Jumps or flatness?*

Aleksey Kolokolov† Roberto Renò‡

February 26, 2018

Abstract

Even moderate amounts of zero returns in financial data, associated with flat prices, are heavily detrimental for reliable jump inference. Price flatness is mistaken for jumps by statistics based on multipower variation. After assuming that flatness is an integral part of the price process, we propose flatness-robust multipower estimators and provide a description of their asymptotic behaviour. We then harness flatness-robust estimators to re-appraise the statistical features of jumps in financial markets. We find that jumps are much less frequent and much less contributing to price variation than what found by the empirical literature so far. In particular, the empirical finding that volatility is driven by a pure jump process is shown to be an artifact due to flatness.

JEL classification: C13; C22; C51; C53

Keywords: Stocks; VIX; Flat Trading; Multipower variation; Jump testing; Jump activity index

*We are grateful to Giuseppe Buccheri, Massimiliano Caporin, Kim Christensen, Fulvio Corsi, Nicola Fusari, Davide Pirino for their comments and research support. We also thanks the participants at the SoFiE Conference 2017 (New York, June 21-23) and to the XIX Workshop In Quantitative Finance (Rome, 24-26 January, 2018). Errors and omissions remain our own.

†SAFE, Goethe University Frankfurt. e-mail: *alexeiuo@gmail.com*, *kolokolov@safe.uni-frankfurt.de*

‡University of Verona, Department of Economics, e-mail: *roberto.reno@univr.it*

1 Introduction

In empirical finance, it is customary to sample prices sparsely on evenly spaced grids, such as one or five minutes. When doing so, prices are often flat, and many price returns are zero. Figure 1 shows the frequency distribution of the daily percentage of zero returns in stocks belonging to the S&P 500 index in more than 17 years. When sampling at one minute, the percentage of zero returns is 36.06%; at five minutes, it is 14.78%; and at the relatively low frequency of ten minutes, it is still 8.97%. Flatness is a non-negligible feature of price data, even for the most liquid stocks, including SPY, the S&P 500 ETF. Nevertheless, the vast majority of models used in finance, both in continuous and in discrete time, postulate the mere impossibility of zero returns in the price dynamics.[1]

In this paper, while being agnostic about the sources of zero returns,[2] we focus on the distortions induced by zeros on popular statistics used in finance. More specifically, we study the distortions of power and multipower estimators, which are used to estimate volatility (Woerner, 2006; Barndorff-Nielsen, Graversen, Jacod, and Shephard, 2006), detect jumps (Barndorff-Nielsen and Shephard, 2004; Lee and Mykland, 2008), and broadly assess the dynamical features of the price process (Andersen, Bollerslev, and Diebold, 2007; Todorov and Tauchen, 2010; Aït-Sahalia and Jacod, 2009a, among many others). Estimates obtained from power and multipower variations on individual stocks are basic ingredients in several financial applications. For example, using 5-minutes high-frequency returns, Zhang, Zhou, and Zhu (2009) show that multipower-based estimates of volatility and jumps explain the variation in the credit default swaps premium; Bradley, Clarke, Lee, and Ornthanalai (2014) use jump tests to show the relevance of analysts' recommendations on stock price movements; and Amaya, Christoffersen, Jacobs, and Vasquez (2015) use realized power variation to estimate skewness and kurtosis and study their impact in pricing the cross section of expected stock returns. In all empirical applications, including the above mentioned ones the presence of zeros is customarily ignored. This paper studies the impact of a general and mathematically tractable form of flatness on all multipower estimators, including realized volatility as a special case.

The reason why it is crucial to understand the impact of flatness on multipower variation is the following. Jumps and flatness are antithetic ingredients of the data generating process. Jumps model price movements, either in the form of sudden, large returns, or in the form of small, but infinitely many, discontinuities (Aït-Sahalia and Jacod, 2012). Flatness models the absence of movement, and materializes in the form of stale prices, or zero returns. Jumps are a widely accepted driver of the price dynamics. Flatness is instead mostly ignored. However, we show here that when looking at the data through the lens of multipower variation, jumps and flatness produce the same effects. Thus, when ignored, flatness can be mistaken for jumps. In our model, flatness can look like jumps since, after a sequence of flat trades, the next observed price cumulates unobserved shocks in what may look as a large returns, but is actually the sum of small, unobserved returns. To restore reliable jump inference, we thus need to assume that flatness is an integral part of the data generating process in the first place,

[1]Exceptions are models with rounded prices (Delattre and Jacod, 1997; Li and Mykland, 2015), the uncertainty zone literature (Robert and Rosenbaum, 2011), and zero-augmented models, as in Hautsch, Malec, and Schienle (2013).

[2]The presence of zeros could stem from price discreteness, absence of trading, low volatility, reluctance to trade due to high transaction costs, asymmetric information (Bandi, Pirino, and Renò, 2017), institutional and technical factors, and from the combined effects of the above.

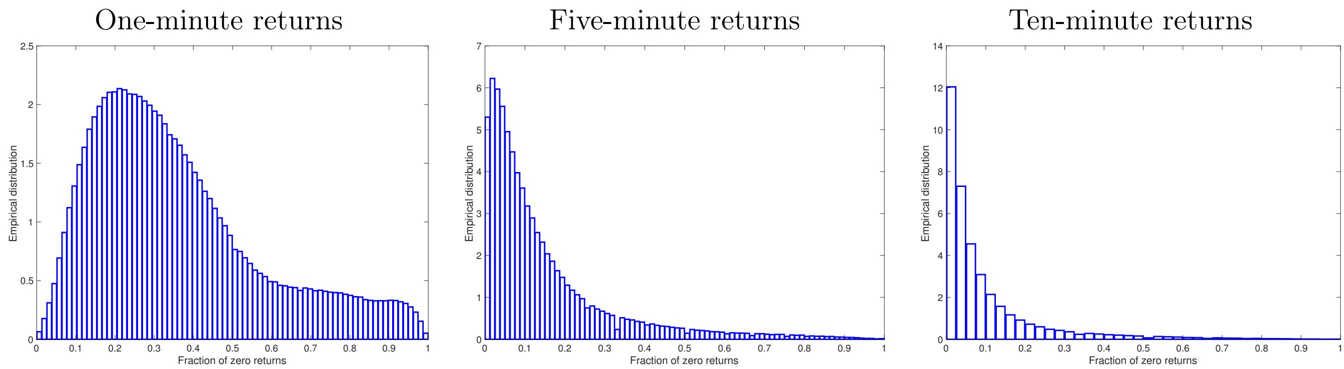

FIGURE 1: Pooled distribution of the daily percentage of zero returns at the one-minute frequency (left panel) and five-minute frequency (center panel) and ten-minute returns (right panel) of S&P 500 stocks. Specifically, we consider here 442 stocks and $4,385$ days, from January 1998 to June 2015. The daily percentage of zeros is computed excluding daily opening and closing periods with zero volume.

and robustify multipower estimators accordingly. The main contribution of this paper is to show that this can be easily done. Importantly, when using flatness-robust estimators, the statistical properties of jumps change dramatically with respect to what routinely found in the empirical literature. Thus, more broadly, our paper shows how compelling it is to have flatness under the null when modeling high-frequency returns.

Borrowing from the existing literature, we introduce a simple, still general form of flatness under the null and provide two theoretical contributions. We first calculate the bias of multipower variation under independent flatness, showing that realized volatility is the unique case which turns out to be consistent independently from the level of flatness, generalizing the results of Phillips and Yu (2009). In all other cases a bias appears, which consists in a multiplicative factor.[3] Our second contribution is then to propose a robust multipower estimator, inspired by the work of Hayashi, Jacod, and Yoshida (2011) and Levine, Wang, and Zou (2016), which is unbiased under the general form of flatness we assume, and provide its asymptotic distribution. Our flatness-robust multipower estimator can be used to obtain unbiased estimates of desired jumps and volatility functionals which are relevant for applications. The correction we propose is handy and requires no further computational effort. The flatness-robust multipower estimator coincides with the traditional one if there are no zeros in the time series under investigation.

We then show some detrimental consequences of using traditional multipower estimators on simulated data which are, as the real ones, contaminated by flatness. First, traditional jump tests are strongly distorted toward rejection of the null, inducing a large number of false positives which are actually due to flat prices. We focus, in particular, on tests based on the difference between realized volatility and bipower variation. Under the null in which the price moves continuosly, this difference is (asymptotically) normally distributed with a variance that can be estimated consistently (Barndorff-Nielsen, Graversen, Jacod, and Shephard, 2006). Under the alternative, the difference is asymptotically positive and represents the contribution of jumps to total quadratic variation (Andersen, Bollerslev, and

[3]In principle, this factor could be easily estimated and wiped out. However, this approach would deliver consistent estimates only under the restrictive assumption of independent flatness.

Diebold, 2007; Corsi, Pirino, and Renò, 2010; Busch, Christensen, and Nielsen, 2011). However, we show here that this difference is also positive when the alternative is flatness without jumps. This implies that tests based on this difference are not able to disentangle the two alternatives, and that the contribution of jumps to total variation using this difference is overestimated if there is flatness in the data. The bias toward rejection is also shared by other popular tests, as for example the one by Lee and Mykland (2008). The robustification we propose is able to correct for these biases and is correctly sized under a null which includes the presence of flatness.

Second, returns driven by Brownian shocks appear to be spuriously driven by a pure jump process when they are contaminated by flatness. The reason for this misleading effect is that traditional measures of jump activity on data which contain zero returns are negatively biased, suggesting that rejection of the Brownian motion could be actually due to the presence of flatness. We mainly consider two estimators of the jump activity index (Aït-Sahalia and Jacod, 2009a), the two-scale estimator of Todorov and Tauchen (2010), based on the ratio of power variations at different frequencies, and the multipower ratio estimator of Kolokolov (2017).[4] We show that both estimators are negatively biased in the presence of zero returns, with the bias increasing with their fraction in the data. Again, our robustified estimator is able to correct for these biases and restore consistent inference of the jump activity index under the assumed null (which, again, includes flatness), thus allowing to truly assess the nature of the driving force of the returns' shocks.

Our empirical contribution is then to reconsider jump properties in financial time series using our flatness-robust multipower estimator. We use data on individual stocks, the market portfolio index, and the volatility index. The examined assets are stocks belonging to the S&P 500 index, the SPY exchange-traded fund (that is, the SPDR S&P 500 trust), and the VIX index as computed by CBOE, all over a long time span. We first show that, after correcting the estimators for flatness, the number of jumps detected in stocks and the contribution of jumps to total quadratic variation are three and six time smaller, respectively, with respect to what found without correcting. These findings help to solve the puzzle put forward by Christensen, Oomen, and Podolskij (2014), who also advocate a much lower contribution of jumps to price variation using ultra-high frequency data. Our results point out that the discrepancy in measures of jump variation at different frequency found in the literature is a spurious by-product of the presence of flatness in high-frequency returns.

We then show that estimates of the activity index on all our assets are not significantly different from 2 (the value implied by the Brownian motion) after correcting for flatness. This supports the presence of the Brownian motion in asset prices. In particular, traditional estimators of the jump activity index estimate a value lower than 2 for the VIX index, which would imply that the dynamics of volatility is driven by a pure, infinite activity jump process (Todorov and Tauchen, 2011b), in contrast with traditional stochastic volatility models typically assumed in asset pricing. However, our data analysis is very sharp in showing that traditional activity estimates are affected by the fraction of zero returns in the data, being strongly negatively correlated with this fraction. We thus argue that the alleged evidence of pure jumps in volatility is also an artifact due to the unaccounted presence of flatness. Using

[4]Alternative estimators of the jump activity index are proposed by Jing, Kong, and Liu (2011); Todorov (2015). A test for a jump activity index less than two is proposed by Jing, Kong, and Liu (2012). These approaches, which are also impacted by flatness, are not based on multipower variation only, so the correction we propose would be insufficient to remove their biases.

robustified estimators, the relation between the level of flatness and the estimated activity disappears, and the activity index of the VIX is invariably estimated to be indistinguishable from 2 for every day in the sample, in keeping with the inescapable presence of a Brownian motion in the volatility dynamics.

The paper is structured as follows. Section 2 lays down the theory for multipower estimators and their robustified counterparts. Section 3 shows the distortions of traditional jump statistics based on multipower estimators on realistic simulations of the price process. Section 4 containes our empirical application which reconsiders the estimates of jump features in financial data. Section 5 concludes.

2 Multipower variation under flatness

2.1 Model setting

We assume that all stochastic processes considered in the paper are defined on a rich enough probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in(-\infty,+\infty)}, \mathcal{P})$ with a filtration $(\mathcal{F}_t)_{t\in[0,t]}$ satisfying the usual conditions (Protter, 1990). We assume that the logarithmic price, starting at $X_0 \in \mathcal{F}_0$, evolves according to a continuous Itō semimartingale with mild conditions on the driving coefficients. We do not include jumps under the null because, throughout the paper, their inclusion is considered our main alternative.

Assumption 2.1 (***The unobserved price process***)**.** *The real-valued logarithmic latent price process* $X = \{X_t;\ t \in [0,T]\}$ *is a Brownian semimartingale*

$$dX_t = \mu_t dt + \sigma_t dW_t, \tag{2.1}$$

where μ_t *and* σ_t *are adapted processes, and* σ_t *is a.s. strictly positive,* W_t *is a Brownian motion. We assume* μ_t *to be locally bounded, and* σ_t *to be càdlàg. Further, there exists* $\gamma > 0$ *and* $C_3 > 0$ *so that*

$$\mathbf{E}\left[(\mu_s - \mu_t)^2 + (\sigma_s - \sigma_t)^2\right] \leq C_3(s-t)^\gamma,$$

for all $s, t \in [0,T]$ *and for* $s-t$ *small enough.*

The model (2.1) is the traditional continuous-time/continuous-price assumption in financial economics, ruling out the possibility of zero returns in the data generating process. The assumptions on the coefficients μ_t and σ_t are extremely mild and encompass virtually all parametric models used in practice. We assume to observe the price process over an equally spaced mesh $t_i = i\Delta_n$, $i = 0, \dots n$, of $n+1$ points, with $\Delta_n = T/n$.

We explicitly introduce flatness in the model under the null hypothesis using triangular arrays of potentially dependent Bernoulli variates, as in Bandi, Pirino, and Renò (2017). Notice that typical modifications of the model (2.1) to add discontinuities and/or market microstructure noise which is independent from price would still not generate any zero return in the data.

Assumption 2.2 (***Flat trades***)**.** *The observed logarithmic price process* $X' = \{X'_t;\ t \in (-\infty, +\infty)\}$ *is observed in the interval* $[0,T]$ *over a grid* $\{t_0, \dots, t_n\}$ *with* $t_i = iT/n$*, and is such that,*

$$X'_{t_i} = X_{t_i}(1 - B_{i,n}) + B_{i,n} X'_{t_{i-1}}, \tag{2.2}$$

where $B_{i,n}$ is a triangular array of Bernoulli random variables, for $i = 1, \dots, n$, such that, as $n \to \infty$,

$$\frac{1}{T}\sum_{i=1}^{n} \Delta_n B_{i,n} \xrightarrow{p} p^{\emptyset}, \tag{2.3}$$

with $p^{\emptyset} \in [0,1]$.

The above assumption is rather general since it does not restrict the dependence structure of the Bernoulli variates. Equation (2.3) is a simple law of large numbers which allows to interpret $p^{\emptyset}$ as the average probability of flat trading. In the simplest dependence structure, the Bernoulli variates are independent and identically distributed, as in Phillips and Yu (2009) and in Section 2.2, and each of them has $\mathcal{P}(B_{i,n} = 1) = p^{\emptyset}$. More generally, they can be time-varying and/or correlated among themselves, but still converging to a fixed average value in the limit. We do not allow for correlation of the Bernoulli variates with the price process X_t.

In the next Section, we start by studying the independent and identically distributed (i.i.d.) case. While restrictive, the simplicity of the i.i.d. case allows to compute the bias of multipower variation explicitly, and understand the impact of flat trading on multipower estimates clearly. In principle, if zeros were i.i.d., it would also be possible to correct for the bias in a straightforward and econometrically feasible way. After making these points, in Section 2.3 we propose a simple modification of the multipower estimator which delivers consistency and asymptotic normality under the general Assumption 2.2 plus natural restrictions on the dependence structure of the Bernoulli variates.

