TV^N \pm q_{\alpha/2}\sqrt{\frac{3.06\hat{IQ}}{N}}, \tag{21}$$

where $q_{\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution and $\hat{IQ}$ is a consistent estimator of IQ.

Figure 12: IV point estimate and 95% CI.



Truncation

An alternative procedure, which can handle the influence of jumps, is to filter returns that are abnormally large compared to what was to be expected, if they were truly drawn from a diffusion process.

The threshold RV:

$$TRV^N = \sum_{i=1}^{N} |\Delta_i^N X|^2 \mathbf{1}_{\{|\Delta_i^N X| < \tau_N\}}, \tag{22}$$

$\mathbf{1}$: indicator function $\quad \tau_N$: positive real-valued function.

The strategy in words: inspect the magnitude of the increment. If it is "too large" (i.e., exceeds τ_N) kill it, otherwise include it in the sum.

Idea was proposed by, e.g., Aït-Sahalia and Jacod (2009a,b); Mancini (2004, 2009). With an appropriate selection of τ_N, all the jumps returns can be annihilated, while the diffusive returns are preserved.

To achieve this, τ_N needs to fulfill:

$$\lim_{N\to\infty} \tau_N = 0 \qquad \text{and} \qquad \lim_{N\to\infty} \frac{\log N}{N\tau_N} = 0, \tag{23}$$

That is, τ_N should vanish at a rate, which is slower than the modulus of continuity of Brownian motion. This ensures that, as $N \to \infty$, τ_N shrinks at a lower pace than the diffusive terms, so that the filtering does not impact those terms that are not influenced by the jump component.

Common approach:

$$\tau_N = q_{1-\alpha/2} * \hat{\sigma} * N^{-\varpi}, 0 < \varpi < 0.5, \tag{24}$$

where $q_{1-\alpha/2}$ is a quantile from the standard normal distribution and $\hat{\sigma}$ is an estimator of the local standard deviation (e.g., $\sqrt{BV^N}$).

$\hat{\sigma}$: scale parameter.

$q_{1-\alpha}$: number of std.'s.

ϖ : rate parameter.

Note that, as such, this requires one to "pre-estimate" volatility before setting a truncation level, leading to an iterative procedure.

Asymptotic distribution of TRV:

$$\sqrt{N}\left(TRV^N - \int_0^1 \sigma_s^2 \mathrm{d}s\right) \xrightarrow{d_s} MN\left(0, 2\int_0^1 \sigma_s^4 \mathrm{d}s\right) \tag{25}$$

Truncation estimator of IQ:

$$TRQ^N = \frac{N}{3}\sum_{i=1}^{N} |\Delta_i^N X|^4 \mathbf{1}_{\{|\Delta_i^N X| < \tau_N\}} \xrightarrow{p} \int_0^1 \sigma_s^4 \mathrm{d}s, \tag{26}$$

i.e., a realized (multi-)power variation estimator with truncation on top.

The TRV is asymptotically efficient (in the Cramer-Rao lower bound sense). The intuition is that, because τ_N decays slowly enough, it doesn't do anything asymptotically, if the process never jumps.

This CREATES a problem for jump testing, because the asymptotic distribution $\sqrt{N}(RV^N - TRV^N)$ is degenerate. Podolskij and Ziggel (2010) propose a wild bootstrap procedure to deal with this problem.

Corsi, Pirino, and Renò (2010) propose to combine the force of bipower variation with truncation in order to do inference about the JV.

Microstructure noise

In practice, assets are not traded in a frictionless market. Prices are affected by market imperfections, such as bid-ask spreads, price discreteness, etc. (e.g., Black, 1986; Niederhoffer and Osborne, 1966; Roll, 1984)

The combination of these effects leads to marked differences between real data and those generated by a jump-diffusion model.

We model this as:

$$Y_{i/N} = X_{i/N} + \epsilon_{i/N}, \tag{27}$$

where X is defined as in (1), while $\epsilon_{i/N}$ is a i.i.d. noise process (independent of X) with finite variance.

Figure 13: Noisy log-return $\Delta_i^N Y$.

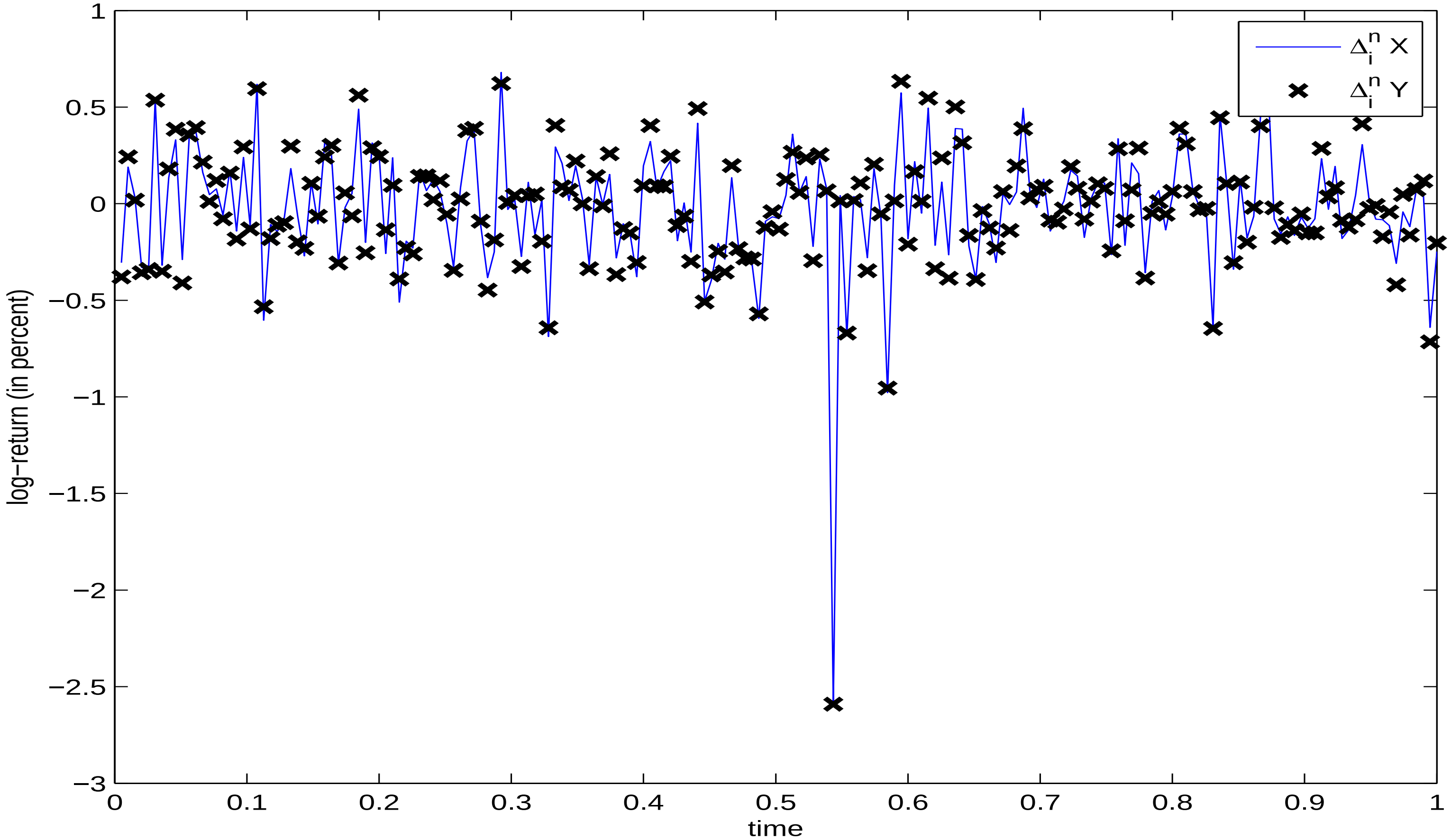


Figure 14: Sample path of noisy log-price Y_t.

