Jumps or staleness?*

DOI:
10.1080/07350015.2023.2203207

Document Version
Accepted author manuscript

Jumps or staleness? *

Aleksey Kolokolov† Roberto Renò‡

February 3, 2023

Abstract

Even moderate amounts of zero returns in financial data, associated with stale prices, are heavily detrimental for reliable jump inference. We harness staleness-robust estimators to reappraise the statistical features of jumps in financial markets. We find that jumps are much less frequent and much less contributing to price variation than what found by the empirical literature so far. In particular, the empirical finding that volatility is driven by a pure jump process is actually shown to be an artifact due to staleness.

JEL classification: C58; C22

Keywords: Staleness; Multipower variation; Jump testing; Jump activity index

*We thank Torben Andersen, Federico Bandi, Massimiliano Caporin, Cecilia Mancini, Dobrislav Dobrev, Marcelo Medeiros, the participants at the SoFiE Conference 2017 (New York, June 21-23), the SoFiE Conference 2018 (Lugano, June 12-14), Quantitative Finance and Financial Econometrics – QFFE 2018 (Marseille, 30 May - 1 June), 10th World Congress of the Bachelier Finance Society (Dublin, 16-20 July, 2018) and the XIX Quantitative Finance Workshop (Rome, 24-26 January, 2018), and seminar partecipants at the Pontifical Catholic University, Rio de Janeiro (2019) and BI Norwegian Business School, Oslo (2022). R. Renò acknowledges PRIN funding from MIUR, Italy, HiDEA Project 2017RSMPZZ. Errors and omissions remain our own.

†Alliance Manchester Business School, Booth St W, Manchester M15 6PB, phone: +44 161 2750436, e-mail: Aleksey.kolokolov@manchester.ac.uk

‡ESSEC Business School, e-mail: reno@essec.edu

1 Introduction

In empirical finance, it is customary to sample prices on evenly spaced grids, such as one or five minutes. When doing so many price returns are zero. For the stocks belonging to S&P500 in our sample, the percentage of zero returns at one minute is 33.96%; at five minutes, it is 16.08%; and at the relatively low frequency of ten minutes, it is still 11.52%. Zeros are a non-negligible feature of high-frequency data, even for the most liquid stocks. Their omnipresence is caused by price staleness, an economically justified phenomenon (Bandi, Pirino, and Renò, 2017), and price discreteness (Bandi, Kolokolov, Pirino, and Renò, 2020 discuss how to separate the two phenomena). Nevertheless, the vast majority of models used in finance postulate the mere impossibility of zero returns in the price dynamics. Exceptions are models with rounded prices (Delattre and Jacod, 1997; Li and Mykland, 2015), the uncertainty zone literature (Robert and Rosenbaum, 2011), and zero-augmented models, as in Hautsch, Malec, and Schienle (2013) and, more recently, Catania, Di Mari, and Santucci de Magistris (2020), Sucarrat and Grønneberg (2022) and Francq and Sucarrat (2022). What are the dangers of ignoring the presence of zeros in the data generating process?

This paper answers this question by focusing on the impact of price staleness on jumps inference. Jumps and staleness are both essential ingredients of the data generating process. Jumps are discontinuities in the price, which materialize either in the form of sudden, large returns, or in the form of small, but infinitely many, discontinuities (see, e.g., Aït-Sahalia and Jacod, 2012). Staleness refers to the absence of movement, and materializes in the form of zero returns. The presence of zero returns has two effects on the return distribution. The first is inflating the distribution at zero. The second, which is more subtle but inherently linked to the linear scaling law of the variance of the Brownian motion, is that non-zero returns following a zero have a larger variance, which inflates the tails of the distribution of observed returns. Both effects concur in making the distribution more leptokurtic than it actually is, which is the same effect of jumps. It is then not surprising that jump-diffusion models which do not account for staleness require more (or more active) jumps to fit the data. We provide formal support for this intuition by exploring the bias induced by staleness on multipower variation, a successful tool in financial econometrics (see, e.g., Woerner, 2006; Barndorff-Nielsen and Shephard, 2004; Barndorff-Nielsen, Graversen, Jacod, and Shephard, 2006), as well as on alternative tools used for jump inference.

Our paper provides several contributions to the existing literature. On the theoretical side, we provide limit theorems for multipower variation under a flexible data generating process which includes staleness. We then introduce new multipower estimators which are designed to remove the biases due to the presence of staleness. We borrow the return scaling from Hayashi, Jacod, and Yoshida (2011) and Levine, Wang, and Zou (2016). With respect to Hayashi, Jacod, and Yoshida (2011), we extend the analysis to multipower variation. Multipower variation with random sampling is also considered in Levine, Wang, and Zou (2016), but under restrictive assumptions on the duration between observation times which are relaxed here. The correction we propose is handy and requires no further computational effort. The staleness-robust multipower estimator coincides with the traditional one if there are no zeros in the time series under investigation.

Further, we study jump inference in finite samples using simulations. We extend the work of Theodosiou and Zikes (2011) and Maneesoonthorn, Martin, and Forbes (2020) by adding staleness in the data generating process, showing that this changes the ranking of various jump tests considerably, and studying various estimators of the jump activity index (JAI, see Todorov and Tauchen, 2010 for a formal definition of this quantity). Similarly to Theodosiou and Zikes (2011), we show that popular jump tests (such as those in Barndorff-Nielsen and Shephard, 2006, Lee and Mykland, 2008 and Aït-Sahalia and Jacod, 2009b) are severely biased toward rejection of the null (that is, detecting more jumps) when zeros are included in the data generating process. However, we also show that tests based on wild bootstrap (Podolskij and Ziggel, 2010; Dovonon, Gonçalves, Hounyo, and Meddahi, 2019) as well as tests based on the correction we propose are robust to the presence of zeros. Regarding existing JAI estimators, we show that they are negatively biased in the presence of zero returns, and increasingly so when zeros are more frequent. When we apply the proposed correction to the considered estimators we remove this bias, thus allowing to truly assess the nature of the main driving force of the returns' shocks.

