Realized moments: identification and pricing*

Federico M. Bandi[†] Aleksey Kolokolov[‡] Davide Pirino[§] Roberto Renò[¶]

Preliminary draft

Abstract

We study the properties of realized high-order moments under a data generating process accounting for key stylized features: infrequent discontinuities in unobserved equilibrium prices and staleness in observed prices, a phenomenon linked to volume dynamics. Our focus is on identification and pricing. In terms of identification, we show how the interplay between price discontinuities and prices staleness will, in general, lead to biased and/or noisy high-order moment estimates. We also show how a combination of thresholding and corrections for staleness-induced biases can be deployed to extract reliable information about high-order continuous and discontinuous variation. Regarding pricing, the use of thresholding and de-biasing leads to ample evidence about the negative cross-sectional pricing of idiosyncratic price discontinuities at high frequency. We show that accounting for staleness is (1) important for the correct identification of high-order moments, (2) revealing about these moments' cross-sectional pricing and (3) informative about the pricing of illiquidity, for which staleness is a rich proxy.

Keywords: Staleness, jumps, high-order moments.

*We thank participants at the conference "Econometrics in the Arena", Verona, Italy, September 12-13 2019, and the "Fourth International Workshop in Financial Econometrics", Jatiúca Resort, Maceió, Brazil, October 6-8 2019. We are particularly grateful to the discussant Alejandro Bernales.

[†]Johns Hopkins University, Carey Business School, 100 International Drive, Baltimore MD 20202, USA and Edhec-Risk Institute, France. E-mail: fbandi1@jhu.edu.

[‡]Alliance Manchester Business School, Booth St W, Manchester M15 6PB, UK. E-mail: alexeiuo@gmail.com.

[§]Università degli Studi di Roma "Tor Vergata," Dipartimento di Economia e Finanza, Via Columbia 2, 00173, Roma, Italy, and Scuola Normale Superiore, Piazza dei Cavalieri 7, 56123, Pisa, Italy. E-mail: davide.pirino@gmail.com.

[¶]Università di Verona, Dipartimento di Scienze Economiche, Via Cantarane 24, 37129, Verona, Italy. E-mail: roberto.reno@univr.it.

1 Introduction

The identification of volatility has been central to the econometric literature and has provided support to empirical work in an array of fields, from finance to macroeconomics. Less emphasis has been placed on higher-order moments. Yet, economic modeling is now beginning to be more attentive to the full distributional properties of economic time series - to tail properties, in particular.

In light of this premise, we study the large sample features of empirical higher-order moments computed from high-frequency (price) data, i.e., *realized high-order moments*. As is standard in the high-frequency literature, we assume prices evolve continuously in time and work with equally-spaced data on a sampling grid spanning the trading day. We obtain daily measures by suitably aggregating intradaily price changes on the assumed grid. Asymptotics are derived by letting the number of intradaily observations grow without bound and, therefore, by making the grid less and less coarse as the sample size increases, i.e., *infill asymptotics*. In addition to other more standard assumptions, the proposed data generating process incorporates two first-order features of the price data. First, the observed process (on the sampling grid) can be stale, and therefore *not* move, due to lack of (or limited) volume (price staleness), an empirical phenomenon illustrated by Bandi et al. (2019). Second, the underlying (unobservable, in the presence of price staleness) price process may display infrequent discontinuities associated with unexpected news arrivals (price jumps). The impact of jumps on realized moment measures is studied in detail in the econometrics literature (see Aït-Sahalia and Jacod, 2015, and the references therein). However, much less is known about the effect of staleness and, in particular, about the theoretical interplay between jumps and staleness. We show that this interplay - a central theme of the present paper - is key to the analysis of realized moments, asymptotically and in any finite sample.

In large samples, it is important to distinguish between trajectories with and without jumps. In the absence of jumps, the properties of realized moments depend on whether they are constructed using *even* functions (the fourth moment case, for instance) or *odd* functions (as in the third moment case). In general, even functions lead to staleness-induced asymptotic biases and limiting zero-mean mixed normal distributions (the realized variance case being a notable exception for which the staleness-induced bias is zero). Odd functions, instead, do not generate a bias but the limiting distributions are uncentered mixed normal. We propose a simple methodology to bias-correct realized moments, when needed, and show consistency and mixed normality of the resulting bias-corrected estimates. On discontinuous trajectories, importantly, the realized moments have to be re-scaled. When suitably re-scaled, they are shown to be consistent for sums of (odd or even) functions of the jumps with limiting mixed normal variates whose variances are augmented (relative to the case without staleness) by the presence of staleness.

In a finite sample, the interaction between staleness and jumps can be understood as follows. Extended periods of price staleness ought to be associated with an ever-changing underlying (unobservable) efficient price process. When staleness comes to an end, observed prices revert to prevailing efficient prices but the latter may be far from observed prices. This effect may generate *spurious* jumps even on a purely continuous trajectory. These spurious jumps have the potential to influence the finite sample properties of the realized moment estimates, particularly in the continuous case and with even functions. The proposed asymptotic bias-correction for the case with discontinuous trajectories is shown to alleviate this issue drastically.

The presence of staleness leads to one key modification in the evaluation of realized moments. Because it generates random periods of time over which the price changes are zero, it modifies the sampling grid (by making it random and coarser than the original grid) and, as a consequence, the sample size (the *effective* sample size being smaller than the original one). Given that the limiting order of the new sample size is the same as that of the old sample size, staleness does not yield modifications of the convergence rates. It only leads to upward biases (with even functions, in the continuous case) and increased estimation uncertainty. We provide methods to characterize both and account for them.

Realized variance is a well-posed estimator in the presence of both staleness and jumps. It is consistent for the overall quadratic variation and mixed normally distributed with an asymptotic distribution whose variance is enhanced by staleness. Higher-order moments are more delicate, the interaction between staleness and jumps playing a more critical role in this more general case. First, differently from realized variance, for the limit of higher-order moments to be meaningful about the tail properties of the price series, higher-order moments have to be re-scaled differently on trajectories with and without jumps. Second, should we establish that a trajectory is continuous, depending on the nature of the moment as defined by the corresponding function (odd or even), the required (to alleviate the impact of staleness) bias-correction should be different. In all cases, staleness affects limiting precision and should, therefore, be accounted for when evaluating limiting variances and performing asymptotic inference.

Realized high-order moments are often used in the literature as inputs in traditional definitions of skewness and kurtosis. We find that, on continuous trajectories, realized kurtosis may considerably overstate true kurtosis due to staleness. Similarly, while unbiased, continuous skewness is noisier in the presence of staleness. We also show that realized skewness and kurtosis diverge with the sample size in the presence of discontinuous trajectories. While realized skewness is a signed measure of jumps, more negative (res. positive) values being associated with more negative (res. positive) jumps, realized kurtosis increases with the number of intra-daily observations (irrespective of the magnitude and number of jumps on the continuous trajectory of the process). Both properties, the latter in particular, are undesirable.

Thresholding, i.e., the use of truncation to identify small and large variation in asset prices, provides a natural way to address these issues. We define (truncated) continuous notions of skewness and kurtosis and show how bias-correcting kurtosis leads to accurate assessments of variation due to uncertainty in price volatility. Similarly, we may define (truncated) discontinuous notions of skewness and kurtosis, as well as various moments of the positive and negative jumps, not affected by the presence of staleness.

We re-evaluate the existing empirical evidence on the cross-sectional pricing of idiosyncratic skewness and kurtosis (c.f., Amaya et al., 2015 and Bollerslev et al., 2019). In order to do so, we use a large cross-section of US stocks and a large sample of high-frequency price data spanning 20 years. Consistent with the economic logic laid out in Amaya et al., 2015 and Bollerslev et al., 2019, we expect measures that are positively correlated with positive price discontinuities and negatively correlated with negative price discontinuities (like the skewness measure of Amaya et al., 2015 or the relative signed jump variation of Bollerslev et al., 2019) to be priced negatively. Negative pricing would, in fact, be symptomatic of aversion to negative skewness (or negative jumps) and attraction to positive skewness (or positive jumps). We provide evidence in support of this logic but emphasize - both in theory and with data - that simple measures based on sums of extracted positive/negative jumps using truncation contain cleaner pricing signal than measures like skewness or relative signed jump variation. The latter are, in fact, contaminated by continuous variation and, therefore, by staleness. Regarding kurtosis, we show that its documented positive pricing (c.f. Amaya et al., 2015) is due to the above-mentioned staleness-induced upward bias. When controlling for staleness, either by bias-correcting directly or by inserting it as a control, the pricing of idiosyncratic kurtosis becomes *negative* in much the same way in which the price of idiosyncratic volatility is negative. This result may be suggestive of a volatility *uncertainty* pricing puzzle analogous to the more classical volatility *level* puzzle of Ang et al. (2006). Finally, we document that a direct measure of staleness, like the percentage of zero returns, commands a positive, and highly statistically-significant, illiquidity premium in expected returns. Because staleness also affects the centering of the distribution of realized kurtosis and the spread of the distribution of realized skewness, we conclude by stressing that accounting for staleness is key to understanding the pricing of higher-order moments in addition to the pricing of illiquidity. In sum, the interaction between illiquidity, as represented by staleness, and *estimated* efficient price dynamics, as represented by the time series of the high-order price moments, has both econometric and pricing implications. These implications are the focus of our analysis.

The paper proceeds as follows. Section 2 presents the model, the general family of estimators and limiting results for realized high-order moments in the continuous case (Subsection 2.2) and in the discontinuous case (Subsection 2.3). Section 3 proposes a simple bias-correction for realized moment estimates in the presence of staleness and studies its limiting properties. In Section 4 we study traditional realized skewness and kurtosis estimates and use their limiting features to further characterize the interplay between staleness and jumps. Section 5 separates continuous from discontinuous variation by thresholding. The section offers a simple strategy to evaluate continuous and discontinuous higher-order moments and, in the former case, perform effective bias-correction to alleviate the impact of staleness. The pricing relevance of discontinuous and continuous variation, as well as that of illiquidity (as proxied by staleness), is re-evaluated in Section 6 using a large cross-section of US stocks and a long span of high-frequency price observations. Section 7 concludes. Proofs are in the Appendix.

2 Limit theory for realized moments

2.1 Setting and assumptions

We assume the underlying (sometimes unobserved) efficient price process follows a continuous-time semimartingale with stochastic volatility and jumps while the observed price process (recorded on a specific sampling grid over the finite interval $[0,t]$) is contaminated by the presence of staleness.

Formally, let X_t be the efficient logarithmic price process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathcal{P})$. The observed logarithmic price process $\widetilde{X}_t$ is defined as:

$$\widetilde{X}_{i\Delta_n} = X_{i\Delta_n}(1-\mathbb{B}_{i,n}) + \widetilde{X}_{(i-1)\Delta_n}\mathbb{B}_{i,n}, \tag{1}$$

where $\Delta_n = \frac{1}{n}$, $\{t_{i,n} = i\,\Delta_n \mid i = 0, \ldots, n\}$ is a partition of the interval $[0,t]$ and $\mathbb{B}_{i,n}$ is a triangular array of $\mathcal{F}_{i\Delta_n}$-measurable Bernoulli random variables defined on the same probability space as the efficient price X_t.

When volumes are zero (or limited, with zero price impacts), prices repeat themselves ($\mathbb{B}_{i,n} = 1$).[1] When volumes are present ($\mathbb{B}_{i,n} = 0$), observed prices coincide with underlying efficient prices. In the model, the probability of staleness is $\mathbb{P}(\mathbb{B}_{i,n} = 1) = p_n^{\varnothing}$. When working with dependent Bernoulli variates and a frequency-specific probability of staleness, the model specification in Eq. (1) coincides with that in Bandi et al. (2017). The case of independent Bernoulli variates with a constant probability of staleness $\mathbb{P}(\mathbb{B}_{i,n} = 1) = p_n^{\varnothing} = p^{\varnothing}$ is studied in the work of Phillips and Yu (2009). Rich empirical evidence for the proposed specification is contained in Bandi et al. (2019).

The realized high-order moment estimators are generalized power variation estimators defined as

$$\mathsf{PV}(f;\widetilde{X}) = n^{r/2-1}\sum_{i=1}^{n} f(\Delta_i \widetilde{X}), \tag{2}$$

where either $f(x) = |x|^r$ or $f(x) = x^r$, for some $r > 0$. Naturally, their asymptotics depend both on the dynamics of the efficient price process and on those of the Bernoulli variates $\mathbb{B}_{i,n}$, to which we now turn.

Assumption 1. Assume X_t evolves as

$$X_t = X_0 + \int_0^t a_s\,ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{|\kappa(s,\delta)|\leqslant 1} \kappa(s,\delta)(\mu-\nu)(ds,d\delta) + \int_0^t \int_{|\kappa(s,\delta)|>1} \kappa(s,\delta)\mu(ds,d\delta), \tag{3}$$

where μ is a Poisson random measure on $R^+ \times E$ with predictable compensator $\nu(dt,dx) = dt\lambda(dx)$,

[1] According to the nomenclature in Bandi et al. (2019), the term "staleness" defines zero returns, whereas the term "idleness" defines zero returns with strictly zero volumes. Consistent with this terminology, we employ the term "staleness" to include all zeros. This choice can be viewed as being less restrictive than solely focusing on idleness and is economically-meaningful - if one views the probability of zeros as an illiquidity proxy - because the volumes associated with staleness are either zero (idleness) or limited (c.f., Bandi et al., 2019).

λ is a σ-finite measure on the set $(E, \mathcal{E})$ and the volatility function σ satisfies the equation

$$\sigma_t = \sigma_0 + \int_0^t a'_s \, ds + \int_0^t \sigma'_s dW_s + \int_0^t \int_{|\widetilde{\kappa}(s,\widetilde{\delta})| \leqslant 1} \widetilde{\kappa}(s, \widetilde{\delta})(\widetilde{\mu} - \widetilde{\nu})(ds, d\widetilde{\delta}) + \int_0^t \int_{|\widetilde{\kappa}(s,\widetilde{\delta})| > 1} \widetilde{\kappa}(s, \widetilde{\delta})\widetilde{\mu}(ds, d\widetilde{\delta}), \quad (4)$$

where, again, $\widetilde{\mu}$ is a Poisson random measure with compensator $\widetilde{\nu}(dt, dx) = dt\widetilde{\lambda}(dx)$, W_t and V_t are independent Brownian motions, and a, a', σ' and v' are adapted càdlàg bounded processes. Given a sequence of increasing stopping times τ_u and deterministic, nonnegative functions γ_u defined on E so that $\int \gamma_u^2(\delta)\lambda(d\delta) < \infty$, we have that $|\kappa(\omega, t, \delta)| \wedge 1 \leqslant \gamma_u(\delta)$ for all u and all (ω, t, δ) with $t \leqslant \tau_u$.

Assumption 2. Assume $\mathbb{B}_{i,n}$ is a triangular array of i.i.d. Bernoulli random variates, independent of X_t, with $\mathbb{E}[\mathbb{B}_{i,n}] = p^{\varnothing} \in [0, 1)$.

Assumption (1) is a standard sufficient condition required for the proof of the types of central limit theorems (CLTs) obtained in high-frequency econometrics. It postulates that the efficient price, as well as its spot volatility, evolve as Brownian semi-martingales with jumps. While we allow for infinite activity (small) jumps, our main interest is in economically-informative large discontinuities. The coefficients of the semi-martingales are assumed to be bounded, a condition which can be relaxed by standard localization arguments. As discussed above, Assumption (2) is consistent with Phillips and Yu (2009).

In order to distinguish between continuous and discontinuous trajectories, we consider two complementary subsets of Ω defined as:

$$\Omega^c_{[0,t]} = \{\omega \in \Omega : X_s(\omega) \text{ is continuous on } [0, t]\}, \quad (5)$$

$$\Omega^j_{[0,t]} = \{\omega \in \Omega : X_s(\omega) \text{ is discontinuous on } [0, t]\}. \quad (6)$$

Below, we establish $\mathcal{F}^\star$-stable convergence in law (denoted by "$\xrightarrow{stably}$") of the power variation estimates, where $\mathcal{F}^\star = \bigvee_{t \geqslant 0} \mathcal{F}^\star_t$, and $(\mathcal{F}^\star_t)_{t\geqslant 0}$ is the sub-filtration of $(\mathcal{F}_t)_{t\geqslant 0}$, such that the process X_t is adapted to $(\mathcal{F}^\star_t)_{t\geqslant 0}$, any $\mathcal{F}^\star_t$-martingale is a $\mathcal{F}_t$-martingale, and $\mathcal{F}^\star_t$ is independent of the σ-algebras generated by the Bernoulli variates.