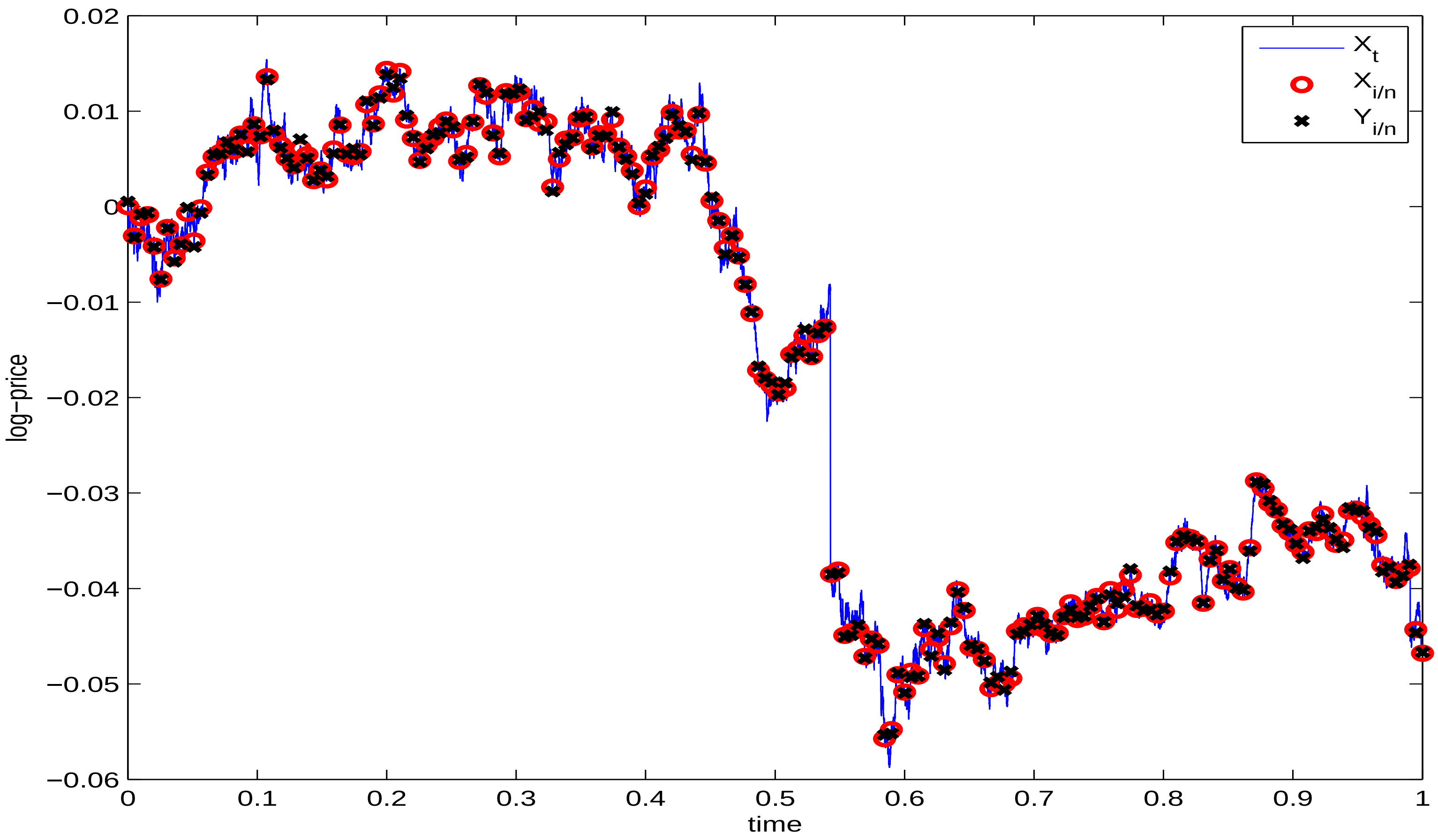


Figure 15: Microstructure noise in SPY

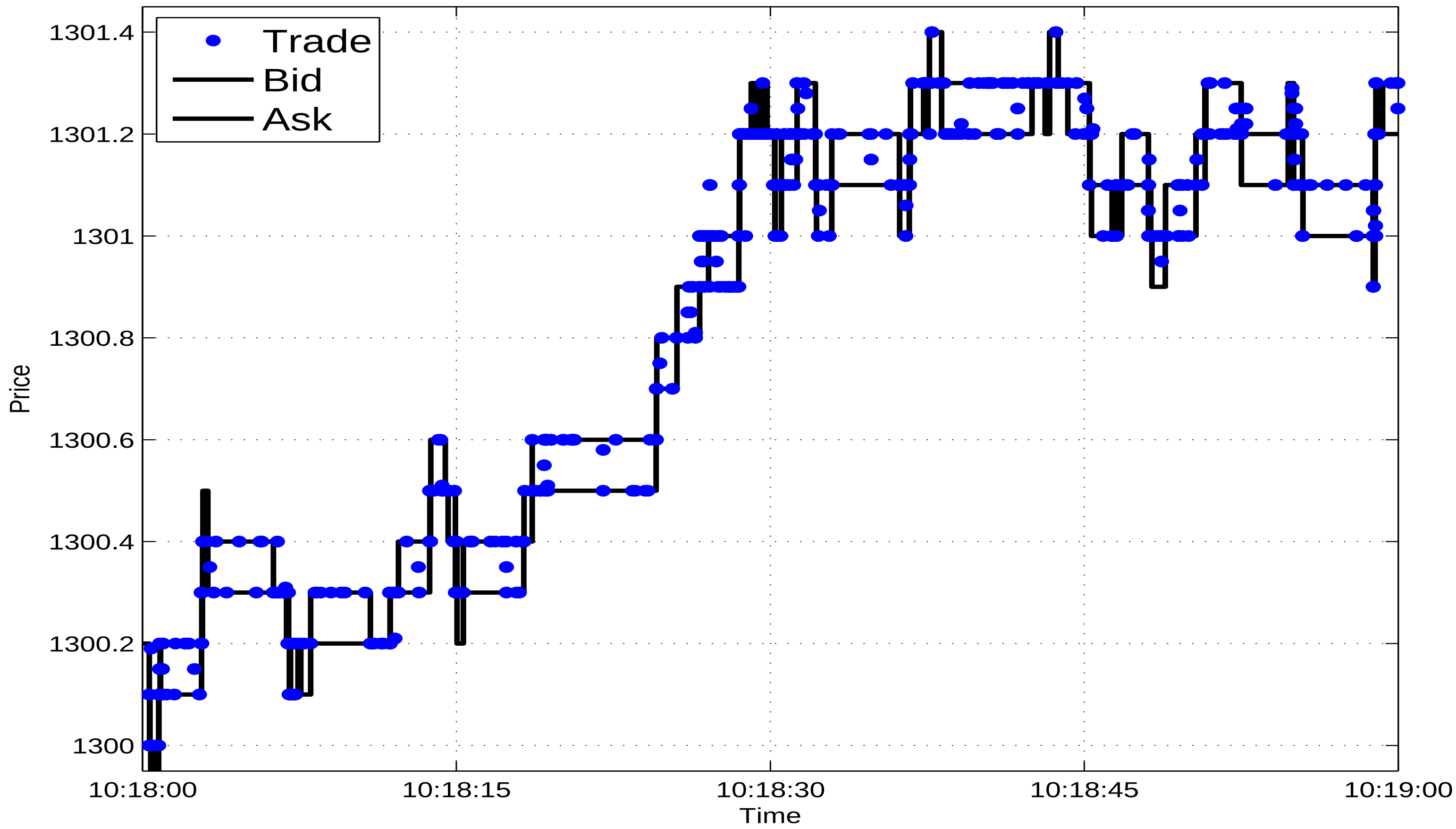
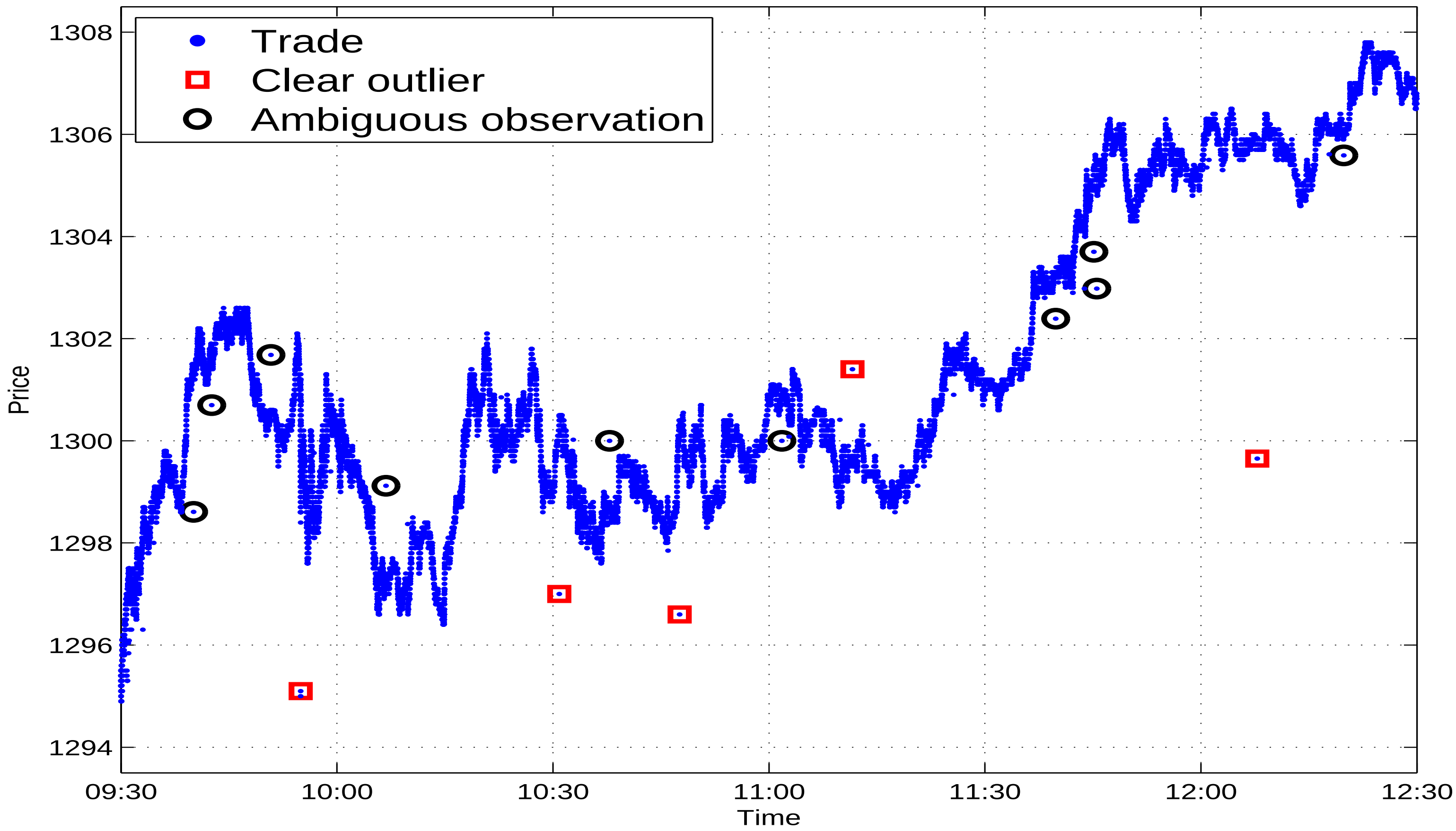


Figure 16: Outliers in SPY



Noise CREATES spurious return variation by inducing serial correlation in the observed log-returns, $\Delta_i^N Y$. It implies $RV^N \to \infty$ as $N \to \infty$... also true for BV^N (e.g., Hansen and Lunde, 2006; Jiang and Oomen, 2008).

To combat the noise, we use the notion that if we locally smooth $Y_{i/N}$ in the vicinity of i/N, we retrieve an estimate, say $\bar{Y}_{i/N}$, which tends to be closer to $X_{i/N}$, because the noise is largely averaged away.

This is called pre-averaging (e.g., Jacod, Li, Mykland, Podolskij, and Vetter, 2009; Podolskij and Vetter, 2009a,b).

Averaging our discrete sample of noisy high-frequency data this way leads to a new set of increments, say $\Delta\bar{Y}_i^N$, based on pre-averaged prices.

To implement pre-averaging, we need a window size $K = K(N)$:

$$K = \theta\sqrt{N} + o\Big(N^{-1/4}\Big), \quad \theta > 0. \tag{28}$$

... and a weight function:

$$w : \mathbb{R} \mapsto \mathbb{R} \tag{29}$$

In practice, $K = [\theta\sqrt{N}]$ (or $\lceil \cdot \rceil$) is used. θ is a tuning-parameter, so anything goes, but $\theta = 1/3$ or $\theta = 1$ has been advocated in several studies (Christensen, Kinnebrock, and Podolskij, 2010; Hautsch and Podolskij, 2013; Christensen, Oomen, and Podolskij, 2014).

$w(x) = \min(x, 1 - x)$ (i.e., a triangular kernel) is standard.