Our empirical contribution is then to reconsider jump features in financial time series using our staleness-robust multipower estimator. The examined assets are stocks belonging to the S&P 500 index, the SPY exchange-traded fund (that is, the SPDR S&P 500 trust), and the VIX index as computed by CBOE. We show that, after correcting for staleness, (i) much fewer jumps are detected, and (ii) the contribution of jumps to total quadratic variation is much smaller. Our results point out that the discrepancy in measures of jump variation at different frequency found in the literature (Christensen, Oomen, and Podolskij, 2014) is a spurious by-product of the presence of staleness in high-frequency returns. We further show that estimates of the jump activity index on all our assets, including the VIX index, are not significantly different from 2 (the value implied by the presence of the Brownian motion) after taking staleness properly into account. In particular, traditional estimators of the jump activity index estimate a value lower than 2 for the VIX index (Todorov and Tauchen, 2011; Andersen, Bondarenko, Todorov, and Tauchen, 2015; Todorov, 2015), which would imply that the dynamics of volatility is driven by a pure, infinite activity jump process, in contrast with traditional stochastic volatility models typically assumed in asset pricing. However, our empirical analysis is very sharp in showing that traditional activity estimates are affected by the fraction of zero returns in the data, being strongly negatively correlated with this fraction. We thus argue that the alleged evidence of pure jumps in volatility is an artifact due to the unaccounted presence of staleness. Using robustified estimators, the relation between the level of staleness and the estimated activity disappears, and the activity index of the VIX is invariably estimated to be indistinguishable from 2 for every day in the sample, in keeping with the inescapable presence of a Brownian motion in the price and volatility dynamics.

The main limitation of our proposed correction is that it may deliver less precise estimates of the quantities of interest when the level of price discreteness (rounding) is too aggressive, as we show using simulations. Indeed, rounding is a form of friction which generates zeros but with different statistical properties with respect to our staleness model. Theoretical treatment of rounding is particularly challenging (see, e.g. Delattre and Jacod, 1997; Li and Mykland, 2015). To overcome this difficulty, we follow Bandi, Kolokolov, Pirino, and Renò (2020) and measure the strength of the rounding in the data using so-called *Rounding Impact Ratio* (RIR) prior to applying the correction. We show that, for realistic values of the RIR, the correction we propose is more reliable than traditional estimators. Moreover, we provide an heuristic solution to the rounding problem, which is dependent on the observed RIR. This improves the precision of the estimator we propose especially in cases in which rounding is particularly aggressive.

The rest of the paper is structured as follows. Section 2 lays down the theory for multipower estimators and their robustified counterparts in the presence of staleness. Section 3 shows the distortions of traditional jump statistics based on multipower estimators on realistic simulations of the price process. Section 4 studies similar distortions on alternative jump statistics which do not use multipower. Section 5 studies the impact of price discreteness on our new staleness-robust estimator. Section 6 containes our empirical application which reconsiders the estimates of jump features in financial data. Section 7 concludes. Three appendices contain mathematical proofs, central limit theorems and a correction robust to the presence of rounding respectively.

2 Multipower variation under staleness

We assume the efficient log-price process X^e_t is a Brownian semimartingale, evolving (in the standard probabilistic setting) as

$$X^e_t = X^e_0 + \int_0^t \mu_s\, ds + \int_0^t \sigma_s dW_s, \tag{2.1}$$

and the volatility process σ satisfies the equation:

$$\sigma_t = \sigma_0 + \int_0^t \mu'_s\, ds + \int_0^t \sigma'_s dW_s + \int_0^t \nu'_s dV_s, \tag{2.2}$$

where W_t and V_t are independent Brownian motions, μ, μ', σ' and ν' are adapted càdlàg bounded processes. The observations are recorded on an equally spaced mesh $t_i = i\Delta_n$, $i = 0, \dots n$, of $n+1$ points, with $\Delta_n = T/n$. Price staleness is introduced as in the model in Bandi, Pirino, and Renò (2017), that is we assume that the observed log-price follows:

$$X_{t_i} = X^e_{t_i}(1 - B_{t_i}) + B_{t_i} X_{t_{i-1}}, \tag{2.3}$$

where (B_{t_i}) is a triangular array of Bernoulli random variables, for $i = 1, \dots, n$, such that, as $n \to \infty$, $\frac{1}{T}\sum_{i=1}^{n} \Delta_n B_{t_i} \xrightarrow{p} p^{\emptyset}$, with $p^{\emptyset} \in [0, 1]$. We define the returns of the observed price process as $\Delta_i X = X_{(i+1)T/n} - X_{iT/n}$. In what follows, without loss of generality, we set $T = 1$. For a vector of m positive real numbers $r = [r_1, ..., r_m]$, and a stochastic process X, we define realized multipower variation as

$$\mathsf{MV}(X; r) = \frac{1}{n} \sum_{i=1}^{n-m+1} \left|\sqrt{n} \cdot \Delta_i X\right|^{r_1} \dots \left|\sqrt{n} \cdot \Delta_{i+m-1} X\right|^{r_m}. \tag{2.4}$$