2.2 The case of i.i.d. flatness

The case of i.i.d flatness is particularly appealing given its simplicity, and it serves perfectly to illustrate the source and the magnitude of the bias when estimating integrated volatility powers. This case has been studied by Phillips and Yu (2009) to compute the distribution of realized volatility and the bias of quarticity under flat trading. We generalize their result by calculating explicitly the bias for all multipower estimators.

Assumption 2.3 (***i.i.d. flat trading***)**.** *The Assumption 2.2 holds with $B_{i,n}$ being a triangular array of independent and identically distributed Bernoulli random variables, independent from the latent price process, with*

$$\mathcal{P}\left(B_{i,n} = 1\right) = p^{\emptyset} \in]0,1[. \tag{2.4}$$

We define the returns of the observed price process as:

$$\Delta_i X' = X'_{(i+1)T/n} - X'_{iT/n}.$$

Without loss of generality, in what follows we set $T = 1$. For a vector of m positive real numbers $r = [r_1, ..., r_m]$, and a general stochastic process X, we define realized multipower variation as

$$\mathsf{MV}(X;r) = \frac{1}{n}\sum_{i=1}^{n-m+1} \left|\sqrt{n}\cdot\Delta_i X\right|^{r_1} \dots \left|\sqrt{n}\cdot\Delta_{i+m-1} X\right|^{r_m}. \tag{2.5}$$

Power variation is obtained as a restriction of $\mathsf{MV}(X;r)$ when the vector r has a single power. The next theorem quantifies the bias multipower variation under iid flat trading. It turns out that the corresponding volatility integral is estimated up to a correction factor which depends only on the probability of flat trading.

Theorem 2.4 (***Law of Large Numbers for multipower variation under iid flat trading***)**.** *Under Assumptions 2.1 and 2.3, as $n \to \infty$,*

$$\mathsf{MV}(X';r) \xrightarrow{p} \left(\prod_{j=1}^{m} \mu_{r_j}\right) \frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}} \mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) \int_0^t |\sigma_s|^{r_+}\, ds, \tag{2.6}$$

where $r_+ = r_1 + ... + r_m$, and $\mathsf{Li}_s(z)$ denotes the polylogarithm function of order s and argument z, and where $\mu_s = \mathbf{E}\left[|U|^s\right] = \frac{2^{s/2}\Gamma((1+s)/2)}{\Gamma(1/2)}$ with U being a standard normal random variable.

The polylogarithm function appearing in Eq. (2.6) and below is a special function defined by the power series

$$\mathsf{Li}_r(x) = \sum_{k=1}^{\infty} \frac{1}{k^r} x^k, \tag{2.7}$$

and is available in standard software packages.

Flatness implies a multiplicative bias which depends just on the first power r_1, and the probability of a flat trade $p^{\emptyset}$. The bias is produced by two competing effects. The first part of the bias is the attenuation due to the fact that the product of consecutive returns may be zero because of flatness. This is a downward bias, which in the i.i.d. case is equal to $(1-p^{\emptyset})^m$. The second part of the bias is a volatility inflation effect, due to the fact that when the product of consecutive returns is non-zero, the first return in the multiplication is computed by the difference of prices whose distance depends on the run length of zero returns preceding the first return. This can be an upward or downward bias, depending on the power r_1, and, in the i.i.d. case, this is equal to $\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right)(1-p^{\emptyset})/p^{\emptyset}$. Since only the first return is affected by prior flatness, the bias depends only on the first power r_1, with the remaining powers playing no role.

Clearly, power variation is a special case of multipower variation (when $m = 1$) and is affected by the same bias. As it is well known, power variation will not remain consistent if jumps are introduced in the price dynamics, while the consistency of multipower estimators with $m > 1$ is robust to the presence of jumps.[5]

Corollary 2.5 (***LLN for power variation***)**.** *Under Assumptions 2.1 and 2.3, for any positive real number r, as $n \to \infty$,*

$$\mathsf{PV}(X';r) = \frac{1}{n}\sum_{i=1}^{n} \left|\sqrt{n}\cdot \Delta_i X'\right|^r \xrightarrow{p} \mu_r \frac{(1-p^{\emptyset})^2}{p^{\emptyset}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\emptyset}\right) \int_0^t |\sigma_s|^r\, ds. \tag{2.8}$$

[5]Jumps could arrive together with flatness. In this case, they will be materialized in the next non-flat return, and annihilated by the next return.

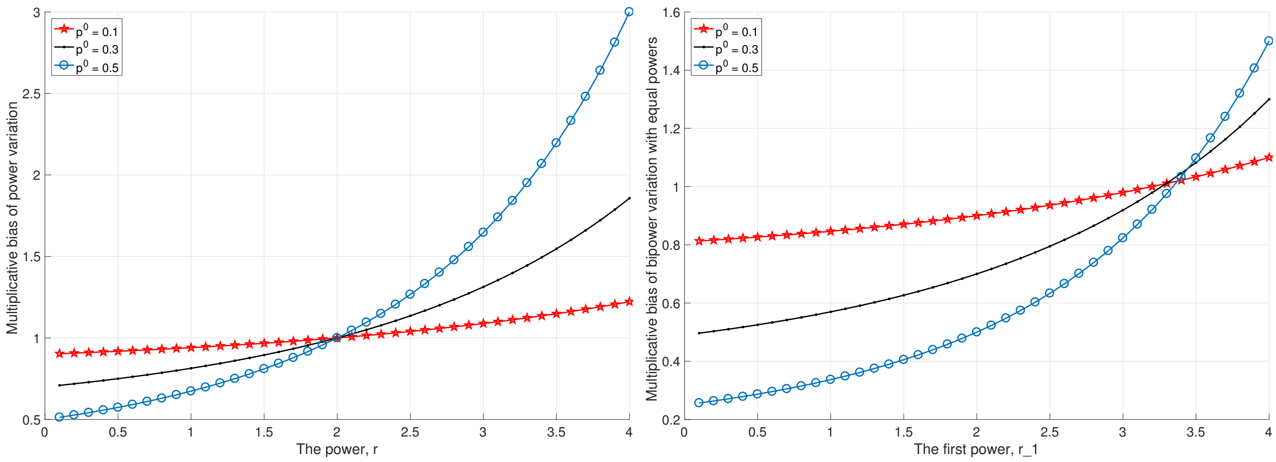

Figure 2: The multiplicative bias of power variation (left panel) and bipower variation with equal powers (right panel) as a function of the powers for different flat trade probabilities, $p^{\emptyset}$.

For the polylogarithm function, it holds:

$$\mathsf{Li}_{-1}\left(p^{\emptyset}\right) = \frac{p^{\emptyset}}{(1-p^{\emptyset})^2} \tag{2.9}$$

which implies that, when $r_1 = 2$, the multiplicative bias takes the form $(1-p^{\emptyset})^{m-1}$. For power variation, for which $m = 1$, thus there is no bias if and only if $r = 2$. Consequently, as also noticed by Phillips and Yu (2009), realized variance, that is $\mathsf{PV}(X', 2)$, does not display any bias due to flat trading, and remains a consistent estimator of integrated variance even under the presence of flat trading. In all other cases, there is a bias. Since

$$\frac{(1-p^{\emptyset})^2}{p^{\emptyset}}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\emptyset}\right)\begin{cases} < 1, & r < 2, \\ = 1, & r = 2, \\ > 1, & r > 2, \end{cases} \tag{2.10}$$

power variation overestimates integrated power variance for powers larger than 2 and underestimates for powers smaller than 2.

Figure 2 displays the multiplicative bias for power variation (left) and bipower variation (right) as a function of the first power, for various choices of the probability of flat trading $p^{\emptyset}$. For power variation, the estimator is asymptotically smaller than the estimation target when $r < 2$, and larger when $r > 2$. For bipower variation, the estimator is unbiased when $r_1 \approx 3.25$ and downward biased for smaller first power. In the typical case used in the empirical literature ($r_1 = r_2 = 1$), bipower variation is downard biased. Larger probability of flatness implies larger bias both for power and multipower estimators.

In the i.i.d. case, we can simply correct the bias after estimating the probability of flat trading, $p^{\emptyset}$.

Under Assumption 2.3, it can be consistently estimated by

$$\widehat{p} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{\Delta_i X' = 0\}}. \tag{2.11}$$

Hence, the *corrected* multipower variation is defined as

$$\mathsf{MV}^{\mathsf{c,iid}}(X'; r) = \left(\frac{(1-\widehat{p})^{m+1}}{\widehat{p}} \mathsf{Li}_{-\frac{r_1}{2}}(\widehat{p}) \right)^{-1} \mathsf{MV}(X'; r). \tag{2.12}$$

However, the requirement of i.i.d. flatness is theoretically and empirically too restrictive. For this reason, in the next Section we follow a different route by replacing the traditional estimator (2.5) with one specifically devised to achieve consistency and asymptotic normality under dependent flatness.

2.3 Relaxing the i.i.d. assumption

When flatness is not i.i.d., the bias illustrated in the above section is still present, but its precise evaluation is more complicated, as it depends on the specific structure of the Bernoulli variates. However, we can take advantage of the fact that, under Assumptions 2.1 and 2.2, the observed price process X' can be viewed as a stochastic process recorded at stochastic sampling times. We can then borrow from the approach of Hayashi, Jacod, and Yoshida (2011) and Levine, Wang, and Zou (2016) to define a flatness-robust multipower estimator and describe its asymptotic properties. The additional condition required to the Bernoulli variates to obtain a cental limit theory for multipower variation under flatness, which will be conveniently written in terms of the moments of the run lengths of the Bernoulli variates conditional to observing a non-zero return.

Given the initial time grid of price observations $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$, we define the stochastic grid of N_n points $\{\tau_0, \ldots, \tau_{N_n}\}$ such that returns are non-zero recursively as

$$\tau_0 = t_k, \; k = \min\{j : \; B_{j,n} = 0\}, \quad \tau_l = t_k, \; k = \min\{j : \; B_{j,n} = 0, \; t_j > \tau_{l-1}\}. \tag{2.13}$$

The (stochastic) set $\{\tau_0, \ldots, \tau_{N_n}\}$ is a subset of $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$. We define the corrected, flatness-robust multipower variation as

$$\mathsf{MV}^{\mathsf{c}}(X; r) = \sum_{i=1}^{N_n - m} \left| \Delta(n,i)^{-1/2} \Delta_{\tau_i} X \right|^{r_1} \ldots \left| \Delta(n, i+m-1)^{-1/2} \Delta_{\tau_{i+m-1}} X \right|^{r_m}, \tag{2.14}$$

where

$$\Delta_{\tau_i} X = X_{\tau_i} - X_{\tau_{i-1}} \tag{2.15}$$

and

$$\Delta(n,i) = \tau_i - \tau_{i-1}. \tag{2.16}$$

Thus, the estimator (2.14) is a multipower estimator applied to non-zero returns rescaled using their inter-trade duration $\Delta(n,i)$. The variables $\Delta(n,i)$ can also be interpreted as the run lengths of the Bernoulli variates, that is the numbers of consecutive zeros after a non-zero return, multiplied by $1/n$.

The estimator in Eq. (2.14) coincides with that in Eq. (2.5) when $B_{i,n} = 0$ identically, that is in the absence of flat trading.

To express the asymptotic variance of the corrected multipower estimator, we need to specify the volatility dynamics. In particular we assume the following:

Assumption 2.6 (***Volatility dynamics)***. *The volatility process σ_t is an Itō semimartingale evolving as:*

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s + \widetilde{\kappa}(\widetilde{\delta}) \star (\underline{\mu} - \underline{\nu})_t + \widetilde{\kappa}'(\widetilde{\delta}) \star \underline{\mu}_t, \tag{2.17}$$

where $\underline{\mu}$ is a Poisson random measure on $(0,\infty) \times \mathbf{R}$ with intensity measure $\underline{\nu}(dt,dx) = dt \otimes \lambda(dx)$; $\widetilde{\kappa}(x)$ is a truncation function and $\widetilde{\kappa}'(x) = 1 - \widetilde{\kappa}(x)$; $\widetilde{\delta}(\omega,t,x)$ is a predictable function on $\Omega \times \mathbf{R}_+ \times \mathbf{R}$. Moreover, for some constant Λ,

$$\left|\tilde{b}_t(\omega)\right| + |\tilde{\sigma}_t(\omega)| \leq \Lambda, \quad \left|\widetilde{\delta}(\omega,t,x)\right| \leq \Lambda\left(\widetilde{\gamma}(x) \vee 1\right), \tag{2.18}$$

where $\widetilde{\gamma}(x)$ is a non-random nonnegative function such that $\int_{\mathbf{R}} \left(\widetilde{\gamma}(x) \vee 1\right) \lambda(dx) < \infty$.

We can now establish a Central Limit Theorem for the estimator (2.14) under mild restriction of the conditional moments of the Bernoulli run lenghts $\Delta(n,i)$.

Theorem 2.7 (***Central Limit Theorem)***. *Assume that Assumptions (2.1), (2.2) and (2.6) hold. Define the deterministic sequence $\varkappa_n = \mathbf{E}\left[\Delta(n,i)\right]$ and assume:*

$$\varkappa_n = O(1/n), \quad N_n = O_p(n), \tag{2.19}$$

$$\mathbf{E}\left[(\Delta(n,i))^q \,|\mathcal{F}_{\tau_{i-1}}\right] \leq C_q \cdot \varkappa_n^q, \quad q \geq 1. \tag{2.20}$$

$$\frac{\mathbf{Var}\left[\Delta(n,i)|\mathcal{F}_{\tau_{i-1}}\right]}{\varkappa_n^2} = \Upsilon + o_p(1/n), \tag{2.21}$$

where $\Upsilon < \infty$ is a non-negative constant and $C_q > 0$ are constants depending on $q \geq 1$. Then, there exists a deterministic sequence $\varkappa_n = \mathbf{E}\left[\Delta(n,i)\right]$, such that as $n \to \infty$,

$$\varkappa_n \,\mathsf{MV}^{\mathsf{c}}(X;r) \xrightarrow{p} \prod_{j=1}^m \mu_{r_j} \int_0^t |\sigma_s|^{r_+} ds, \tag{2.22}$$

$$\frac{1}{\sqrt{\varkappa_n}}\left(\varkappa_n \,\mathsf{MV}^{\mathsf{c}}(X;r) - \prod_{j=1}^m \mu_{r_j} \int_0^t |\sigma_s|^{r_+} ds\right) \xrightarrow{L-st} U_t + Z_t, \tag{2.23}$$

where U_t and Z_t are independent Gaussian random variables with

$$\mathbf{Var}\left[U_t\right] = \left(\prod_{l=1}^m \mu_{2r_l} - (2m-1)\prod_{l=1}^m \mu_{r_l}^2 + 2\sum_{k=1}^{m-1}\prod_{l=1}^k \mu_{r_l} \prod_{l=m-k+1}^m \mu_{r_l} \prod_{l=1}^{m-k} \mu_{r_l + r_{l+k}}\right) \int_0^t |\sigma_s|^{2r_+} ds, \tag{2.24}$$

$$\mathbf{Var}\left[Z_t\right] = \Upsilon \prod_{j=1}^m \left(\mu_{r_j}\right)^2 \int_0^t |\sigma_s|^{2r_+} ds, \tag{2.25}$$

where the convergence in (2.23) is stable in law.