The return series, post pre-averaging:

$$\Delta_i^N \bar{Y} = \sum_{j=1}^{K} w_j^N \Delta_{i+j}^N Y = -\sum_{j=0}^{K} (w_{j+1}^N - w_j^N) Y_{\frac{i+j}{N}}, \tag{30}$$

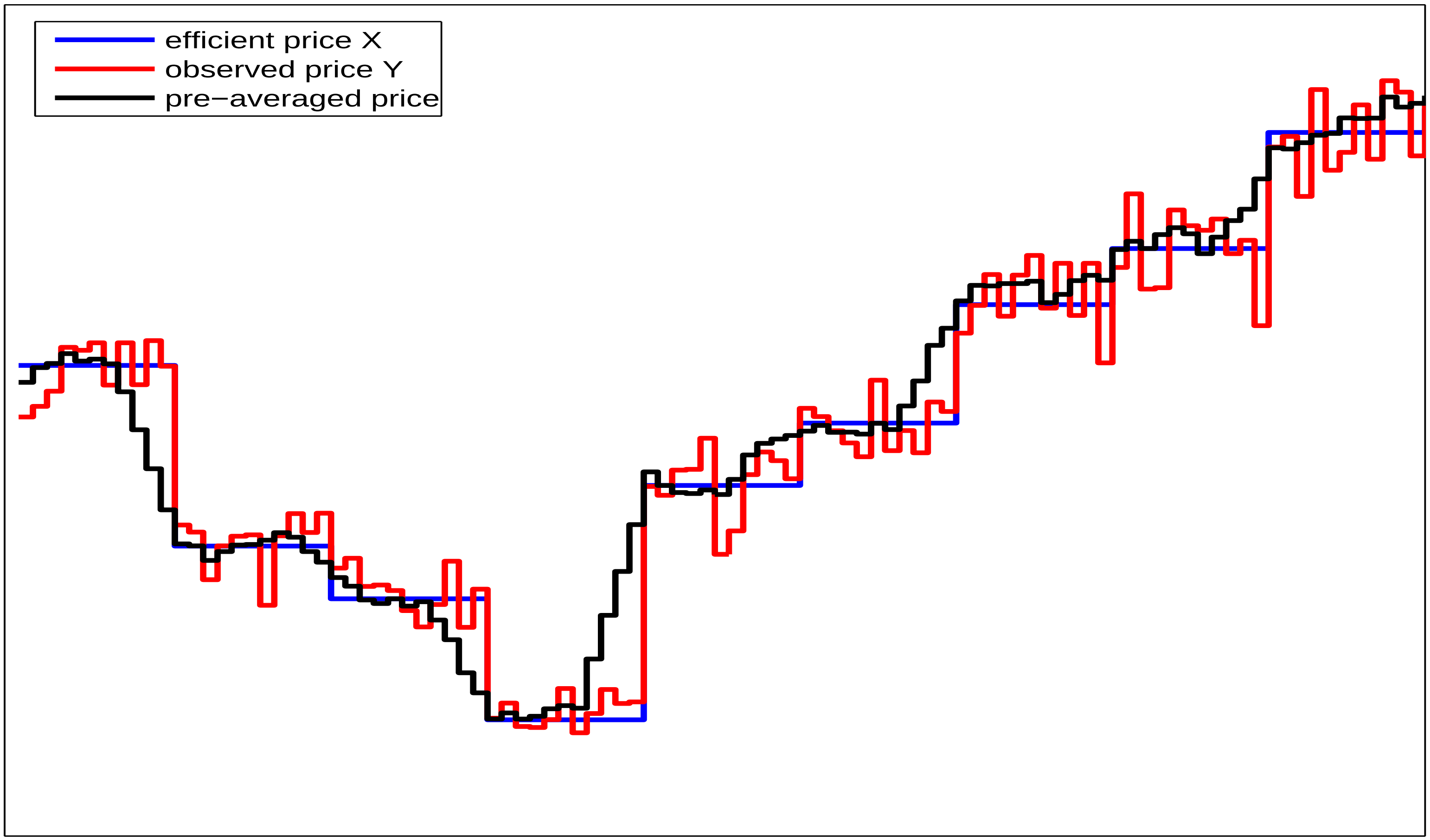
for $i = 1, \ldots, N-K+2$, where $w_j^N = w(j/K)$.

Equivalent formula (for K even and w as above):

$$\Delta_i^N \bar{Y} = \frac{1}{K} \sum_{j=1}^{K/2} Y_{\frac{i+K/2+j}{N}} - \frac{1}{K} \sum_{j=1}^{K/2} Y_{\frac{i+j}{N}}. \tag{31}$$

The sequence $(2\Delta_i^N \bar{Y})_{i=1}^{N-K+2}$ can be interpreted as a new set of increments from a price process that is constructed by simple averaging of the noisy log-price series, $(Y_{i/N})_{i=1}^{N}$, in a neighbourhood of i/N.

Figure 17: Illustration of pre-averaging



Note that the pre-averaged returns are computed from an overlapping noisy return series (a rolling window)...

... an overlapping sample problem.

This CREATES a strong, positive autocorrelation in the series $(\Delta_i^N \bar{Y})_{i=1}^{N-K+2}$, which we need to account for in the inference procedures (e.g., for the computation of standard errors).

Figure 18: ACF of $\Delta_i^N Y$ — SPY data.

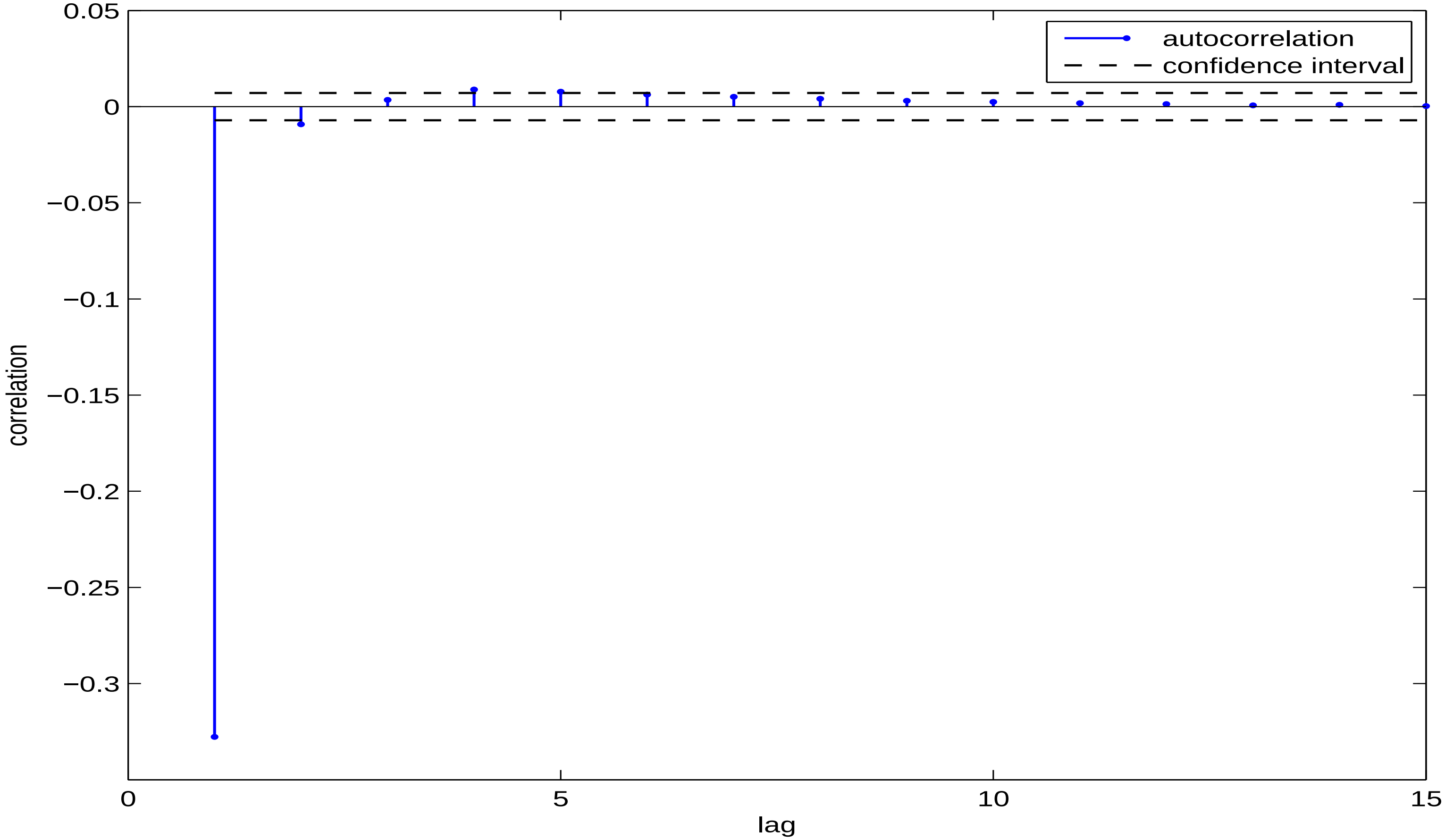
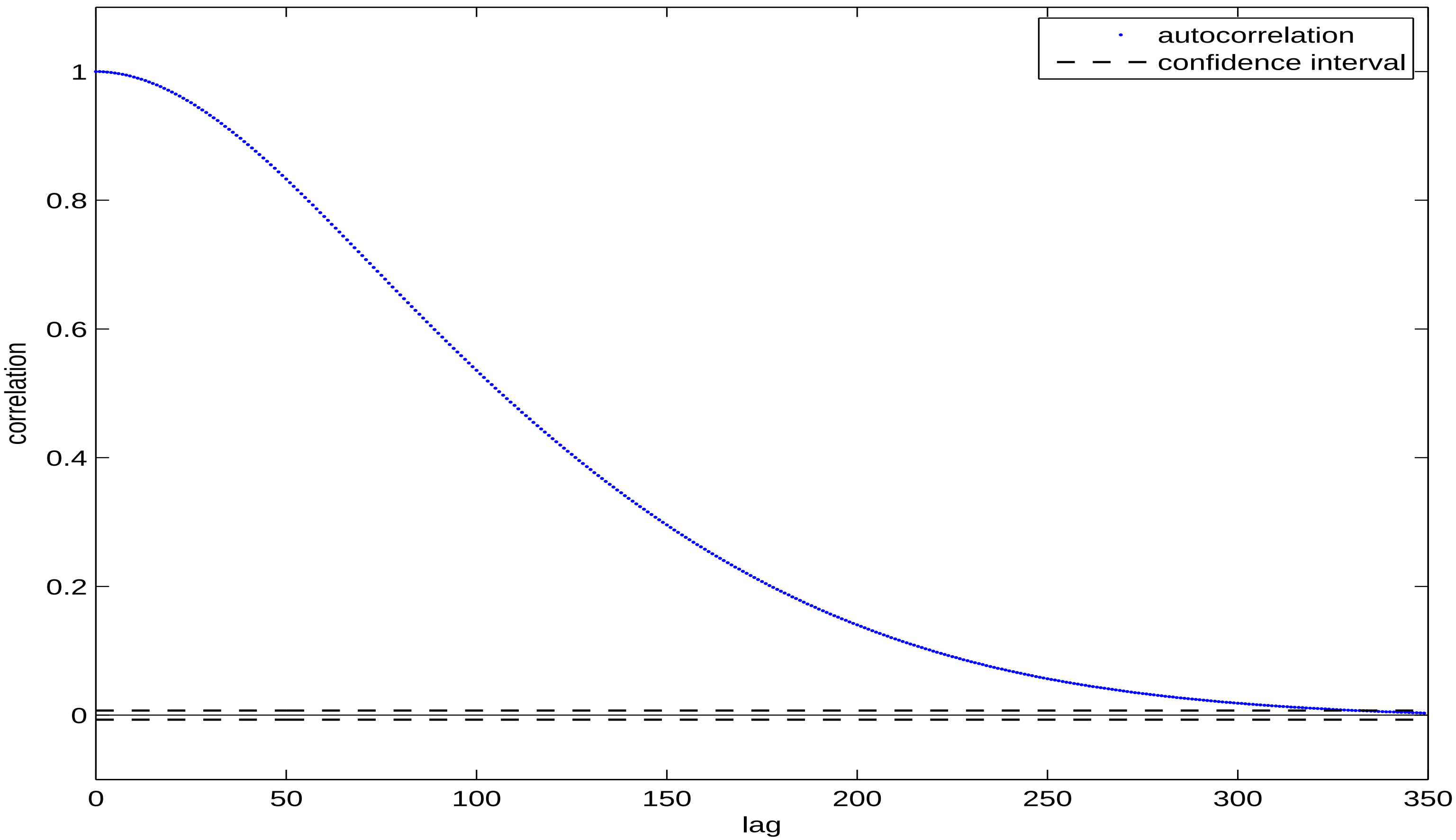


Figure 19: ACF of $\Delta_i^N \bar{Y}$ — SPY data.