In the presence of zeros, multipowers are biased estimators of integrated volatility powers. If the Bernoulli variates are independent and identically distributed (i.i.d.), the bias can be derived explicitly, that is we can prove that:[1]

$$\mathsf{MV}(X; r) \xrightarrow{p} \left(\prod_{j=1}^{m} \mu_{r_j}\right) \frac{(1 - p^{\emptyset})^{m+1}}{p^{\emptyset}} \mathsf{Li}_{-\frac{r_1}{2}}\left(p^{\emptyset}\right) \int_0^1 |\sigma_s|^{r_+} ds, \tag{2.5}$$

where $r_+ = r_1 + ... + r_m$, and $\mathsf{Li}_s(z)$ denotes the polylogarithm function of order s and argument z, and where $\mu_s = \frac{2^{s/2}\Gamma((1+s)/2)}{\Gamma(1/2)}$. The multiplicative bias only depends on the first power r_1, the

[1] Proof is available upon request.

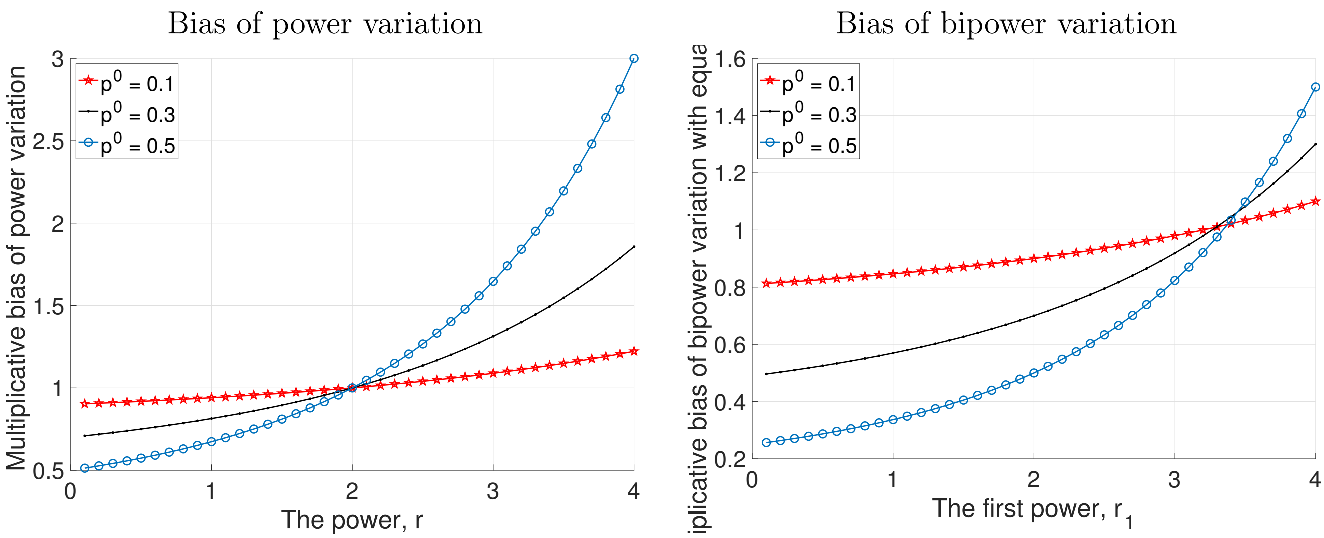

Figure 1: The multiplicative bias of power variation (left panel) and bipower variation with equal powers (right panel) as a function of the powers for different staleness probability, $p^{\emptyset}$.

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

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

While the i.i.d. case is natural to provide intuitions, it is highly unrealistic in practice (Kolokolov, Livieri, and Pirino, 2020; Bandi, Kolokolov, Pirino, and Renò, 2020). Below we propose a correction to multipower estimators which is designed to work for a more realistic specification of Bernoulli variables presented by Assumption 1.

Assumption 1. *Assume that* (B_{t_i}) *is a triangular array of Bernoulli random variables with* $\mathbf{E}[B_{t_i}|p_{t_i}] = p_{t_i}$*, where* p_t *is a stochastic process taking values in* $[0,1)$ *and evolving as*

$$p_t = p_0 + \int_0^t \mu_s'' ds + \int_0^t \nu_s'' dV_s + \int_0^t v_s'' dZ_s, \tag{2.6}$$

where Z_t *is a Brownian motion independent from* W_t *and* V_t *(which appear in equations (2.1) and (2.2)),* μ''*,* ν'' *and* v'' *are adapted càdlàg bounded processes, and conditionally on* p_t *the Bernoulli random variables* B_{t_i} *and* B_{t_j} *are independent* $\forall t_i \neq t_j$*.*

Assumption 1 is general enough to allow for standard features of staleness in the data. For example, the common factor in the shocks to volatility and the shocks to the probability of zeros allows for correlation between stale prices and the volatility level. Our assumptions also allow for the joint presence of a leverage effect. Moreover, the assumption allows introducing a common intraday diurnal pattern in zeros and volatility (typically, more zeros are observed in periods of low volatility around lunch hours). However, Assumption 1 does not allow zeros to be correlated with the Brownian component of the efficient price, which is ruling out, e.g., rounding. Section 5 is dedicated to study the impact of price discreteness on our results.