2.2 Power variation on $\Omega^c_{[0,t]}$

We begin by assuming that the trajectory of the efficient price process is continuous on $[0, t]$, that is, $\Delta X_u = X_u - X_{u-} = 0 \ \forall u \in [0, t]$. In this case, the limiting properties of the realized high-order moments follow from the general limit theorem for realized power variation which we present next.

Theorem 2.1. Let Assumptions (1) and (2) hold and $\Delta X_u = 0 \ \forall u \in [0, t]$. Set $\rho(f) = \mathbb{E}[f(U)]$ and $\rho(f, k) = \mathbb{E}\left[f(U)U^k\right] \forall k > 0$, where $U \sim \mathcal{N}(0, 1)$. Then, as $n \to \infty$,

$$\mathsf{PV}(f; \widetilde{X}) \xrightarrow{u.c.p.} \frac{(1 - p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \int_0^t |\sigma_s|^r \, ds, \quad (7)$$

where $\mathsf{Li}_s\left(x\right) = \sum\limits_{k=1}^{\infty} \frac{1}{k^s} x^k$ is the polylogarithm function. Moreover, as $n \to \infty$,

$$\sqrt{n}\left(\mathsf{PV}(f;\widetilde{X}) - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\rho(f)\int_0^t \left|\sigma_s\right|^r \, ds\right) \overset{stably}{\longrightarrow} \int_0^t \alpha_s(1)\, ds + \int_0^t \alpha_s(2)\, dW_s + \int_0^t \alpha_s(3)\, dW'_s, \tag{8}$$

where W' is a Brownian motion defined on an extension of the original probability space and independent of W,

$$\alpha_s(1) = \sigma_s^{r-1}\left(a_s\rho(f') + \frac{1}{2}\sigma'_s\left(\rho(f',2) - \rho(f')\right)\right)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right), \tag{9}$$

$$\alpha_s(2) = \rho(f,1)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right)\sigma_s^r, \tag{10}$$

and

$$\alpha_s(3) = \sqrt{\left(\rho(f^2)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-r}\left(p^{\varnothing}\right) - (\rho(f))^2\left(2\frac{(1-p^{\varnothing})^4}{(p^{\varnothing})^2}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\mathsf{Li}_{-\frac{r}{2}-1}\left(p^{\varnothing}\right) - \frac{(1-p^{\varnothing})^6}{(p^{\varnothing})^3}\left(\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\right)^2\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\right) - \left(\rho(f,1)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right)\right)^2\right)\sigma_s^{2r}}. \tag{11}$$

Theorem 2.1 shows that both the probability limit and the asymptotic distribution of realized power variation are affected by staleness (i.e., $p^{\varnothing}$). The polylogarithm (or Jonquière's) function, $\mathsf{Li}_s\left(x\right)$, which appears in Eq. (7) and in Eqs. (9)-(11), is a special function defined by the power series

$$\mathsf{Li}_s\left(x\right) = \sum_{k=1}^{\infty} \frac{1}{k^s} x^k.$$

It is available in standard software packages. Using the properties of the polylogarithm function, it is straightforward to show that, if $p^{\varnothing} = 0$, then Theorem 2.1 coincides with the standard stable CLT for generalized power variation derived by Kinnebrock and Podolskij (2008).

In order to reveal the meaning of Theorem 2.1, we specialize the statement to a series of corollaries, depending on whether the function f is odd or even. If f is *odd* (for example, if $f(x) = x^3$), then $\rho(f) = 0$ and the probability limit of the estimator is unbiased, as we formalize below.

Corollary 2.1.1. Assume $f(x) = x^r$, where r is an odd integer. Let Assumptions (1) and (2) hold and $\Delta X_u = 0 \ \forall u \in [0,t]$. Then, as $n \to \infty$,

$$n^{\frac{r}{2}-1}\sum_{i=1}^{n}\left(\Delta_i\widetilde{X}\right)^r \overset{u.c.p.}{\longrightarrow} 0, \tag{12}$$

and

$$n^{\frac{r-1}{2}}\sum_{i=1}^{n}\left(\Delta_i\widetilde{X}\right)^r \overset{stably}{\longrightarrow} \int_0^t \alpha_s(1)\, ds + \int_0^t \alpha_s(2)\, dW_s + \int_0^t \alpha_s(3)\, dW'_s, \tag{13}$$

where W' is a Brownian motion defined on an extension of the original probability space and independent of W,

$$\alpha_s(1) = \sigma_s^{r-1}\left(a_s\mu_{r-1} + \frac{1}{2}\sigma_s'\left(\mu_{r+1} - \mu_{r-1}\right)\right)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right), \tag{14}$$

$$\alpha_s(2) = \frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mu_{r+1}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right)\sigma_s^r, \tag{15}$$

and

$$\alpha_s(3) = \sqrt{\mu_{2r}\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-r}\left(p^{\varnothing}\right)\sigma_s^{2r} - \left(\alpha_s(2)\right)^2}, \tag{16}$$

with $\mu_r = \mathbb{E}\left[|U|^r\right]$ and $U \sim \mathcal{N}(0,1)$.

When f is odd, power variation is asymptotically unbiased for zero. Intuitively, this result reflects the fact that increments of Brownian motion have symmetric distributions and staleness does not affect symmetry (because of the local martingale features of the driving terms of the underlying efficient price process). However, staleness influences convergence in law through the structure of the estimator's limiting variance.

Note that, if f is odd, power variation converges stably in law to the process U_t defined as the sum of three terms:

$$\mathsf{U}_t = \int_0^t \alpha_s(1)\,ds + \int_0^t \alpha_s(2)\,dW_s + \int_0^t \alpha_s(3)\,dW_s'.$$

The process U_t is uncentered mixed normal. Its look differs from that of the standard limit of the realized variance estimator because of the presence of the terms $\int_0^t \alpha_s(1)\,ds$ and $\int_0^t \alpha_s(2)\,dW_s$, which depend on the drift process, i.e., μ_t, and on the driving shock of the local martingale portion of X_t, i.e., the Brownian motion W_t. Importantly, the difference is not driven by staleness, but emerges solely due to the fact that f is an odd function (Kinnebrock and Podolskij (2008) provide details in other contexts). As said, staleness, however, affects both the location and the scale of U_t via the form of the process $\alpha_t(i)$, for $i = 1, 2, 3$.

Consider, now, the case in which f is an *even* function (for example, $f(x) = |x|^r$, for a generic $r > 0$, or $f(x) = x^r$, for an even $r > 0$). If $f(x) = x^r$ with an even $r > 0$, then $\mathsf{PV}(x^r; \widetilde{X}) = \mathsf{PV}(|x|^r; \widetilde{X})$.

Corollary 2.1.2. Let Assumptions (1) and (2) hold and $\Delta X_u = 0$ $\forall u \in [0,t]$. Let $\mu_r = \mathbb{E}\left[|U|^r\right]$ and $U \sim \mathcal{N}(0,1)$. Then, as $n \to \infty$,

$$n^{\frac{r}{2}-1}\sum_{i=1}^n \left|\Delta_i\widetilde{X}\right|^r \xrightarrow{u.c.p.} \frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\mu_r\int_0^t |\sigma_s|^r\,ds, \tag{17}$$

and

$$\sqrt{n}\left(n^{\frac{r}{2}-1}\sum_{i=1}^n \left|\Delta_i\widetilde{X}\right|^r - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\mu_r\int_0^t |\sigma_s|^r\,ds\right) \xrightarrow{stably} \int_0^t \alpha_s\,dW_s', \tag{18}$$

where W' is a Brownian motion defined on an extension of the original probability space and independent of W and

$$\alpha_s = \sqrt{\left[\mu_{2r}\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-r}\left(p^{\varnothing}\right) - (\mu_r)^2\left(2\frac{(1-p^{\varnothing})^4}{(p^{\varnothing})^2}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\mathsf{Li}_{-\frac{r}{2}-1}\left(p^{\varnothing}\right) - \frac{(1-p^{\varnothing})^6}{(p^{\varnothing})^3}\left(\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right)\right)^2\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\right)\right]\sigma_s^{2r}}. \tag{19}$$

Corollary 2.1.2 shows that power variation measures computed from the absolute values of increments of the observed price process are - in general - asymptotically biased due to staleness. The case $r = 2$ is an important exception. If $r = 2$, the properties of the polylogarithm function imply that, for any $p^{\varnothing} \in [0, 1)$, $\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-1}\left(p^{\varnothing}\right) = 1$. This case is considered below separately. If $r > 2$, the bias term takes the form of a function of two variables, r and $p^{\varnothing}$, which is increasing in both arguments. Table 2.2 displays values of the bias of the realized fourth moment, $r = 4$, for different values of $p^{\varnothing}$. As shown, the percentage bias can be substantial even for moderate values of $p^{\varnothing}$. For any $p^{\varnothing} \in (0, 1)$, the percentage bias, $\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) - 1$, is larger than 0 if $r > 2$ and it is smaller than 0 for $r < 2$. In sum, the realized high-order moment are overestimated due to the presence of staleness.

Table 1: We report values of the bias of PV, with $r = 4$, for different $p^{\varnothing}$s.

$p^{\varnothing}$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right)$	1.222	1.500	1.857	2.333	3.000	3.999	5.651	8.558

Finally, if $r = 2$, the power variation $\mathsf{PV}(x^r; \widetilde{X})$ coincides with realized variance, denoted by $\mathsf{RV}(\widetilde{X})$. The asymptotic properties of realized variance under staleness (but without jumps, as in this subsection) are investigated by Phillips and Yu (2009), who show that $\mathsf{RV}(\widetilde{X})$ remains (asymptotically) unbiased for the quadratic variation of the efficient logarithmic price X_t. However, the limiting variance of $\mathsf{RV}(\widetilde{X})$ increases with the frequency of zeros. The results in Phillips and Yu (2009) can be deduced from Theorem 2.1 by specializing it to the case $r = 2$, as shown in the Corollary below.

Corollary 2.1.3. Let Assumptions (1) and (2) hold and $\Delta X_u = 0\ \forall u \in [0, t]$. Then, as $n \to \infty$,

$$\sum_{i=1}^{n}\left|\Delta_i\widetilde{X}\right|^2 \xrightarrow{u.c.p.} \int_0^t |\sigma_s|^2\, ds, \tag{20}$$

and

$$\sqrt{n}\left(\sum_{i=1}^{n}\left|\Delta_i\widetilde{X}\right|^2 - \int_0^t |\sigma_s|^2\, ds\right) \xrightarrow{stably} \int_0^t \alpha_s\, dW_s', \tag{21}$$

with

$$\alpha_s^2 = \frac{4 - 2(1 - p^{\varnothing})}{(1 - p^{\varnothing})} \sigma_s^4. \tag{22}$$

2.3 Power variation on $\Omega^j_{[0,t]}$

We now consider the situation in which the trajectories of the efficient price X_t have discontinuities on the observation interval $[0, t]$. The case without staleness, with a focus on volatility, has been considered at some length in the literature (see, e.g., Barndorff-Nielsen et al., 2006, and Veraart, 2010)

We devote our attention to the limiting behavior of power variation with large powers ($r \geqslant 2$), namely the powers which are used in computing realized high-order moments and are affected by the presence of jumps (as is well-known, jumps do not affect the convergence in probability of power variation with $r < 2$, see -e.g.- Barndorff-Nielsen et al., 2006, and Jacod, 2008).

As is natural in the discontinuous case, we work with re-scaled (or, equivalently, non-normalized) power variations defined as

$$\sum_{i=1}^{n} f\left(\Delta_i \widetilde{X}\right) = n^{1-\frac{r}{2}}\, \mathsf{PV}(f; \widetilde{X}). \tag{23}$$

The standardization $1/n$ in Eq. (44) is now not needed because of convergence to a finite sum (that of functions of the jumps over the interval $[0, t]$). The standardization $n^{\frac{r}{2}}$ in Eq. (44) is also not needed because one does not have to offset the limiting probability order of the driving Brownian increments and we can let them vanish to zero to identify the genuine jump futures. Hence, we multiply PV by $n^{1-\frac{r}{2}}$.

First, we consider convergence in probability. In contrast to the continuous case, the probability limit of power variation on $\Omega^j_{[0,t]}$ is robust to staleness.

Theorem 2.2. Let Assumptions (1) and (2) hold. Let either $f(x) = |x|^r$ or $f(x) = x^r$, for some $r > 2$. Then, as $n \to \infty$,

$$\sum_{i=1}^{n} f\left(\Delta_i \widetilde{X}\right) \xrightarrow{u.c.p.} \sum_{s \leqslant t} f(\Delta X_s), \tag{24}$$

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

On $\Omega^j_{[0,t]}$, power variation with powers larger than 2 depends asymptotically only on jumps. For example, the quantities $\sum_{i=1}^n \left(\Delta_i \widetilde{X}\right)^3$ and $\sum_{i=1}^n \left(\Delta_i \widetilde{X}\right)^3$, which will be employed to compute realized skewness and realized kurtosis in Section 4, converge to the sums of the jumps of X on the interval $[0, t]$ raised to the third and the fourth power, respectively. The limiting behavior of realized variance (i.e., the case $r = 2$) is different and is, therefore, considered separately at the end of this subsection.

In the absence of staleness, a general CLT for power variation on $\Omega^j_{[0,t]}$ was obtained by Jacod and Protter (2012) for $r > 3$ and by Koike and Liu (2019) for $r = 3$. In the first case, power variation converges stably in law to a complex mean-zero limiting process, which depends on the jumps of X. In the second case, the centering of the limiting distribution depends on the features of the continuous variation of the process. Our results are, as expected, contaminated by zeros, thereby resulting in different limiting processes, as follows from the two theorems below.

Theorem 2.3. Let Assumptions (1) and (2) hold. Let either $f(x) = |x|^r$ or $f(x) = x^r$, for some $r > 3$. Then, as $n \to \infty$,

$$\sqrt{n}\left(\sum_{i=1}^{n} f\left(\Delta_i \widetilde{X}\right) - \sum_{s \leqslant t} f(\Delta X_s)\right) \overset{stably}{\longrightarrow} \sum_{p:S_p \leqslant t} f'(\Delta X_{S_p})\left(\sigma_{S_p -}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right), \tag{25}$$

where $(S_p)_{p \leqslant 1}$ is a sequence of stopping times which exhausts the jumps of X and U_p^-, U_p^+, ξ_p^-, ξ_p^+ are four sequences of independent random variables defined on an extension $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{P}}\right)$ of the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. The random quantities U_p^- and U_p^+ follow standard normal distribution, $\xi_p^- = u_p L_p$ and $\xi_p^+ = (1 - u_p) L_p$, where the u_ps have uniform distributions on $[0, 1]$ and the L_ps are discrete random variables on $\{1, 2, \ldots\}$ with $\widetilde{\mathcal{P}}(L_p = \ell) = (1 - p^{\varnothing})\left(p^{\varnothing}\right)^{\ell - 1}$ for $\ell \in \{1, 2, \ldots\}$.

The limiting process defined in Eq. (25) is a square-integrable $\mathcal{F}^\star$-martingale. The existence of such process follows from the same argument as in the proof of Proposition 5.1.1 in Jacod and Protter (2012). Without staleness, the limiting process takes the form

$$\sum_{p:S_p \leqslant t} f'(\Delta X_{S_p})\left(\sigma_{S_p -}\sqrt{u_p}U_p^- + \sigma_{S_p}\sqrt{1 - u_p}U_p^+\right). \tag{26}$$

Hence, staleness increases the asymptotic variance, thereby yielding an information loss when measuring jump moments by virtue of power variation. Finally, we note that the stable convergence in Theorem 2.3 holds for a fixed t but it does not hold for the Skorokhod topology (see Remark 5.1.3 in Jacod and Protter (2012) for details). We now turn to the case $r = 3$.

Theorem 2.4. Let Assumptions (1) and (2) hold. Let either $f(x) = |x|^3$ or $f(x) = x^3$. Then, as $n \to \infty$,

$$\sqrt{n}\left(\sum_{i=1}^{n} f\left(\Delta_i \widetilde{X}\right) - \sum_{s \leqslant t} f(\Delta X_s)\right) \overset{stably}{\longrightarrow} \frac{(1 - p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{3}{2}}\left(p^{\varnothing}\right) \rho(f) \int_0^t |\sigma_s|^3 \, ds + \sum_{p:S_p \leqslant t} f'(\Delta X_{S_p})\left(\sigma_{S_p -}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right), \tag{27}$$

where $\rho(f) = \mathbb{E}[f(U)]$ with $U \sim \mathcal{N}(0, 1)$, as above, and the second term on the right-hand side is defined as in Theorem 2.3.