Pre-averaged realized variance and bipower variation

With the above notation in place, we introduce a noise- and outlier-robust version of the RV and BV:

$$RV^* = \frac{N}{N-K+2}\frac{1}{K\psi_K}\sum_{i=1}^{N-K+2}|\Delta_i^N \bar{Y}|^2 - \frac{\hat{\omega}^2}{\theta^2\psi_K}, \tag{32}$$

$$BV^* = \frac{N}{N-2K+2}\frac{1}{K\psi_K}\frac{\pi}{2}\sum_{i=1}^{N-2K+2}|\Delta_i^N \bar{Y}||\Delta_{i+K}^N \bar{Y}| - \frac{\hat{\omega}^2}{\theta^2\psi_K}, \tag{33}$$

where $\psi_K = (1+2K^{-2})/12$.

The construction is almost identical to the no noise setting, apart from the fact that $\Delta_i^N Y$ has been replaced by $\Delta_i^N \bar{Y}$.

The intuition behind the bipower construction is that $(\Delta_i^N \bar{Y})$ is autocorrelated (up to the Kth lag), which is broken by multiplying pre-averaged returns that are K terms apart.

This leads to a lower, effective sample of size $N - 2K + 2$.

$\frac{\widehat{\omega}^2}{\theta^2 \psi_K}$ is a bias-correction, which compensates for the residual microstructure noise that remains after pre-averaging. It appears, as we are balancing the order of the signal $(\Delta_i^N \bar{X})$ and the noise $(\Delta_i^N \bar{\epsilon})$...

... leads to optimal rates of convergence!

We can reduce the noise further by taking a pre-averaging window of larger order than $O(\sqrt{N})$. This can potentially wipe out the noise completely (it also makes the framework more robust to non-i.i.d. noise), but it results in suboptimal rates of convergence.

Note that the bias-correction drops out when we compute $RV^* - BV^*$, so it is of limited importance for jump measurement and testing.

$\widehat{\omega}^2$ is an estimator of the noise variance $\omega^2 = E(\epsilon^2)$. It can be estimated in a number of ways [see, e.g., Gatheral and Oomen (2010) for a comparison of estimators]. Here, we use the estimator proposed by Oomen (2006):

$$\hat{\omega}^2 = -\frac{1}{N-1}\sum_{i=2}^{N} \Delta_{i-1}^{N} Y \Delta_i^N Y. \tag{34}$$

Asymptotic theory

Assume Y follows Eq. (27) and $E(u^4) < \infty$. As $N \to \infty$,

$$RV^* \xrightarrow{p} [X]_1, \qquad BV^* \xrightarrow{p} \int_0^1 \sigma_s^2 \mathrm{d}s. \tag{35}$$

Moreover, suppose $E(u^8) < \infty$, and that X is a continuous semimartingale, i.e., X follows Eq. (1) but with $N_t \equiv 0$ for all t. Then, under suitable assumptions on σ, as $N \to \infty$, it further holds that

$$N^{1/4} \begin{pmatrix} RV^* - \int_0^1 \sigma_s^2 \mathrm{d}s \\ BV^* - \int_0^1 \sigma_s^2 \mathrm{d}s \end{pmatrix} \xrightarrow{d_s} MN(0, \Sigma^*), \tag{36}$$

a mixed normal distribution with conditional covariance matrix Σ^*, where Σ^* is defined in Podolskij and Vetter (2009b).

Pre-averaging slows down the rate of convergence. $N^{-1/4}$ is nonetheless the fastest rate in noisy diffusion models (Gloter and Jacod, 2001a,b).

In general, Σ^* has a complicated structure (even with i.i.d., independent noise), which is 1) typically not known in closed-form and 2) does not "factorize" as in the noise-free setting. Thus, in contrast to the frictionless world, turning (36) into a feasible theory (via estimation of Σ^*) is difficult — at least until recently.

Problem: the estimator of Σ^* proposed by Podolskij and Vetter (2009b) was not guaranteed (and often failed) to be positive semi-definite. This leads to a huge proportion of negative variance estimates of $RV^* - BV^*$, which makes the standardization for jump testing slightly complex (no pun intended). Recent work, which addresses this problem: Christensen, Podolskij, Thamrongrat, and Veliyev (2016); Mykland and Zhang (2016).

Table 1: Proportion of ill-conditioned covariance matrix estimates, $\tilde{\Sigma}_n^*$.

	general noise				i.i.d. noise				
	BM		SV		BM		SV		SPY
$n =$	2,340	23,400	2,340	23,400	2,340	23,400	2,340	23,400	n_{actual}
2-dimensional setting									
Panel A: Nonpositive definite									
$\theta =$ 0.33	0.177	0.094	0.180	0.101	0.178	0.099	0.181	0.111	0.111
1.00	0.340	0.242	0.343	0.251	0.348	0.241	0.347	0.249	0.233
Panel B: Negative variance									
$\theta =$ 0.33	0.141	0.066	0.149	0.072	0.140	0.070	0.146	0.077	0.064
1.00	0.286	0.199	0.290	0.210	0.292	0.201	0.289	0.206	0.184
Panel C: Condition number ≥ 20									
$\theta =$ 0.33	0.073	0.053	0.071	0.058	0.069	0.057	0.071	0.062	0.045
1.00	0.080	0.096	0.081	0.101	0.079	0.095	0.081	0.100	0.088

Note. We show the proportion of ill-conditioned covariance matrix estimates, when the Podolskij and Vetter (2009a) estimator $\tilde{\Sigma}_n^*$ of Σ^* is used.

Figure 20: Properties of the standardized $RV^* - BV^*$.

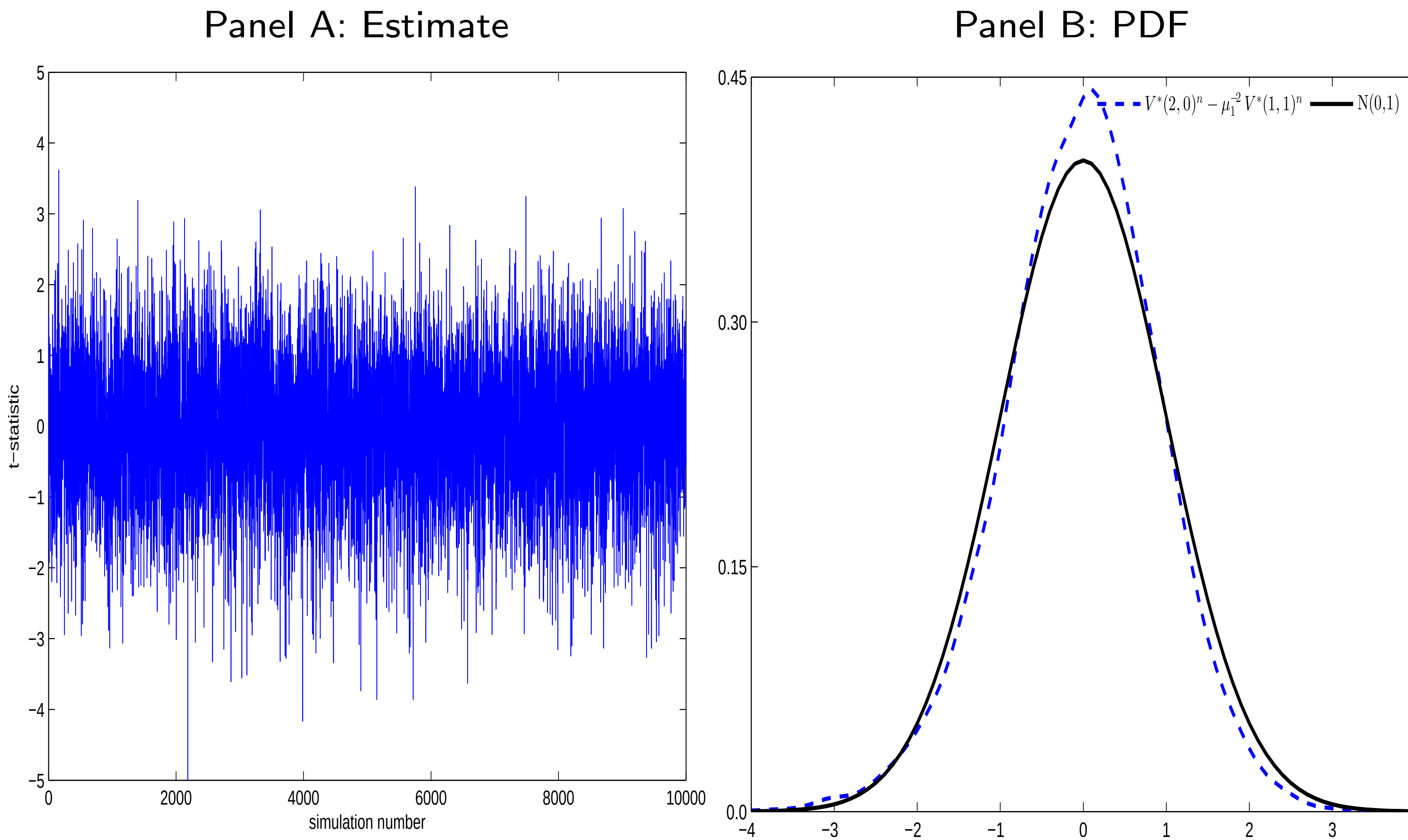


Figure 21: IV estimates and standard error.