Given the initial time grid of price observations $\{t_0, t_1, \ldots, t_n\}$, we define the new stochastic grid of N_n points $\{\tau_0, \ldots, \tau_{N_n}\}$ such that returns are non-zero: $\tau_0 = t_k,\ k = \min\{j : \ B_{t_j} = 0\}$, and $\tau_l = t_k,\ k = \min\{j : \ B_{t_j} = 0,\ t_j > \tau_{l-1}\}$. Let $\Delta_{\tau_i} X = X_{\tau_i} - X_{\tau_{i-1}}$ and $\Delta(n,i) = \tau_i - \tau_{i-1}$ denote respectively non-equispaced non-zero returns and their durations (the variables $\Delta(n,i)$ can also be interpreted as the run lengths of the Bernoulli variates, that is the numbers of consecutive zeros after a non-zero return, multiplied by Δ_n). Staleness-robust multipower variation are defined using a generalization of the power estimator of Hayashi et al. (2011). However, we need to adapt their theory to allow for dependency between the Bernoulli variates and price volatility (which violates Assumption C in their paper, see also the discussion in Li, Mykland, Renault, Zhang, and Zheng, 2014). This is done in the following theorem.

Theorem 1. *Assume that* X_t *is defined by equations (2.1) and (2.2) and Assumption 1 holds. Let* f *be a continuous function on* $\mathbf{R}^m$ *for some* $m \geq 1$*, which satisfies*

$$|f(x_1, \ldots, x_m)| \leq C \prod_{j=1}^{m} (1 + \|x_j\|^p), \tag{2.7}$$

for some $p > 0$ and $C > 0$. Define the robustified multipower variation as

$$\mathsf{V}'(f)^n = \overline{\Delta}_n \sum_{i=1}^{N_n-m+1} f\left(\Delta(n,i)^{-1/2}\Delta_{\tau_i}X, \ldots, \Delta(n,i+m-1)^{-1/2}\Delta_{\tau_{i+m-1}}X\right), \tag{2.8}$$

where $\overline{\Delta}_n$ denotes a deterministic sequence of numbers, such that, for $q = 1, 2$,

$$\sum_{i=1}^{N_n} \mathbf{E}\left[\left(\Delta(n,i) - \overline{\Delta}_n\right)^q |\mathcal{F}_{\tau_{i-1}}\right] \xrightarrow{p} 0.$$

Then,

$$\mathsf{V}'(f)^n \xrightarrow{p} \int_0^1 \rho_{\sigma_u}(f)du, \tag{2.9}$$

where $\rho_{\sigma_u}(f) = \mathbf{E}\left[f(u_1, \ldots, u_m)\right]$ with $u_1, \ldots, u_m$ being i.i.d. $\mathcal{N}(0, \sigma_u^2)$ random variables.

The estimator (2.8) does not exactly coincide with the Hayashi et al. (2011) estimator in the case of power variation because of the scaling factor. The scaling factor we suggest, borrowed from Levine, Wang, and Zou (2016), has the simple advantage that the covariance among different robustified multipowers is such that the asymptotic distribution of the test-statistics of Barndorff-Nielsen and Shephard (2006) (used below in our empirical application) needs not to be changed. In practice, one can estimate $\overline{\Delta}_n$, e.g., with the sample average of $\Delta(n,i)$'s. However, estimating $\overline{\Delta}_n$ is not required for jump-testing and JAI estimation as the constant appears in both numerator and denominator of the tests statistics and cancels when dividing. Further discussion and associated Central Limit Theorems are presented in Appendix B.

In the case of multipowers, the estimator (2.8) takes the form:

$$\mathsf{MV}^{\mathsf{c}}(X;r) = \overline{\Delta}_n \sum_{i=1}^{N_n-m+1} \left|\Delta(n,i)^{-1/2}\Delta_{\tau_i}X\right|^{r_1} \ldots \left|\Delta(n,i+m-1)^{-1/2}\Delta_{\tau_{i+m-1}}X\right|^{r_m}, \tag{2.10}$$

and it coincides with that in Eq. (2.4) when $B_{t_i} = 0$ $\forall i$ identically, that is in the absence of zeros. However, under the presence of price staleness, $\mathsf{MV}^{\mathsf{c}}(X'; [2])$ differs from standard realized variance, although it converges to the same limit. The objects of econometric interest are the integrated volatility powers of model (2.1). When the model is contaminated by staleness, as in Eq. (2.3), traditional estimated volatility powers (2.4) are distorted as shown by Theorem 1. Inference about the object of interest (volatility powers of the efficient price) is restored by the estimator (2.10).

2.1 Comparison to tick-time sampling

Our model for staleness can be viewed as a specific form of implicit random sampling. In our model, the price is observed on a (random) subset of a pre-specified deterministic grid, and the randomness in the observation times is driven by the Bernoulli variables according to Eq. (2.3). The bias in multipower estimators is thus generated by the randomness of the sampling times, as it is the case with realized estimators when random sampling is explicit and observation grids are not equally spaced. Explicit random sampling is more typically referred as tick-time or event-time sampling in the literature, see e.g. Aït-Sahalia and Mykland, 2003 and Griffin and Oomen, 2008 for a discussion. In the case of explicit random sampling, it is natural to rescale returns with the exact distance between observations, as we do in Eq. (2.10). This is done, for example, in Hayashi, Jacod, and Yoshida (2011) to study power variation, and in Levine, Wang, and Zou (2016) to study multipower variation under iid sampling times.