Theorem 2.4 shows that, if $r = 3$, the power variation converges stably to an uncentered limit distribution. The bias is of the form $\frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{3}{2}}\left(p^{\varnothing}\right) \rho(f) \int_0^t |\sigma_s|^3 \, ds$, which coincides with the limit in probability of power variation on $\Omega_{[0,t]}^c$. Notice that, in the odd case $f(x) = x^3$, the bias disappears, as $\rho(x^3) = 0$.

Finally, consider realized variance on $\Omega_{[0,t]}^j$. As in the continuous case, $\mathsf{RV}(\widetilde{X})$ is a robust estimator of the quadratic variation (now inclusive of the sum of the squared jumps) of the efficient price process. However, the limit distribution of $\mathsf{RV}(\widetilde{X})$ is contaminated by staleness, as evidenced by the theorem below.

Theorem 2.5. Let Assumption (1) and (2) hold. Then, as $n \to \infty$,

$$\mathsf{RV}(\widetilde{X}) \xrightarrow{u.c.p.} \int_0^t |\sigma_s|^2 \, ds + \sum_{s \leqslant t} (\Delta X_s)^2 \tag{28}$$

and

$$\sqrt{n}\left(\mathsf{RV}(\widetilde{X}) - \int_0^t |\sigma_s|^2 \, ds - \sum_{s \leqslant t} (\Delta X_s)^2\right) \xrightarrow{stably} \int_0^t \alpha_s \, dW'_s + 2 \sum_{p: S_p \leqslant t} \Delta X_{S_p} \left(\sigma_{S_p -}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right), \tag{29}$$

where $\alpha_s^2 = \frac{4-2(1-p^{\varnothing})}{(1-p^{\varnothing})}\sigma_s^4$ and the second term on the right-hand side is defined as in Theorem 2.3.

3 Correcting for staleness

Theorem 2.1 implies that, on $\Omega^c_{[0,t]}$, odd power variation is robust to staleness (in the sense of convergence in probability) while even power variation is asymptotically biased.

We note that the presence of zeros has the effect of turning the original deterministic sampling grid into a new random (and coarser) grid. Hence, the limiting bias may be corrected by utilizing the logic in Hayashi et al. (2011) for dealing with power variation on irregular grids.

Let $t^*_{1,n}, \ldots, t^*_{N^n_t,n}$ be the (random) partition of the interval $[0, t]$ constructed from the original (deterministic) partition by removing the points for which $\mathbb{B}_{i,n} = 1$. Set $\Delta(i, n) = t^*_{i,n} - t^*_{i-1,n}$ and denote by $\Delta^*_i X = X_{t^*_{i,n}} - X_{t^*_{i-1,n}}$ the increments of the process X_t over the new partition. Note that $\Delta^*_i X = \Delta^*_i \widetilde{X}$, by construction.

We define the *corrected power variation* as:

$$\mathsf{PV}^c(f, \widetilde{X}) = \sum_{i=1}^{N^n_t} \Delta(i, n) f\left((\Delta(i, n))^{-1/2} \Delta^*_i X\right). \tag{30}$$

Naturally, the term $\Delta(i, n)^{-1/2}$ in the f function accounts for the new grid (and the new, effective number of observations) by replacing the original standardization (i.e., $n^{1/2}$). Similarly, the previous standardization (of the sum) by the number of observations $1/n = \Delta_n$ is now replaced by $\Delta(i, n)$.

The limiting properties of Eq. (30) can be derived from the work of Hayashi et al. (2011) provided the random partition $t^*_{1,n}, \ldots, t^*_{N^n_t,n}$ generated by the Bernoulli variables satisfies suitable regularity conditions discussed below.

First, due to the independence of the triangular array of Bernoulli variates $\mathbb{B}_{i,n}$ of the efficient price process, the random durations $\Delta(i, n)$ are independent of X_t, which guarantees that condition (**C**) of Hayashi et al. (2011) holds. Next, the mutual independence of $\mathbb{B}_{i,n}$ implies that $\sup_{i=1,\ldots N^n_t+1} \Delta(n, i)$

converges in probability to zero as $n \to \infty$. Finally, under Assumption (2), the "power variations" of the durations $\Delta(i, n)$ converge uniformly in probability, as follows from the following lemma.

Lemma 3.1. Let Assumption (2) hold. Then, for any $q > 0$, as $n \to \infty$,

$$A(q)_t^n = n^{q-1} \sum_{i=1}^{N_t^n} |\Delta(n,i)|^q \xrightarrow{u.c.p.} \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}} \mathsf{Li}_{-q}\left(p^{\varnothing}\right) t. \tag{31}$$

Lemma 3.1 implies that condition $(\mathbf{D}(q))$ of Hayashi et al. (2011) holds (with $r_n = n$ in their notations) for every $q > 0$. We note that, by imposing economically-motivated structure on data sparsity (due to staleness), our proposed approach provides a data generating process which permits theoretical verification of the condition $(\mathbf{D}(q))$ in Hayashi et al. (2011).

We can now establish the limiting properties of the corrected power variation on $\Omega^c_{[0,t]}$.

Theorem 3.2. Let Assumptions (1) and (2) hold and $f(x) = |x|^r$ with $r > 0$. Then, on $\Omega^c_{[0,t]}$, as $n \to \infty$,

$$\mathsf{PV}^c(f; \widetilde{X}) \xrightarrow{u.c.p.} \mu_r \int_0^t |\sigma_s|^r \, ds. \tag{32}$$

Moreover, as $n \to \infty$,

$$\sqrt{n}\left(\mathsf{PV}^c(f; \widetilde{X}) - \mu_r \int_0^t |\sigma_s|^r \, ds\right) \xrightarrow{stably} \left(\mu_{2r} - \mu_r^2\right) \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}} \mathsf{Li}_{-2}\left(p^{\varnothing}\right) \int_0^t \sigma_s^r \, dW_s'. \tag{33}$$

Theorem 3.2 shows that corrected power variation is robust to staleness: $\mathsf{PV}^c(f, \widetilde{X})$ converges in probability to the same limit as $\mathsf{PV}^c(f, X)$ and $\mathsf{PV}(f, X)$. The asymptotic variance of $\mathsf{PV}^c(f, \widetilde{X})$ is, however, contaminated by staleness. Specifically, it takes the form $\left(\mu_{2r} - \mu_r^2\right) \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}} \mathsf{Li}_{-2}\left(p^{\varnothing}\right) \int_0^t \sigma_s^{2r} \, ds$ and, since $\frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}} \mathsf{Li}_{-2}\left(p^{\varnothing}\right) > 1 \; \forall p^{\varnothing} \in (0,1)$, it is larger than the asymptotic variance of $\mathsf{PV}(f, X)$ which takes the simpler form $\left(\mu_{2r} - \mu_r^2\right) \int_0^t \sigma_s^{2r} \, ds$.

Due to Lemma 3.1, we can establish an interesting convergence result, which can be used to estimate the asymptotic variance of $\mathsf{PV}^c(f; \widetilde{X})$:

Theorem 3.3. Let Assumptions (1) and (2) hold and $r > 0$. Then, on $\Omega^c_{[0,t]}$, as $n \to \infty$,

$$n \sum_{i=1}^{N_t^n} \Delta(n,i)^{2-r} \left|\Delta_i^\star X\right|^{2r} \xrightarrow{u.c.p.} \mu_{2r} \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}} \mathsf{Li}_{-2}\left(p^{\varnothing}\right) \int_0^t |\sigma_s|^{2r} \, ds. \tag{34}$$

In light of Theorem 3.3, in spite of the information loss due to staleness, the asymptotic variance of $\mathsf{PV}^c(f; \widetilde{X})$ can be estimated as:

$$\frac{\mu_{2r} - \mu_r^2}{\mu_{2r}} n \sum_{i=1}^{N_t^n} \Delta(n,i)^{2-r} \left|\Delta_i^\star X\right|^{2r}. \tag{35}$$

4 Realized skewness and kurtosis

This section applies the previous limiting results to skewness (S) and kurtosis (K) measures. Ignoring means, write

$$\mathsf{S} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^3}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^2\right)^{\frac{3}{2}}} = \frac{\sqrt{n}\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^3}{\left(\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^2\right)^{\frac{3}{2}}} \text{ and } \mathsf{K} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^4}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^2\right)^{2}} = \frac{n\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^4}{\left(\sum_{i=1}^{n}\left(\Delta_i \widetilde{X}\right)^2\right)^{2}}. \tag{36}$$

In terms of power variation, the quantities in Eq. (36) can also be expressed as

$$\mathsf{S} = \frac{\mathsf{PV}(x^3;\widetilde{X})}{\left(\mathsf{RV}(\widetilde{X})\right)^{\frac{3}{2}}} \quad \text{and} \quad \mathsf{K} = \frac{\mathsf{PV}(x^4;\widetilde{X})}{\left(\mathsf{RV}(\widetilde{X})\right)^{2}}. \tag{37}$$

Eq. (36) is classical in the sense that it presents textbook notions of skewness and kurtosis broadly applied in the literature. We will show that these typical notions have, in general, *atypical* limits when doing infill asymptotics, as reasonable in the presence of high-frequency data.

Table 2 provides the summary of the limiting values (in the sense of convergence in probability) for both measures, with and without the correction for zeros and with and without jumps.

Table 2: The table reports the limits in probability of realized skewness and kurtosis in the absence (on Ω^c) and the presence (on Ω^j) of jumps and in the presence ($p^{\varnothing} > 0$) or absence ($p^{\varnothing} = 0$) of return zeros

	$p^{\varnothing} > 0$		$p^{\varnothing} = 0$	
	on $\Omega^c_{[0,t]}$	on $\Omega^j_{[0,t]}$	on $\Omega^c_{[0,t]}$	on $\Omega^j_{[0,t]}$
S	0	∞	0	∞
$\frac{1}{\sqrt{n}}\mathsf{S}$	0	$\frac{\sum_{s\leqslant t}(\Delta X_s)^3}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{\frac{3}{2}}}$	0	$\frac{\sum_{s\leqslant t}(\Delta X_s)^3}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{\frac{3}{2}}}$
K	$\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\frac{3\int_0^t\lvert\sigma_s\rvert^4\,ds}{\left(\int_0^t\lvert\sigma_s\rvert^2\,ds\right)^2}$	$+\infty$	$\frac{3\int_0^t\lvert\sigma_s\rvert^4\,ds}{\left(\int_0^t\lvert\sigma_s\rvert^2\,ds\right)^2}$	$+\infty$
$\frac{1}{n}\mathsf{K}$	0	$\frac{\sum_{s\leqslant t}(\Delta X_s)^4}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{2}}$	0	$\frac{\sum_{s\leqslant t}(\Delta X_s)^4}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{2}}$
S^c	0	∞	0	∞
$\frac{1}{\sqrt{n}}\mathsf{S}^c$	0	$\frac{\frac{1-p^{\varnothing}}{p^{\varnothing}}\mathsf{Li}_{-1}\left(p^{\varnothing}\right)\sum_{p:S_p\leqslant t}\left(\phi_p\Delta X_{S_p}\right)^3}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{\frac{3}{2}}}$	0	$\frac{\sum_{s\leqslant t}(\Delta X_s)^3}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{\frac{3}{2}}}$
K^c	$\frac{3\int_0^t\lvert\sigma_s\rvert^4\,ds}{\left(\int_0^t\lvert\sigma_s\rvert^2\,ds\right)^2}$	$+\infty$	$\frac{3\int_0^t\lvert\sigma_s\rvert^4\,ds}{\left(\int_0^t\lvert\sigma_s\rvert^2\,ds\right)^2}$	$+\infty$
$\frac{1}{n}\mathsf{K}^c$	0	$\frac{\frac{1-p^{\varnothing}}{p^{\varnothing}}\mathsf{Li}_{-1}\left(p^{\varnothing}\right)\sum_{p:S_p\leqslant t}\left(\phi_p\Delta X_{S_p}\right)^4}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{2}}$	0	$\frac{\sum_{s\leqslant t}(\Delta X_s)^4}{\left(\int_0^t \lvert\sigma_s\rvert^2\,ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{2}}$

Theorem 2.1 implies that realized skewness is robust to the presence of staleness on continuous

trajectories. On discontinuous trajectories, irrespective of staleness, Theorems 5.2 and 2.5 (because of the quadratic term in the denominator) require re-scaling (by $n^{-1/2}$) to identify jump skewness. For a large n, on $\Omega^j_{[0,t]}$, we have

$$\mathsf{S} \approx \sqrt{n}\frac{\sum_{s\leqslant t}(\Delta X_s)^3}{\left(\int_0^t |\sigma_s|^2\, ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^{\frac{3}{2}}} \longrightarrow \begin{cases} +\infty, & \text{if } \sum_{s\leqslant t}(\Delta X_s)^3 > 0, \\ -\infty, & \text{if } \sum_{s\leqslant t}(\Delta X_s)^3 < 0. \end{cases} \tag{38}$$

Therefore, in the absence of re-scaling, realized skewness may *only* be interpreted as a signed measure of jumps on the interval $[0,t]$. Large positive (negative) values of S support the presence of sufficiently many and/or sufficiently large (relative to the quadratic variation) positive (or negative) jumps on $[0,t]$. Small values of S are symptomatic of absence of large jumps on $[0,t]$.

Realized kurtosis behaves differently. In the absence of jumps, the value of realized kurtosis depends on volatility and the number of zeros. The bias associated with zeros takes the form

$$\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right) > 1, \tag{39}$$

which increases with $p^{\varnothing}$ and may reach large values, as illustrated in Table 1. In the presence of jumps, for large n values, we have:

$$\mathsf{K} \approx \frac{\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right) 3\int_0^t |\sigma_s|^4\, ds + n\sum_{s\leqslant t}(\Delta X_s)^4}{\left(\int_0^t |\sigma_s|^2\, ds + \sum_{s\leqslant t}(\Delta X_s)^2\right)^2} \longrightarrow \infty, \tag{40}$$

indicating that kurtosis ought to be large on $\Omega^j_{[0,t]}$. In essence, kurtosis may result in large values on $\Omega^c_{[0,t]}$ as well as on $\Omega^j_{[0,t]}$.

We now consider corrected (as in Section 4) versions of realized skewness and kurtosis defined as

$$\mathsf{S}^c = \frac{\mathsf{PV}^c(x^3;\widetilde{X})}{\left(\mathsf{RV}(\widetilde{X})\right)^{\frac{3}{2}}} \quad \text{and} \quad \mathsf{K}^c = \frac{\mathsf{PV}^c(x^4;\widetilde{X})}{\left(\mathsf{RV}(\widetilde{X})\right)^{2}}. \tag{41}$$

Note that we do not correct the realized variance in the denominators of S^c and K^c as $\mathsf{RV}(\widetilde{X})$ is robust to staleness.

As expected, correcting realized skewness for the presence of staleness does not lead to any asymptotic improvement of the measure. In the case of kurtosis, correcting for staleness eliminates bias on $\Omega^c_{[0,t]}$ at the cost of re-scaling the jumps on $\Omega^j_{[0,t]}$. On $\Omega^c_{[0,t]}$, for large n, corrected kurtosis takes the usual form:

$$\mathsf{K}^c \approx 3\frac{\int_0^t |\sigma_s|^4\, ds}{\left(\int_0^t |\sigma_s|^2\, ds\right)^2}. \tag{42}$$

In other words, if volatility is constant over the interval $[0,t]$, $\mathsf{K}^c = 3$ as for the Gaussian distribution. However, on $\Omega^j_{[0,t]}$, we have

$$\mathsf{K}^c \approx \frac{3\frac{\int_0^t |\sigma_s|^4 \, ds}{\left(\int_0^t |\sigma_s|^2 \, ds\right)^2} + n\frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-1}\left(p^{\varnothing}\right) \sum\limits_{p:S_p \leqslant t} \left(\phi_p \Delta X_{S_p}\right)^4}{\left(\int_0^t |\sigma_s|^2 \, ds + \sum\limits_{s \leqslant t} \left(\Delta X_s\right)^2\right)^2}. \tag{43}$$

Hence, asymptotically, corrected realized kurtosis ought to take on large values on $\Omega^j_{[0,t]}$ due to the exploding term produced by the jumps. In a finite sample, there is a chance of underestimation because the random variables ϕ_p are such that $|\phi_p| < 1$ almost surely (even though $\frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-1}\left(p^{\varnothing}\right) = \frac{1}{1-p^{\varnothing}}$).