The asymptotic limit of the $\mathsf{MV}^c(X', r)$, the flatness-robust multipower estimators, can be written as the sum of two random variables. The first one, U_t, is the "traditional" one, that is the only one that would appear when $B_{i,n} = 0$ identically. The additional term Z_t is originated by the presence of flatness. It is the limit of the difference:

$$\frac{1}{\sqrt{\varkappa_n}} \prod_{j=1}^{m} \mu_{r_j} \left(\sum_{i=1}^{N_n} \left| \sigma_{\tau_{i-1}} \right|^{r_+} \varkappa_n - \int_0^t |\sigma_s|^{r_+} \, ds \right), \tag{2.26}$$

which would be 0 if $B_{i,n} = 0$ identically. The statement of the CLT can be expressed as:

$$\frac{1}{\sqrt{\varkappa_n}} \left(\varkappa_n \left(\prod_{j=1}^{m} \right)^{-1} \mathsf{MV}^{\mathsf{c}}(X; r) - \int_0^t |\sigma_s|^{r_+} \, ds - \sqrt{\varkappa_n} \widetilde{Z}_t \right) \xrightarrow{L-st} \widetilde{U}_t(r), \tag{2.27}$$

where the random variable $\widetilde{Z}_t = \left(\prod_{j=1}^{m} \right)^{-1} Z_t$ depends on the sum of powers r_+, but not on individual powers, while $\widetilde{U}_t(r) = \left(\prod_{j=1}^{m} \right)^{-1} U_t(r)$ depends on the individual powers $r_1, \ldots, r_m$. So, if we consider the difference between two multipowers with the same r_+ but different powers (as in the case of realized volatility and bipower variation), the asymptotic distribution will not depend on $\widetilde{Z}_t$, and the correlation between the corresponding random variables $\widetilde{U}_t(r)$ is computed as in the "usual" case without flatness.

The conditions (2.19)-(2.21) can be interpreted as conditions on the dependence structure of the Bernoulli variates necessary for the CLT to hold. To better understand their meaning, we re-write them in the i.i.d. case when, the distribution of $\Delta(n, i)$ conditionally on $\mathfrak{F}_{\tau_{i-1}}$ can be represented as

$$\mathcal{P} \left(\Delta(n, i) = x | \mathfrak{F}_{\tau_{i-1}} \right) = \begin{cases} \left(1 - p^{\emptyset}\right), & x = \Delta_n, \\ \left(1 - p^{\emptyset}\right) p^{\emptyset}, & x = 2\Delta_n, \\ \left(1 - p^{\emptyset}\right) \left(p^{\emptyset}\right)^{k-1}, & x = k\Delta_n, \quad k = 3, 4, \ldots \end{cases} \tag{2.28}$$

where $\Delta_n = \frac{1}{n}$. Consequently, for every $q \geq 1$,

$$\mathbf{E} \left[(\Delta(n, i))^q \, | \mathcal{F}_{\tau_{i-1}} \right] = C_q^{iid} \left(\Delta_n \right)^q, \tag{2.29}$$

where the q-dependent constant C_q^{iid} can be simply evaluated as:

$$C_q^{iid} = \left(1 - p^{\emptyset}\right) \sum_{k=1}^{\infty} k^q \left(p^{\emptyset}\right)^{k-1} = \frac{\left(1 - p^{\emptyset}\right)}{p^{\emptyset}} \mathsf{Li}_{-q} \left(p^{\emptyset}\right). \tag{2.30}$$

Similarly, in the i.i.d. case,

$$\varkappa_n = C_1^{iid} \Delta_n = \frac{\Delta_n}{1 - p^{\emptyset}}$$

so that

$$\Upsilon = p^{\emptyset}.$$

The above computation clarifies that the condition (2.20) for the non i.i.d. case requires that the moments of the run lenghts have the same decay rate of the i.i.d. case.

For example, consider a simple non i.i.d. model with dependent Bernoulli variates. The model has two parameters $p^{\emptyset\emptyset}$ and $p^{\emptyset 1}$, and its dynamics is defined as:

$$\begin{aligned} \mathcal{P}\left(B_{i,n}=1|B_{i-1,n}=1\right) &= p^{\emptyset\emptyset}, \\ \mathcal{P}\left(B_{i,n}=1|B_{i-1,n}=0\right) &= p^{\emptyset 1}. \end{aligned} \tag{2.31}$$

The unconditional probability of flat trading is given by

$$\overline{p}^{\emptyset}=\frac{p^{\emptyset 1}}{1-p^{\emptyset\emptyset}+p^{\emptyset 1}}. \tag{2.32}$$

We finally assume that $B_{0,n}$ has probability $\overline{p}^{\emptyset}$. In this case we are modeling the dependence allowing the probability of a zero to depend only on the previous value of the Bernoulli trial. The model coincides with the i.i.d. case when $p^{\emptyset\emptyset}=p^{\emptyset 1}=\overline{p}^{\emptyset}$. In this case, we have:

$$\mathcal{P}\left(\Delta(n,i)=x|\mathfrak{F}_{\tau_{i-1}}\right)=\begin{cases}\left(1-p^{\emptyset 1}\right), & x=\Delta_n,\\ p^{\emptyset 1}\left(1-p^{\emptyset\emptyset}\right), & x=2\Delta_n,\\ p^{\emptyset 1}p^{\emptyset\emptyset}\left(1-p^{\emptyset\emptyset}\right), & x=3\Delta_n,\\ p^{\emptyset 1}(p^{\emptyset\emptyset})^{(k-2)}\left(1-p^{\emptyset\emptyset}\right), & x=k\Delta_n, \quad k=4,5,...\end{cases} \tag{2.33}$$

so that:

$$\mathbf{E}\left[(\Delta(n,i))^q\,|\mathcal{F}_{\tau_{i-1}}\right]=C_q'^{iid}\left(\Delta_n\right)^q, \tag{2.34}$$

where now

$$C_q'^{iid}=1-p^{\emptyset 1}+\frac{(1-p^{\emptyset\emptyset})p^{\emptyset 1}}{(p^{\emptyset\emptyset})^2}\left(\mathsf{Li}_{-q}\left(p^{\emptyset\emptyset}\right)-p^{\emptyset\emptyset}\right), \tag{2.35}$$

$$\varkappa_n=C_1'^{iid}\Delta_n=\frac{\Delta_n}{1-\overline{p}^{\emptyset}},$$

and

$$\Upsilon=p^{\emptyset 1}\frac{1+p^{\emptyset\emptyset}-p^{\emptyset 1}}{(1-p^{\emptyset\emptyset}+p^{\emptyset 1})^2}.$$

which shows that the restrictions (2.19)-(2.21) are also satisfied in this case. More complicated dependent structures similar to (2.31) can be accommodated accordingly. However, notice that the asymptotic variance of the estimator can be estimated without assuming any specific dependence structure of the Bernoulli variates, using Eq. (2.20).

3 Distortions to jump inference

In this section we show, using simulated data, that the bias in estimating integrated volatility powers using multipower estimators, originated by the presence of flatness in the data generating process, distorts popular jump statistics and jeopardizes jump inference. We then show that using the flatness-robust multipower estimator (2.14) is a solution to this issue.

The model we simulate for the observed price process X_t' is given by Eq. (2.2), where the latent price dynamics is driven by

$$\begin{aligned} dX_t =& \mu\, dt + \gamma_t \sigma_t dW_t \\ d\log\sigma_t^2 =& (\alpha - \beta \log\sigma_t^2)\, dt + \eta dW_t^\sigma, \end{aligned} \tag{3.1}$$

where W and W^σ are standard Brownian motions with $\operatorname{corr}(dW, dW^\sigma) = \rho$, σ_t is a stochastic volatility factor and γ_t is an intraday effect. We use parameters estimated by Andersen, Benzoni, and Lund (2002) on S&P500 prices and expressed in daily units for percentage returns: $\mu = 0.0304$, $\alpha = -0.012$, $\beta = 0.0145$, $\eta = 0.1153$, $\rho = -0.6127$, and we use:

$$\gamma_t = \frac{1}{0.1033}(0.1271t^2 - 0.1260t + 0.1239),$$

as estimated by us on S&P500 intraday returns, where t is expressed as a fraction of one day.

We discretize the model for X in the interval $[0,1]$ with the Euler scheme using a time step of 1/78, corresponding to 5 minutes returns in a 6.5 hours trading day, or 1/390, corresponding to 1 minute. We then simulate flatness through Eq. (2.2). The dynamics of zero returns id generated using Bernoulli variates $B_{i,n}$ given by the model (2.31), which allows for dependency. The model has two parameters, expressing the probability of a zero return after a zero and after a non-zero respectively. Of course, the parameters depend on the sampling frequency. When we simulate returns at the 5-minutes frequency, unless otherwise stated, we set $p^{\emptyset\emptyset} = 18.52\%$ and $p^{\emptyset 1} = 12.77\%$ which are the sample average of estimates obtained on stock data using sample conditional expectations of zeros at five-minute frequency. The fact that, in stock data, $p^{\emptyset\emptyset} > p^{\emptyset 1}$ significantly means that zeros are persistent (that is, a zero is more likely after a zero than otherwise) and suggests that the simple i.i.d dynamics postulated in Phillips and Yu (2009) is unrealistic. Our Monte Carlo simulation accommodates this feature of the data.[6] For all simulation sets, we use $10,000$ replications.

3.1 Distortions to jump detection

Multipower variation is extensively used for jump testing. Here we mainly study a paradigmatic example, namely the BNS test of Barndorff-Nielsen and Shephard (2006). The BNS test is defined, as suggested in Huang and Tauchen (2005), as:

$$\mathsf{BNS} = \frac{1 - \frac{\mathsf{BPV}}{\mathsf{RV}}}{\sqrt{\frac{\theta}{n}\max\left(1, \frac{\mathsf{TriPV}}{\mathsf{BPV}^2}\right)}}, \tag{3.2}$$

where $\mathsf{RV} = \mathsf{MV}(X', [2])$, $\mathsf{BPV} = (\mu_1)^{-2}\,\mathsf{MV}(X', [1,1])$, $\mathsf{TriPV} = \left(\mu_{4/3}\right)^{-3}\mathsf{MV}(X', [4/3, 4/3, 4/3])$, and $\theta \approx 0.61$. The test is shown to be asymptotically (as $n \to \infty$) standard normal under Assumption 2.1, when $X' = X$. However, under flatness, the test behaves quite differently. Intuitively, the presence of unexpected zeros inflates the numerator since, while realized variance is still estimated consistently,

[6] Alternatively, we could also impose a deterministic intraday effect in the frequency of zeros, which also appears quite pronounced in the data. However the results provided in this Section would be qualitatively and quantitatively the very same.

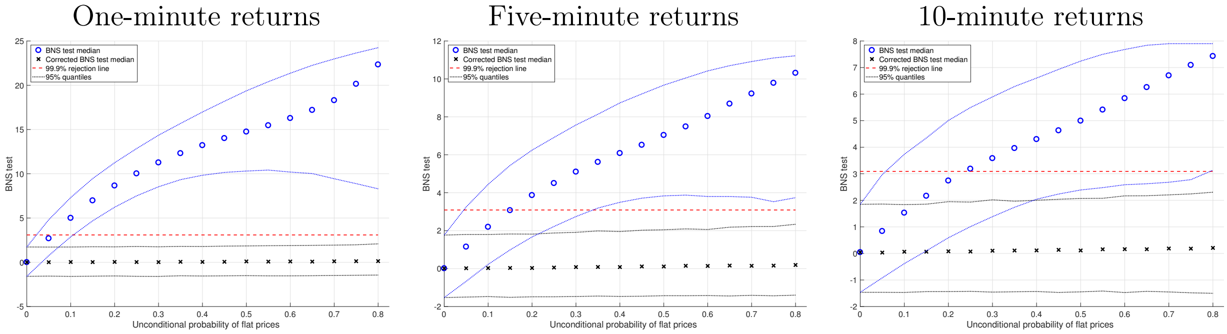

FIGURE 3: Quantiles (median, 5% and 95%) of the standard BNS test and the BNS^c version of the test corrected for flatness on simulated data, as a function of the generated probability of flatness. Left panel: $n = 390$ (1 minute returns). Center panel: $n = 78$ (five-minute returns). Right panel: $n = 39$ (ten-minute returns).

bipower variation is downard biased. Additionally, flatness deflates the denominator since the quarticity estimate is also downard biased. Both these effects inflate the BNS test, distorting it toward rejection of the null, with the distortion increasing with the unconditional probability of flatness. The solution to this issue is to use the flatness-robust multipower estimators in place of standard ones in (3.2). We write

$$\mathsf{BNS}^c = \frac{1 - \frac{\mathsf{BPV}^c}{\mathsf{RV}^c}}{\sqrt{0.61 \cdot \varkappa_n \cdot \max\left(1, \frac{\mathsf{TriPV}^c}{(\mathsf{BPV}^c)^2}\right)}} \xrightarrow{L} \mathcal{N}(0,1), \tag{3.3}$$

where $\varkappa_n = \mathbf{E}\left[\Delta(n,i)\right]$, $\mathsf{RV}^c = \mathsf{MV^c}(X', [2])$, $\mathsf{BPV}^c = (\mu_1)^{-2}\,\mathsf{MV^c}(X', [1,1])$, $\left(\mu_{4/3}\right)^{-3}\mathsf{TriPV}^c = \mathsf{MV^c}(X', [4/3, 4/3, 4/3])$. By Theorem 2.7, we have that BNS^c is asymptotically standard normal under Assumption 2.2, that is under flatness. Note that, despite of the fact that RV is a consistent estimator of the quadratic variation even under the presence of flat trades, it is still necessary to replace it with the corrected version RV^c in Eq. (3.3) in order to preserve the asymptotic distribution of the test statistic.

Figure 3 shows median and 95% quantiles of the distribution of the two competing test statistics (BNS and BNS^c) under our simulation setting, which includes flatness but in which jumps are absent. We analyze daily simulations sampled at two frequencies: one minute ($n = 390$, left panel) and five minutes ($n = 78$, right panel). Since we want to study the impact on the level of flatness, we use different levels of the unconditional probability of flatness $\overline{p}^{\emptyset}$, as defined by Eq. (2.32). For each set of simulations, we first set the value of $\overline{p}^{\emptyset}$, we then solve for the parameters $p^{\emptyset\emptyset}$ and $p^{\emptyset 1}$ after imposing the restriction $p^{\emptyset 1} = \frac{2}{3}p^{\emptyset\emptyset}$. This restriction reflects the heightened probability of consecutive flat trades observed in five-minute stock prices.

Figure 3, to be evaluated in conjunction with Figure 1, clearly shows the danger of ignoring zeros. When sampling at the one-minute frequency, even moderate levels of flatness such as 10% would imply extremely large distortions of the test statistics, such that more than 95% of the generations would appear as false positives. At five minutes, the impact is less destructive but still extremely worrying. For example, for an average flatness of 10% we have 25.92% of false positives. On the other hand, the corrected version of the estimator in which flatness is fully taken into account presents no distortions

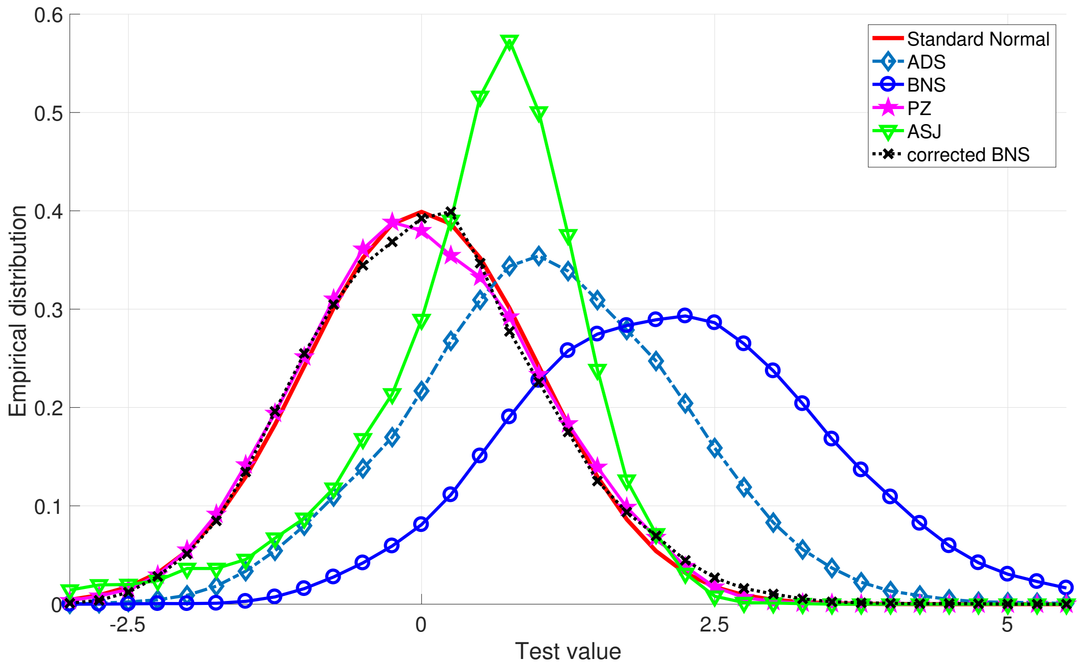

Figure 4: Distribution of several test statistics on simulated daily samples at the five minutes frequency ($n = 78$ returns). Flatness is generated with the conditional dependency specified in (2.31), with $p^{\emptyset\emptyset} = 18.52\%$ and $p^{\emptyset 1} = 12.77\%$. The considered tests are: the median test in Andersen, Dobrev, and Schaumburg (2012) (ADS); the BNS test in Eq. (3.2) (BNS); the Podolskij and Ziggel (2010) test (PZ); the Aït-Sahalia and Jacod (2009b) test (ASJ), and the corrected BNS test in Eq. (3.3).

at both frequencies, even for extremely large probability of flatness, with a median centered in zero and a distribution close to a standard normal.