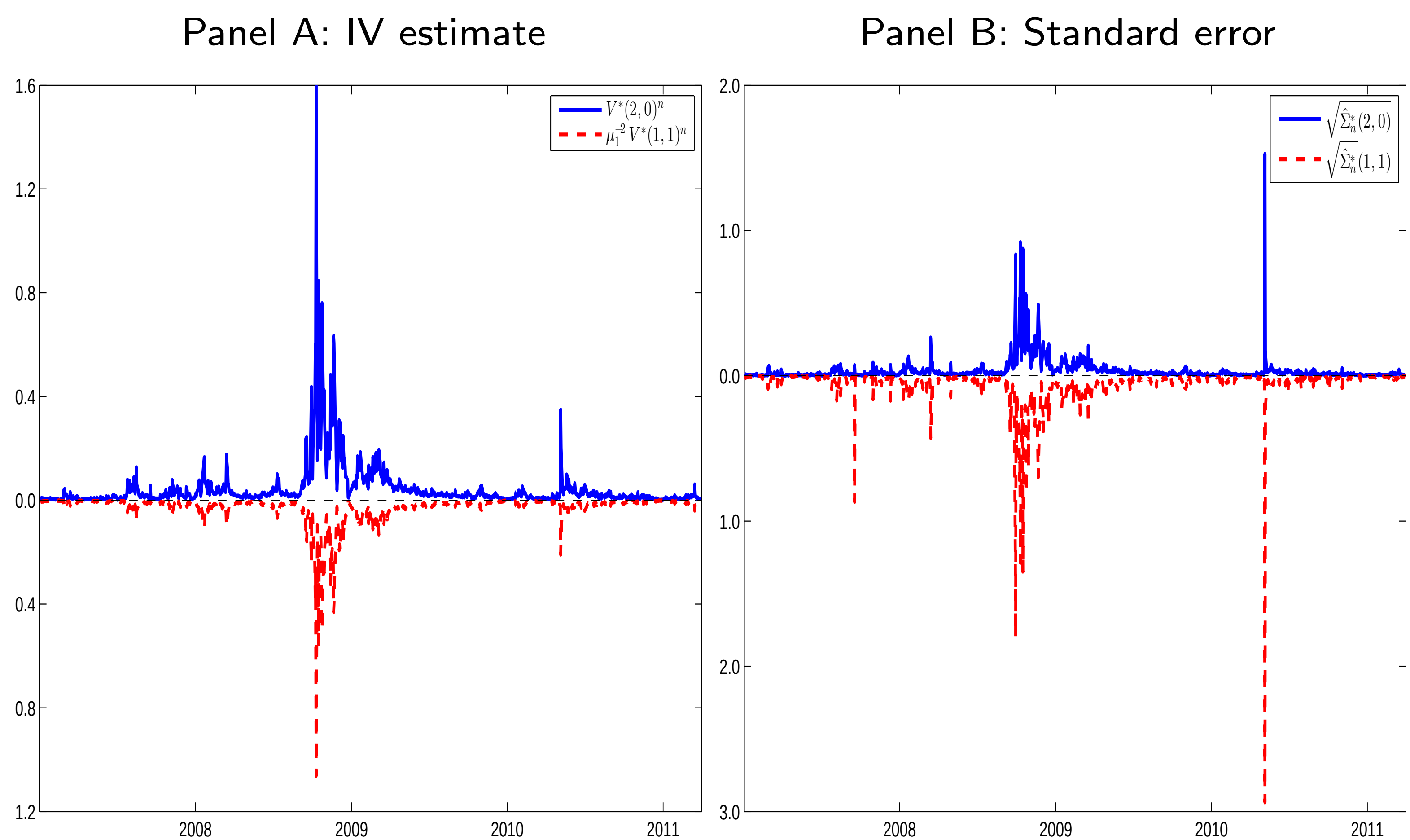
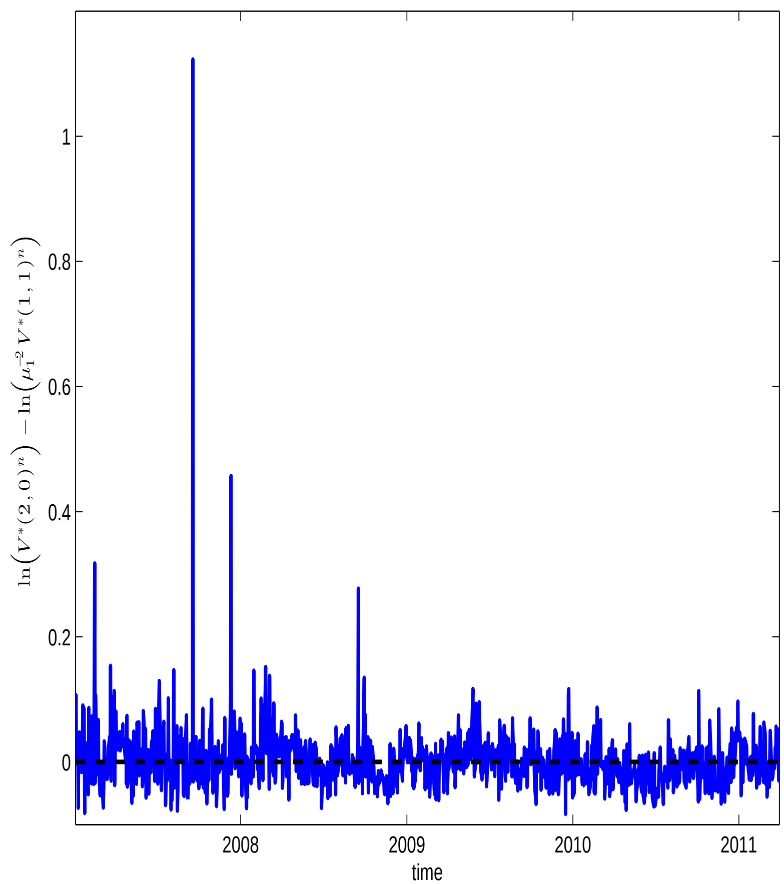
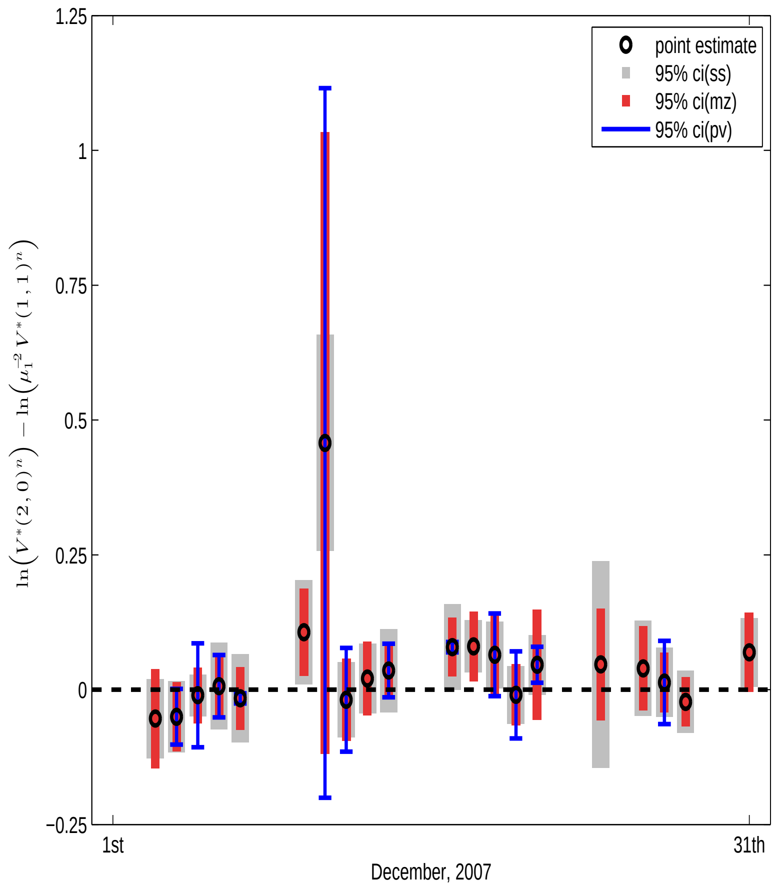


Figure 22: Inference : $\ln\left(RV^*\right) - \ln\left(BV^*\right)$.

Panel A: Point estimate



Panel B: Confidence interval



Jump variation in practice

Christensen, Oomen, and Podolskij (2014) apply the toolbox developed above to examine the importance of the jump component in practice.

They report the controversial finding that the jump variation, as measured from noisy tick data, is about 1%.

$\Rightarrow$ There are very few (large) jumps, compared to the many small diffusive shocks driven by the continuous influx of news to the market.

This is an order of magnitude smaller than what has been found in the extant literature from "low-frequency" noise-free data (often based on 5- to 15-minute sampling of the price process), where the JV across a huge strand of papers is about 10% on average (see Table 1 in the paper).

The (possible) explanation follows.

Figure 23: Jump variation as a function of N.

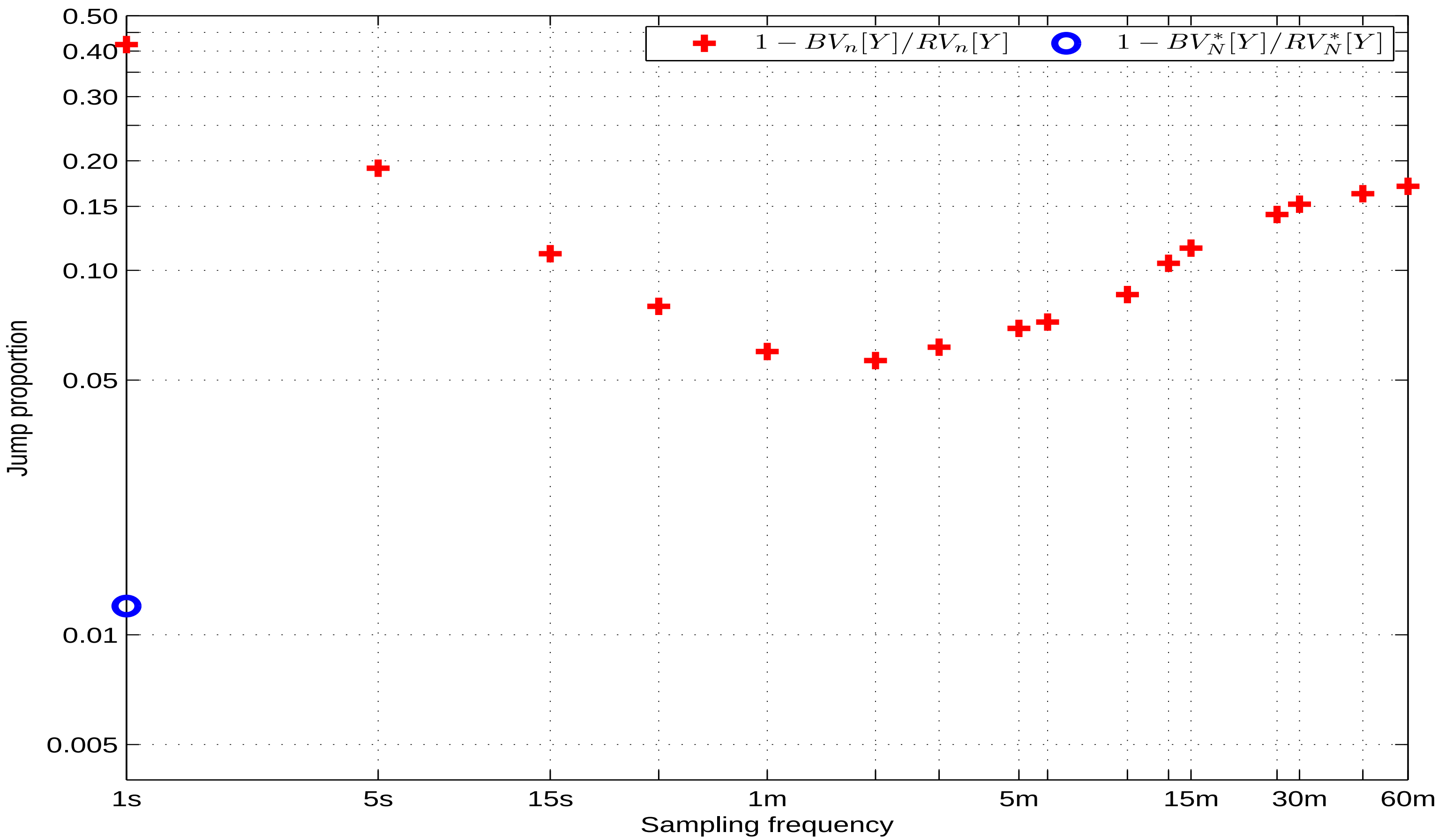


Figure 24: Fukushima earthquake: Jump in USDJPY?