In sum, whether one accounts for staleness or not, during jump days the limiting properties of standard skewness and kurtosis measures depend on the sample size. Arguably, this is not a favorable property of these classical measures when applied to high-frequency data. Below, we provide a solution.

5 Disentangling continuous variation from jumps

In the absence of staleness, the continuous variation of X_t can be separated from jumps using truncated power variation, as in Mancini (2009). The truncation techniques remains valid under the presence of zeros, provided staleness is properly taken into account, as explained below.

Realized *truncated* power variation is defined as

$$\mathsf{TPV}(f; \widetilde{X}) = n^{r/2-1} \sum_{i=1}^{n} f(\Delta_i \widetilde{X}) \cdot I_{\{|\Delta_i \tilde{X}| < \theta_n\}}, \tag{44}$$

where $I_{\{.\}}$ denotes the indicator function and θ_n is a (stochastic) sequence such that $\theta_n \to 0$ with $\frac{\theta_n}{\sqrt{\Delta_n \log(\Delta_n)}} \to \infty$ a.s. Classical arguments as in Mancini (2009) imply that, irrespective of the presence of jumps (i.e., on both $\Omega^c_{[0,t]}$ and $\Omega^j_{[0,t]}$), for sufficiently large n, we have:

$$\mathsf{TPV}(f; \widetilde{X}) \approx \mathsf{TPV}(f; \widetilde{X}'), \tag{45}$$

where $\widetilde{X}'$ denotes the continuous portion of $\widetilde{X}$.

In order to account for the presence of zeros, the corrected version of truncated power variation is now defined as:

$$\mathsf{TPV}^c(f, \widetilde{X}) = \sum_{i=1}^{N_t^n} \Delta(i,n) f\left(\left(\Delta(i,n)\right)^{-1/2} \Delta_i^* X\right) \cdot I_{\left\{|\Delta_i^* X| < \alpha_i |\Delta(i,n)|^{\rho}\right\}}, \tag{46}$$

where the α_is are some positive bounded random variables and $\rho \in (0, 1/2)$.

Theorem 5.1. Let Assumptions (1) and (2) hold, let $f(x) = |x|^r$ with $r > 0$ and let the functions γ_u in Assumption (1) satisfy $\int \gamma_u^\beta(\delta)\lambda(d\delta) < \infty$ for some $0 < \beta < 2$. If either $r \leqslant 2$ or $r > 2$ and $\rho \geqslant \frac{r-2}{2r-2\beta}$, as $n \to \infty$,

$$\mathsf{TPV}^c(f, \widetilde{X}) - \mathsf{PV}^c(f, \widetilde{X}') \xrightarrow{u.c.p.} 0. \tag{47}$$

In addition, if $\beta \in (0, 1]$, as $n \to \infty$,

$$\sqrt{n}\left(\mathsf{TPV}^c(f, \widetilde{X}) - \mathsf{PV}^c(f, \widetilde{X}')\right) \xrightarrow{u.c.p.} 0. \tag{48}$$

Thus, the variation of the jump components can be measured by (non-normalized) jump power variation defined as

$$\mathsf{J}(f, \widetilde{X}) = \sum_{i=1}^{n} f\left(\Delta_i^* X\right) \cdot I_{\{\left|\Delta_i^* X\right| \geqslant \alpha_i |\Delta(i,n)|^\rho\}}. \tag{49}$$

On $\Omega^j_{[0,t]}$, $\mathsf{J}(f, \widetilde{X})$ behaves as the non-normalized power variation in Theorem 5.2. However, the use of truncation allows us to relax the assumption $r > 2$ of the theorem. In other words, we have the following theorem.

Theorem 5.2. Let Assumptions (1) and (2) hold and $f(x) = O\left(|x|^r\right)$ as $x \to 0$ for some $r > 0$. Then, as $n \to \infty$,

$$\mathsf{J}(f, \widetilde{X}) \xrightarrow{u.c.p.} \sum_{s \leqslant t} f(\Delta X_s). \tag{50}$$

This result allows us to simply employ the sum of the jumps ($\mathsf{J}(x, \widetilde{X})$) when measuring the (unsigned or signed) strength of the jump contribution.

In order to account for the sign of the jump variation, we consider the sum of the positive and negative jumps measured respectively as:

$$\mathsf{J}^+(x, \widetilde{X}) = \sum_{i=1}^{n} (\Delta_i^* X) \cdot I_{\{\Delta_i^* X \geqslant \alpha_i |\Delta(i,n)|^\rho\}}, \quad \text{and} \quad \mathsf{J}^-(x, \widetilde{X}) = \sum_{i=1}^{n} (\Delta_i^* X) \cdot I_{\{\Delta_i^* X \leqslant -\alpha_i |\Delta(i,n)|^\rho\}}. \tag{51}$$

In applications, these measures may represent less noisy assessments of the contributions of jumps than the more volatile jump skewness and kurtosis measures which could readily be obtained from Eq. (49) above.

6 High-order moments and asset prices

The cross-sectional pricing of higher-order moments has been the subject of recent literature (e.g., Amaya et al., 2015 and Bollerslev et al., 2019) The work of Amaya et al. (2015) points to the negative pricing of idiosyncratic skewness, as measured by K in Eq. (36), an empirical finding justified by market participants' attention (resp. aversion) to large positive (resp. negative) payoffs. They also report on the positive pricing of idiosyncratic kurtosis, a somewhat less robust result possibly explained by aversion to diffusive/jump variation in volatility. Just like idiosyncratic skewness, the difference between positive semi-variance and negative semi-variance ("relative signed jumps" or RSJ) is a measure of the relative contribution of positive and negative jumps to price variation. It is expected to be positive (resp. negative) when positive (resp. negative) jumps dominate. Bollerslev et al. (2019) argue that controlling for RSJ reduces drastically the impact of idiosyncratic skewness, thereby implying that RSJ constitutes a superior measure of jump variation than idiosyncratic skewness, at least for the purpose of asset pricing.

We re-evaluate this literature, and expand on its scope, along several dimensions. We have previously shown that skewness and kurtosis are dominated by jumps in days in which jumps are present (c.f. Section 4). In days without jumps, skewness is a noisy measure of zero, while kurtosis captures stochastic volatility, inclusive of jumps in volatility. As discussed, continuous kurtosis suffers from severe upward biases in the presence of staleness (c.f. Section 4). We have also shown that, more generally (i.e., for every day in the sample), one may use thresholding to define continuous and jump analogues of skewness and kurtosis (c.f. Section 5). Continuous (truncated) skewness should, again, be a noisy measure of zero (and, therefore, not affect pricing). Continuous (truncated) kurtosis will, once more, be upward biased due to staleness. Thus, the positive link between kurtosis and illiquidity, something which is emphasized by Amaya et al., 2015, may, at least in part, be a by-product of staleness-induced biases. In light of these observations, relying on the theory laid out earlier, we study the pricing of continuous and discontinuous variation through truncated (standardized or unstandardized) high-order moments. The importance of bias-correcting continuous (truncated) kurtosis will be evaluated using a sensible pricing metric.

We consider a large cross-section of intraday prices for 4809 NYSE-listed stock. The data are recorded on a one-minute grid from 9:30 a.m. to 4 p.m., from January 1, 1998, to June 29, 2018, thereby amounting to a long sample of intra-daily data ideally suited for a pricing study. In order to soften the effect of market microstructure noise, we aggregate returns up to the five-minute frequency (Δ_n, in our previous notation). We remove from consideration days with more than 70% of zero intraday returns and for which rounding (as represented by the *rounding impact ratios* of Bandi et al., 2019) is estimated to be relatively larger. These quality cuts give us almost 12 million daily observations of the realized quantities of interest.

We begin by illustrating the impact of staleness on skewness and kurtosis measures in our data. Figure 1 shows percentiles (the 10^{th}, the 50^{th} and the 90^{th}) of the difference between traditional realized skewness and kurtosis and the same measures corrected for the presence of zeros. The figure fully supports our theoretical predictions. Regarding skewness, the presence of zeros does not induce biases but makes the measure noisier (an implication of Corollary 2.1.1), as shown by the increase in the width of the distribution of the difference as the percentage of zeros increases. Regarding kurtosis, the presence of zeros yields a large positive bias which increases with the percentage of zeros (an implication of Corollary 2.1.2). In essence, Figure 1 is a visual representation of the

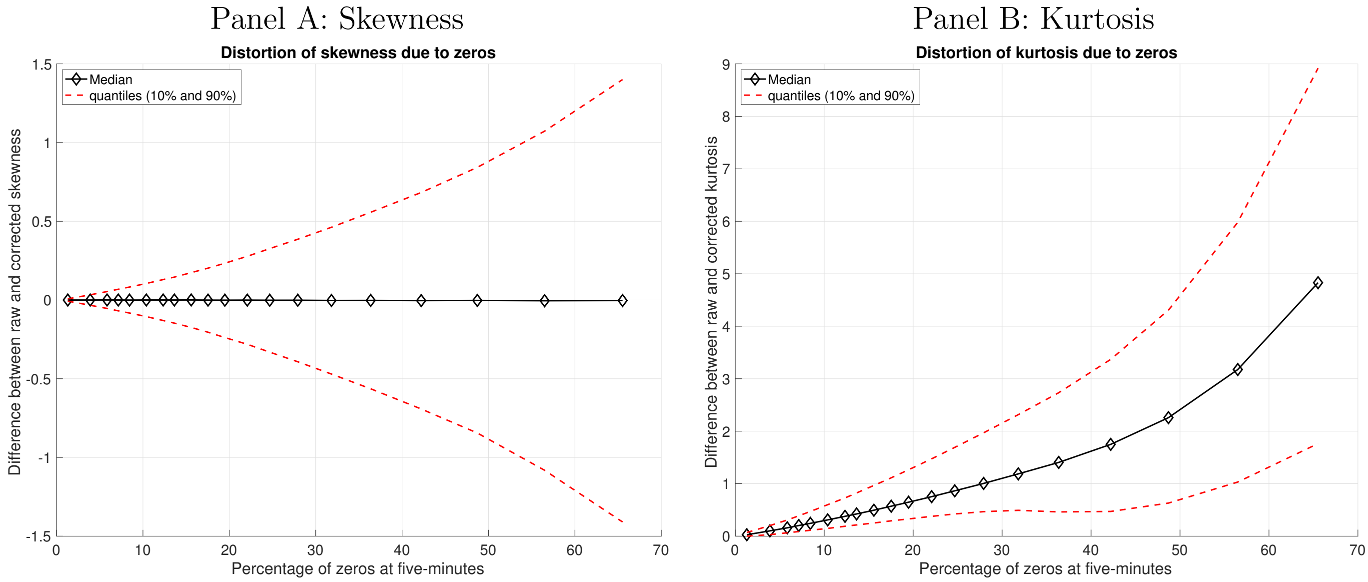

Figure 1: Panel A: median difference between daily raw skewness and daily corrected (for staleness) skewness across the stocks in our sample (a total of 11,892,327 daily observations). We report it along with the 10^{th} and the 90^{th} percentile, as a function of the percentage of zeros during the day. Panel B: median difference between daily raw kurtosis and daily corrected (for staleness) kurtosis across the same stocks, again as a function of the percentage of zeros during the day.

need for bias-correcting realized kurtosis, a measure whose documented correlation with illiquidity proxies is, at least in part, a by-product of its correlation with zeros. It also provides a justification for obtaining more precise skewness estimates by bias-correcting realized skewness.

6.1 Cross-sectional regressions

Next, we turn to asset pricing implications. We assess the relation between future excess returns and past realized skewness and kurtosis by carrying out (over time) cross-sectional regressions, as in Fama and MacBeth (1973). For each time t, we estimate the following model:

$$r_{i,t+1} = \alpha_i + \beta_t' X_{i,t} + \epsilon_{i,t+1}, \tag{52}$$

where $r_{i,t+1}$ denotes the weekly excess return of the i-th stock over week $t, t+1$ and $X_{i,t}$ is a vector of weekly *characteristics* for the i-th stock over week $t-1, t$. The vector includes the weekly (averaged, over five consecutive trading days) measures of realized skewness and kurtosis (S and K), the weekly measures of *truncated* skewness and kurtosis (TS and TK), weekly *bias-corrected* truncated skewness and kurtosis (TS^c and TK^c), past weekly returns (R), the square-root of weekly truncated realized variance ($\sqrt{\mathsf{TRV}}$), the weekly average of jumps (J), the weekly average of positive jumps (J^+), the weekly average of negative jumps (J^-), the weekly average percentage of zeros (RZ), the weakly illiquidity ratio of Amihud (2002) (Ami) and the standardized (by realized variance) difference between positive semi-variance and negative semi-variance (RSJ). Table 3 reports the time series averages of the estimated slopes $\widehat{\beta}_t$s as well as the associated t-statistics.

Model (1) is the same as the main model in Amaya et al. (2015) and confirms that, when employing S, realized skewness is negatively priced. This finding is usually interpreted as being symptomatic of market's preference over positively skewed stocks and, symmetrically, aversion to negative skew. Because positively skewed stocks usually display higher idiosyncratic volatility, the negative price of skewness is sometimes invoked as a justification for the idiosyncratic volatility puzzle of Ang et al. (2006), i.e., the negative dependence between average cross-sectional stock returns and idiosyncratic volatility.[2] Realized kurtosis is, instead, priced with a positive sign representing aversion to volatility uncertainty. The corresponding coefficient is, however, insignificant. In light of the positive correlation between illiquidity and realized kurtosis, the positive pricing of realized kurtosis may be induced by the positive association between illiquidity and expected returns, as justifiable by compensation for illiquidity risk. We will show that at least a portion of this risk compensation is induced by contaminations (i.e., biases) in estimated kurtosis.

Model (2) considers truncated (and, hence, continuous) versions of realized skewness and kurtosis. While statistical significance decreases, thereby pointing to a critical, separate role for jumps in cross-sectional pricing, results are qualitatively similar to what was found in model (1). Because continuous skewness is, in principle, a noisy measure of zero, its pricing ability is surprising. We return to it, and explain it, in model (5).

As emphasized, realized kurtosis is biased in the presence of staleness. It is, therefore, interesting to add the weekly percentage of zeros to truncated skewness and kurtosis. We do so in model (3). Consistent with theory, truncated skewness is hardly impacted. Truncated kurtosis, however, now has a statistically-significant and negative partial effect on expected returns. In other words, the positive partial effect of truncated kurtosis on expected returns has been entirely absorbed by its previous source of bias once that source of bias, i.e., the extent of staleness, is accounted for explicitly. In essence, since staleness provides information about volume dynamics (c.f., see Bandi et al., 2019), the statistical significance of zeros reflects illiquidity pricing in expected returns. The positive impact of uncorrected kurtosis on expected returns may therefore also be the consequence of illiquidity pricing. When controlling for zeros explicitly, kurtosis impacts expected returns negatively. This is reminiscent of the idiosyncratic volatility puzzle of Ang et al. (2006). Not only does higher idiosyncratic volatility lead to lower expected returns, an unexpected result with several recent justifications, but higher *volatility* of idiosyncratic volatility also seems to lead to lower expected returns.

In order to provide direct evidence about the impact of staleness on realized measures, we now bias-correct the truncated measures. In agreement with theory and the logic of model (3), in model (4) we show that bias-corrected truncated kurtosis has, indeed, a statistically-significant, negative

[2] The association between low prices, high idiosyncratic volatility and positive skewness is, for instance, a feature of lottery stocks, see Kumar, 2009.

partial effect on cross-sectional returns. Bias-corrected truncated skewness also continues to have a statistically-significant negative impact on expected returns, a result which was deemed earlier to be surprising in light of the fact that truncated skewness should be a noisy measure of zero carrying no pricing information. We address this issue next.

It is arguably the case that the previous specifications omit natural controls. The first is a measure of the efficient price variance. The second is a measure of trend, something which may have (and will be shown to have) a strong impact on continuous skewness in small sample. The third is a measure of discontinuous variation, such variation being ruled out by truncation.

In model (5), we add R (the total return over the previous week) and $\sqrt{\mathsf{TRV}}$. The loading on TS^c becomes insignificant, thereby supporting the idea that continuous skewness was - in previous specifications - likely proxying for a trend. The trend variable has a strongly significant negative coefficient, representing *reversals*. Truncated realized variance has a statistically-significant negative coefficient in line with the negative pricing of idiosyncratic volatility first put forward by Ang et al. (2006).