This flatness-induced distortion toward spurious jump detection is typically shared by other jump testing strategies. Figure 4 shows the estimated distribution of different tests on daily simulations of five minute returns, using the average parameters for conditional flatness estimated on stocks. These tests are directly comparable since their asymptotic distribution should be standard normal under the assumed null in which jumps are absent. In addition to the BNS and BNS^c test, reported as benchmarks, we consider: i) the ADS median test proposed by Andersen, Dobrev, and Schaumburg (2012), which uses the same principle of the BNS test but propose a solution explicitly meant soften the impact of zero returns; ii) the ASJ test of Aït-Sahalia and Jacod (2009b), which uses the ratio of power variations at different frequency; and iii) the PZ test of Podolskij and Ziggel (2010), which is based on truncated variation.[7] The ADS test compares realized variance with an estimator of integrated variance based on the median of a number of consecutive absolute intraday returns. The usage of the median estimator is purposedly proposed to deal with the problem of zero returns. Indeed, as shown in Figure 4, the test is

[7]The ADS test we implement in Figure 4 is based on Eq. (6) in Andersen, Dobrev, and Schaumburg (2012), where IQ is estimated with the $MedRQ$ estimator. The ASJ test is implemented with power variation with power $p = 4$ calculated over the two scales (1/78 and 2/78); the variance of ASJ test statistics is estimated without truncating large returns, since we do not simulate jumps under the null. The PZ test is implemented with power $p = 2$, and the test statistics is perturbed by random draws from the distribution $\mathcal{P}^\eta = \frac{1}{2}(\delta_{1-\tau} + \delta_{1+\tau})$, where δ is the Dirac measure and $\tau = 0.05$. The threshold used in the PZ test is computed as in Corsi, Pirino, and Renò (2010).

less distorted than the BNS test. However, under flatness, even the median estimator is biased for the same reasons of multipower variation. Consequently, ADS is still distorted toward the rejection of the null, even if less than BNS. The ASJ test procedure in Aït-Sahalia and Jacod (2009b) is contaminated by flatness as well. Under flatness, the bias appears in both the numerator and the denominator of the proposed test statistics. Since the bias of power variations is larger for the large powers, the ASJ test statistic is still distorted. The PZ test is based on the difference between power variation and its truncated version, and makes use of random perturbation of returns (wild bootstrap) in order to set up a confidence region. As a result, despite of the bias in the power variation, the distribution of PZ test remains standard normal under the null by construction. Indeed, Figure 4 shows that the distribution of the PZ test under the null is as close to standard normal as the distribution of the BNS^c test.

The distortions due to flatness affect similarly other popular jump test statistics which were not reported here, and which are based on multipower variations as well. For example, the tests of Lee and Mykland (2008) and Lee and Hannig (2010) compare the magnitude of the absolute value of the returns to an estimate of spot volatility obtained with multipower variation. Under flatness, the spot volatility estimate will result to be artificially smaller, and the test again inflated and distorted toward rejection. The test proposed by Corsi, Pirino, and Renò (2010), who combine multipower variation and truncation, suffers of exactly the same bias of the BNS test. Alternative tests proposed by Andersen, Dobrev, and Schaumburg (2012) are based on the $minRQ$ and $minRV$ estimators, and are more distorted than that based on the med estimators and reported in Figure 4. Using a different Monte Carlo setting, Theodosiou and Zikes (2011) analyze the performance of several jump tests under flatness, also showing that all the ones they consider, including the tests mentiond here and the test based on variance swaps of Jiang and Oomen (2008), are distorted toward finding more jumps, with the exception of the Podolskij and Ziggel (2010) test.

3.2 Distortion to jump activity estimation

We now consider estimation of the jump activity index. Estimating the index is very important for specifying the data generating process observed at high frequency: the index is equal to 2 for all paths of Brownian semimartingales with jumps; it is generally less than 2 for pure jump processes. Consequently, estimating the index allows to decide whether Brownian motion is needed to describe financial data.

Formally, the jump activity index (from now on denoted BG, sinces it coincides with the Blumenthal-Getoor index for pure jump Lévy processes) of an Itō semimartingale X is defined as

$$BG = \inf \left\{ p \geq 0 : \sum_{s \leq t} |\Delta X_s|^p < \infty \right\}. \tag{3.4}$$

Under regular observation scheme without flat trading, the BG index can be estimated using power and multipower variation. We concentrate on two methodologies. The first method is due to Todorov and Tauchen (2010) and makes use of power variations analyzed at different frequencies. The BG estimator is defined as:

$$\widehat{\beta}_{TT} = \frac{\ln(2)\cdot p}{\ln(2)+\ln\left(V_n(X,p,2\Delta_n)\right)-\ln\left(V_n(X,p,\Delta_n)\right)}, \quad p>0, \tag{3.5}$$

where

$$V_n(X,p,\Delta_n) = \sum_{i=1}^{n} |\Delta_i X|^p .$$

Since flatness is more frequent at the highest frequency Δ_n than at the lowest frequency $2\Delta_n$, the denominator is expected to become bigger in the presence of zero returns, thus lowering the estimate of the BG index.

An alternative method is studied in Kolokolov (2017). A consistent estimator of the BG index can be defined as the solution of the equation:

$$Q(\widehat{\beta}; X, r_+) = 0, \tag{3.6}$$

where

$$Q(\beta; X, r_+) = \frac{\mathsf{MV}(X;[r_+])}{\mathsf{MV}(X;[r_+/2\ r_+/2])} - \frac{\mu_{r_+}(\beta)}{\mu_{r_+/2}(\beta)\mu_{r_+/2}(\beta)}, \tag{3.7}$$

where

$$\mu_r(\beta) = \frac{2^r\Gamma((1+r)/2)\Gamma(1-r/\beta)}{\Gamma(1-r/2)\Gamma(1/2)}, \tag{3.8}$$

the parameter β denotes the genuine unknown value of the BG index and r_+ is a positive number, which is strictly less that β. This latter method is more precise than the method of Todorov and Tauchen (2010), which is the reason why we analize it here. However, this is also distorted in the presence of flatness and, as for his competitor (3.5), it underestimates the BG index. The danger of ignoring flatness in this case is thus to estimate a jump "vibrancy" that is higher (that is with a smaller BG) than the real one, thus concluding that the price process is driven by pure jumps without Brownian motion shocks.

Again, our solution to this issue is simply to replace power/multipower variations in the definitions of the estimators by the corrected versions. Hence, the corrected BG estimator based on two-scale power variation, $\widehat{\beta}^c_{TT}$, is defined as

$$\widehat{\beta}^c_{TT} = \frac{\ln(2)\cdot p}{\ln(2)+\ln\left(V^c_n(X,p,2\Delta_n)\right)-\ln\left(V^c_n(X,p,\Delta_n)\right)}, \quad p>0, \tag{3.9}$$

where $V^c_n(X,p,\Delta_n)$ is the corrected version of $V_n(X,p,\Delta_n)$.

The corrected multipower-based BG estimator $\widehat{\beta}^c$, is accordingly defined as the solution of the following equation:

$$0 = \frac{\mathsf{MV}^{\mathsf{c}}(X;[r_+])}{\mathsf{MV}^{\mathsf{c}}(X;[r_+/2\ r_+/2])} - \frac{\mu_{r_+}(\widehat{\beta}^c)}{\mu_{r_+/2}(\widehat{\beta}^c)\mu_{r_+/2}(\widehat{\beta}^c)}. \tag{3.10}$$

We compare the estimates on simulated data at with a discretization step of five minutes. The left panel of Figure 5 shows the distribution on estimates of the BG index obtained with the original method ($\widehat{\beta}_{TT}$ in Eq. (3.5)) with that obtained with the flatness-robust estimator ($\widehat{\beta}^c_{TT}$ in Eq. (3.9)).

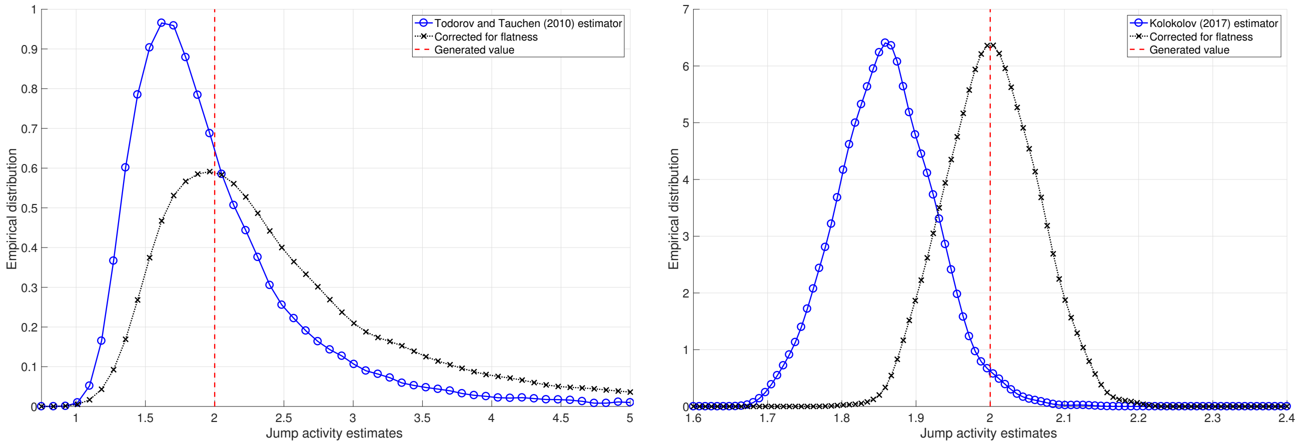

FIGURE 5: Distribution of the BG estimates on simulated data in which the jump activity is equal to 2. Left panel: Todorov and Tauchen (2010) estimator, original versus corrected version, implemented on one month of five-minutes returns ($n = 78 \cdot 22$). Right panel: Kolokolov (2017) estimator, original versus corrected version, implemented on one day of five-minutes returns ($n = 78$). The figure illustrates i) the bias of traditional estimator under flatness; ii) that robustified estimators using corrected multipowers provide unbiased estimates of the generated jump activity index; iii) that the Kolokolov (2017) estimator is more precise than that of Todorov and Tauchen (2010).

For each replication, we use a time-span of one month of five-minutes observations. The estimators are implemented with $p = 1$, which is optimal for estimating BG index in Brownian motion case (see Figure 1 in Todorov and Tauchen, 2011a). In our simulation settings jumps are absent and the true BG index is 2. Figure 5 clearly shows that ignoring the presence of flatness would result in an artificially lower estimate of the BG index, which would induce to reject the presence of a Brownian motion in the vast majority of replications. When correcting for flatness using the flatness-robust multipower estimators, the distribution is instead closer to the right value of 2, as it should be. However, Figure 5 also shows that the estimator of Todorov and Tauchen (2010) is quite noisy even when using one month of five-minutes observations ($78 \cdot 22$ returns). Both the corrected and the original estimator present a wide right tail and are far from resembling the Gaussian distribution which is predicted asymptotically.[8] This makes inherently difficult to disentangle a BG index which is exactly 2 from one that is slight lower using this estimator.

This is the reason why, given the importance of the issue, we also analyze an alternative estimator. The right panel of Figure 5 shows the analogous distribution of the estimates obtained with the multipower ratio of Kolokolov (2017). We implement the estimators with $r_+ = 1.5$, which is suitable for estimating jump activity of infinite variation pure jump processes as well as of Brownian semimartingales, and is consistent with the value used in the empirical application. We now use a shorter time-span of $n = 78$ five-minutes observations, corresponding to a single trading day. The Figure shows indeed that this estimator is much less disperse across replications, and that one day of observations is enough

[8]When increasing the sampling frequency of the returns, both the estimators become more symmetric and closer to a Gaussian distribution, as predicted by the theory, with the corrected estimator being centered around 2 and the traditional distribution being centered around a lower value which depends on the unconditional probability of flatness. However, the strategy of increasing the sampling frequency in the data is prevented by the unpredictable impact of market microstructure noise at the higher frequencies.

to estimate the BG index with sufficient precision. Again, ignoring the presence of flatness would result in a spuriously lower BG estimate, which would significantly reject the presence of the Brownian motion even on simulated data driven by Brownian motion. When correcting multipower estimators for flatness, the estimates are instead simmetrically centered around 2 across replications.

4 Revising the evidence on jumps in financial markets

The Monte Carlo experiments in the above Section show in a clear-cut way that neglecting the presence of zero returns in the data generating process would lead to identify spurious jumps. Reliable inference on jumps can be restored by using flatness-robust multipower variations. This Section is then devoted to re-evaluate the empirical evidence on jumps in financial markets using flatness-robust estimators. The Section is divided into two parts. The first is dedicated to the amount of jumps in financial markets, and the second to the estimation of the jump activity index.

4.1 How many times do stocks jump really?

The data set we use here is the collection of 442 stocks belonging to the S&P500 index and quoted on the New York Stock Exchange (NYSE). The sample ranges from January 2, 1998 to June 8, 2015, for a total of $4,385$ trading days, and consists of one-minute returns from 9:30 to 16:00, New York GMT. Moreover, we use prices of SPY (the SPDR S&P 500 trust exchange-traded fund) sampled at the one-minute frequency in the same daily time span, from January 1, 2001 to December 28, 2012, for a total of $3,017$ trading days. The data, coming from multiple exchanges and electronic networks, went trough a standard filtering procedure. Data are cross-checked, tested and verified so that outliers and bad ticks are removed. In our application, we mainly use 5-minute grids. The 5-minute frequency represents a customary tradeoff between achieving enough statistical power and avoiding distortions which could potentially arise from microstructure noise. We trim data at the beginning and at the end of the day with zero volume, so that we do not include zero returns resulting from delayed opening or anticipated market closure. We compute the tests only in days in which we have at least 20 returns, that is, 100 minutes of trading activity. In total, we compute $N = 1,817,932$ daily tests (plus $3,017$ daily tests on SPY).