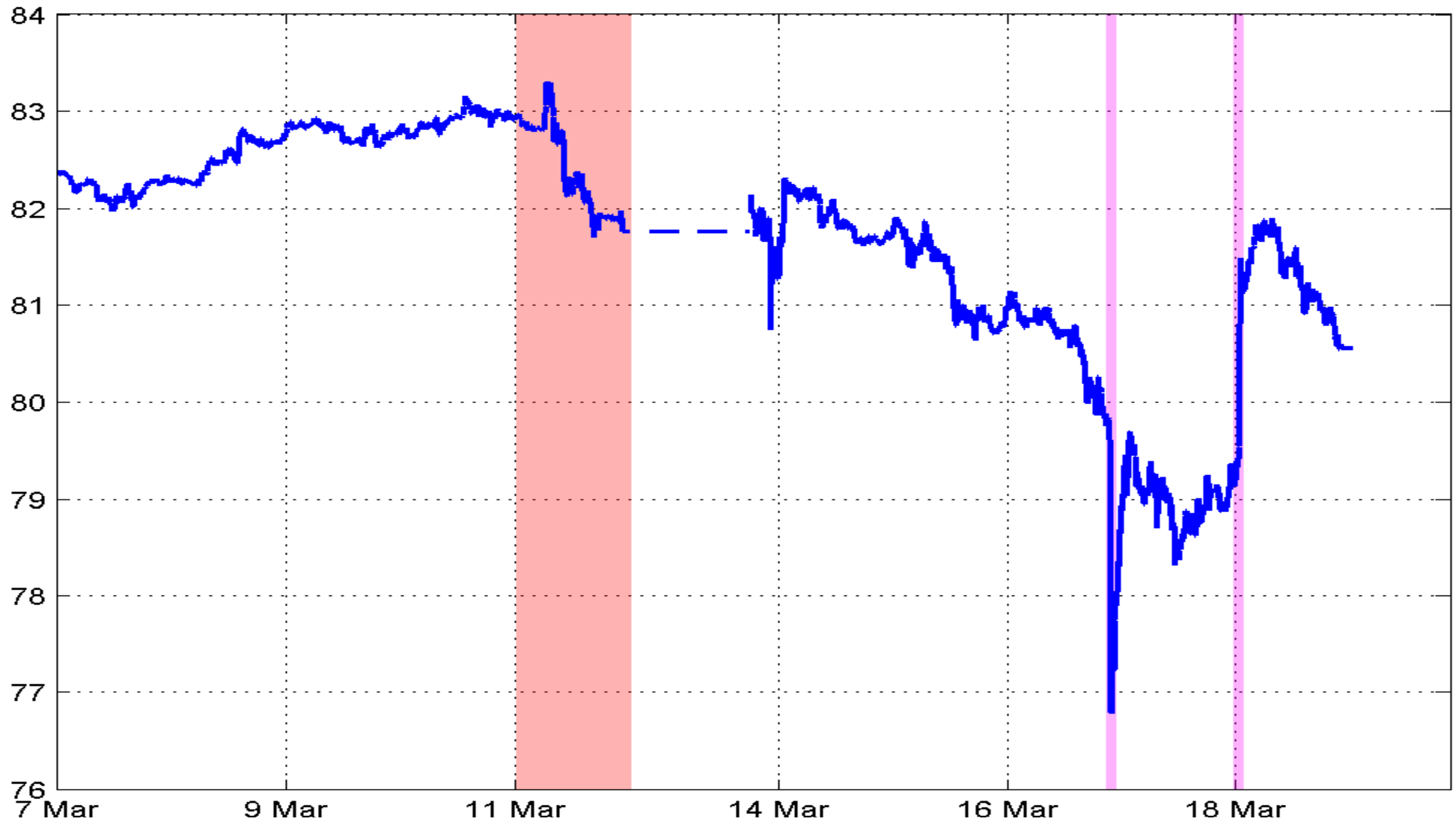


Figure 25: Fukushima earthquake: Jump in USDJPY?

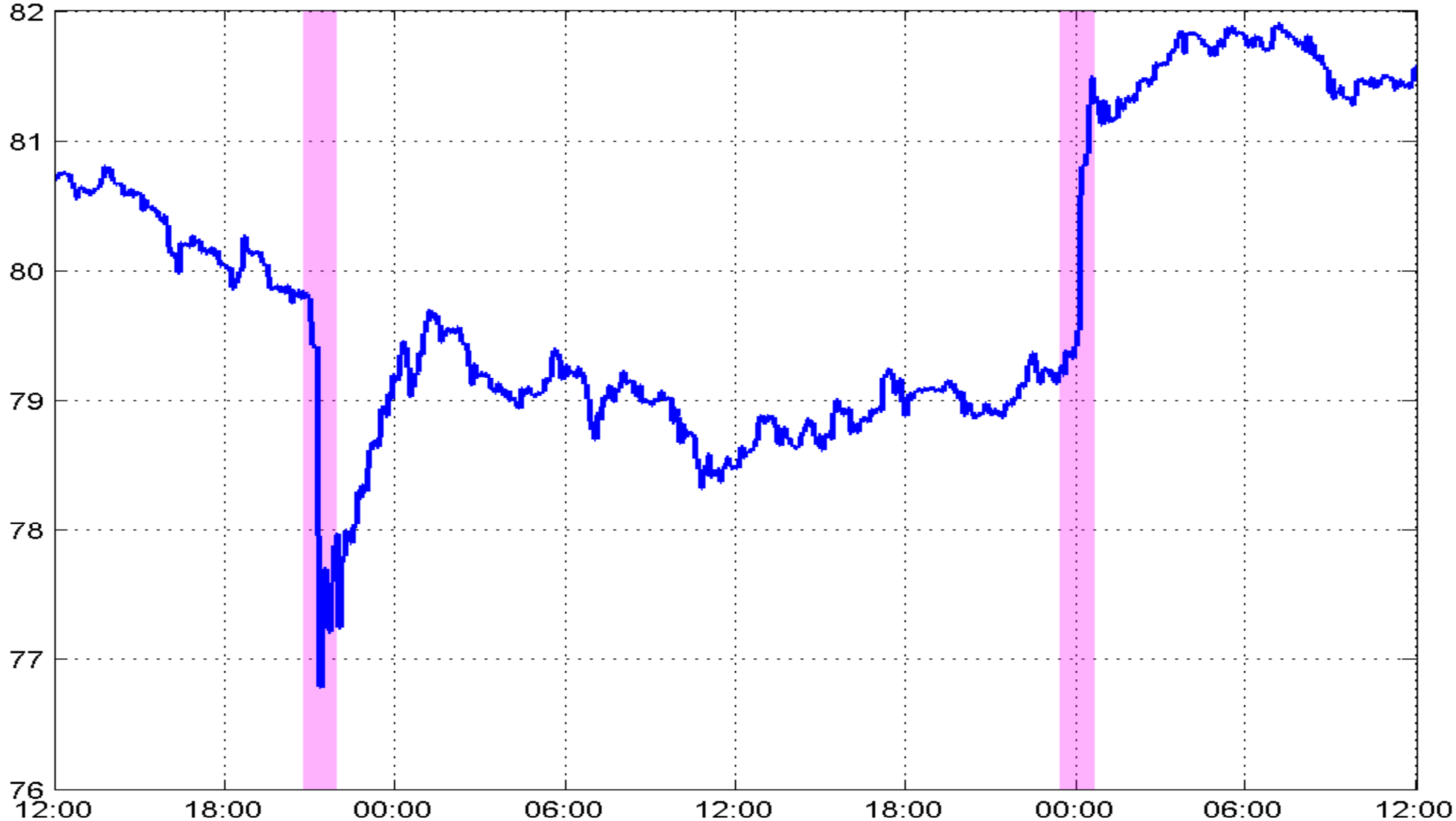


Figure 26: Fukushima earthquake: Jump in USDJPY?

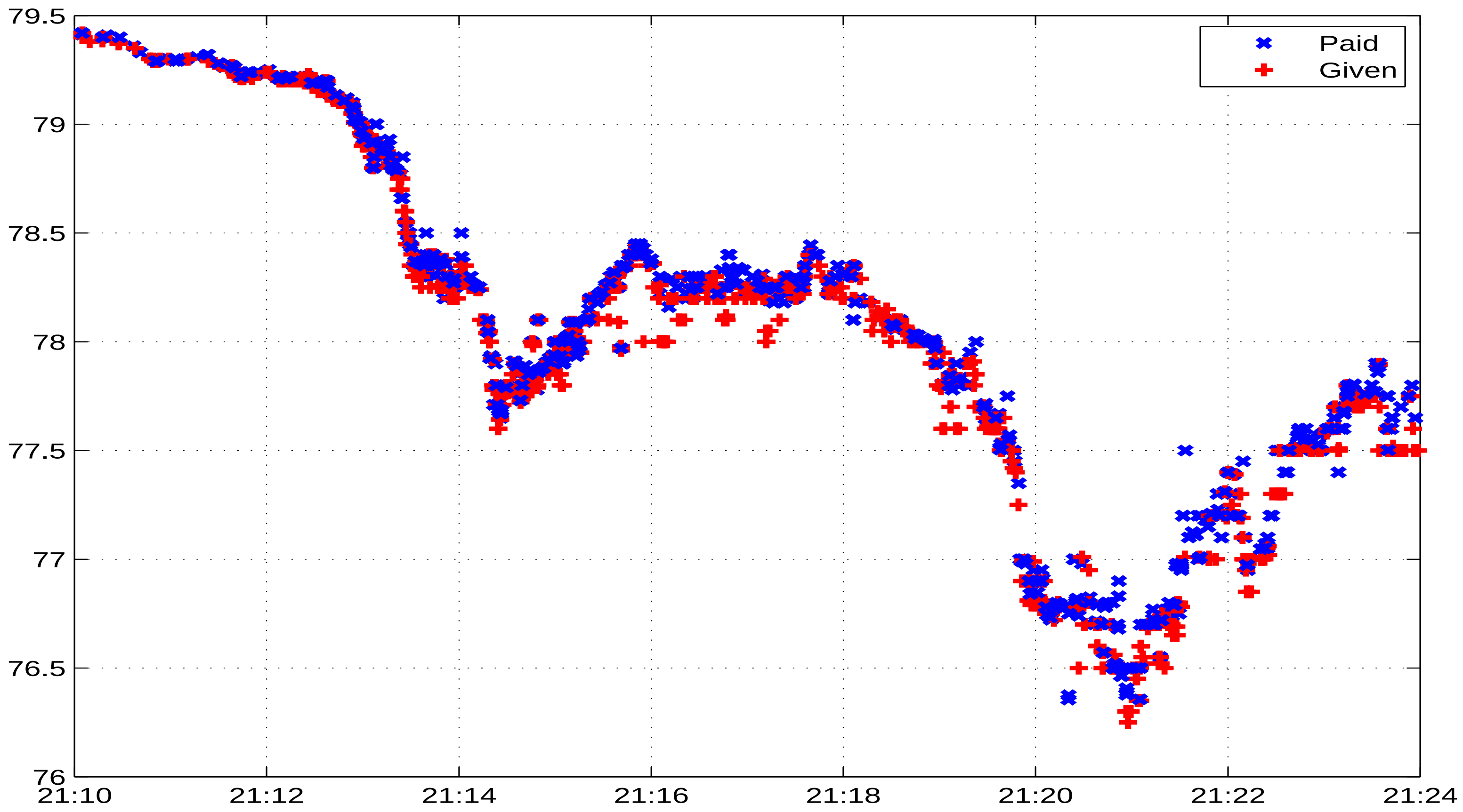


Figure 27: Fukushima earthquake: Jump in USDJPY?