In model (6), we add jumps. The measure J is defined as the weekly sum of intra-daily returns greater (in absolute value) than the truncation level. Jumps are strongly significant and with an economically-sound negative sign. The economic interpretation of this negative sign is analogous to that on the negative sign of skewness: market participants are naturally conjectured to fear large negative returns and like large positive returns. These predispositions appear to be reflected in equilibrium expected returns. By the same argument, because the justification for the pricing ability of skewness (as reported, e.g., in Amaya et al., 2015) purely comes from jumps, it appears reasonable to capture discontinuities directly (i.e., using J) rather than indirectly by way of measures (like K) which, as shown, are more likely to be contaminated by continuous variation and staleness. In model (7), and other specifications, we distinguish between positive and negative jumps. Both are generally statistically-significant and consistently with a negative sign.

Our measure of illiquidity (RZ) is added in model (8) (with jumps) and model (9) (with both positive and negative jumps) along with Amihud's measure (Amihud, 2002), Ami. As emphasized, the former has been found in Bandi et al., 2019 to capture intra-daily volume dynamics (volume levels and volume clusters) and is internally *consistent* given the assumed data generating process, the latter is a key benchmark in the literature defined as price impact per unit volume. We find - once more - that the impact of realized zeros is strongly positive and significant, which is revealing of an illiquidity premium. The Amihud measure is, instead, insignificant. When illiquidity is added, the impact of continuous kurtosis disappear. Thus, staleness-corrected continuous kurtosis continues to be correlated with illiquidity.

Finally, we evaluate robustness to the use of realized semi-variance. We employ the RSJ measure of Bollerslev et al. (2019). In model (10), we confirm their main result, namely RSJ is a better proxy for jumps than raw skewness. In model (11), where we add the trend measure, we notice that the significance of RSJ is strongly attenuated. In model (12), which uses our measure of jumps based on truncation, separated into its positive and negative components, the coefficient on RSJ becomes insignificant. While we support the logic in Amaya et al. (2015) and Bollerslev et al. (2019) regarding the importance of jumps in cross-sectional pricing, their same logic justifies using measures that are *less* contaminated by continuous variation (and by staleness, as a result). A simple jump measure constructed using (staleness-adjusted) truncation appears to achieve this goal better than both raw skewness and RSJ.

Our final specification (13) summarizes our main conclusions. Continuous (corrected) skewness and kurtosis are insignificant. So is the RSJ measure of Bollerslev et al. (2019), when controlling for jumps directly, and Amihud's illiquidity measure, when controlling for staleness. Instead, we find significant pricing impacts associated with idiosyncratic volatility, the trend variable and jumps, all with a negative sign, and realized zeros, with a positive sign.

6.2 Single-sorted portfolios

We complement, and support, the previous results by constructing long-short (high-low) portfolios and examining their payoffs. After sorting stocks into deciles based on the level of each characteristic, we go long stocks in the highest characteristic decile and short stocks in the lowest characteristic decile. We report both average returns and average returns in excess of the Fama-French 5-factor model (alphas) over the next 5 days (c.f. Table 4) and over the next 22 days (c.f. Table 5).

Over the shorter 5 day horizon, the only unexpected findings (solely due to lack of conditioning on other driving characteristics) have to do with the statistically-significant, negative average returns associated with portfolios constructed on the basis of negative jumps and Amihud's measure. Both the former characteristic (a more direct measure of adverse price moves than skewness) and the latter characteristic (a measure of illiquidity) should be associated with positive - rather than negative, as in the data - long-short returns. Positive jumps and realized zeros, on the other hand, display statistically-significant, negative and positive (respectively) average returns (and alphas), a result which is consistent with economic logic. Over the longer 22 day horizons, zeros continue to be associated with positive average returns and alphas. All other variables generally lose their pricing ability (with the exception of idiosyncratic volatility which leads - as in much of the recent literature following Ang et al. (2006) - to negative average returns).

In order to account for the impact of alternative characteristics, we now turn to double sorts. We focus exclusively on key variables resulting from the previous theoretical and empirical treatment: jumps and zeros.

6.3 Double-sorted portfolios

We double-sort into 25 quantile portfolios. We focus on realized zeros and aggregate jumps, realized zeros and positive jumps and realized zeros and negative jumps and double sort in both directions, first based on one characteristic and then the other (c.f., Table 6 and Table 7).

In agreement with previous findings, the average returns on portfolios long high zero/low jump stocks (irrespective of whether the jumps are positive, negative or aggregate) and short low zero/high jump stocks are positive. So are their 5-factor Fama-French alphas. Similarly, the average returns and alphas on portfolios long low zero/high jump stocks and short high zero/low jump stocks are negative.

We conclude that the interplay between granular features of the efficient price distribution (i.e., price discontinuities) and features of the trading process affecting the way in which efficient prices are revealed (i.e., zeros) has important pricing implications at high frequencies.

7 Conclusions

This paper evaluates the properties of high-frequency high-order moments under a data generating process accounting for two key stylized features, namely infrequent discontinuities in unobserved equilibrium prices and staleness in observed prices. The latter is a phenomenon known to be linked to trading volumes' first and second moments and, therefore, to the level and variability of liquidity (c.f., Bandi et al., 2019).

We study identification and pricing. In terms of identification, we discuss how the interaction between price discontinuities and prices staleness will, in general, lead to biased and/or noisy high-order moment estimates. A combination of thresholding and corrections for staleness-induced biases is, however, shown to be effective in yielding information about high-order variation, both in its continuous and in its discontinuous notion.

In terms of pricing, we document an interesting interaction between genuine features of the equilibrium price process (jumps) and features of the trading mechanism (staleness). Because, in our framework, jumps and staleness are aspects of the same data generating process for observed prices, not only is our measure of liquidity (zeros) natural, it is model-driven. We, therefore, view the proposed approach as a first step in the analysis of the high-frequency pricing of both granular price features and trading frictions in the context of an econometric model which, cohesively, allows for (and permits identification of) both.

Table 3: Reports Fama-MacBeth cross-sectional regressions of weekly stock returns on stock characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
S	−0.098***												
	(−6.415)												
K	0.004												
	(1.143)												
TS		−0.098***	−0.096***							−0.010	0.009		
		(−5.043)	(−4.933)							(−0.710)	(0.639)		
TK		0.004	−0.017**							0.002	0.003		
		(0.668)	(−2.131)							(0.356)	(0.409)		
TS^c				−0.127***	0.017	0.029*	0.031*	0.028*	0.027*				0.033**
				(−5.344)	(1.019)	(1.679)	(1.836)	(1.696)	(1.660)				(1.991)
TK^c				−0.014	−0.007	−0.007	−0.012	−0.007	−0.007				−0.008
				(−1.570)	(−0.931)	(−0.960)	(−1.575)	(−1.033)	(−1.067)				(−1.168)
R					−2.102***	−2.009***	−1.997***	−1.904***	−1.905***		−1.110***	−1.392***	−1.833***
					(−7.277)	(−6.836)	(−6.830)	(−6.700)	(−6.718)		(−3.052)	(−3.909)	(−5.991)
$\sqrt{TRV}$					−0.073***	−0.073***	−0.083***	−0.075***	−0.075***				−0.076***
					(−3.763)	(−3.730)	(−3.975)	(−3.432)	(−3.456)				(−3.482)
J						−0.026**		−0.037***					
						(−2.575)		(−3.682)					
J^+							−0.016		−0.040***			−0.036***	−0.039***
							(−1.434)		(−3.556)			(−2.866)	(−3.347)
J^-							−0.052***		−0.032***			−0.028**	−0.031***
							(−4.240)		(−2.770)			(−2.290)	(−2.629)
RZ			0.384***					0.386***	0.412***				0.409***
			(2.978)					(3.738)	(3.715)				(3.695)
AMI								−0.291	−0.267				−0.265
								(−1.045)	(−0.859)				(−0.856)
RSJ										−0.745***	−0.333***	−0.116	−0.059
										(−6.695)	(−3.367)	(−1.240)	(−0.700)

Table 4: Average returns on long-short (high-low) portfolios over 5 trading days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	$high - low$	$high - low\ \alpha$
S	0.206**	0.178*	0.132	0.071	0.066	0.044	0.012	0.005	−0.016	−0.035	−0.241***	−0.245***
	(2.007)	(1.704)	(1.262)	(0.705)	(0.646)	(0.439)	(0.125)	(0.056)	(−0.162)	(−0.369)	(−6.019)	(−6.205)
K	0.055	0.026	0.083	0.064	0.060	0.073	0.060	0.095	0.070	0.079	0.024	−0.002
	(0.537)	(0.257)	(0.843)	(0.641)	(0.599)	(0.722)	(0.593)	(0.942)	(0.702)	(0.816)	(0.496)	(−0.044)
TS	0.200*	0.133	0.129	0.086	0.069	0.050	0.015	0.026	−0.022	−0.023	−0.224***	−0.226***
	(1.929)	(1.260)	(1.247)	(0.847)	(0.681)	(0.503)	(0.154)	(0.272)	(−0.228)	(−0.250)	(−5.681)	(−5.884)
TK	0.054	0.032	0.070	0.092	0.068	0.057	0.060	0.086	0.082	0.063	0.009	0.001
	(0.532)	(0.313)	(0.712)	(0.935)	(0.684)	(0.572)	(0.589)	(0.854)	(0.806)	(0.647)	(0.189)	(0.027)
TS^c	0.187*	0.158	0.123	0.095	0.051	0.054	0.015	0.007	0.003	−0.029	−0.216***	−0.223***
	(1.785)	(1.499)	(1.189)	(0.935)	(0.507)	(0.548)	(0.156)	(0.072)	(0.034)	(−0.313)	(−5.202)	(−5.654)
TK^c	0.121	0.054	0.078	0.056	0.062	0.090	0.047	0.066	0.048	0.043	−0.077*	−0.073*
	(1.209)	(0.552)	(0.792)	(0.567)	(0.621)	(0.900)	(0.464)	(0.652)	(0.470)	(0.430)	(−1.816)	(−1.808)
R	−0.054	0.101	0.154	0.144	0.112	0.088	0.093	0.048	0.013	−0.034	0.020	−0.006
	(−0.388)	(0.908)	(1.553)	(1.579)	(1.281)	(1.020)	(1.066)	(0.521)	(0.134)	(−0.282)	(0.284)	(−0.090)
$\sqrt{\mathsf{TRV}}$	0.128**	0.114*	0.130*	0.135	0.101	0.106	0.085	−0.013	−0.015	−0.107	−0.235	−0.241***
	(2.220)	(1.655)	(1.697)	(1.609)	(1.116)	(1.058)	(0.767)	(−0.100)	(−0.103)	(−0.625)	(−1.640)	(−3.155)
J	0.159	0.072	0.083	0.107	0.072	0.070	0.072	0.025	0.027	−0.057	−0.216***	−0.224***
	(1.278)	(0.643)	(0.834)	(1.170)	(0.810)	(0.782)	(0.822)	(0.263)	(0.246)	(−0.477)	(−4.849)	(−5.097)
J^+	0.068	0.091	0.099	0.076	0.066	0.069	0.036	0.047	0.040	−0.046	−0.113**	−0.096**
	(0.719)	(1.072)	(1.066)	(0.774)	(0.649)	(0.642)	(0.330)	(0.418)	(0.347)	(−0.375)	(−1.983)	(−2.113)
J^-	0.089	0.064	0.057	0.079	0.037	0.067	0.080	0.097	0.116	0.037	−0.051	−0.071
	(0.713)	(0.538)	(0.511)	(0.713)	(0.341)	(0.643)	(0.814)	(1.050)	(1.355)	(0.397)	(−0.836)	(−1.411)
RZ	−0.045	0.002	0.068	0.062	0.098	0.074	0.084	0.095	0.155	0.147*	0.192**	0.144***
	(−0.377)	(0.022)	(0.705)	(0.618)	(0.953)	(0.679)	(0.782)	(0.926)	(1.575)	(1.661)	(2.439)	(2.586)
AMI	0.161*	0.107	0.099	0.090	0.029	0.057	0.043	0.033	0.107	−0.061	−0.222***	−0.222***
	(1.905)	(1.180)	(1.041)	(0.925)	(0.284)	(0.554)	(0.404)	(0.309)	(0.969)	(−0.532)	(−3.526)	(−5.518)
RSJ	0.252**	0.206*	0.113	0.089	0.060	0.034	0.035	−0.023	−0.048	−0.057	−0.309***	−0.311***
	(2.322)	(1.941)	(1.083)	(0.867)	(0.600)	(0.351)	(0.356)	(−0.234)	(−0.513)	(−0.611)	(−5.911)	(−6.216)

Table 5: Average returns on long-short (high-low) portfolios over 22 trading days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	$high - low$	$high - low\ \alpha$
S	0.431	0.174	0.220	0.220	0.264	0.221	0.290	0.202	0.288	0.399	−0.032	0.123
	(0.906)	(0.357)	(0.476)	(0.475)	(0.592)	(0.501)	(0.674)	(0.474)	(0.673)	(0.958)	(−0.208)	(0.842)
K	0.168	0.089	0.186	0.236	0.157	0.334	0.305	0.332	0.282	0.617	0.450**	0.372**
	(0.397)	(0.210)	(0.431)	(0.522)	(0.348)	(0.713)	(0.642)	(0.714)	(0.610)	(1.388)	(2.071)	(2.311)
TS	0.310	0.138	0.179	0.281	0.258	0.319	0.238	0.308	0.296	0.380	0.070	0.217
	(0.631)	(0.280)	(0.376)	(0.624)	(0.587)	(0.734)	(0.550)	(0.732)	(0.702)	(0.926)	(0.428)	(1.354)
TK	0.119	0.114	0.225	0.161	0.235	0.334	0.307	0.288	0.406	0.518	0.400*	0.357**
	(0.288)	(0.268)	(0.520)	(0.362)	(0.527)	(0.724)	(0.654)	(0.595)	(0.849)	(1.162)	(1.746)	(2.126)
TS^c	0.237	0.116	0.278	0.365	0.335	0.236	0.157	0.331	0.247	0.406	0.169	0.303*
	(0.476)	(0.237)	(0.598)	(0.808)	(0.757)	(0.536)	(0.358)	(0.784)	(0.599)	(0.991)	(0.939)	(1.756)
TK^c	0.424	0.312	0.282	0.191	0.185	0.291	0.361	0.247	0.226	0.189	−0.235	−0.270
	(1.002)	(0.745)	(0.661)	(0.433)	(0.414)	(0.627)	(0.794)	(0.528)	(0.470)	(0.398)	(−1.057)	(−1.465)
R	−0.515	0.363	0.373	0.486	0.304	0.457	0.394	0.340	0.217	0.289	0.803**	0.754**
	(−0.740)	(0.712)	(0.851)	(1.198)	(0.782)	(1.179)	(1.030)	(0.858)	(0.507)	(0.529)	(2.068)	(2.050)
$\sqrt{\mathsf{TRV}}$	0.633**	0.551*	0.566*	0.530	0.509	0.519	0.260	0.013	−0.347	−0.525	−1.157*	−1.355***
	(2.418)	(1.796)	(1.645)	(1.424)	(1.284)	(1.201)	(0.521)	(0.022)	(−0.510)	(−0.651)	(−1.703)	(−3.851)
J	0.228	0.285	0.307	0.370	0.353	0.340	0.430	0.228	0.018	0.131	−0.097	−0.049
	(0.406)	(0.580)	(0.702)	(0.940)	(0.930)	(0.898)	(1.081)	(0.517)	(0.036)	(0.246)	(−0.725)	(−0.362)
J^+	0.391	0.299	0.430	0.248	0.201	0.116	0.149	0.111	0.226	0.411	0.020	0.122
	(1.039)	(0.795)	(1.113)	(0.608)	(0.460)	(0.247)	(0.294)	(0.209)	(0.427)	(0.723)	(0.062)	(0.596)
J^-	0.442	0.294	0.205	0.140	0.088	0.244	0.359	0.284	0.269	0.331	−0.112	−0.170
	(0.766)	(0.541)	(0.402)	(0.276)	(0.189)	(0.564)	(0.879)	(0.712)	(0.716)	(0.888)	(−0.356)	(−0.850)
RZ	−0.184	−0.087	0.164	0.206	0.219	0.355	0.536	0.256	0.456	0.811*	0.995***	0.723***
	(−0.335)	(−0.206)	(0.393)	(0.482)	(0.508)	(0.794)	(1.122)	(0.509)	(0.959)	(1.832)	(2.627)	(3.164)
AMI	0.625*	0.491	0.395	0.289	0.274	0.135	0.246	0.175	0.313	−0.236	−0.860***	−0.849***
	(1.719)	(1.248)	(0.956)	(0.680)	(0.611)	(0.289)	(0.518)	(0.352)	(0.622)	(−0.439)	(−2.827)	(−4.887)
RSJ	0.213	0.180	0.161	0.239	0.342	0.191	0.478	0.351	0.246	0.307	0.093	0.337*
	(0.413)	(0.364)	(0.331)	(0.514)	(0.767)	(0.440)	(1.123)	(0.839)	(0.605)	(0.781)	(0.438)	(1.749)

Table 6: Table reports the average 1-week ahead returns sorted on double-sorted portfolios. For each panel, entitled as A → B, all stocks in the sample are first sorted into 5 quintiles on the basis of the first variable A (columns). Within each quintile, the stocks are then sorted into 5 quintiles according to the second variable B (rows).