An alternative route to the implicit random sampling we propose would be to define multipower estimators under explicit random sampling and study the properties of such estimators. However, our approach based on (randomized) calendar time is motivated by several reasons. The first is that, in many instances, the empirical analysis is constrained by data availability, which typically comes on pre-specified grids in calendar time (e.g., one-minute data). The second is that our theoretical treatment stresses that the usual grids used in empirical analysis, defined in calendar time, are actually random grids because of price staleness, which originates biases (due to random sampling) which need to be corrected. The third one is that calendar time is the norm in empirical work. Relevant exceptions that use random sampling for realized measures are Li, Mykland, Renault, Zhang, and Zheng (2014), who use power variations to test for endogeneity of sampling times; Andersen, Dobrev, and Schaumburg (2009); Hong, Nolte, Taylor, and Zhao (2021), who use price durations to estimate integrated volatility powers; Andersen, Dobrev, and Schaumburg (2012) who use robust filtering based on nearest neighbors; and Li, Nolte, Nolte, and Yu (2022), who propose a jump test based on event time. The fourth reason is that, even if we have the full transaction record at our disposal, it is common to do *sparse sampling* to soften the impact of frictions (market microstructure noise, price discreteness, time endogeneity) which are absent in our theoretical treatment. Sparse sampling is less common in tick time, since it may be influenced by the properties of the arrival times (for example, when trading switches from intense to infrequent). Sparse sampling in calendar time is more natural. It transforms trade inaction into staleness, a feature that is accommodated by our theory.

The theoretical properties of the implicit random sampling scheme associated with staleness are determined by the assumptions we impose on the triangular array of Bernoulli variables. For example, iid Bernoulli satisfy the conditions of Theorem 1 and Theorem 2 of Li, Mykland, Renault, Zhang, and Zheng (2014) (the iid case is Example 3 in their paper), so we can use their test for endogeneity.[2] Since we assume that the Bernoulli variables are independent from the Brownian motion in the efficient price, our implicit random sampling is non-endogenous in their terminology. The Li, Mykland, Renault, Zhang, and Zheng (2014) test on five-minute grids does not reject the null of non-endogenity. This is consistent with the idea, discussed above, that sparse sampling is a simple solution against frictions that contaminate the standard price diffusion. In Section 5, we will specifically discuss the impact of price discreteness.

3 Distortions to jump inference

We now assess the distortion to jump inference produced by staleness in finite samples. To this purpose we simulate 5-minute returns on the latent price process X_t for 6.5 hours of trading using a stochastic volatility model. Our settings are borrowed from Caporin, Kolokolov, and Renò (2017), who adopted the stochastic volatility model from Andersen, Benzoni, and Lund (2002) enlarged by the presence of intraday effect in volatility. The efficient log-price process X_t is specified as:

$$dX_t = \mu dt + \gamma_t \sqrt{V_t} dW_{1,t}, \tag{3.1}$$

$$d \log V_t = (\alpha - \beta \log V_t)dt + \eta dW_{2,t}, \tag{3.2}$$

where $W_{1,t}$ and $W_{2,t}$ are correlated Brownian motions and γ_t is the adjustment for the intraday effect. The parameters of the stochastic volatility model are from Table IV of Andersen et al. (2002), corresponding to the column $SV_1 J, \rho \neq 0$. The intraday effect takes the following form:

$$\gamma_t = \frac{1}{1033} \left(0.1271t^2 - 0.1260t + 0.1239\right), \tag{3.3}$$

as estimated by Caporin, Kolokolov, and Renò (2017) on S&P500 index data.

We add zeros with Eq. (2.3). The generated probability of observing a zero return is time varying, and it follows a deterministic inverted U-shape intraday pattern. We model time-varying probability of zeros as:

$$p_t = p^U \cdot \gamma_t^{(p)}, \tag{3.4}$$

[2] The generalization of Theorem 2 of Li, Mykland, Renault, Zhang, and Zheng (2014) under Assumption 1 is left for further research. Indeed, we cannot apply directly Theorem 2 as it is, since sampling times are not independent from the efficient price, due to the asssumed dependence between Bernoulli variates and volatility shocks.

where p^U is the constant giving the unconditional probability of staleness and $\gamma_t^{(p)}$ is the deterministic function of time which provides the inverse U-shape pattern. $\gamma_t^{(p)}$ is defined as $\gamma_t^{(p)} = \frac{\gamma_0}{\gamma_t}$, where γ_t is the intraday effect in volatility (3.3) and γ_0 is the normalizing constant which ensures that $\int \gamma_t^{(p)} = 1$ (the integral is taken over one trading day), so on average the daily number of zeros is close to p^U. For better illustrating the effect of staleness we consider different values of the unconditional probability p^U.

We start by looking at two paradigmatic examples, namely the BNS test of Barndorff-Nielsen and Shephard (2006) and LM test of Lee and Mykland (2008). The BNS test is defined, as suggested in Huang and Tauchen (2005), as:

$$\mathsf{BNS} = \frac{1 - \frac{\mathsf{MV}(X;[1\ 1])}{\mathsf{MV}(X;[2])}}{\sqrt{\frac{\theta}{n} \max\left(1, \frac{\mathsf{MV}(X;[4/3\ 4/3\ 4/3])}{\mathsf{MV}(X;[1\ 1])^2}\right)}}, \tag{3.5}$$

where $\theta \approx 2.61$. Intuitively, the presence of zeros due to staleness inflates the numerator since, while realized variance is unbiased, bipower variation is downard biased. Additionally, staleness deflates the denominator since the quarticity estimate is also downard biased. The joint combination of these two effects inflates the BNS test dramatically, distorting it toward rejection of the null, with the distortion increasing with the unconditional probability of staleness. The solution to this issue is to use the staleness-robust multipower estimators in place of standard ones in (3.5).