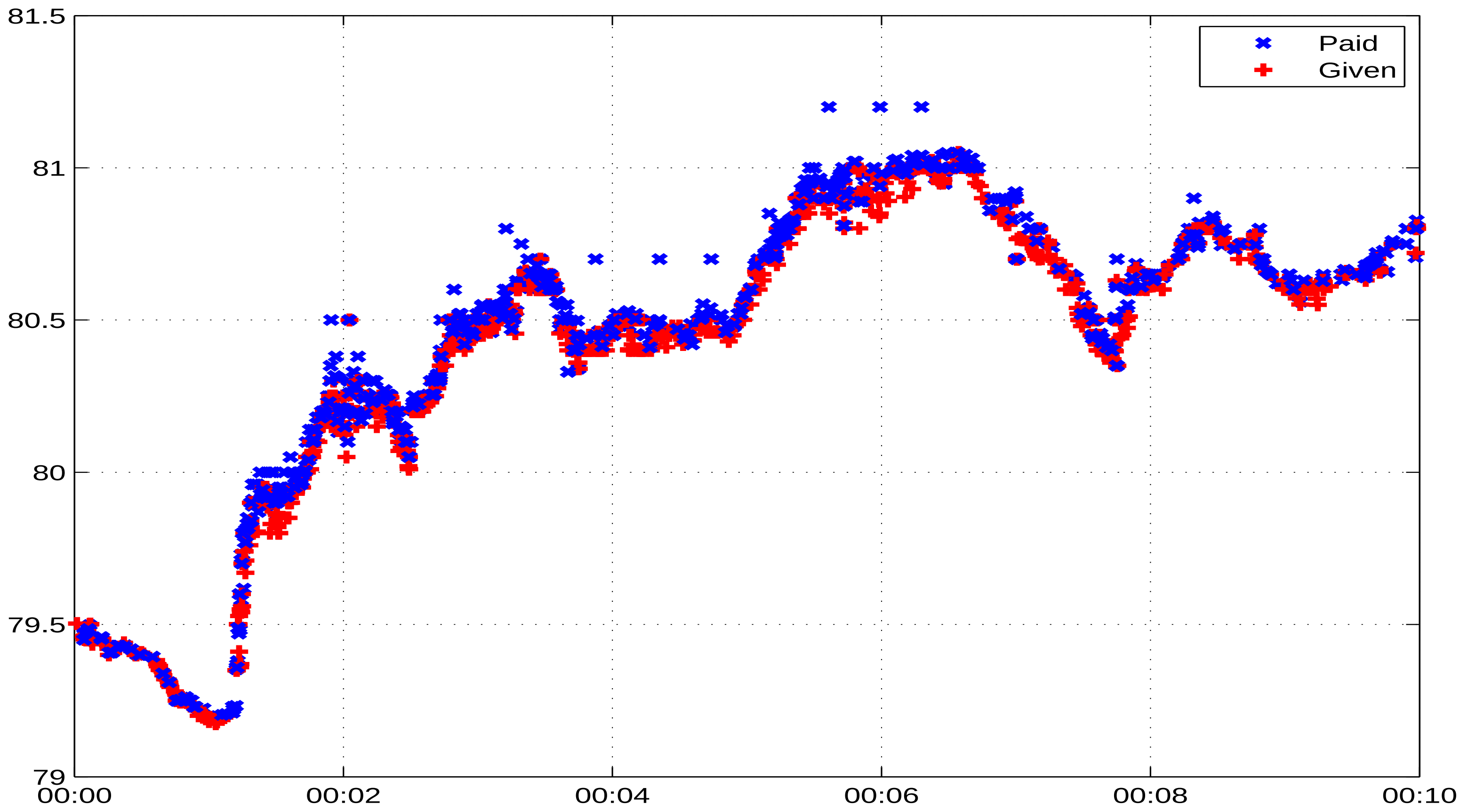


Figure 28: Jump in Merck?

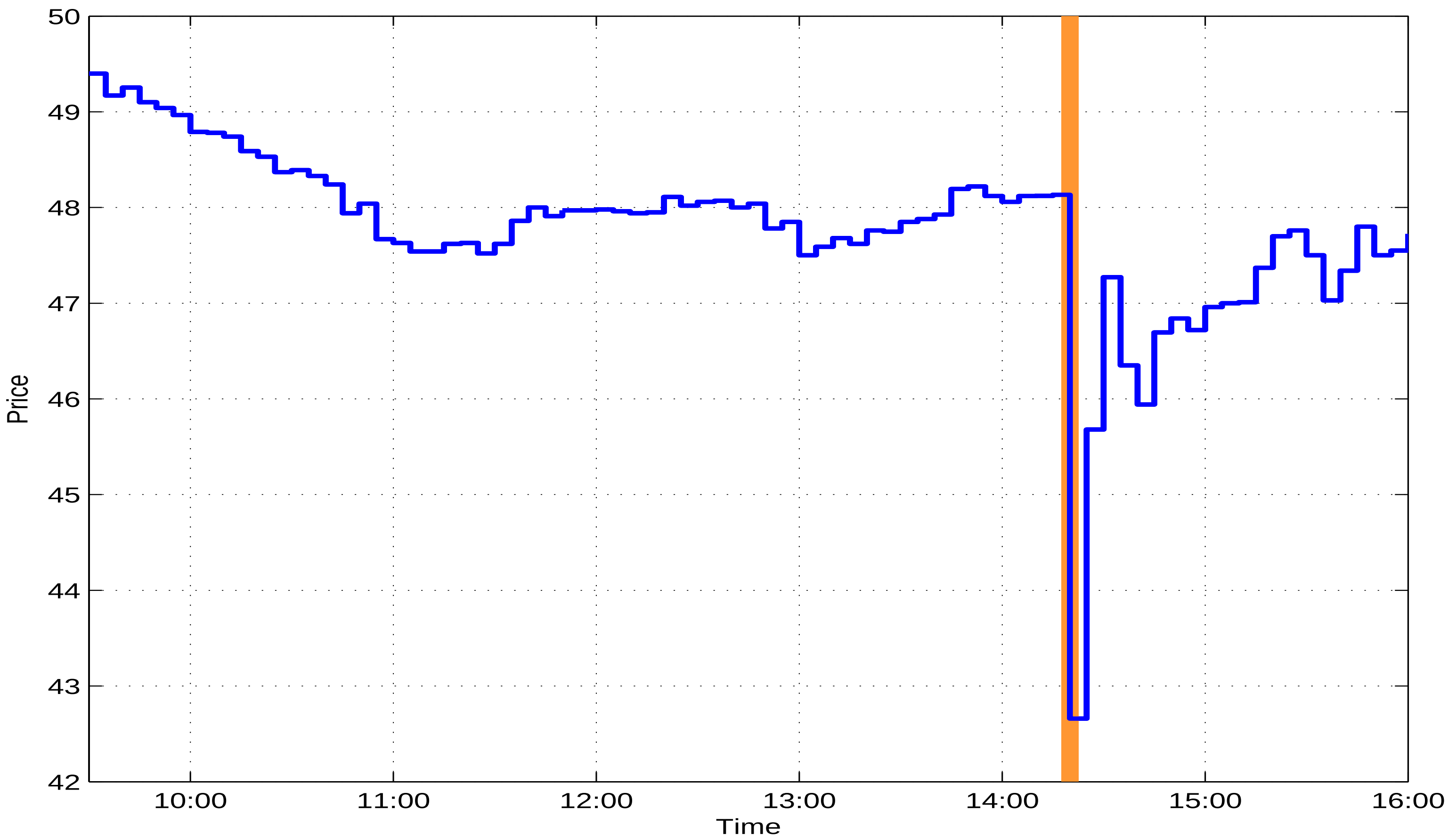
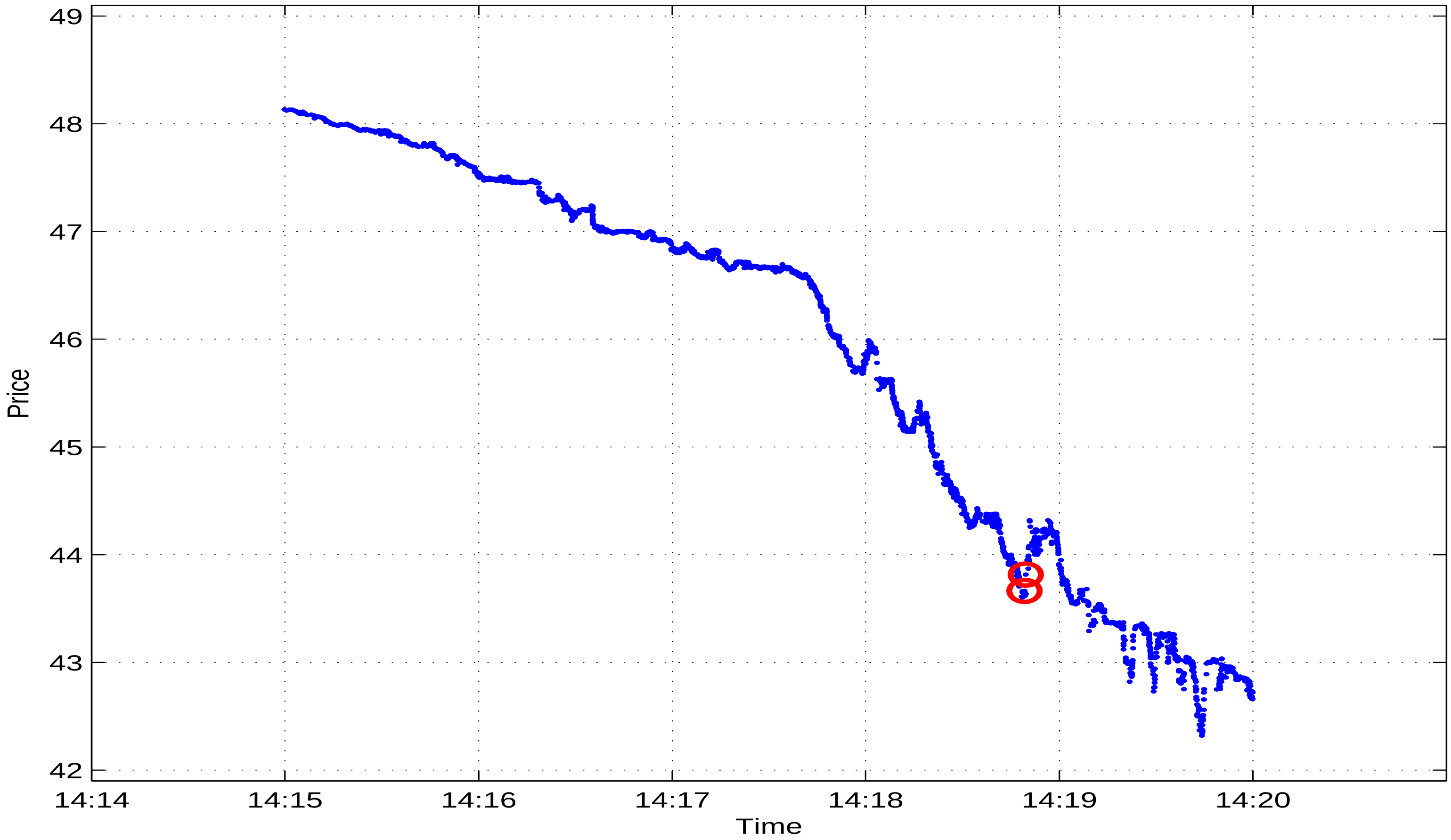


Figure 29: Jump in Merck?