	RZ→J						J→RZ					
(1)	−0.015	0.042	0.178	0.199	0.213**	0.123	−0.038	0.054	−0.019	−0.041	−0.183	−0.045
	(−0.112)	(0.343)	(1.377)	(1.572)	(1.998)	(1.049)	(−0.268)	(0.494)	(−0.167)	(−0.390)	(−1.321)	(−0.386)
(2)	0.062	0.128	0.110	0.123	0.203**	0.125	0.131	0.102	0.063	0.005	−0.056	0.049
	(0.612)	(1.361)	(1.093)	(1.197)	(2.197)	(1.339)	(0.981)	(1.070)	(0.698)	(0.054)	(−0.446)	(0.471)
(3)	−0.011	0.091	0.074	0.065	0.115	0.067	0.152	0.094	0.074	0.097	0.005	0.084
	(−0.109)	(1.001)	(0.775)	(0.663)	(1.294)	(0.740)	(1.204)	(0.956)	(0.820)	(1.020)	(0.043)	(0.827)
(4)	−0.005	0.072	0.107	0.094	0.143	0.082	0.170	0.093	0.109	0.088	0.075	0.107
	(−0.049)	(0.788)	(1.092)	(0.958)	(1.517)	(0.901)	(1.517)	(0.886)	(1.168)	(0.883)	(0.657)	(1.064)
(5)	−0.216	−0.067	−0.052	0.011	0.081	−0.048	0.201**	0.141	0.078	0.094	0.098	0.123
	(−1.627)	(−0.547)	(−0.419)	(0.092)	(0.757)	(−0.425)	(2.089)	(1.506)	(0.824)	(1.029)	(1.016)	(1.352)
high − low	−0.201***	−0.108**	−0.230***	−0.188***	−0.133**	−0.172***	0.239***	0.086	0.096	0.135**	0.281***	0.168***
	(−3.548)	(−2.236)	(−4.384)	(−3.014)	(−2.313)	(−5.392)	(2.796)	(1.234)	(1.401)	(1.973)	(3.099)	(2.597)
high − low α	−0.190***	−0.118**	−0.226***	−0.179***	−0.141**	−0.171***	0.219***	0.039	0.038	0.082	0.248***	0.125***
	(−3.376)	(−2.466)	(−4.213)	(−2.905)	(−2.446)	(−5.418)	(3.404)	(0.739)	(0.750)	(1.514)	(3.578)	(2.931)
	RZ→J^{+}						J^{+}→RZ					
(1)	0.015	0.098	0.115	0.121	0.199**	0.110	−0.016	−0.022	−0.063	−0.115	−0.350**	−0.113
	(0.146)	(1.064)	(1.192)	(1.211)	(2.255)	(1.197)	(−0.142)	(−0.193)	(−0.486)	(−0.805)	(−2.132)	(−0.898)
(2)	0.019	0.140	0.140	0.120	0.167*	0.117	0.064	0.045	0.037	−0.028	−0.013	0.021
	(0.190)	(1.584)	(1.415)	(1.228)	(1.926)	(1.317)	(0.693)	(0.453)	(0.331)	(−0.215)	(−0.093)	(0.192)
(3)	−0.023	0.085	0.123	0.108	0.182*	0.095	0.098	0.150	0.100	0.104	0.154	0.121
	(−0.208)	(0.835)	(1.164)	(1.033)	(1.891)	(0.973)	(1.093)	(1.543)	(0.930)	(0.911)	(1.280)	(1.202)
(4)	−0.067	0.010	0.099	0.123	0.175*	0.068	0.077	0.143	0.131	0.126	0.107	0.117
	(−0.532)	(0.091)	(0.832)	(1.108)	(1.668)	(0.633)	(0.814)	(1.428)	(1.242)	(1.205)	(1.027)	(1.203)
(5)	−0.256*	−0.123	−0.104	0.026	−0.003	−0.092	0.142	0.108	0.142	0.158*	0.102	0.130
	(−1.697)	(−0.853)	(−0.724)	(0.191)	(−0.024)	(−0.707)	(1.436)	(1.161)	(1.523)	(1.792)	(1.022)	(1.444)
high − low	−0.272***	−0.221***	−0.219***	−0.095	−0.202***	−0.202***	0.158**	0.130*	0.205**	0.273***	0.452***	0.244***
	(−3.695)	(−2.662)	(−2.657)	(−1.253)	(−2.987)	(−3.378)	(2.326)	(1.687)	(2.325)	(2.897)	(4.292)	(3.424)
high − low α	−0.237***	−0.186***	−0.200***	−0.077	−0.192***	−0.178***	0.094*	0.057	0.146**	0.237***	0.463***	0.199***
	(−4.067)	(−3.454)	(−3.537)	(−1.295)	(−3.095)	(−5.251)	(1.911)	(0.970)	(2.327)	(3.493)	(5.778)	(4.462)
	RZ→J^{-}						J^{-}→RZ					
(1)	−0.089	−0.037	0.045	0.190	0.150	0.052	−0.065	−0.006	−0.052	−0.020	−0.039	−0.036
	(−0.585)	(−0.258)	(0.303)	(1.367)	(1.251)	(0.390)	(−0.388)	(−0.043)	(−0.395)	(−0.169)	(−0.346)	(−0.282)
(2)	−0.026	0.069	0.118	0.067	0.100	0.065	0.097	0.065	0.074	0.130	0.024	0.078
	(−0.199)	(0.603)	(1.007)	(0.595)	(0.963)	(0.599)	(0.699)	(0.505)	(0.647)	(1.318)	(0.263)	(0.718)
(3)	−0.004	0.058	0.054	0.054	0.127	0.058	0.162	0.041	0.064	0.089	0.074	0.086
	(−0.032)	(0.577)	(0.488)	(0.515)	(1.326)	(0.590)	(1.314)	(0.360)	(0.569)	(0.897)	(0.821)	(0.838)
(4)	0.034	0.099	0.112	0.059	0.205**	0.102	0.119	0.058	0.088	0.092	0.085	0.088
	(0.343)	(1.080)	(1.147)	(0.594)	(2.297)	(1.133)	(1.133)	(0.556)	(0.852)	(0.940)	(0.908)	(0.917)
(5)	−0.027	0.076	0.082	0.106	0.159*	0.079	0.143	0.178**	0.088	0.140	0.113	0.132
	(−0.265)	(0.828)	(0.869)	(1.090)	(1.816)	(0.875)	(1.427)	(2.062)	(0.911)	(1.507)	(1.162)	(1.473)
high − low	0.062	0.114	0.037	−0.084	0.009	0.028	0.207*	0.184*	0.140	0.160**	0.152**	0.169**
	(0.822)	(1.329)	(0.434)	(−1.088)	(0.132)	(0.434)	(1.955)	(1.892)	(1.628)	(2.025)	(2.157)	(2.363)
high − low α	0.048	0.078	0.015	−0.095	−0.009	0.007	0.225***	0.165**	0.080	0.081	0.086*	0.127***
	(0.793)	(1.341)	(0.248)	(−1.550)	(−0.146)	(0.190)	(2.731)	(2.408)	(1.283)	(1.442)	(1.709)	(2.937)

Table 7: Performance of *high* − *low* double-sorted portfolios.

	(1)	(2)	(3)	(4)	(5)	(6)
mean	0.429***	−0.384***	0.456***	−0.492***	0.177**	−0.182**
	(4.796)	(−4.364)	(4.078)	(−4.801)	(2.304)	(−2.341)
α	0.383***	−0.358***	0.402***	−0.465***	0.175**	−0.163**
	(5.018)	(−5.258)	(4.990)	(−6.352)	(2.452)	(−2.486)
conditional mean	0.101**		0.006		−0.050	
	(2.425)		(0.096)		(−0.941)	

Note. (1) – long high RZ and low J stocks and short low RZ and high J stocks. (2) – long low RZ and high J stocks and short high RZ and low J stocks. (3) – long high RZ and low J^+ stocks and short low RZ and high J^+ stocks. (4) – long low RZ and high J^+ stocks and short high RZ and low J^+ stocks. (5) – long high RZ and low J^- stocks and short low RZ and high J^- stocks. (6) – long low RZ and high J^- stocks and short high RZ and low J^- stocks. The last row reports the constant after regressing the returns on portfolios (1), (3) and (5) on each other.

References

Aït-Sahalia, Y. and J. Jacod (2015). *High Frequency Financial Econometrics.* Princeton University Press.

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.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *The Journal of Finance 61*(1), 259–299.

Bandi, F., A. Kolokolov, D. Pirino, and R. Renò (2019). Zeros. *Management Science*. Forthcoming.

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

Barndorff-Nielsen, O. E., E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*, pp. 33–68.

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., N. Shephard, and M. Winkel (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and Their Applications 116*, 798–806.

Bibinger, M. and M. Vetter (2015). Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps. *Ann. Inst. Statist. Math. 67*(4), 707–743.

Bollerslev, T., S. Z. Li, and B. Zhao (2019). Good volatility, bad volatility, and the cross section of stock returns. *Journal of Financial and Quantitative Analysis*.

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.

Jacod, J. (2008). Asymptotic properties of realized power variations and associated functionals of semimartingales. *Stochastic processes and their applications 118*, 517–559.

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

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes.* Springer.

Kinnebrock, S. and M. Podolskij (2008). A note on the central limit theorem for bipower variation of general functions. *Stochastic Processes and their Applications 118*(6), 1056 – 1070.

Koike, Y. and Z. Liu (2019). Asymptotic properties of the realized skewness and related statistics. *Annals of the Institute of Statistical Mathematics 71*(4), 703–741.

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance 64*(4), 1889–1933.

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

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

Veraart, A. (2010). Inference for the jump part of quadratic variation of Ito semimartingales. *Econometric Theory 26*(2), 331–368.

A Appendix: Preliminary Proofs

Proof of Theorem 2.1. In order to shorten the proofs, and to simplify notation, we present the proofs under the additional assumption that there are no jumps in volatility. That is, we assume the volatility process σ_t satisfies the equation

$$\sigma_t = \sigma_0 + \int_0^t a'_s \, ds + \int_0^t \sigma'_s dW_s + \int_0^t v'_s dV_s, \tag{53}$$

where W_t and V_t are independent Brownian motions while a', σ' and v' are adapted càdlàg bounded processes.

Let $t^*_{1,n}, \dots, t^*_{N^n_t,n}$ be the partition of the interval $[0, t]$, constructed from the original partition by removing the points for which $\mathbb{B}_{i,n} = 1$. Thus, the power variation can be expressed as

$$\mathsf{PV}(f, \widetilde{X}) = n^{r/2-1} \sum_{i=1}^{N^n_t} f(\Delta^*_i X), \tag{54}$$

where $\Delta^*_i X = X_{t^*_{i,n}} - X_{t^*_{i-1,n}}$ are the increments of the process X_t over the new partition. Let $\beta^n_i = \sigma_{t^*_{i-1,n}} \Delta^*_i W$ be an approximation of the increments $\Delta^*_i X$ and set $\Delta(i, n) = t^*_{i,n} - t^*_{i-1,n}$.

We start with preliminary computations of the conditional moments of $f(\beta^n_i)$, $\Delta(i, n)$ and related quantities, which are used in the subsequent proofs. First, notice that, for $k = 1, 2, \dots,$

$$\mathbb{P}\left\{\Delta^*_i W = W_{t^*_{i-1,n}+k\,\Delta_n} - W_{t^*_{i-1,n}}\right\} = (1 - p^{\varnothing})(p^{\varnothing})^{k-1}. \tag{55}$$

By the law of iterated expectations, this implies, that

$$\mathbb{E}\left[f(\Delta^*_i W) \,\Big|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta^{r/2} \rho(f) \sum_{k=1}^{\infty} k^{r/2}(1 - p^{\varnothing}) p^{k-1}. \tag{56}$$

Hence, we have

$$\begin{aligned}
\mathbb{E}\left[f(\beta^n_i) \,\Big|\, \mathcal{F}_{t^*_{i-1,n}}\right] &= \sigma^r_{t^*_{i-1,n}} \mathbb{E}\left[f(\Delta^*_i W) \,\Big|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \sigma^r_{t^*_{i-1,n}} \Delta^{r/2}_n \rho(f) \sum_{k=1}^{\infty} k^{r/2}(1 - p^{\varnothing}) \left(p^{\varnothing}\right)^{k-1} \\
&= \Delta^{r/2}_n \sigma^r_{t^*_{i-1,n}} \rho(f) \frac{1 - p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right).
\end{aligned} \tag{57}$$

Similarly, we obtain

$$\mathbb{E}\left[(f(\beta^n_i))^2 \,\Big|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta^r_n \sigma^{2r}_{t^*_{i-1,n}} \rho(f^2) \frac{1 - p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r}\left(p^{\varnothing}\right), \tag{58}$$

$$\mathbb{E}\left[\Delta(i, n) \,\Big|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta_n \frac{1 - p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-1}\left(p^{\varnothing}\right), \tag{59}$$

and more generally, for $r > 0$,

$$\mathbb{E}\left[(\Delta(i, n))^r \,\Big|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta^r_n \frac{1 - p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r}\left(p^{\varnothing}\right). \tag{60}$$

Now, notice that, for any $p^{\varnothing} \in [0,1)$,

$$(1-p^{\varnothing})\mathbb{E}\left[\Delta(i,n) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta_n \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}} \mathsf{Li}_{-1}\left(p^{\varnothing}\right) = \Delta_n. \tag{61}$$

Finally, observe that

$$\Delta(i,n) f(\beta_i^n) \stackrel{d}{=} \Delta(i,n)^{1+r/2} \sigma^r_{t^*_{i-1,n}} f(U), \tag{62}$$

where $U \sim \mathcal{N}(0,1)$ and $\stackrel{d}{=}$ denotes equality in law. Therefore, we have

$$\mathbb{E}\left[\Delta(i,n) f(\beta_i^n) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta_n^{1+r/2} \sigma^r_{t^*_{i-1,n}} \rho(f) \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r/2-1}\left(p^{\varnothing}\right). \tag{63}$$

Next, consider the four quantities defined as

$$\begin{aligned}
A_1 &= \sqrt{n} \sum_{i=1}^{N_t^n} \left(n^{\frac{r}{2}-1} f(\beta_i^n) - n^{\frac{r}{2}-1} \mathbb{E}\left[f(\Delta_i^* X) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right), \\
A_2 &= \sqrt{n} \sum_{i=1}^{N_t^n} n^{\frac{r}{2}-1} \left(\left(f(\Delta_i^* X) - \mathbb{E}\left[f(\Delta_i^* X) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right) - \left(f(\beta_i^n) - \mathbb{E}\left[f(\beta_i^n) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right)\right), \\
A_3 &= \sqrt{n} \sum_{i=1}^{N_t^n} n^{\frac{r}{2}-1} \left(\mathbb{E}\left[f(\Delta_i^* X) - f(\beta_i^n) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right), \\
A_4 &= \sqrt{n} \sum_{i=1}^{N_t^n} \left(n^{\frac{r}{2}-1} \mathbb{E}\left[f(\Delta_i^* X) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \int_0^t |\sigma_s|^r \, ds \right).
\end{aligned} \tag{64}$$

The proof of the main result follows from three statements:

$$\sqrt{n} \sum_{i=1}^{N_t^n} \left(n^{\frac{r}{2}-1} f(\beta_i^n) - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \int_0^t |\sigma_s|^r \, ds \right) \xrightarrow{stably} \int_0^t \alpha_s(2)\, dW_s + \int_0^t \alpha_s(3)\, dW'_s, \tag{65}$$

$$\sqrt{n} \sum_{i=1}^{N_t^n} n^{\frac{r}{2}-1} \left(\left(f(\Delta_i^* X) - \mathbb{E}\left[f(\Delta_i^* X) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right) - \left(f(\beta_i^n) - \mathbb{E}\left[f(\beta_i^n) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right)\right) \xrightarrow{P} 0, \tag{66}$$

$$\sqrt{n} \sum_{i=1}^{N_t^n} n^{\frac{r}{2}-1} \left(\mathbb{E}\left[f(\Delta_i^* X) - f(\beta_i^n) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]\right) \xrightarrow{u.c.p.} \int_0^t \alpha_s(1)\, ds, \tag{67}$$

which are proven in three sequential steps.