The LM test represent a different class of jump detection methodologies including the tests of Lee and Mykland (2008), Andersen, Bollerslev, and Dobrev (2007), and Lee and Hannig (2010), based on comparing the magnitude of the absolute value of the returns to an estimate of spot volatility. The statistic $\mathcal{L}(i)$, which tests at time t_i whether there was a jump on an interval (t_{i-1}, t_i), is defined as:

$$\mathcal{L}(i) = \frac{\Delta_i X}{\frac{1}{K-2} \sum_{j=i-K+2}^{i-1} |\Delta_j X|\,|\Delta_{j-1} X|}, \quad i = K, \dots, n. \tag{3.6}$$

where K is a tuning parameter determining a window size over which the instantaneous volatility is estimated. As soon as zeros appear among the K increments preceding t_{i-1}, the spot volatility in the denominator of $\mathcal{L}(i)$ is underestimated. Hence, under staleness, many of $\mathcal{L}(i)$'s (and, as a consequence, $\max_i (\mathcal{L}(i))$) are distorted toward rejection of the null. We again correct by replacing the bipower estimator in the denominator of (3.6) with its corrected version.

Figure 2 clearly shows the danger of ignoring zeros. Both the BNS and the LM test statistics are heavily distorted. For the BNS test, with an average staleness of 10% we have 25.92% of

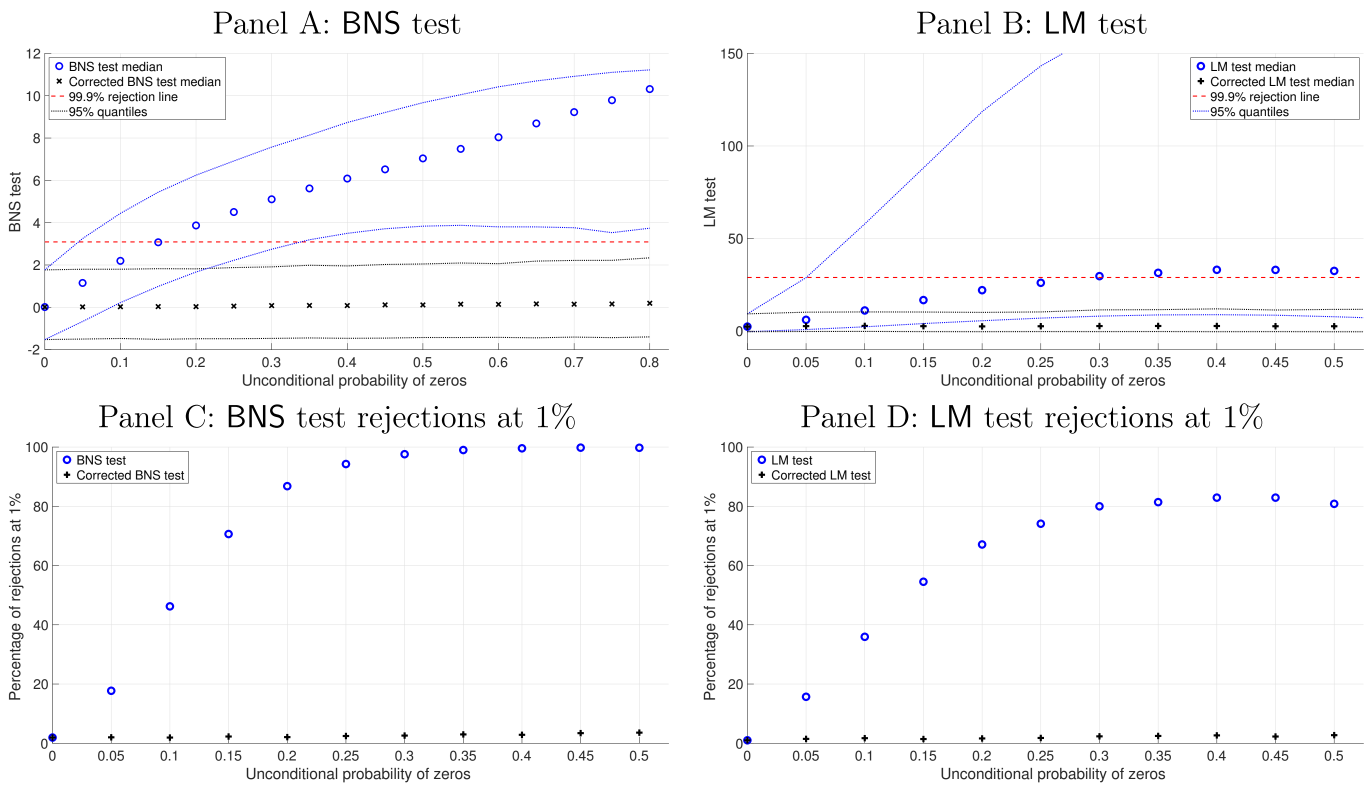

Figure 2: Quantiles (median, 5% and 95%) of daily jump tests in their original and corrected version, with the 99.9% rejection line, on simulated 5-minute returns, as a function of the generated probability of staleness. Panel A: BNS test. Panel B: LM test. Panel C and D report rejection rates of BNS test and LM test respectively at confidence level 1%.

false positives; and a staleness larger than 40% would imply a detection of spurious jumps almost automatically. The situation for the LM test is not any better. On the other hand, the corrected tests in which staleness is fully taken into account present no distortions at both frequencies, even for extremely large probability of staleness, with a median centered in zero and a distribution close to a standard normal.