Figure 6 shows the distribution of the BNS test and of its version corrected for flatness, BNS^c, on daily stocks, pooled across days and stocks. The BNS test distribution is far from a standard normal. It is centered around a positive value and largely skewed to the right. As a consequence, it would reveal jumps in 10.91% of the cases at the 99.9% confidence interval. This is the classic puzzle of the BNS test (Huang and Tauchen, 2005), that is that jumps appear to be too many in the data. Ironically, our Monte Carlo experiments suggest that this happens not because prices are too active, but because prices are too stale. Figure 7 shows an iconic example, that is a day in which the Bank of America (BAC) stocks moves extremely smoothly from the opening price of \$16.67 to the closing price of \$16.85. The largest 5-minute change in the data in this day is a modest \$0.10. The observed dynamics presents a large number of zeros, 21.79% of 5-minute changes, mainly clustered in the middle of the day. The BNS test value is 5.22, with a p-value of 10^{-8}, which would imply a highly significant jump at any

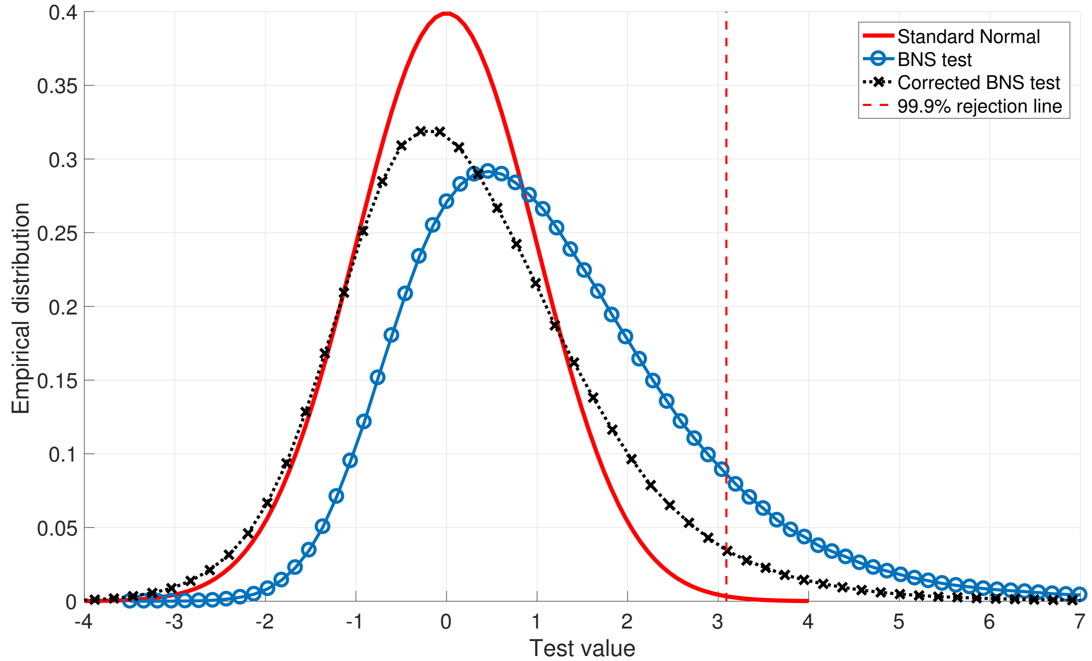

Figure 6: Pooled distribution, for all the considered S&P500 stocks from January 1998, to June 2015, of the daily BNS tests and corrected BNS^c tests using five-minutes returns. In total, we compute $1,817,932$ daily tests.

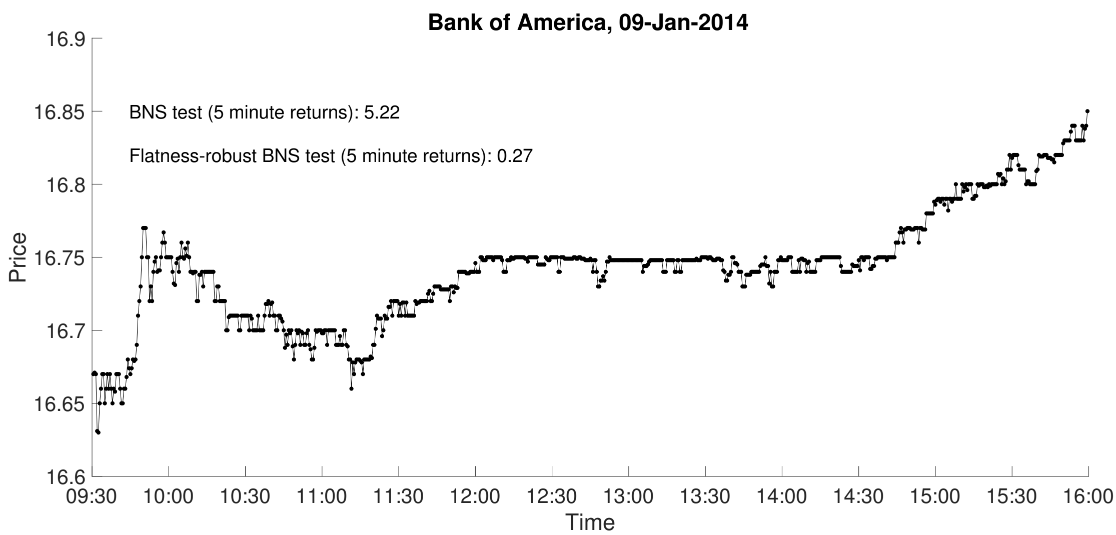

Figure 7: Time series of traded prices of the BAC stock on January 9, 2014, with reported BNS and BNS^c statistics computed on five-minutes returns. The standard BNS test would detect a jump at any reasonable confidence interval. The flatness-robust BNS test would not.

reasonable confidence interval. As argued, the BNS test is tricked by flatness in the data. When we use the flatness-robust BNS^c test, the value of the test is just 0.27, consistent with what we see in the Figure, that is no visible jumps.

When computed an all stocks for all days, the distribution of the flatness-corrected BNS test, also shown in Figure 6, is indeed more symmetric around zero and still skewed to the right, but with a less pronounced right tail. The robust test would detect jumps in just 3.40% of the cases at the 99.9% confidence interval.[9] Of course, this does not even mean that we have jumps in roughly more than 3% of the days, because of the multiple testing problem. If we follow the approach of Bajgrowicz, Scaillet, and Treccani (2016) and we use the universal threshold $\sqrt{2\log N} = 5.369$ suggested in their paper to deal with the multiple testing issue, where N is the total number of tests performed, we would have jumps only in 0.33% of the days when we use the flatness-robust test, corresponding to 0.83 jumps per stock every year. Thus, jumps are a non-negligible feature of stock data; but they happen roughly once a year, not once a month or more.

The distortion toward spurious detection of jumps is much more severe for less liquid stocks. This is to be expected since the distortion of the traditional jump tests depends on the amount of flatness, and the latter is a measure of illiquidity (Bandi, Pirino, and Renò, 2017). Figure 8 shows the percentage of detected jumps as a function of the average stock trading volume, obtained by conditional averages of detected jumps with respect to average logarithmic volume quantiles. The Figure shows clearly that without correcting for the presence of flatness, one would be induced to think that jumps are more frequent on less liquid stocks. This spurious relation would completely disappear when using the flatness-robust BNS^c statistics.

We then analyze results on the SPY, which represents the market portfolio. Despite the high liquidity of the asset, the percentage of zeros is still substantial at 5 minutes and equal to 4.42%, smaller than what observed for the average stock, but still not negligible.[10] The standard BNS test detects jumps (at 5 minutes) in 2.85% of the days at the 99.9% confidence intervals, thus less frequently than stocks, consistently with the spurious detection of more jumps on less liquid assets documented before. The corrected version of the test reveals jumps only in 1.03% of the days, roughly one third, and a rate consistent with what found for individual stocks. Thus, the impact of flatness on jump detection is substantial even for the market index (as represented by its exchange-traded fund) sampled at 5 minutes.

Also the relative contribution of jumps to total quadratic variation is overestimated by traditional bipower variation. If we define:

$$\mathsf{JV} = \mathsf{RV} - \mathsf{BPV},$$

the average relative contribution $\mathsf{JV}\,/\,\mathsf{RV}$ at the 5 minutes frequency would be estimated to be 12.18%. However, this is again an artificially inflated result due to flatness: when estimating the average relative contribution of jumps to total quadratic variation using flatness-robust multipowers $\mathsf{JV}^c = \mathsf{RV}^c - \mathsf{BPV}^c$, the outcome is just 2.04%, a figure six time smaller. On SPY, the corresponding figures are estimated to be 5.99% with $\mathsf{JV}\,/\,\mathsf{RV}$ and 2.31% with $\mathsf{JV}^c\,/\,\mathsf{RV}^c$, in line with the estimate on individual stocks,

[9] If we test using 1-minute data, the BNS test would reveal massive jumps (in 46.22% of the days), while the corrected BNS would detect jumps in 3.47% of the days, fairly consistently with the 5-minute results. However, the 1-minute data are also contaminated by market microstructure noise, which distorts the test statistics and requires a microstructure-robust test.

[10] These averages are computed only before the first and the last non-zero returns observed in the day, thus excluding periods of delayed openings and early closing of the market. At 1 minute, the percentage of observed zeros on S%P500 futures averages at 10.39%; at 10 minutes it averages at 3.11%. These figures are consistent with the level of zeros observed on individual stocks displayed in Figure 1.

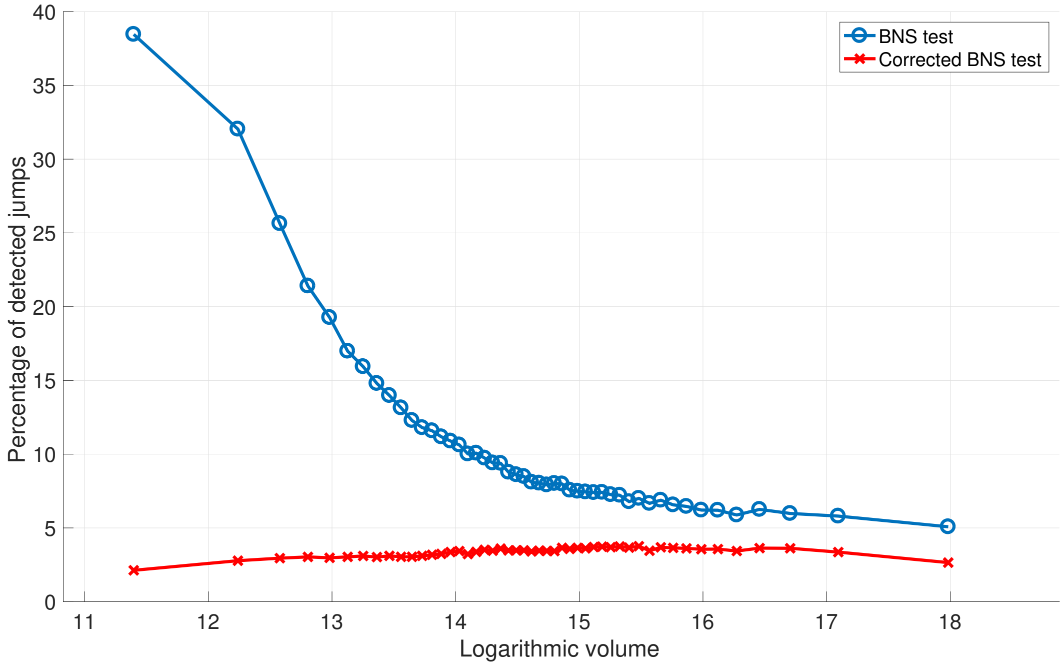

Figure 8: Percentage of detected daily jumps, using the 99.9% confidence intervals, using the BNS and the flatness-robust BNS^c statistics, computed in 40 quantiles of daily trading volume, as a function of trading volume itself.

and consistent with the relatively smaller frequency of zero returns for SPY.

The discrepancy between the amount of jump variation measured at different frequency has been highlighted by Christensen, Oomen, and Podolskij (2014) as well. Using ultra-high frequency data at the tick frequency, they report a jump variation which is, also, roughly one sixth than what found at five minutes (1.3% versus 7.3% for DJIA constituents). The discrepancy is ascribed to a finite sample effect induced by the changes of the volatility level in the sampling window. This effect is thus more pronounced when volatility changes rapidly (volatility bursts). A similar reduction is generated by flatness and these two effects are concurrent in artificially increasing jump variation at lower frequencies. When traditional estimators are robustified for the presence of flatness, the contribution of jumps to price variation is small even at five minutes. The fact that, after robustification, the jump variation obtained at five minutes and that at the tick-by-tick frequency are roughly the same suggests that flatness is the main source of the effect and volatility bursts play a minor role.

Summarizing, we showed that ignoring the presence of flatness in the data would result in a nearly three times large estimate of the number of jumps in financial markets, and nearly six times larger estimate of the contribution of jumps to price variation. Our results indicate that jumps are rare events in the data. Their role is still extremely important in finance, since these rare events can still be extremely impactful on option prices, especially with short maturity (Andersen, Fusari, and Todorov, 2017), and related to fundamental news (Bajgrowicz, Scaillet, and Treccani, 2016; Caporin, Kolokolov, and Renò, 2017). However, their size, distribution, frequency and dynamic properties can easily be artificially inflated, and by a large amount, when using estimators which are not robust to flatness.

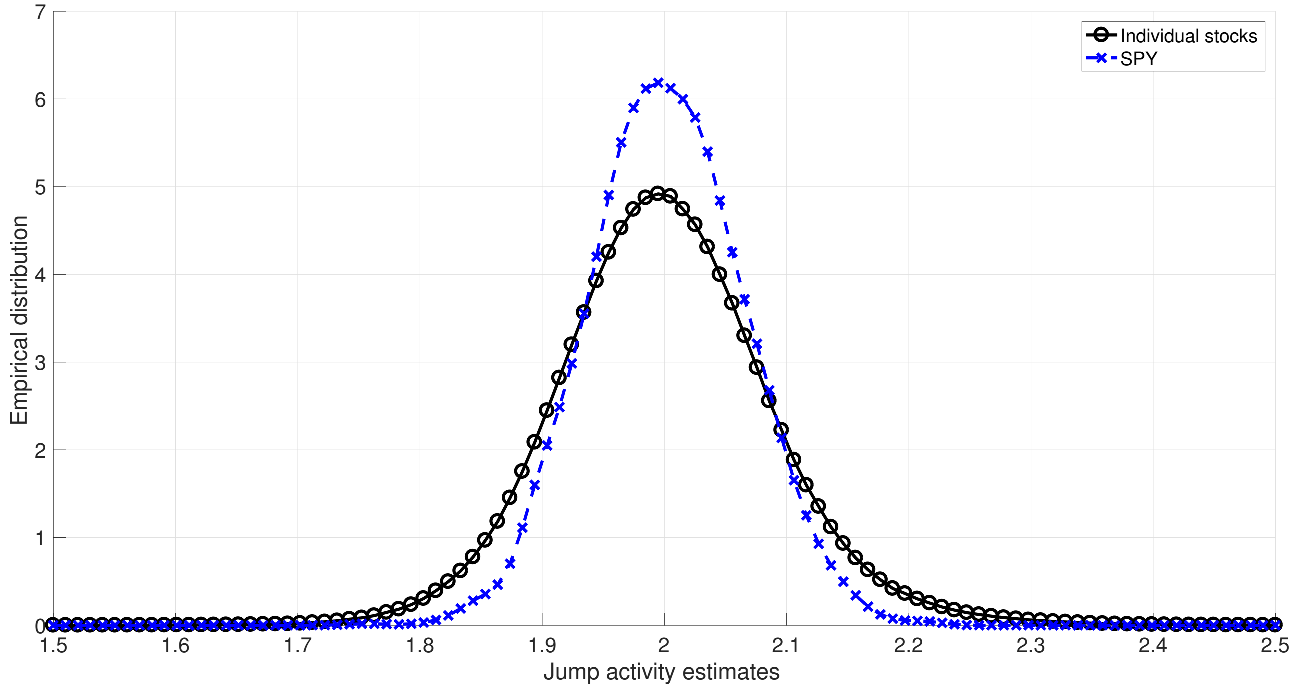

Figure 9: Pooled distribution of jump activity estimates on individual stocks and SPY, as measured every day for every considered asset. An estimated value of 2 implies presence of the Brownian motion in the considered model.

4.2 What is the activity of the volatility of the stock index?

Figure 9 shows the empirical distribution of the jump activity index estimates on individual stocks and SPY for every day in our sample (roughly 2 million estimates). We use the flatness-robust Kolokolov's estimator (3.10) with $r_+ = 1.5$. The figure shows very clearly, and pervasively, that the jump activity index of stock prices, as well as that of the index, does not depart from the value of 2, implying the omnipresence of the Brownian motion in the price dynamics. The literature is slightly ambiguous on this point. Todorov and Tauchen (2011b) support the presence of Brownian motion in the S&P500 futures, while Jing, Kong, and Liu (2012) reject the presence of Brownian motion on one year of Microsoft high frequency prices. Our results strongly support the presence of Brownian motion in stock prices.[11] On the other hand, Todorov and Tauchen (2011b) find that the jump activity index of the VIX index is significantly smaller than 2, advocating that the volatility process is a pure jump process without Brownian motion. Indeed, as they show in their paper, the activity of the VIX is the same as that of the volatility of the stock index.

Do we really need to dispense with the Brownian motion when modeling volatility? To answer this question, in this section we consider five-minute returns on the VIX index recorded from September 22, 2003 to June 2, 2015, for a total of $2,943$ trading days. We drop zeros at the beginning and at the end of the day, and after this we drop days with less than 60 observed returns. The data for VIX, as in the case of individual stocks, are heavily contaminated by the presence of flat trading: on average,

[11] The test of Jing, Kong, and Liu (2012) boils down to counting the excessive number of returns below a pre-determined threshold. Of course, the presence of zero returns in the time series of stock prices documented in Figure 1 impacts severely their test toward rejection of the Brownian motion null.