Step 1. We start by proving the stable convergence of Eq. (65). Express the right-hand side of Eq. (65) as follows:

$$\sqrt{n} \sum_{i=1}^{N_t^n} \left(n^{\frac{r}{2}-1} f(\beta_i^n) - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \int_0^t |\sigma_s|^r \, ds \right) = A_1 + A_2, \tag{68}$$

where

$$A_1 = \sqrt{n} \sum_{i=1}^{N_t^n} \left(n^{\frac{r}{2}-1} f(\beta_i^n) - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \sigma^r_{t^*_{i-1,n}} \Delta(i,n) \right), \tag{69}$$

$$A_2 = \sqrt{n}\left(\sum_{i=1}^{N_t^n} \sigma^r_{t^*_{i-1,n}} \Delta(i,n) - \int_0^t |\sigma_s|^r \, ds\right) \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f). \tag{70}$$

By the same arguments as in Barndorff-Nielsen et al. (2006), $A_2 \xrightarrow{u.c.p.} 0$. Therefore, it is sufficient to prove that

$$A_1 \xrightarrow{stably} \int_0^t \alpha_s(2)\, dW_s + \int_0^t \alpha_s(3)\, dW'_s. \tag{71}$$

Express A_1 as the sum of $\mathcal{F}_{t^*_{i,n}}$-measurable random variables:

$$A_1 = \sum_{i=1}^{N_t^n} \zeta_i^n, \tag{72}$$

where

$$\zeta_i^n = \sqrt{n}\left(n^{\frac{r}{2}-1} f(\beta_i^n) - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \sigma^r_{t^*_{i-1,n}} \Delta(i,n)\right). \tag{73}$$

As pointed out in Hayashi et al. (2011), since we are proving only $\mathcal{F}^\star$-stable convergence, it is sufficient to verify the conditions of Theorem IX.7.28 of Jacod and Shiryaev (2003) with respect to the filtration $\left(\mathcal{F}_{t^*_{i,n}}\right)$. Using Eq. (57) and Eq. (59) we obtain:

$$\begin{aligned}
\mathbb{E}\left[\zeta_i^n \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] &= \sqrt{n}\left(\Delta_n \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r/2}\left(p^{\varnothing}\right) \rho(f) \sigma^r_{t^*_{i-1,n}} - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) \rho(f) \sigma^r_{t^*_{i-1,n}} \Delta_n \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-1}\left(p^{\varnothing}\right)\right) \\
&= \sqrt{n}\Delta_n \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r/2}\left(p^{\varnothing}\right) \rho(f) \sigma^r_{t^*_{i-1,n}} \left(1 - \frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-1}\left(p^{\varnothing}\right)\right) = 0,
\end{aligned} \tag{74}$$

where the last equality follows from Eq. (61). Next, using Eq. (58), Eq. (60) and Eq. (63), we obtain

$$\begin{aligned}
\mathbb{E}\left[(\zeta_i^n)^2 \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] &= \rho(f^2) \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r}\left(p^{\varnothing}\right) \sigma^{2r}_{t^*_{i-1,n}} \Delta_n \\
&+ \left(\frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-r/2}\left(p^{\varnothing}\right) \rho(f)\right)^2 \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-2}\left(p^{\varnothing}\right) \sigma^{2r}_{t^*_{i-1,n}} \Delta_n \\
&- 2\frac{(1-p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-r/2}\left(p^{\varnothing}\right) \rho(f) \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r/2-1}\left(p^{\varnothing}\right) \rho(f) \sigma^{2r}_{t^*_{i-1,n}} \Delta_n.
\end{aligned} \tag{75}$$

Hence,

$$\sum_{i=1}^{N_t^n} \mathbb{E}\left[(\zeta_i^n)^2 \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \tilde{\Upsilon} \sum_{i=1}^{N_t^n} \sigma^{2r}_{t^*_{i-1,n}} \Delta_n, \tag{76}$$

where $\tilde{\Upsilon}$ is a constant of the form

$$\tilde{\Upsilon} = \rho(f^2) \frac{1-p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-r}\left(p^{\varnothing}\right) - (\rho(f))^2 \left(2\frac{(1-p^{\varnothing})^3}{(p^{\varnothing})^2} \mathsf{Li}_{-r/2}\left(p^{\varnothing}\right) \mathsf{Li}_{-r/2-1}\left(p^{\varnothing}\right) - \frac{(1-p^{\varnothing})^5}{(p^{\varnothing})^3} \left(\mathsf{Li}_{-r/2}\left(p^{\varnothing}\right)\right)^2 \mathsf{Li}_{-2}\left(p^{\varnothing}\right)\right). \tag{77}$$

Now, consider the sum $\sum\limits_{i=1}^{N_t^n} \sigma^{2r}_{t^*_{i-1,n}} \Delta_n$. We have

$$\sum_{i=1}^{N_t^n} \sigma^{2r}_{t^*_{i-1,n}} \Delta_n = \sum_{i=1}^{n} \sigma^{2r}_{t_{i,n}-k(n,i)\Delta_n} \left(1-\mathbb{B}_{i,n}\right) \Delta_n \xrightarrow{P} (1-p^{\varnothing}) \int_0^t \sigma_s^{2r}\, ds, \tag{78}$$

where $k(n,i)$ is the number of consecutive events $\{\mathbb{B}_{i-k,n}=1\}$ prior to the instant $t_{i,n}$ and the convergence follows from the fact that

$$\sum_{i=1}^{n} \mathbb{E}\left[\sigma^{2r}_{t_{i,n}-k(n,i)\Delta_n} \left(1-\mathbb{B}_{i,n}\right) \Delta_n \;\middle|\; \mathcal{F}_{t_{i-1,n}}\right] = \sum_{i=1}^{n} \sigma^{2r}_{t_{i,n}-k(n,i)\Delta_n} (1-p^{\varnothing})\Delta_n \longrightarrow (1-p^{\varnothing}) \int_0^t \sigma_s^{2r}\, ds. \tag{79}$$

Consequently,

$$\sum_{i=1}^{N_t^n} \mathbb{E}\left[(\zeta_i^n)^2 \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right] \xrightarrow{P} \Upsilon \int_0^t \sigma_s^{2r}\, ds, \tag{80}$$

where Υ is a constant of the form

$$\Upsilon = \rho(f^2)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-r}\left(p^{\varnothing}\right) - (\rho(f))^2\left(2\frac{(1-p^{\varnothing})^4}{(p^{\varnothing})^2}\mathsf{Li}_{-r/2}\left(p^{\varnothing}\right)\mathsf{Li}_{-r/2-1}\left(p^{\varnothing}\right) - \frac{(1-p^{\varnothing})^6}{(p^{\varnothing})^3}\left(\mathsf{Li}_{-r/2}\left(p^{\varnothing}\right)\right)^2\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\right). \tag{81}$$

Next, since $\mathbb{E}\left[\Delta(n,i)\, \Delta_i^* W \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right] = 0$, using Eq. (55) we obtain

$$\mathbb{E}\left[\zeta_i^n\, \Delta_i^* W \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right] = \mathbb{E}\left[n^{\frac{r-1}{2}} f(\beta_i^n)\Delta_i^* W \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right] = \frac{1-p^{\varnothing}}{p^{\varnothing}}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right)\rho(f,1)\, \sigma^r_{t^*_{i-1,n}} \Delta_n, \tag{82}$$

which implies (analogously to the proof of Eq. (80)) that

$$\sum_{i=1}^{N_t^n} \mathbb{E}\left[\zeta_i^n\, \Delta_i^* W \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right] \xrightarrow{P} \rho(f,1)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right) \int_0^t \sigma_s^r\, ds. \tag{83}$$

Finally, for any bounded martingale N, which is orthogonal to W and defined on the same probability space, we immediately deduce that:

$$\mathbb{E}\left[\zeta_i^n\, \Delta_i^* N \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right] = 0. \tag{84}$$

Now, the assertion of *Step 1* follows from Eqs. (74), (80), (83), (84) and Theorem IX.7.28 in Jacod and Shiryaev (2003).

Step 2. The proof of the convergence in Eq. (66) is the same as that in Barndorff-Nielsen et al. (2006), hence it is omitted.

Step 3. Now, we prove the statement in Eq. (67). Write

$$\sqrt{n}\sum_{i=1}^{N_t^n} n^{\frac{r}{2}-1}\left(\mathbb{E}\left[f(\Delta_i^* X) - f(\beta_i^n) \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right]\right) = \sum_{i=1}^{N_t^n} \xi_i^n, \tag{85}$$

where

$$\xi_i^n = n^{\frac{r-1}{2}}\mathbb{E}\left[f(\Delta_i^* X) - f(\beta_i^n) \;\middle|\; \mathcal{F}_{t^*_{i-1,n}}\right]. \tag{86}$$

Notice that $\Delta_i^* X - \beta_i^n = \zeta_i^n(1) + \zeta_i^n(2)$, where

$$\zeta_i^n(1) = \int_{t^*_{i-1,n}}^{t^*_{i,n}} \left(a_u - a_{t^*_{i-1,n}}\right) du + \int_{t^*_{i-1,n}}^{t^*_{i,n}} \left[\int_{t^*_{i-1,n}}^{u} a'_s \, ds + \int_{t^*_{i-1,n}}^{u} \left(\sigma'_s - \sigma'_{t^*_{i-1,n}}\right) dW_s + \int_{t^*_{i-1,n}}^{u} \left(v'_s - v'_{t^*_{i-1,n}}\right) dV_s\right] dW_u, \tag{87}$$

$$\zeta_i^n(2) = \int_{t^*_{i-1,n}}^{t^*_{i,n}} a_{t^*_{i-1,n}} \, du + \sigma'_{t^*_{i-1,n}} \int_{t^*_{i-1,n}}^{t^*_{i,n}} \left(W_u - W_{t^*_{i-1,n}}\right) dW_u + v'_{t^*_{i-1,n}} \int_{t^*_{i-1,n}}^{t^*_{i,n}} \left(V_u - V_{t^*_{i-1,n}}\right) dW_u. \tag{88}$$

Consequently, ξ_i^n can be decomposed as

$$\xi_i^n = \xi_i^n(1) + \xi_i^n(2) + \xi_i^n(3), \tag{89}$$

with

$$\xi_i^n(1) = \mathbb{E}\left[n^{\frac{r-1}{2}} f'(\beta_i^n)\zeta_i^n(2) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right], \tag{90}$$

$$\xi_i^n(2) = \mathbb{E}\left[n^{\frac{r-1}{2}} f'(\beta_i^n)\zeta_i^n(1) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right], \tag{91}$$

$$\xi_i^n(3) = \mathbb{E}\left[n^{\frac{r-1}{2}} \left(f'(\gamma_i^n) - f'(\beta_i^n)\right)\left(\zeta_i^n(1) + \zeta_i^n(2)\right) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right]. \tag{92}$$

where γ_i^n is a random variable satisfying $|\gamma_i^n - \beta_i^n| \leqslant |\Delta_i^* X - \beta_i^n|$. The same arguments as in Barndorff-Nielsen et al. (2006) imply that

$$\sum_{i=1}^{N_t^n} \left(\xi_i^n(2) + \xi_i^n(3)\right) \xrightarrow{P} 0.$$

Hence, we are left to prove that $\sum_{i=1}^{N_t^n} \xi_i^n(1) \xrightarrow{P} \int_0^t \alpha_s(1)\, ds$. Conditionally on the event $A_k = \{\Delta(i,n) = k\,\Delta_n\}$, for some $k = 1, 2, \dots$, we have

$$\mathbb{E}\left[f'(\beta_i^n)\zeta_i^n(2) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}} \cap A_k\right] = \left(a_{t^*_{i-1,n}} \sigma^{r-1}_{t^*_{i-1,n}} \rho(f') + \frac{1}{2}\sigma'_{t^*_{i-1,n}} \sigma^{r-1}_{t^*_{i-1,n}} \left(\rho(f',2) - \rho(f')\right)\right) (\Delta_n\, k)^{\frac{r+1}{2}}. \tag{93}$$

Therefore, using Eq. (60) and the law of iterated expectations, we obtain

$$\begin{aligned}
\mathbb{E}\left[n^{\frac{r-1}{2}} f'(\beta_i^n)\zeta_i^n(2) \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] &= \Delta_n \left(a_{t^*_{i-1,n}} \sigma^{r-1}_{t^*_{i-1,n}} \rho(f') + \frac{1}{2}\sigma'_{t^*_{i-1,n}} \sigma^{r-1}_{t^*_{i-1,n}} \left(\rho(f',2) - \rho(f')\right)\right) \mathbb{E}\left[(\Delta(i,n))^{\frac{r+1}{2}} \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] \\
&= \Delta_n \left(a_{t^*_{i-1,n}} \sigma^{r-1}_{t^*_{i-1,n}} \rho(f') + \frac{1}{2}\sigma'_{t^*_{i-1,n}} \sigma^{r-1}_{t^*_{i-1,n}} \left(\rho(f',2) - \rho(f')\right)\right) \frac{1 - p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right).
\end{aligned} \tag{94}$$

Analogously to the proof of Eq. (80), the expression in Eq. (94) implies that

$$\sum_{i=1}^{N_t^n} \xi_i^n(1) \xrightarrow{P} \int_0^t \alpha_s(1)\, ds, \tag{95}$$

where

$$\alpha_s(1) = \sigma_s^{r-1} \left(a_s \rho(f') + \frac{1}{2}\sigma'_s \left(\rho(f',2) - \rho(f')\right)\right) \frac{(1 - p^{\varnothing})^2}{p^{\varnothing}} \mathsf{Li}_{-\frac{r+1}{2}}\left(p^{\varnothing}\right). \tag{96}$$

$\square$

Proof of Corollary 2.1.1. The proof is straightforward. □

Proof of Corollary 2.1.2. The proof is straightforward. □

Proof of Corollary 2.1.3. Corollary 2.1.3 is a particular case of the Corollary 2.1.2 with $r = 2$. Indeed, if $r = 2$, $n^{\frac{r}{2}-1}\sum_{i=1}^{n}\left|\Delta_i\widetilde{X}\right|^r = \sum\limits_{i=1}^{n}\left|\Delta_i\widetilde{X}\right|^2$, $\mu_r = 1$, $\mu_{2r} = 3$ and $\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-\frac{r}{2}}\left(p^{\varnothing}\right) = \frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-1}\left(p^{\varnothing}\right) = 1$, $\forall p^{\varnothing}$. Hence, Corollary 2.1.2 implies that

$$\sum_{i=1}^{n}\left|\Delta_i\widetilde{X}\right|^2 \xrightarrow{P} \int_0^t |\sigma_s|^2\, ds, \tag{97}$$

and

$$\sqrt{n}\left(\sum_{i=1}^{n}\left|\Delta_i\widetilde{X}\right|^2 - \int_0^t |\sigma_s|^2\, ds\right) \xrightarrow{stably} \int_0^t \alpha_s\, dW'_s, \tag{98}$$

where

$$\begin{aligned}
\alpha_s^2 &= \left[\mu_{2r}\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right) - 2\frac{(1-p^{\varnothing})^4}{(p^{\varnothing})^2}\mathsf{Li}_{-1}\left(p^{\varnothing}\right)\mathsf{Li}_{-2}\left(p^{\varnothing}\right) + \frac{(1-p^{\varnothing})^6}{(p^{\varnothing})^3}\left(\mathsf{Li}_{-1}\left(p^{\varnothing}\right)\right)^2\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\right]\sigma_s^{2r}\\
&= \left[\mu_{2r}\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right) - 2\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right) + \frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\right]\sigma_s^{2r}\\
&= \left(\mu_{2r}-\mu_r^2\right)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\mathsf{Li}_{-2}\left(p^{\varnothing}\right)\sigma_s^{2r} = \left(\mu_{2r}-\mu_r^2\right)\frac{(1-p^{\varnothing})^2}{p^{\varnothing}}\frac{p^{\varnothing}(1+p^{\varnothing})}{(1-p^{\varnothing})^3}\sigma_s^{2r} = 2\frac{(1+p^{\varnothing})}{(1-p^{\varnothing})}\sigma_s^{2r}\\
&= \frac{4-2(1-p^{\varnothing})}{(1-p^{\varnothing})}.
\end{aligned} \tag{99}$$

□

Proof of Theorem 5.2. We notice that

$$\sum_{i=1}^{n} f\left(\Delta_i\widetilde{X}\right) = \sum_{i=1}^{N_t^n} f(\Delta_i^* X), \tag{100}$$

where $\Delta_i^* X = X_{t^*_{i,n}} - X_{t^*_{i-1,n}}$ are the increments of the process X_t over the new (random) partition $t^*_{1,n}, \ldots, t^*_{N_t^n,n}$ of the interval $[0, t]$, constructed from the original partition by removing the points for which $\mathbb{B}_{i,n} = 1$. The statement of the Theorem holds by Theorem 3.3.1 of Jacod & Protter (2004). □

Proof of Theorem 2.3. Once we write the problem as in Eq. (100), the proof is similar to that of Theorem 5.5.1 in Jacod and Protter (2012). Therefore, we present the general structure of the proof and omit technical details contained in Jacod and Protter (2012).