This staleness-induced distortion toward spurious jump detection is typically shared by other jump testing strategies. Figure 3 compares the distributions of different tests estimated on our simulations with 10% of zeros. These tests are directly comparable since their asymptotic distribution should be standard normal under the assumed null in which jumps are absent. In addition to the BNS, reported as benchmarks, we consider: i) the ADS median test proposed by Andersen, Dobrev, and Schaumburg (2012); ii) the ASJ test of Aït-Sahalia and Jacod (2009b), which uses the ratio of power variations at different frequency; iii) the PZ test of Podolskij and Ziggel (2010), which is based on truncated variation; iv) the DGHM test of Dovonon, Gonçalves, Hounyo, and Meddahi (2019), which is based on wild bootstrap of bipower variation.[3] The ADS test compares

[3] The ADS test we implement in Figure 3 is based on Eq. (6) in Andersen, Dobrev, and Schaumburg (2012),

realized variance with an estimator of integrated variance based on the median of a number of consecutive absolute intraday returns. The median estimator was purposely proposed to deal with outliers. Indeed, as shown in Figure 3, the test is less distorted than the BNS test. However, under staleness, even the median estimator is biased for the same reasons of multipower variation. Consequently, ADS is still distorted toward the rejection of the null, even if less than BNS. The ASJ test procedure in Aït-Sahalia and Jacod (2009b) is contaminated by staleness as well. Under staleness, the bias appears in both the numerator and the denominator of the proposed test statistics. Since the bias of power variations is larger for the large powers, the ASJ test statistic is still distorted. The PZ test is based on the difference between power variation and its truncated version, and makes use of random perturbation of returns (wild bootstrap) in order to set up a confidence region. Thus, despite of the bias in the power variation, the distribution of PZ test remains standard normal under the null by construction. Similarly, the randomness of the wild-bootstrap DGHM test comes from the artificially generated random variables, which preserves normality under the null. Indeed, Figure 3 shows that the distribution of the wild-bootstrap tests under the null is as close to standard normal as the distribution of the corrected BNS test, if not closer.

Distortions due to staleness similarly affect alternative popular jump tests which were not reported here, and which are based on multipower variations as well. The test proposed by Corsi, Pirino, and Renò (2010), who combine multipower variation and truncation, suffers of exactly the same bias of the BNS test. Alternative tests proposed by Andersen, Dobrev, and Schaumburg (2012) are based on the $minRQ$ and $minRV$ estimators, and are more distorted than that based on the median estimators and reported in Figure 3. Using simulations, Theodosiou and Zikes (2011) analyze the performance of several jump tests under staleness, also showing that all the ones they consider, including the tests mentiond here and the test based on variance swaps of Jiang and Oomen (2008), are distorted toward finding more jumps, with the exception of the PZ test.

We now consider estimation of the Jump Activity Index (JAI) using two different multipower estimators. The first is a modification of Todorov and Tauchen (2010) estimator used in Andersen, Bondarenko, Todorov, and Tauchen (2015) that makes use of power variations computed over different frequencies. The second is introduced in Kolokolov (2022), and is based on the ratio of

where IQ is estimated with the $MedRQ$ estimator. The ASJ test is implemented with power variation with power $p = 4$ calculated over the two scales (1/78 and 2/78); the variance of ASJ test statistics is estimated without truncating large returns, since we do not simulate jumps under the null. The PZ test is implemented with power $p = 2$, and the test statistics is perturbed by random draws from the distribution $\mathcal{P}^\eta = \frac{1}{2}(\delta_{1-\tau} + \delta_{1+\tau})$, where δ is the Dirac measure and $\tau = 0.05$. The threshold used in the PZ test is computed as $c_\theta^2 \widehat{V}_t$, where $\widehat{V}_t$ is an estimator of the local standard deviation computed as in Corsi, Pirino, and Renò (2010) with bandwidth parameter $L = 10$. In this simulation exercise, we use $c_\theta = 5$.