Table 1: Reports mean, median and median absolute deviation ($MAD = med|\widehat{\beta} - med(\widehat{\beta}|)$ on the monthly estimates of the jump activity index using the Todorov and Tauchen (2010) method.

	without correction			*corrected for flatness*		
	$MEAN$	$MEDIAN$	MAD	$MEAN$	$MEDIAN$	MAD
$p = 0.5$	1.57	1.58	0.19	2.13	1.92	0.29
$p = 0.7$	1.64	1.64	0.17	1.91	1.80	0.22
$p = 0.95$	1.69	1.68	0.17	1.82	1.73	0.19

the daily fraction of zero returns at the five-minute frequency is 10.17% , a figure which is consistent with the amount of flatness observed in stock prices. Thus, we reappraise the estimation of the activity index of the VIX after robustifying the estimators for the presence of flatness.

We start by estimating the jump activity of the VIX with the two-scale power estimator, following closely the approach in Todorov and Tauchen (2010). To this purpose, we use monthly non-overlapping observation windows, e.g. in order to compute each estimate we use intraday five-minutes first differences over 22 days. We truncate differences larger, in absolute value, than 1.5 to soften the impact of large outliers. The estimator, presented in Eq. (3.5), is based on the ratio of two power variations sampled at different frequencies. As in Section 3, we consider the original estimator, and the estimator flatness-corrected estimator in which power variations are replaced by the corrected estimator (2.14). Table 1 reports summary statistics on the estimates. Ignoring the presence of flat trading leads to substantial underestimation of the jump activity index. Consistently with the potential presence of a bias due to flatness, the estimates are lower when implenting the estimator with small powers: the median of the estimates ranges from 1.57 (for the power $p = 0.5$) to 1.69 (for the power $p = 0.95$). When corrected multipowers are applied, the median of the jump activity index estimates becomes close to 2 for all powers, and virtually indistinguishable from that value.

While supporting the view that the jump activity index might have been underestimated in the recent literature because of the unaccounted presence of flatness, the inherent noise of the estimator (3.5) makes the presented evidence rather inconclusive. For this reason, and to shed light on this crucial issue, we now refine the analysis using the more precise estimator suggested by Kolokolov (2017). As shown in Section 3, this estimator, based on the ratio of multipower variations with different powers and defined in Eq. (3.10), allows sufficiently precise estimation of the jump activity index for each day in the sample. For illustration we apply also the estimator defined by (3.6), non-robust to the presence of flat prices. In both cases we use $r_+ = 1.5$. Figure 10 shows the empirical densities of the daily BG estimates for the two cases. If we do not correct for flat trading, we again get an estimated jump activity index which is significantly lower than 2. When instead we correct for flat trading, the distribution of the BG estimates is symmetrically centered around the value of 2, which does not allow to reject for the presence of a Brownian component in the volatility dynamics.

Figure 11 clarifies that the bias of the non-corrected BG estimator is strongly correlated with the frequency of zero returns. The two panels show the scatter plot of daily BG estimates with the

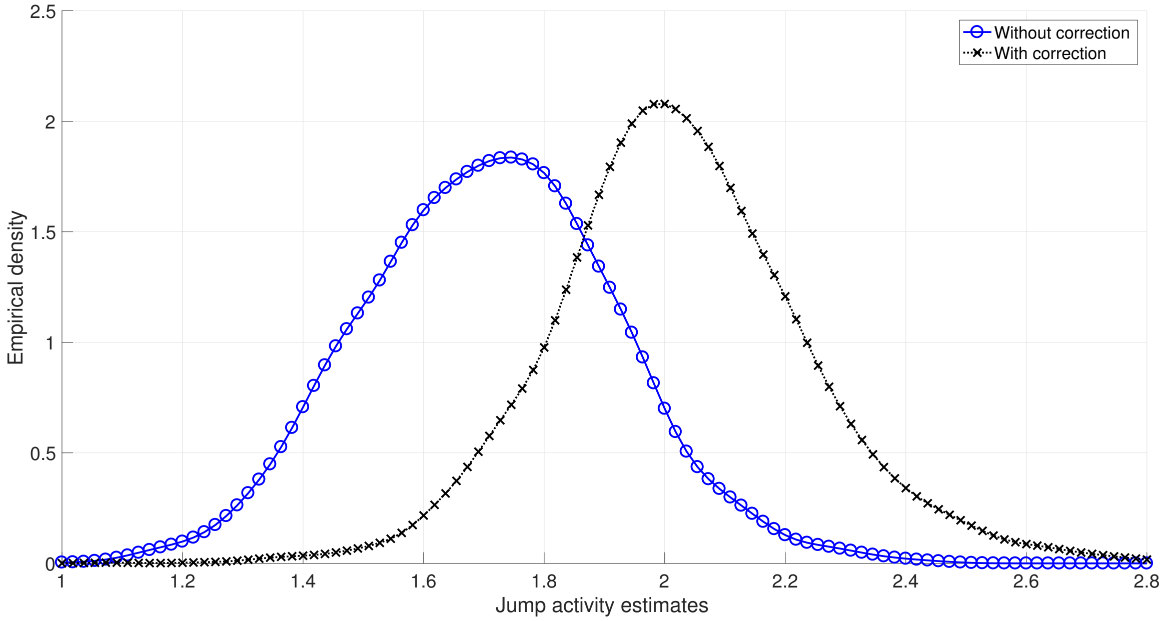

FIGURE 10: Empirical density of daily estimates of the BG index on VIX data using the Kolokolov (2017) method, original versus flatness-corrected. An estimated value of 2 implies presence of the Brownian motion in the considered model.

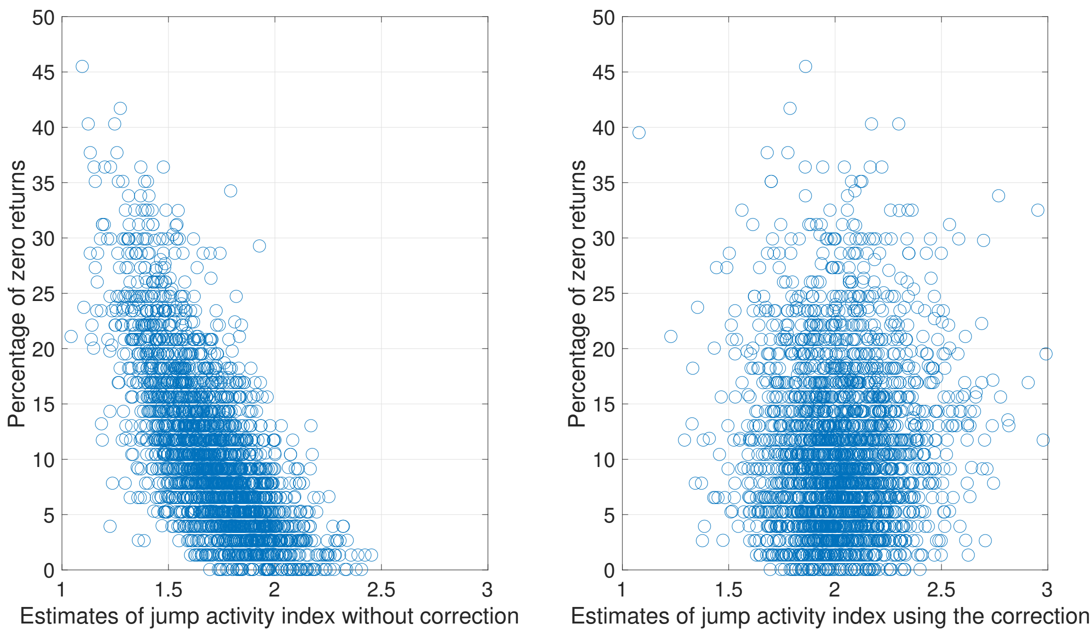

FIGURE 11: Scatter plot of daily estimates of the BG index on VIX data using the Kolokolov (2017) method versus the percentage of zero returns in each day. Left panel: original estimator. Right panel: flatness-corrected estimator.

percentage of zero returns in that day. The left panel considers the estimator based on traditional multipowers, and shows that the higher the percentage of zeros, the lower the estimate of the BG index, as predicted by the theory (while, of course, flatness should have nothing to do with jump activity). The right panel consider the estimators obtained with corrected multipowers, and shows no relation between the two quantities, with the jump activity index estimates invariably centered around 2.

Our conclusion, based on the proposed evidence, is that we cannot reject Brownian motion as the driving factor not only of stock prices (and, of course, of the market index which is a portfolio of them), but also of of the stock index volatility, in agreement with the assumptions made by traditional stochastic volatility models in continuous-time finance. We also conclude that the rejection of the presence of the Brownian motion found in the recent literature is actually an artifact due to ignoring the presence of flatness in the data.

5 Conclusions

Ignoring the presence of zero returns in financial data would result in incorrect inference about the number, frequency, contribution to price variation and actitivity of the underlying jump process in a time series. After robustification of multipower variation to the presence of flatness, desirable statistical properties of the proposed estimators are restored. The robustification we propose is straightforward to implement and does not require any additional computational cost.

When applying the robustified estimator to stock and VIX data, we find results which are very different from those obtained in the literature so far. Jumps are much less frequent, contribute much less to jump variation and are much less vibrant than what suggested by the existing empirical literature. Rough estimates obtained by us on nearly two decades for S&P 500 stocks are: a frequency of roughly one jump per stock per year, and contribution of jumps to quadratic variation of roughly 2%. For the volatility of the stock index, our findings indicate that its activity is compatible with that assumed by standard stochastic volatility models, whose shocks are driven by Brownian motion, undermining the empirical relevance of pure jump processes in finance.

More generally, our paper strongly advocates for the inclusion of flatness in the primitive assumptions for the data generating process of high-frequency data, since the distortions of not including it might not be limited to multipower estimators, but to general inference in financial econometrics. Of course, our paper also suggests that the understanding of the nature and dynamics feature of flatness in the data constitute a rich research agenda in finance which needs to be further explored.

References

Aït-Sahalia, Y., and J. Jacod, 2009a, "Estimating the Degree of Activity of Jumps in High Frequency Data," *Annals of Statistics*, 37, 2202–2244.

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

Aït-Sahalia, Y., and J. Jacod, 2012, "Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data," *Journal of Economic Literature*, 50(4), 1007–1050.

Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez, 2015, "Does realized skewness predict the cross-section of equity returns?," *Journal of Financial Economics*, 118(1), 135–167.

Andersen, T., L. Benzoni, and J. Lund, 2002, "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239–1284.

Andersen, T., T. Bollerslev, and F. X. Diebold, 2007, "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89, 701–720.

Andersen, T., D. Dobrev, and E. Schaumburg, 2012, "Jump Robust Volatility Estimation using nearest neighbor truncation," *Journal of Econometrics*, 169(1), 75–93.

Andersen, T., N. Fusari, and V. Todorov, 2017, "Short-Term Market Risks Implied by Weekly Options," *Journal of Finance*, 72(3), 1335–1386.

Bajgrowicz, P., O. Scaillet, and A. Treccani, 2016, "Jumps in high-frequency data: Spurious detections, dynamics, and news," *Management Science*, 62(8), 2198–2217.

Bandi, F., D. Pirino, and R. Renò, 2017, "EXcess Idle Time," *Econometrica*, 85(6), 1793–1846.

Barndorff-Nielsen, O. E., E. Graversen, J. Jacod, and N. Shephard, 2006, "Limit theorems for bipower variation in financial econometrics," *Econometric Theory*, 22, 677–719.

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

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

Bradley, D., J. Clarke, S. Lee, and C. Ornthanalai, 2014, "Are analysts? recommendations informative? Intraday evidence on the impact of time stamp delays," *The Journal of Finance*, 69(2), 645–673.

Busch, T., B. Christensen, and M. Nielsen, 2011, "The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets," *Journal of Econometrics*, 160, 48–57.

Caporin, M., A. Kolokolov, and R. Renò, 2017, "Systemic Co-jumps," *Journal of Financial Economics*, 126(3), 563–591.

Christensen, K., R. C. Oomen, and M. Podolskij, 2014, "Fact or friction: Jumps at ultra high frequency," *Journal of Financial Economics*, 114(3), 576–599.

Corsi, F., D. Pirino, and R. Renò, 2010, "Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting," *Journal of Econometrics*, 159, 276–288.

Delattre, S., and J. Jacod, 1997, "A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors," *Bernoulli*, pp. 1–28.

Hautsch, N., P. Malec, and M. Schienle, 2013, "Capturing the zero: a new class of zero-augmented distributions and multiplicative error processes," *Journal of Financial Econometrics*, 12(1), 89–121.

Hayashi, T., J. Jacod, and N. Yoshida, 2011, "Irregular sampling and central limit theorems for power variations: The continuous case," *Annales de l'institut Henri Poincaré (B)*, 47(4), 1197–1218.

Huang, X., and G. Tauchen, 2005, "The relative contribution of jumps to total price variance," *Journal of Financial Econometrics*, 3(4), 456–499.

Jacod, J., 2012, "Statistics and high-frequency data," in *Statistical Methods for Stochastic Differential Equations*, ed. by M. Kessler, A. Lindner, and M. Sorensen. Chapman & Hall/CRC Press/Taylor and Francis, pp. 191–309.

Jacod, J., and P. Protter, 2012, *Discretization of Processes*. Springer-Verlag.

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

Jing, B.-Y., X.-B. Kong, and Z. Liu, 2011, "Estimating the jump activity index under noisy observations using high-frequency data," *Journal of the American Statistical Association*, 106, 558–568.

——— , 2012, "Modeling high-frequency financial data by pure jump processes," *The Annals of Statistics*, 40, 759–784.

Kolokolov, A., 2017, "Estimating jump activity using multipower variation," Working Paper.

Lee, S., and P. Mykland, 2008, "Jumps in financial markets: A new nonparametric test and jump dynamics," *Review of Financial Studies*, 21(6), 2535.

Lee, S. S., and J. Hannig, 2010, "Detecting jumps from Lévy jump diffusion processes," *Journal of Financial Economics*, 96(2), 271–290.

Levine, M., X. Wang, and J. Zou, 2016, "Asymptotic properties of functionals of increments of a continuous semimartingale with stochastic sampling times," Working paper.

Li, Y., and P. A. Mykland, 2015, "Rounding errors and volatility estimation," *Journal of Financial Econometrics*, 13(2), 478–504.

Phillips, P. C. B., and J. Yu, 2009, "Information loss in volatility measurement with flat price trading," Working paper.

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

Protter, P., 1990, *Stochastic Integration and Differential Equations*. Springer-Verlag.

Robert, C. Y., and M. Rosenbaum, 2011, "A New Approach for the Dynamics of Ultra-High-Frequency Data: The Model with Uncertainty Zones," *Journal of Financial Econometrics*, 9(2), 344–366.

Theodosiou, M. G., and F. Zikes, 2011, "A comprehensive comparison of nonparametric tests for jumps in asset prices," Working Paper.

Todorov, V., 2015, "Jump activity estimation for pure-jump semimartingales via self-normalized statistics," *Annals of Statistics*, 43(4), 1831–1864.

Todorov, V., and G. Tauchen, 2010, "Activity Signature Functions for High-Frequency Data Analysis," *Journal of Econometrics*, 54, 125–138.

Todorov, V., and G. Tauchen, 2011a, "Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation," *Annals of Applied Probability*, 21, 546–588.

——— , 2011b, "Volatility jumps," *Journal of Business and Economic Statistics*, 3(29), 356–371.

Woerner, J., 2006, "Power and Multipower Variation: inference for high frequency data," in *Stochastic Finance*, ed. by A. N. Shiryaev, M. d. R. Grossinho, P. Oliveira, and M. Esquivel. Springer, pp. 343–364.

Zhang, B. Y., H. Zhou, and H. Zhu, 2009, "Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms," *Review of Financial Studies*, 22(12), 5099–5131.