For any semimartingale, say Z_t, define

$$\mathsf{V}(f;Z)_t^n = \sum_{i=1}^{N_t^n} f(\Delta_i^* Z), \quad \overline{\mathsf{V}}(f;Z)_t^n = \sqrt{n}\left(\mathsf{V}(f;Z)_t^n - \sum_{s\leqslant t^*_{N_t^n,n}} f(\Delta Z_s)\right). \tag{101}$$

For any $m \geqslant 1$, denote by $(T(m,r) : r \geqslant 1)$ the successive jump times of the process $N^m = \mathbf{1}_{\{1/m<\gamma\leqslant 1/(m-1)\}} \star \mu$ and let S_p be the reordering of the double sequence $(T(m,r) : r \geqslant 1)$ into a single sequence. Now, fix $m \geqslant 1$

and let P_m be the set of all p such that $S_p = T(m', r)$ for some $r \geqslant 1$ and some $m' \leqslant m$. Then, define

$$X(m)_t = X_t - \sum_{p \in P_m : S_p \leqslant t} \Delta X_{S_p} = X_t - \left(\delta \mathbf{1}_{\{\gamma > 1/m\}}\right) \star \mu. \tag{102}$$

Next, denote by $\Omega_n(t, m)$ the set of all ω such that each interval $(t^*_{i-1,n}, t^*_{i,n}]$ contains at most one jump of $\left(\delta \mathbf{1}_{\{\gamma > 1/m\}}\right) \star \mu$ and such that $|\Delta^*_i X(m)| \leqslant 2/m$ for all i. For each S_p, denote the edges of the interval $(t^*_{i-1,n}, t^*_{i,n}]$, which contains S_p, by $S_-(n,p) = t^*_{i-1,n}$ and $S_+(n,p) = t^*_{i,n}$. Then, for all $t > 0$ and $m \geqslant 1$, $\Omega_n(t,m) \to \Omega$ almost surely as $n \to \infty$. On $\Omega_n(t,m)$, we have

$$\mathsf{V}(f;X)^n_t = \sum_{i=1}^{N^n_t} f(\Delta^*_i X) = \sum_{i=1}^{N^n_t} f(\Delta^*_i X(m)) + \sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} f(\Delta^*_{S_+(n,p)} X(m) + \Delta X_{S_p}) - \sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} f(\Delta^* X(m)), \tag{103}$$

and

$$\sum_{s \leqslant t^*_{N^n_t,n}} f(\Delta X_s) = \sum_{s \leqslant t^*_{N^n_t,n}} f(\Delta X(m)_s) + \sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} f(\Delta X_{S_p}). \tag{104}$$

Therefore, on $\Omega_n(t,m)$, we obtain the decomposition

$$\overline{\mathsf{V}}(f;X)^n_t = \overline{\mathsf{V}}(f;X(m))^n_t + \sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} \zeta^n_p, \tag{105}$$

with

$$\zeta^n_p = \sqrt{n}\left(f(\Delta^*_{S_+(n,p)} X(m) + \Delta X_{S_p}) - f(\Delta X_{S_p}) - f(\Delta^*_{S_+(n,p)} X(m))\right). \tag{106}$$

Set

$$R_-(n,p) = \sqrt{n}(X_{S_p -} - X_{S_-(n,p)}), \quad R_+(n,p) = \sqrt{n}(X_{S_+(n,p)} - X_{S_p}). \tag{107}$$

With this notations, ζ^n_p can be expressed as:

$$\begin{aligned} \zeta^n_p &= \sqrt{n}\left(f\left(\Delta X_{S_p} + \frac{1}{\sqrt{n}}(R_-(n,p) + R_+(n,p))\right) - f(\Delta X_{S_p}) - f\left(\frac{1}{\sqrt{n}}(R_-(n,p) + R_+(n,p))\right)\right) \\ &= f'\left(\Delta X_{S_p} + \tilde{R}(n,p)\right)(R_-(n,p) + R_+(n,p)) - \sqrt{n} f\left(\frac{1}{\sqrt{n}}(R_-(n,p) + R_+(n,p))\right), \end{aligned} \tag{108}$$

where $\tilde{R}(n,p)$ lies between ΔX_{S_p} and $\frac{1}{\sqrt{n}}(R_-(n,p) + R_+(n,p))$. Exactly the same arguments as in Proposition 4.4.10 of Jacod and Protter (2012) gives that the sequence $(R_-(n,p) + R_+(n,p))$ is bounded in probability (for a fixed p),

$$R_-(n,p) - \sigma_{S_p -} \alpha_-(n,p) \xrightarrow{p} 0, \quad R_+(n,p) - \sigma_{S_p} \alpha_+(n,p) \xrightarrow{p} 0, \tag{109}$$

where

$$\alpha_-(n,p) = \sqrt{n}(W_{S_p -} - W_{S_-(n,p)}), \quad \alpha_+(n,p) = \sqrt{n}(W_{S_+(n,p)} - W_{S_p}), \tag{110}$$

and

$$(\alpha_-(n,p), \alpha_+(n,p)) \xrightarrow{stably} \left(\sqrt{\xi^-_p} U^-_p, \sqrt{\xi^+_p} U^+_p\right). \tag{111}$$

Thus,

$$\zeta_p^n - f'\left(\Delta X_{S_p}\right)\left(R_-(n,p) + R_+(n,p)\right) \xrightarrow{p} 0, \tag{112}$$

and

$$\left(R_-(n,p) + R_+(n,p)\right) \xrightarrow{stably} \sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+, \tag{113}$$

which implies that

$$\zeta_p^n \xrightarrow{stably} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right). \tag{114}$$

Since on $\Omega_n(t,m)$ the sum $\sum_{p\in P_m: S_p \leqslant t^*_{N_t^n,n}} \zeta_p^n$, contains finitely many entries, Eq. (114) implies that

$$\sum_{p\in P_m: S_p \leqslant t^*_{N_t^n,n}} \zeta_p^n \xrightarrow{stably} \sum_{p\in P_m: S_p \leqslant t^*_{N_t^n,n}} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right). \tag{115}$$

Then, since for $p \notin P_m$ all jumps ΔX_{S_p} are smaller than $1/m$ in absolute value (and by the boundedness of σ_t), we obtain

$$\sum_{p\in P_m: S_p \leqslant t^*_{N_t^n,n}} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right) - \sum_{p: S_p \leqslant t^*_{N_t^n,n}} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right) \xrightarrow{u.c.p.} 0. \tag{116}$$

Now, by the same arguments as in step 4 of the proof of Theorem 5.5.1 in Jacod and Protter (2012), for any $\epsilon > 0$, we obtain

$$\lim_{m\to\infty}\limsup_{n\to\infty} \mathcal{P}\left(\Omega_n(t,m) \cap \left\{\sup_{s\leqslant t^*_{N_t^n,n}} \sqrt{n}\left|\overline{\mathsf{V}}(f;X(m))_s^n\right| > \epsilon\right\}\right) = 0, \tag{117}$$

which implies that

$$\overline{\mathsf{V}}(f;Z)_t^n \xrightarrow{stably} \sum_{p: S_p \leqslant t^*_{N_t^n,n}} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right). \tag{118}$$

Since t is not a fixed time of discontinuity of the process $\mathsf{V}(f;X)_t^n$, for any fixed t the above convergence implies that

$$\overline{\mathsf{V}}(f;Z)_t^n \xrightarrow{stably} \sum_{p: S_p \leqslant t} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p-}\sqrt{\xi_p^-}U_p^- + \sigma_{S_p}\sqrt{\xi_p^+}U_p^+\right). \tag{119}$$

Finally, in order to complete the proof, it is enough to show that, for a fixed t,

$$U_n = \sqrt{n}\left|\sum_{s\leqslant t^*_{N_t^n,n}} f(\Delta X_s) - \sum_{s\leqslant t} f(\Delta X_s)\right| \xrightarrow{p} 0. \tag{120}$$

Let Ω_n be the set on which there are no jumps in X bigger than 1 between the time $t^*_{N_t^n,n}$ and t. On Ω_n, $|f(x)| \leqslant K|x|^2$, for a constant K. Hence, on Ω_n, $U_n \leqslant K\sqrt{n}\sum_{t^*_{N_t^n,n} < s \leqslant t} |\Delta X_s|^2$. Consequently, by conditioning on $t - t^*_{N^n_t,n}$ and using the law of iterated expectations, we obtain

$$\mathbb{E}\left[U_n \mathbf{1}_{\Omega_n}\right] \leqslant K\sqrt{n}\mathbb{E}\left[\sum_{t^*_{N^n_t,n} < s \leqslant t} |\Delta X_s|^2\right] \leqslant K\sqrt{n}\mathbb{E}\left[\int_{t^*_{N^n_t,n}}^t \int |\delta(s,z)|^2 \, \lambda(dz)\, ds\right] \leqslant K'\sqrt{n}\,\mathbb{E}\left[\left(t - t^*_{N^n_t,n}\right)\right], \tag{121}$$

where K' is another constant and $t - t^*_{N^n_t,n}$ is a discrete random variable with probability mass function given by $\mathbb{P}\left(t - t^*_{N^n_t,n} = k\Delta_n\right) = (1 - p^{\varnothing})\left(p^{\varnothing}\right)^{(k-1)}$ for $k \in \{1, 2, \dots\}$. Thus,

$$\mathbb{E}\left[U_n \mathbf{1}_{\Omega_n}\right] \leqslant K'\sqrt{n}\,\Delta_n \sum_{s=1}^{\infty} k^s (1 - p^{\varnothing})\left(p^{\varnothing}\right)^s \longrightarrow 0. \tag{122}$$

Since $\Omega_n \to \Omega$ as $n \to \infty$, $\mathbb{E}\left[U_n \mathbf{1}_{\Omega_n}\right] \to 0$ implies that $U_n \xrightarrow{p} 0$, which completes the proof. $\square$

Proof of Theorem 2.4. As in the proof of Theorem 2.3, on $\Omega_n(t,m)$, we obtain the decomposition:

$$\overline{\mathsf{V}}(f;X)^n_t = \overline{\mathsf{V}}(f;X(m))^n_t + \sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} \zeta^n_p, \tag{123}$$

where

$$\sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} \zeta^n_p \xrightarrow{stably} \sum_{p \in P_m : S_p \leqslant t^*_{N^n_t,n}} f'\left(\Delta X_{S_p}\right)\left(\sigma_{S_p -}\sqrt{\xi^-_p}\, U^-_p + \sigma_{S_p}\sqrt{\xi^+_p}\, U^+_p\right). \tag{124}$$

Next, we notice that: ...

$\square$

Proof of Theorem 2.5. The proof is analogical to the proof of Theorem 2 in Bibinger and Vetter (2015). Hence, it is omitted. $\square$

Proof of Lemma 3.1. Consider the decomposition:

$$A(q)^n_t = A_1 + A_2, \tag{125}$$

where

$$A_1 = \sum_{i=1}^{N^n_t} \zeta_i, \quad \text{with} \quad \zeta_i = n^{q-1}\left(\Delta(n,i)^q - \mathbb{E}\left[\Delta(n,i)^q\right]\right), \tag{126}$$

and

$$A_2 = n^{q-1} \sum_{i=1}^{N^n_t} \mathbb{E}\left[\Delta(n,i)^q\right]. \tag{127}$$

Due to the independence of the Bernoulli variables, we have:

$$\mathbb{E}\left[\Delta(n,i)^q\right] = \mathbb{E}\left[\Delta(i,n)^q \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] = \Delta^q_n \frac{1 - p^{\varnothing}}{p^{\varnothing}} \mathsf{Li}_{-q}\left(p^{\varnothing}\right), \tag{128}$$

where the second inequality follows from equation (60). Consequently, we have:

$$\mathbb{E}\left[\zeta_i \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] = 0, \tag{129}$$

and

$$\mathbb{E}\left[\zeta_i^2 \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] = n^{-2}\left(\frac{1-p^{\varnothing}}{p^{\varnothing}}\mathsf{Li}_{-2q}\left(p^{\varnothing}\right) - \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}}\left(\mathsf{Li}_{-q}\left(p^{\varnothing}\right)\right)^2\right). \tag{130}$$

Hence, ζ_i is $\mathcal{F}_{t^*_{i-1,n}}$-martingale and

$$\sum_{i=1}^{N_t^n}\mathbb{E}\left[\zeta_i^2 \,\middle|\, \mathcal{F}_{t^*_{i-1,n}}\right] \leqslant C\, n^{-2}N_t^n \longrightarrow 0, \tag{131}$$

which implies that A_1 is asymptotically negligible.

Then, substituting (128) in A_2 we obtain:

$$A_2 = n^{q-1}\sum_{i=1}^{N_t^n}\Delta_n^q\frac{1-p^{\varnothing}}{p^{\varnothing}}\mathsf{Li}_{-q}\left(p^{\varnothing}\right) = \frac{1-p^{\varnothing}}{p^{\varnothing}}\mathsf{Li}_{-q}\left(p^{\varnothing}\right)n^{-1}N_t^n. \tag{132}$$

It can be easily seen that:

$$n^{-1}N_t^n = \frac{1}{n}\sum_{i=1}^{n}\left(1-\mathbb{B}_{i,n}\right) \overset{u.c.p.}{\longrightarrow} \left(1-p^{\varnothing}\right)t. \tag{133}$$

Consequently,

$$A_2 \overset{u.c.p.}{\longrightarrow} \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}}\mathsf{Li}_{-q}\left(p^{\varnothing}\right)\, t, \tag{134}$$

which completes the proof.

□

Proof of Theorem 3.2. By Theorems 3.1 and 3.2 of Hayashi et al. (2011), on $\Omega^c_{[0,t]}$, as $n\to\infty$,

$$\mathsf{PV}^c(f;\widetilde{X}) \overset{u.c.p.}{\longrightarrow} \mu_r\int_0^t \left|\sigma_s\right|^r\, ds, \tag{135}$$

and

$$\sqrt{n}\left(\mathsf{PV}^c(f;\widetilde{X}) - \mu_r\int_0^t \left|\sigma_s\right|^r\, ds\right) \overset{stably}{\longrightarrow} \left(\mu_{2r}-\mu_r^2\right)\int_0^t \sigma_s^r\sqrt{a(2)_s}\, dW'_s, \tag{136}$$

where $a(q)_t$ is a stochastic process, such that for all t, as $n\to\infty$,

$$A(q)_t^n = n^{q-1}\sum_{i=1}^{N_t^n}\left|\Delta(n,i)\right|^q \overset{u.c.p.}{\longrightarrow} \int_0^t a(q)_s\, ds, \tag{137}$$

provided that such a process exists. By Lemma 3.1, the process $a(q)_t$ exists for every $q>0$ and takes the following form:

$$a(q)_t = \frac{\left(1-p^{\varnothing}\right)^2}{p^{\varnothing}}\mathsf{Li}_{-q}\left(p^{\varnothing}\right)\, t. \tag{138}$$

Hence, the statement of the present Theorem follows from straightforward algebraic computations. □

Proof of Theorem 3.3. By Theorems 3.1 of Hayashi et al. (2011) and Lemma 3.1,

$$n \sum_{i=1}^{N_t^n} \Delta(n,i)^{2-r} \left|\Delta_i^{\star} X\right|^{2r} \xrightarrow{u.c.p.} \mu_{2r} \int_0^t |\sigma_s|^{2r} \, a(2)_s \, ds, \tag{139}$$

where $a(q)_t$ is a stochastic process defined as in the proof of Theorem 3.2, which completes the proof. $\square$