The volatility burst hypothesis

The above suggests that many of the large, significant jumps identified with low-frequency 5-minute sampling are spurious, and they appear as much more gradual price changes, when we zoom in the price process.

To paraphrase Aït-Sahalia and Jacod (2009b), jumps can only be identified with certainty in the continuous-time limit, i.e. as $\tau \to 0$

$$X_t - X_{t-\tau} = \int_{t-\tau}^{t} \sigma_s \mathrm{d}W_s + \sum_{i=N_{t-\tau}}^{N_t} J_i, \tag{37}$$

otherwise, a burst of (edit: large increase in) volatility can be observationally equivalent to a jump at lower frequency.

To provide support for the burst of volatility hypothesis, we draw noise-free prices from a linear Brownian motion:

$$\mathrm{d}X_t = \sigma_t \mathrm{d}W_t, \quad \text{for } t \in [0, 1], \tag{38}$$

where $\sigma_t = 3\sigma^*$ for $t \in [16/32, 17/32]$ and $\sigma_t = \sigma^*$ otherwise. σ^* is fixed at a level corresponding to 40% in annualized terms.

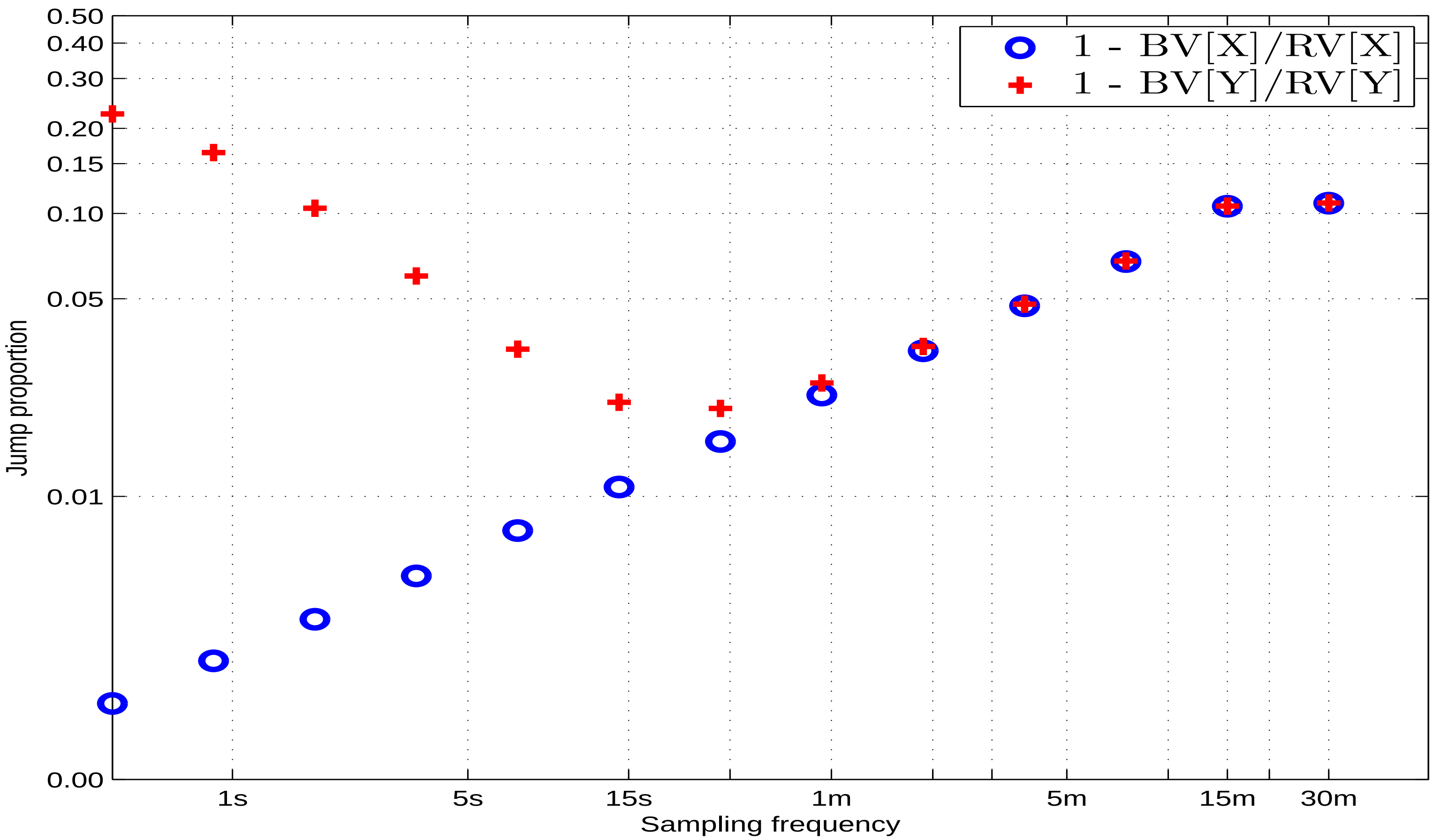
σ_t is piecewise constant and increases three-fold over a short interval of the day (equivalent to 15-minutes for an eight-hour trading session).

The price path is still continuous, so the true JV is zero.

We simulate noisy log-prices as above, $Y = X + u$, using a realistic level of i.i.d. noise and then round Y to the nearest cent to induce price discreteness, based on a starting price of $50 in levels.

Finally, we calculate RV and BV across a range of sampling frequencies and report the average implied JV over 10,000 independent simulation runs.

Figure 30: Jump variation from simulation.



What is going on?

We show that, under suitable conditions:

$$E\left(BV - \int_0^1 \sigma_s^2 \mathrm{d}s\right) = -\frac{1}{N} E\left(\frac{1}{12}\int_0^1 \frac{\upsilon_s^2}{\sigma_s^2}\mathrm{d}s\right) + o(N^{-1}), \qquad (39)$$

where υ_s^2 is the (spot) variance of variance (or vol-of-vol).

Thus, BV is downward biased in finite samples, which translates into an inflated JV measure.

The bias gets more pronounced with a lowering of the sampling frequency and in periods with a high volatility of volatility.

In a recent paper, Christensen, Oomen, and Renò (2016) provide an alternative explanation via local drift explosions (the drift burst hypothesis).

References

Aït-Sahalia, Y., and J. Jacod, 2009a, "Estimating the degree of activity of jumps in high frequency data," *Annals of Statistics*, 37(5), 2202–2244.

———, 2009b, "Testing for jumps in a discretely observed process," *Annals of Statistics*, 37(1), 184–222.

Andersen, T. G., D. Dobrev, and E. Schaumburg, 2012, "Jump-robust volatility estimation using nearest neighbour truncation," *Journal of Econometrics*, 169(1), 75–93.

Back, K., 1991, "Asset prices for general processes," *Journal of Mathematical Economics*, 20(4), 371–395.

Barndorff-Nielsen, O. E., and N. Shephard, 2004, "Power and bipower variation with stochastic volatility and jumps," *Journal of Financial Econometrics*, 2(1), 1–48.

———, 2006, "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, 4(1), 1–30.

Black, F., 1986, "Noise," *Journal of Finance*, 41(3), 529–543.

Christensen, K., S. Kinnebrock, and M. Podolskij, 2010, "Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data," *Journal of Econometrics*, 159(1), 116–133.

Christensen, K., R. C. A. Oomen, and M. Podolskij, 2010, "Realised quantile-based estimation of the integrated variance," *Journal of Econometrics*, 159(1), 74–98.

——— , 2014, "Fact or friction: Jumps at ultra high frequency," *Journal of Financial Economics*, 114(3), 576–599.

Christensen, K., R. C. A. Oomen, and R. Renò, 2016, "The drift burst hypothesis," Working paper, Aarhus University.

Christensen, K., M. Podolskij, N. Thamrongrat, and B. Veliyev, 2016, "Inference from high-frequency data: A subsampling approach," *Journal of Econometrics*, (Forthcoming).

Corsi, F., D. Pirino, and R. Renò, 2010, "Threshold bipower variation and the impact of jumps on volatility forecasting," *Journal of Econometrics*, 159(2), 276–288.

Delbaen, F., and W. Schachermayer, 1994, "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, 300(1), 463–520.

Gatheral, J., and R. C. A. Oomen, 2010, "Zero-intelligence realized variance estimation," *Finance and Stochastics*, 14(2), 249–283.

Gloter, A., and J. Jacod, 2001a, "Diffusions with measurement errors. I - local asymptotic normality," *ESAIM: Probability and Statistics*, 5, 225–242.

——— , 2001b, “Diffusions with measurement errors. II - measurement errors,” *ESAIM: Probability and Statistics*, 5, 243–260.

Hansen, P. R., and A. Lunde, 2006, “Consistent ranking of volatility models,” *Journal of Econometrics*, 131(1-2), 97–121.

Hautsch, N., and M. Podolskij, 2013, “Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence,” *Journal of Business and Economic Statistics*, 31(2), 165–183.

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(7), 2249–2276.

Jiang, G. J., and R. C. A. Oomen, 2008, “Testing for jumps when asset prices are observed with noise – A ”swap variance” approach,” *Journal of Econometrics*, 144(2), 352–370.

Mancini, C., 2004, “Estimation of the characteristics of jump of a general Poisson-diffusion model,” *Scandinavian Actuarial Journal*, 2004(1), 42–52.

——— , 2009, “Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps,” *Scandinavian Journal of Statistics*, 36(2), 270–296.

Mykland, P. A., N. Shephard, and K. Sheppard, 2012, “Efficient and feasible inference for the components of financial variation using blocked multipower variation," Working paper, University of Chicago.

Mykland, P. A., and L. Zhang, 2016, "Assessment of uncertainty in high frequency data: The observed asymptotic variance," *Econometrica*, (Forthcoming).

Niederhoffer, V., and M. F. M. Osborne, 1966, "Market making and reversal on the stock exchange," *Journal of the American Statistical Association*, 61(316), 897–916.

Oomen, R. C. A., 2006, "Properties of realized variance under alternative sampling schemes," *Journal of Business and Economic Statistics*, 24(2), 219–237.

Podolskij, M., and M. Vetter, 2009a, "Bipower-type estimation in a noisy diffusion setting," *Stochastic Processes and their Applications*, 119(9), 2803–2831.

——— , 2009b, "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," *Bernoulli*, 15(3), 634–658.

Podolskij, M., and D. Ziggel, 2010, "New tests for jumps in semimartingale models," *Statistical Inference for Stochastic Processes*, 13(1), 15–41.

Protter, P. E., 2004, *Stochastic Integration and Differential Equations*. Springer-Verlag, Heidelberg, 1st edn.

Roll, R., 1984, "A simple implicit measure of the effective bid-ask spread in an efficient market," *Journal of Finance*, 39(4), 1127–1139.

Veraart, A. E. D., 2011, "How precise is the finite sample approximation of the asymptotic distribution of realised variation measures in the presence of jumps?," *Advances in Statistical Analysis*, 95(3), 252–291.

Vetter, M., 2010, "Limit theorems for bipower variation of semimartingales," *Stochastic Processes and their Applications*, 120(1), 22–38.