A Proofs of theorems

In what follows, we assume the existence of infinitely many returns and Bernoulli variates sampled before t_0 for every fixed value of n. In this way, we can safely ignore end-effects since, as $n \longrightarrow \infty$, there is no need any more of observations before t_0. By a standard localization procedure (as in Jacod and Protter (2012, Section 4.4.1)), we can assume in the proofs that μ_t and σ_t are bounded.

Proof of Theorem 2.4. Write:

$$\mathsf{MV}(X';r) = \frac{1}{n}\sum_{i=1}^{n} n^{\frac{r_+}{2}} \left|\sigma_{t_{i^*-1}}\right|^{r_+} \xi_i + R_n, \tag{A.1}$$

where $i^* \leq i$ is the (random) largest integer not larger than i such that $B_{i^*,n} = 0$, and

$$\xi_i = \left|W'_{t_i} - W'_{t_{i-1}}\right|^{r_1} \dots \left|W'_{t_{i+m-1}} - W'_{t_{i+m-2}}\right|^{r_m} = \left|\Delta_i W'\right|^{r_1} \dots \left|\Delta_{i+m-1} W'\right|^{r_m}, \tag{A.2}$$

with

$$W'_{t_i} = W_{t_i}(1 - B_{i,n}) + B_{i,n} W'_{t_{i-1}}. \tag{A.3}$$

The random variables ξ_i have the following structure:

$$\xi_i = \begin{cases} 0, & \text{if for at least one } j \in \{0,1,...,m-1\},\ B_{i+j,n} = 1 \\ |\Delta_i W|^{r_1} |\Delta_{i+1} W|^{r_2} \dots |\Delta_{i+m-1} W'|^{r_m}, & \text{if } B_{i,n} = ... = B_{i+m-1,n} = 0 \text{ and } B_{i-1,n} = 0 \\ |\Delta_i W + \Delta_{i-1} W|^{r_1} |\Delta_{i+1} W|^{r_2} \dots |\Delta_{i+m-1} W'|^{r_m}, & \text{if } B_{i,n} = ... = B_{i+m-1,n} = 0 \text{ and } B_{i-1,n} = 1,\ B_{i-2,n} = 0 \\ ... & \\ \left|\sum_{j=0}^{k-1} \Delta_{i-j} W\right|^{r_1} |\Delta_{i+1} W|^{r_2} \dots |\Delta_{i+m-1} W'|^{r_m}, & \text{if } B_{i,n} = ... = B_{i+m-1,n} = 0 \text{ and } B_{i-1,n} = ... = B_{i-k,n} = 1,\ B_{i-k-1,n} = 0 \\ ... & \end{cases} \tag{A.4}$$

Hence, conditionally on the whole path of the Brownian motion W_t, the random variables ξ_i have a discrete distribution which can be represented as:

$$\xi_i = \begin{cases} |\Delta_i W|^{r_1} |\Delta_{i+1} W|^{r_2} \dots |\Delta_{i+m-1} W'|^{r_m}, & \text{with probability } (1-p^{\emptyset})^{m+1} \\ |\Delta_i W + \Delta_{i-1} W|^{r_1} |\Delta_{i+1} W|^{r_2} \dots |\Delta_{i+m-1} W'|^{r_m}, & \text{with probability } (1-p^{\emptyset})^{m+1} p^{\emptyset} \\ ... & \\ \left|\sum_{j=0}^{k-1} \Delta_{i-j} W\right|^{r_1} |\Delta_{i+1} W|^{r_2} \dots |\Delta_{i+m-1} W'|^{r_m}, & \text{with probability } (1-p^{\emptyset})^{m+1} (p^{\emptyset})^{k-1} \\ ... & \\ 0, & \text{with probability } 1 - \sum_{k=1}^{+\infty} (1-p^{\emptyset})^{m+1} (p^{\emptyset})^{k-1} \end{cases} \tag{A.5}$$

Consequently,

$$
\begin{aligned}
\mathbf{E}\left[n^{\frac{r_+}{2}}\left|\sigma_{t_{i^*-1}}\right|^{r_+}\xi_i \mid \mathcal{F}_{i^*-1}\right] &= \left|\sigma_{t_{i^*-1}}\right|^{r_+}\prod_{j=1}^{m}\mu_{r_j}\sum_{k=1}^{+\infty}k^{r_1/2}(1-p^{\emptyset})^{m+1}(p^{\emptyset})^{k-1} \\
&= \left|\sigma_{t_{i^*-1}}\right|^{r_+}\prod_{j=1}^{m}\mu_{r_j}\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}}\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right).
\end{aligned}
\tag{A.6}
$$

Using the Law of iterative expectations, equation (A.6) implies:

$$
\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\overline{\xi}_i\right] = 0, \tag{A.7}
$$

where

$$
\overline{\xi}_i = n^{\frac{r_+}{2}}\left|\sigma_{t_{i^*-1}}\right|^{r_+}\xi_i - \left|\sigma_{t_{i^*-1}}\right|^{r_+}\prod_{j=1}^{m}\mu_{r_j}\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}}\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right). \tag{A.8}
$$

Next we show that

$$
\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\overline{\xi}_i\right)^2\right] \longrightarrow 0. \tag{A.9}
$$

Due to the structure of ξ_i, the structure of the product of ξ_i and ξ_{i+j} is the following:

$$
\xi_i\xi_{i+j} = \begin{cases}
0, & \text{if } \min\left(\overline{\xi}_i, \overline{\xi}_{i+j}\right) = 0, \\
\left|\sum\limits_{v=0}^{k^*(i)-1}\Delta_{i-v}W\right|^{r_1}\prod\limits_{l=1}^{j-1}|\Delta_{i+l}W|^{r_{l+1}}\prod\limits_{l=1}^{m-j}|\Delta_{i+l}W|^{r_{j+l}+r_l}\prod\limits_{l=m-j+1}^{m}|\Delta_{i+l}W|^{r_l}, & \text{if } j \le m-1, \\
\left|\sum\limits_{v=0}^{k^*(i)-1}\Delta_{i-v}W\right|^{r_1}\prod\limits_{l=2}^{m-1}|\Delta_{i+l}W|^{r_l}\left|\sum\limits_{v=0}^{k^*(i+j)-1}\Delta_{i+j-v}W\right|^{r_1}\prod\limits_{l=1}^{m-1}|\Delta_{i+j+l}W|^{r_l}, & \text{if } j \ge m,
\end{cases}
\tag{A.10}
$$

where $k^*(i)$ and $k^*(i+j)$ are respectively the (random) numbers of consecutive zero increments right before the times t_i and t_{i+j}, and $k^*(i+j) \le (i+j-(i+m-1))$ almost surely. This implies that $\mathbf{E}[\xi_i\xi_j] \le K$ for all i and j and some constant K.

Consider the case $j \ge m$. If $B_{i,n} = ... = B_{i+m-1,n} = 0$ and $B_{i-1,n} = ... = B_{i-k,n} = 1$, $B_{i-k-1,n} = 0$, the product $\xi_i\xi_{i+j}$ can take $j-m+2$ values:

$$
\xi_i\xi_{i+j} = \begin{cases}
\left|\sum\limits_{j=0}^{k-1}\Delta_{i-j}W\right|^{r_1}\prod\limits_{v=1}^{m-1}|\Delta_{i+v}W|^{r_{v+1}}|\Delta_{i+j}W|^{r_1}\prod\limits_{l=1}^{m-1}|\Delta_{i+j+l}W|^{r_{l+1}} & \text{if } B_{i+j,n} = ... = B_{i+j+m-1,n} = 0 \text{ and } B_{i+j-1,n} = 0 \\
\left|\sum\limits_{j=0}^{k-1}\Delta_{i-j}W\right|^{r_1}\prod\limits_{v=1}^{m-1}|\Delta_{i+v}W|^{r_{v+1}}|\Delta_{i+j-1}W + \Delta_{i+j}W|^{r_1}\prod\limits_{l=1}^{m-1}|\Delta_{i+j+l}W|^{r_{l+1}} & \text{if } B_{i+j,n} = ... = B_{i+j+m-1,n} = 0,\ B_{i+j-1,n} = 1 \text{ and } B_{i+j-2,n} = 0 \\
... \\
\left|\sum\limits_{j=0}^{k-1}\Delta_{i-j}W\right|^{r_1}\prod\limits_{v=1}^{m-1}|\Delta_{i+v}W|^{r_{v+1}}\left|\sum\limits_{v=0}^{s-1}\Delta_{i+j-v}W\right|^{r_1}\prod\limits_{l=1}^{m-1}|\Delta_{i+j+l}W|^{r_{l+1}} & \text{if } B_{i+j,n} = ... = B_{i+j+m-1,n} = 0,\ B_{i+j-1,n} = ... = B_{i+m,n} = 1 \\
0 & \text{if for at least one } l \in \{0,1,...,m-1\},\ B_{i+j+l,n} = 1
\end{cases}
\tag{A.11}
$$

Hence, conditionally on the whole path of the Brownian motion W_t and on the event

$$\{B_{i,n} = ... = B_{i+m-1,n} = 0 \text{ and } B_{i-1,n} = ... = B_{i-k,n} = 1,\ B_{i-k-1,n} = 0\},$$

the distribution of the random variables $\xi_i\xi_{i+j}$ can be represented as:

$$\xi_i\xi_{i+j} = \begin{cases} \left|\sum_{j=0}^{k-1} \Delta_{i-j}W\right|^{r_1} \prod_{v=1}^{m-1} |\Delta_{i+v}W|^{r_{v+1}} |\Delta_{i+j}W|^{r_1} \prod_{l=1}^{m-1} |\Delta_{i+j+l}W|^{r_{l+1}} & \text{with probability } (1-p^{\emptyset})^{m+1} \\ \left|\sum_{j=0}^{k-1} \Delta_{i-j}W\right|^{r_1} \prod_{v=1}^{m-1} |\Delta_{i+v}W|^{r_{v+1}} |\Delta_{i+j-1}W + \Delta_{i+j}W|^{r_1} \prod_{l=1}^{m-1} |\Delta_{i+j+l}W|^{r_{l+1}} & \text{with probability } (1-p^{\emptyset})^{m+1}p^{\emptyset} \\ ... & \\ \left|\sum_{j=0}^{k-1} \Delta_{i-j}W\right|^{r_1} \prod_{v=1}^{m-1} |\Delta_{i+v}W|^{r_{v+1}} \left|\sum_{v=0}^{s-1} \Delta_{i+j-v}W\right|^{r_1} \prod_{l=1}^{m-1} |\Delta_{i+j+l}W|^{r_{l+1}} & \text{with probability } (1-p^{\emptyset})^{m+1}\left(p^{\emptyset}\right)^{j-m} \\ 0 & \text{with probability } 1 - \sum_{k=1}^{j-m+1} (1-p^{\emptyset})^{m+1}\left(p^{\emptyset}\right)^{k-1} \end{cases} \tag{A.12}$$

Consequently,

$$\begin{aligned} \mathbf{E}\left[\xi_i\xi_{i+j}\right] &= n^{-r_+} \prod_{l=1}^{m} \mu_{r_l}^2 \sum_{k=1}^{+\infty} \left(\sum_{l=1}^{j-m+1} l^{r_1/2}(1-p^{\emptyset})^{m+1}(p^{\emptyset})^{l-1} \right) k^{r_1/2}(1-p^{\emptyset})^{m+1}(p^{\emptyset})^{k-1} \\ &= n^{-r_+} \prod_{l=1}^{m} \mu_{r_l}^2 \frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}} \mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) \left(\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}} \mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) - \sum_{l=j-m+2}^{\infty} l^{r_1/2}(1-p^{\emptyset})^{m+1}(p^{\emptyset})^{l-1} \right) \\ &= n^{-r_+} \prod_{l=1}^{m} \mu_{r_l}^2 \left(\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}} \right)^2 \mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) \left(\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) - \sum_{l=j-m+2}^{\infty} l^{r_1/2}(p^{\emptyset})^{l} \right). \end{aligned} \tag{A.13}$$

Hence,

$$\begin{aligned} \left|\mathbf{E}\left[\xi_i\xi_{i+j}\right] - \mathbf{E}\left[\xi_i\right]\mathbf{E}\left[\xi_{i+j}\right]\right| &= \left| n^{-r_+} \left(\prod_{l=1}^{m} \mu_{r_l}^2 \right) \left(\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}} \right)^2 \mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) \sum_{l=j-m+2}^{\infty} l^{r_1/2}(p^{\emptyset})^{l} \right| \\ &\leq n^{-r_+} K \times \sum_{l=j-m+2}^{\infty} l^{r_1/2}(p^{\emptyset})^{l}, \end{aligned} \tag{A.14}$$

for some constant K. Since σ_t is bounded,

$$\begin{aligned} \mathbf{E}\left[\xi_i\xi_{i+j}\right] = \mathbf{E}\left[\mathbf{E}\left[\xi_i\xi_{i+j} \middle| (\sigma_t)\right]\right] &\leq \mathbf{E}\left[\left|\sigma_{t_{i^*-1}}\right|^{r_+} \left|\sigma_{t_{(i+j)^*-1}}\right|^{r_+} K \sum_{l=j-m+2}^{\infty} l^{r_1/2}(p^{\emptyset})^{l} \right] \\ &\leq K_2 \sum_{l=j-m+2}^{\infty} l^{r_1/2}(p^{\emptyset})^{l}, \end{aligned} \tag{A.15}$$

for a constant K_2. Than,

$$\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\bar{\xi}_i\right)^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{E}\left[\bar{\xi}_i^2\right] + \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{m-1}\mathbf{E}\left[\bar{\xi}_i\bar{\xi}_{i+j}\right] + \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=m}^{n-i}\mathbf{E}\left[\bar{\xi}_i\bar{\xi}_{i+j}\right] \tag{A.16}$$

Since $\mathbf{E}\left[\bar{\xi}_i\bar{\xi}_j\right] \leq K$ for all i and j and some constant K, the first two terms converge to zero:

$$\frac{1}{n^2}\sum_{i=1}^{n}\mathbf{E}\left[\bar{\xi}_i^2\right] + \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{m-1}\mathbf{E}\left[\bar{\xi}_i\bar{\xi}_{i+j}\right] \leq \frac{1}{n^2}\sum_{i=1}^{n}K \longrightarrow 0. \tag{A.17}$$

For the third term we have:

$$\frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=m}^{n-i}\mathbf{E}\left[\bar{\xi}_i\bar{\xi}_{i+j}\right] \leq K\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=m}^{n-i}\sum_{l=j-m+2}^{\infty} l^{r_1/2}(p^{\emptyset})^l \longrightarrow 0. \tag{A.18}$$

Consequently, $\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\bar{\xi}_i\right)^2\right] \longrightarrow 0$. Using Chebyshev's inequality and the definition of the convergence in probability,

$$\frac{1}{n}\sum_{i=1}^{n}\bar{\xi}_i \xrightarrow{p} 0. \tag{A.19}$$

By Riemann integrability of σ_t, Assumption 2.1 and the result (A.31),

$$\sum_{i=1}^{n}\frac{1}{n}\left|\sigma_{t_{i^*-1}}\right|^{r_+}\prod_{j=1}^{m}\mu_{r_j}\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}}\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) \xrightarrow{p} \prod_{j=1}^{m}\mu_{r_j}\frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}}\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right)\int_0^t|\sigma_s|^{r_+}\,ds. \tag{A.20}$$

Consequently,

$$\frac{1}{n}\sum_{i=1}^{n}n^{\frac{r_+}{2}}\left|\sigma_{t_{i^*-1}}\right|^{r_+}\xi_i \xrightarrow{p} \frac{(1-p^{\emptyset})^{m+1}}{p^{\emptyset}}\mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right)\prod_{j=1}^{m}\mu_{r_j}\int_0^t|\sigma_s|^{r_+}\,ds. \tag{A.21}$$

We now prove that the reminder, R_n, is asymptotically negligible. This part of the proof is similar to the proof of Lemma 6.9 in Jacod (2012). Consider the decomposition:

$$\Delta_i X' = \frac{\beta_i'}{\sqrt{n}} + \frac{\chi_i'}{\sqrt{n}} \tag{A.22}$$

with

$$\beta_{i,l}' = \frac{\sigma_{t_{i^*-1}}\Delta_{i+l}W'}{\sqrt{n}}, \quad \chi_{i,l}' = \frac{\sqrt{(K_{i+l,n}+1)/n}\displaystyle\int_{(i+l-K_{i+l,n}-1)/n)}^{(i+l)/n}\left(\mu_s ds + (\sigma_s - \sigma_{t_{i^*-1}})dW_s\right)}{\sqrt{n}}. \tag{A.23}$$

Than, the reminder can be expressed as:

$$R_n = \frac{1}{n}\sum_{i=1}^{n-m+1}\eta_i - \frac{1}{n}\sum_{i=n-m+2}^{n}\left|\beta'_{i,0}\right|^{r_1}\dots\left|\beta'_{i,m-1}\right|^{r_m}, \tag{A.24}$$

with

$$\eta_i = \left|\sqrt{n}\cdot\Delta_i X'\right|^{r_1}\dots\left|\sqrt{n}\cdot\Delta_{i+m-1} X'\right|^{r_m} - \left|\beta'_{i,0}\right|^{r_1}\dots\left|\beta'_{i,m-1}\right|^{r_m}. \tag{A.25}$$

Denoting by C a constant, which may change from line to line, and denoting by $K_{i,n} = i - i^*$, we have the following estimate:

$$\begin{aligned}\mathbf{E}_{i+l-K_{i+l,n}-1}\left[\left|\beta'_{i,l}\right|^q\right] &= \mathbf{E}_{i+l-K_{i+l,n}-1}\left[\mathbf{E}_{i+l-K_{i+l,n}-1}\left[\left|\beta'_{i,l}\right|^q|K_{i+l,n}\right]\right]\\ &\le C\mathbf{E}\left[K_{i+l,n}^{q/2}\right]n^{-q/2}\\ &\le C\mathbf{E}\left[K_n^{q/2}\right]n^{-q/2},\end{aligned} \tag{A.26}$$

where K_n denotes the supremum of $K_{i,n}$ and we use the independence of $K_{i,n}$ from $\mathcal{F}_{i-K_{i,n}-1}$. Similarly, we obtain the following estimate:

$$\mathbf{E}_{i+l-K_{i+l,n}-1}\left[\left|\chi'_{i,l}\right|^q\right] \le C\mathbf{E}\left[K_n^{1\vee q/2}\right]n^{-1\vee -q/2}. \tag{A.27}$$

Next, denote by $G_A(\epsilon)$ the supremum of $(\prod_{k=1}^m |x_k + y_k|^{r_k} - \prod_{k=1}^m |x_k|^{r_k})$, over all $|x_k| \le A$ and $|y_k| \le \epsilon$, which converges to 0 as $\epsilon \to 0$. As in Lemma 6.9 of Jacod (2012), for all $A > 1$ and $\epsilon > 0$, we obtain:

$$|\eta_i| \le G_A(\epsilon) + C\sum_{k=1}^m\left(h_{\epsilon,A,n}\left(\beta_{i,k-1},\chi_{i,k-1}\right)\prod_{j=1,\dots,m,\ m\neq k} g(\beta_{i,j-1},\chi_{i,j-1})\right), \tag{A.28}$$

where for a positive number p which depends on powers,

$$h_{\epsilon,A,n}\left(x,y\right) = \frac{|x|^{p+1}}{A} + |x|^p\left(|y|\wedge 1\right) + A^p\frac{|y|\wedge 1}{\epsilon^2} + \frac{|y|^{p+1}}{A}, \tag{A.29}$$

$$g(x,y) = 1 + |x|^p + |y|^p. \tag{A.30}$$

Now notice that, under the iid case,

$$K_n = O_p\left(\log n\right). \tag{A.31}$$

Using the estimates (A.26) and (A.27) and Cauchy-Schwarz inequality we thus obtain:

$$\mathbf{E}_{i+k-K_{i+l,n}-2}\left[\left|\beta'_{i,k-1}\right|^{p+1}\right] \le C\mathbf{E}\left[K_n^{\frac{p+1}{2}}\right]n^{-\frac{p+1}{2}} \le C\left(\frac{\log n}{n}\right)^{\frac{p+1}{2}}, \tag{A.32}$$

$$\mathbf{E}_{i+k-K_{i+l,n}-2}\left[\left|\chi'_{i,k-1}\right|^{p+1}\right] \le C\mathbf{E}\left[K_n^{\frac{p+1}{2}}\right]n^{-\frac{p+1}{2}} \le C\left(\frac{\log n}{n}\right)^{\frac{p+1}{2}}, \tag{A.33}$$

$$\mathbf{E}_{i+k-K_{i+l,n}-2}\left[\left|\beta'_{i,k-1}\right|^p\left|\chi'_{i,k-1}\right|\right] \le C\sqrt{\mathbf{E}\left[K_n^p\right]\mathbf{E}\left[K_n\right]}n^{-\frac{p+1}{2}} \le C\left(\frac{\log n}{n}\right)^{\frac{p+1}{2}}, \tag{A.34}$$

$$\mathbf{E}_{i+k-K_{i+l,n}-2}\left[\left|\chi'_{i,k-1}\right|\right] \leq C\mathbf{E}\left[K_n\right] n^{-1} \leq C\left(\frac{\log n}{n}\right). \tag{A.35}$$

Consequently,

$$\sup_{i,k}\left(\mathbf{E}_{i+k-K_{i+l,n}-2}\left[h_{\epsilon,A,n}\left(\beta_{i,k-1},\chi_{i,k-1}\right)\prod_{j=1,\ldots,m,\ m\neq k} g(\beta_{i,j-1}\chi_{i,j-1})\right]\right) \leq \psi_n(A,\epsilon), \tag{A.36}$$

where

$$\lim_{A\to\infty}\limsup_{n\to\infty}\psi_n(A,\epsilon) = 0. \tag{A.37}$$

By successive conditioning we obtain:

$$\sup_i\left(\mathbf{E}_{i-K_n-1}\left[|\eta_i|\right]\right) \leq G_A(\epsilon) + C\psi_n(A,\epsilon), \tag{A.38}$$

where the right hand side converges to zero as $\epsilon \to 0$ and $n \to \infty$ (using $G_A(\epsilon) \longrightarrow 0$ as $\epsilon \to 0$). Finally, using successive conditioning we obtain:

$$\mathbf{E}\left[R_n\right] = \frac{1}{n}\sum_{i=1}^{n-m+1}\mathbf{E}\left[\eta_i\right] \leq \frac{n-m+1}{n}\left(G_A(\epsilon) + C\psi_n(A,\epsilon)\right). \tag{A.39}$$

Since $\frac{n-m+1}{n} \to 1$ as $n \to \infty$, the statement of the theorem follows by sending $\epsilon \to \infty$.

$\square$

Proof of Theorem 2.7.

The proof is almost identical to the proofs in the case of the stochastic sampling schemes established by Levine, Wang, and Zou (2016). Set

$$\zeta_i = \left|\Delta(n,i)^{-1/2}\Delta_{\tau_i}X\right|^{r_1}\ldots\left|\Delta(n,i+m-1)^{-1/2}\Delta_{\tau_{i+m-1}}X\right|^{r_m}, \tag{A.40}$$

$$\zeta'_i = \left|\sigma_{\tau_{i-1}}\right|^{r_+}\left|\Delta(n,i)^{-1/2}\Delta_{\tau_i}W\right|^{r_1}\ldots\left|\Delta(n,i+m-1)^{-1/2}\Delta_{\tau_{i+m-1}}W\right|^{r_m}, \tag{A.41}$$

$$\zeta''_i = \zeta_i - \zeta'_i. \tag{A.42}$$

Than, the Central Limit Theorem follows from the following four lemmas:

Lemma A.1.

$$\sqrt{\varkappa_n}\sum_{i=1}^{N_n}\left(\zeta''_i - \mathbf{E}_{\tau_{i-1}}\left[\zeta''_i\right]\right) \xrightarrow{u.c.p.} 0. \tag{A.43}$$

Lemma A.2.

$$\frac{1}{\sqrt{\varkappa_n}}\prod_{j=1}^{m}\mu_{r_j}\left(\sum_{i=1}^{N_n}\left|\sigma_{\tau_{i-1}}\right|^{r_+}\varkappa_n - \int_0^t\left|\sigma_s\right|^{r_+}ds\right) \xrightarrow{L-st} Z_t. \tag{A.44}$$

Lemma A.3.

$$\sqrt{\varkappa_n}\sum_{i=1}^{N_n}\left(\zeta_i' - \prod_{j=1}^{m}\mu_{r_j}\left|\sigma_{\tau_{i-1}}\right|^{r+}\right) \xrightarrow{L-st} U_t. \tag{A.45}$$

Lemma A.4.

$$\sqrt{\varkappa_n}\sum_{i=1}^{N_n}\mathbf{E}_{\tau_{i-1}}\left[\zeta_i''\right] \xrightarrow{u.c.p.} 0. \tag{A.46}$$

□

We provide only the proof of Lemma A.2, which delivers asymptotic term Z_t different from the case considered by Levine, Wang, and Zou (2016).

Proof of Lemma A.2.

Define the function $g(\sigma) = \prod_{j=1}^{m}\mu_{r_j}|\sigma|^{r+}$. By simple algebra,

$$\frac{1}{\sqrt{\varkappa_n}}\prod_{j=1}^{m}\mu_{r_j}\left(\sum_{i=1}^{N_n}\left|\sigma_{\tau_{i-1}}\right|^{r+}\varkappa_n - \int_0^t|\sigma_s|^{r+}\,ds\right) = -\sum_{i=1}^{N_n}\eta_i - \sum_{i=1}^{N_n}\epsilon_i, \tag{A.47}$$

where

$$\eta_i = \frac{1}{\sqrt{\varkappa_n}}\int_{\tau_{i-1}}^{\tau_i}\big(g(\sigma_u) - g(\sigma_{\tau_{i-1}})\big)\,du, \tag{A.48}$$

and

$$\epsilon_i = \frac{1}{\sqrt{\varkappa_n}}\int_{\tau_{i-1}}^{\tau_i}g(\sigma_{\tau_{i-1}})\left(1 - \frac{\varkappa_n}{\Delta(n,i)}\right)du. \tag{A.49}$$

Hence, it is enough to show that

$$\sum_{i=1}^{N_n}\eta_i \xrightarrow{u.c.p.} 0, \tag{A.50}$$

and

$$\sum_{i=1}^{N_n}\epsilon_i \xrightarrow{L-st} Z_t. \tag{A.51}$$

Expand η_i as $\eta_i = \eta_i' + \eta_i''$, where

$$\eta_i' = \frac{1}{\sqrt{\varkappa_n}}g'(\sigma_{\tau_{i-1}})\int_{\tau_{i-1}}^{\tau_i}\big(g(\sigma_u) - g(\sigma_{\tau_{i-1}})\big)\,du, \tag{A.52}$$

$$\eta_i'' = \frac{1}{\sqrt{\varkappa_n}}\int_{\tau_{i-1}}^{\tau_i}\big(g(\sigma_u) - g(\sigma_{\tau_{i-1}}) - g'(\sigma_{\tau_{i-1}})(\sigma_u - \sigma_{\tau_{i-1}})\big)\,du. \tag{A.53}$$

Moreover, expand η_i' as $\eta_i' = \eta_i'(1) + \eta_i'(2)$, where

$$\eta_i'(1) = \frac{1}{\sqrt{\varkappa_n}}g'(\sigma_{i-1})\int_{t_{i-1}}^{t_i}du\int_{t_{i-1}}^{u}\tilde{b}_s ds, \tag{A.54}$$

and

$$\eta_i'(2) = \frac{1}{\sqrt{\varkappa_n}} g'(\sigma_{\tau_{i-1}}) \int_{\tau_{i-1}}^{\tau_i} du \left(\int_{\tau_{i-1}}^{u} \tilde{\sigma}_s dW_s + \int_{\tau_{i-1}}^{u} \int \tilde{\delta}(s,x)(\underline{\mu} - \underline{\nu})(ds,dx)) \right), \tag{A.55}$$

By boundedness of $\sigma_{\tau_{i-1}}$ and $\tilde{b}$, we have $|\eta_i'(1)| \le \Lambda \frac{(\Delta(n,i))^2}{\sqrt{\varkappa_n}}$. Hence,

$$\sum_{i=1}^{N_n} \left|\eta_i'(1)\right| \le \Lambda \sum_{i=1}^{N_n} \frac{(\Delta(n,i))^2}{\sqrt{\varkappa_n}} \longrightarrow 0. \tag{A.56}$$

By similar arguments $\mathbf{E}_{\tau_{i-1}}\left[\eta_i'(2)\right] = 0$. Moreover, using Doob's and Cauchy-Schwarz inequalities $\mathbf{E}_{\tau_{i-1}}\left[(\eta_i'(2))^2\right] \le \mathbf{E}_{\tau_{i-1}}\left[\Delta(n,i)^2\right]$. Hence,

$$\sum_{i=1}^{N_n} \left|(\eta_i'(2))^2\right| \le \Lambda \sum_{i=1}^{N_n} \mathbf{E}_{\tau_{i-1}}\left[\Delta(n,i)^2\right] \longrightarrow 0. \tag{A.57}$$

For η_i'' we have: $\eta_i'' \le \frac{K}{\sqrt{\varkappa_n}} \int_{\tau_{i-1}}^{\tau_i} \left|\sigma_u - \sigma_{\tau_{i-1}}\right|^2 du$. Hence, using the inequality

$$\mathbf{E}_t\left[\|\sigma_{t+s} - \sigma_t\|^q\right] \le K_q s^{1\wedge(q/2)}, \tag{A.58}$$

for a constant K_q which may depend on q, we obtain:

$$\sum_{i=1}^{N^n} \mathbf{E}_{\tau_{i-1}}\left[\left|\eta_i''\right|\right] \le \Lambda \sum_{i=1}^{N^n} \frac{\mathbf{E}_{\tau_{i-1}}\left[\Delta(n,i)^2\right]}{\sqrt{\varkappa_n}} \longrightarrow 0. \tag{A.59}$$

Inequalities (A.56), (A.57) and (A.59) together imply the condition (A.50).

Now, consider $\epsilon_i = \frac{1}{\sqrt{\varkappa_n}} \int_{\tau_{i-1}}^{\tau_i} g(\sigma_{\tau_{i-1}}) \left(1 - \frac{\varkappa_n}{\Delta(n,i)}\right) du = \frac{1}{\sqrt{\varkappa_n}} g(\sigma_{\tau_{i-1}}) \left(\Delta(n,i) - \varkappa_n\right)$. By assumptions we have $\mathbf{E}_{\tau_{i-1}}\left[\epsilon_i\right] \longrightarrow 0$. The conditional variance takes the following form:

$$\mathbf{Var}_{\tau_{i-1}}\left[\epsilon_i\right] = g(\sigma_{i-1})^2 \frac{\mathbf{Var}_{\tau_{i-1}}\left[\Delta(n,i)\right]}{\varkappa_n^2} \varkappa_n. \tag{A.60}$$

Thus,

$$\sum_{i=1}^{N_n} \mathbf{Var}_{\tau_{i-1}}\left[\epsilon_i\right] = s_n^2 + (s_n')^2, \tag{A.61}$$

where

$$s_n^2 = \sum_{i=1}^{N_n} g(\sigma_{\tau_{i-1}})^2 \frac{\mathbf{Var}_{\tau_{i-1}}\left[\Delta(n,i)\right]}{\varkappa_n^2} \Delta(n,i) \xrightarrow{p} \Upsilon \int_0^t g(\sigma_u)^2 du. \tag{A.62}$$

$$(s_n')^2 = \sum_{i=1}^{N_n} g(\sigma_{\tau_{i-1}})^2 \frac{\mathbf{Var}_{\tau_{i-1}}\left[\Delta(n,i)\right]}{\varkappa_n^2} \left(\Delta(n,i) - \varkappa_n\right) \xrightarrow{p} 0. \tag{A.63}$$

Finally, $\sum_{i=1}^{N_n} \frac{\mathbf{E}_{\tau_{i-1}}\left[\Delta(n,i)^4\right]}{\varkappa_n} \xrightarrow{p} 0$ implies that $\sum_{i=1}^{N_n} |\epsilon_i|^4 \xrightarrow{p} 0$. Consequently,

$$\sum_{i=1}^{N_n} \epsilon_i \xrightarrow{L-st} Z_t, \tag{A.64}$$

which completes the proof. □