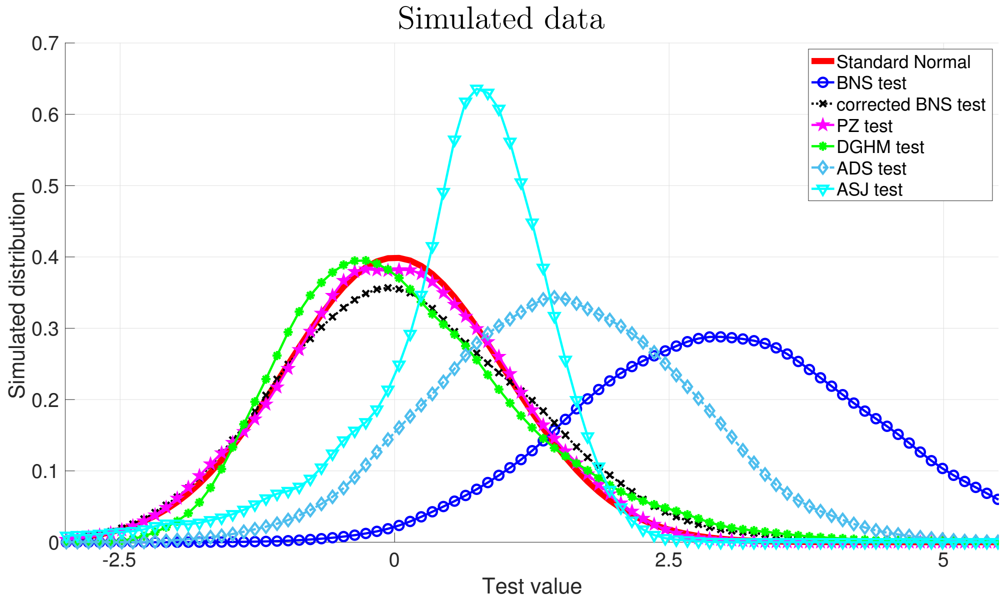

FIGURE 3: Distribution of test statistics on simulated daily samples at the five minutes frequency ($n = 78$ returns). The considered tests are: the median test in Andersen, Dobrev, and Schaumburg (2012) (ADS); the BNS test in Eq. (3.5) (BNS); the Podolskij and Ziggel (2010) test (PZ); Dovonon, Gonçalves, Hounyo, and Meddahi (2019) test (DGHM); the Aït-Sahalia and Jacod (2009b) test (ASJ), and the corrected BNS test.

multipowers with different powers. Both estimators are downward biased in the presence of zeros. Again, our solution to this issue is simply to replace power/multipower variations in the definitions of the estimators by the corrected versions.

For comparison, we additionally consider estimators which do not use multipowers: the Aït-Sahalia and Jacod (2009a) estimator based on counting the overshoots over threshold; the Jing, Kong, and Liu (2012) estimator based on counting small increments; and the Hounyo and Varneskov (2017) wild-bootstrap estimator. We compare the estimators on simulated data in which jumps are absent and the true JAI is 2. The left panel of Figure 4 shows the distribution of JAI estimates using the original estimators in the presence of zeros. It shows that ignoring the presence of staleness would result in an artificially lower JAI estimate, which would induce to reject the presence of a Brownian motion in the vast majority of replications. The effect is evident not only for multipower-based, but for all considered estimators. The right panel of Figure 4 shows the distribution of corrected estimators. It shows that, when correcting for staleness using rescaled returns, the estimates are symmetrically centered around 2 across replications. The right panel also shows that the corrected Kolokolov (2022) estimator is the most precise. Therefore, we use it as our main tool for JAI inferencing in the empirical section.

Price staleness can bias not only for JAI estimation, but also similar tests for pure-jumps processes.

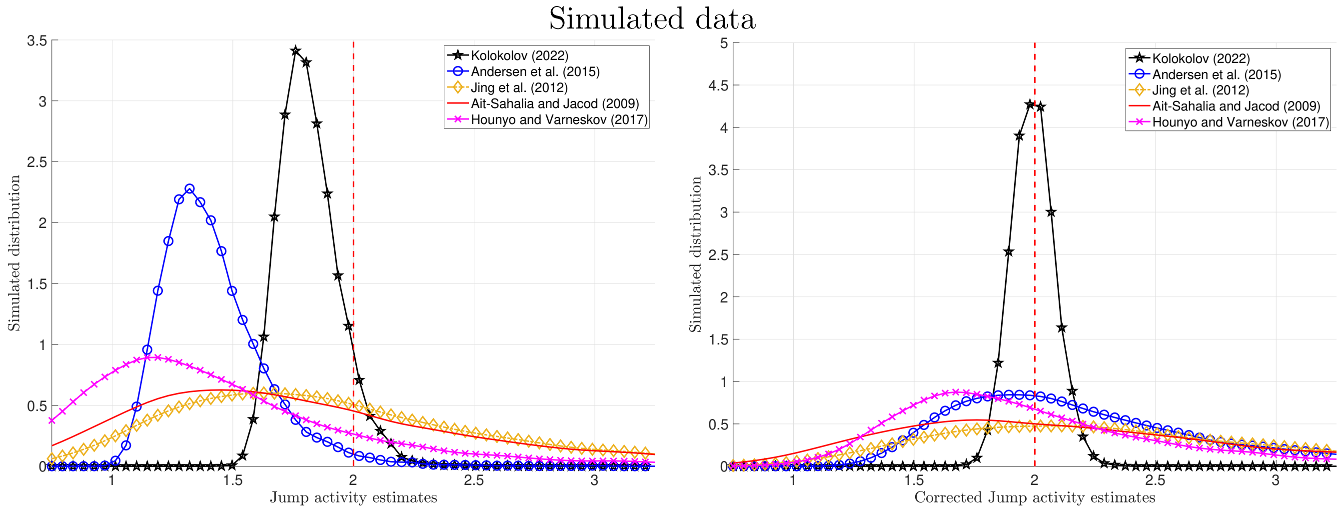

FIGURE 4: Distribution of the JA estimates on simulated data in which the jump activity is equal to 2. Left panel: the estimators under the presence of zeros. Right panel: corrected estimators. The considered estimators are: Andersen et al. (2015), Kolokolov (2022), Aït-Sahalia and Jacod (2009a), Jing, Kong, and Liu (2012) and Hounyo and Varneskov (2017). All estimators are implemented on one day of five-minutes returns ($n = 78$).

For example, we found that zeros severely distort the test of Jing, Kong, and Liu (2012) toward the rejection of the Brownian motion, while the test by Kong, Jing, and Liu (2015) is robust to the presence of staleness. The latter test, however, requires sampling on a substantially higher frequency than the one considered in our paper, therefore it is not used for the empirical analysis. The corresponding simulations are available upon request.

4 Alternatives to multipower

The econometric literature suggests a number of alternative jump robust estimators of integrated variance, which might be affected by jumps and/or staleness less than multipower variation. Here we analyze threshold realized variance (Mancini, 2009) and robust neighborhood truncation (Andersen, Dobrev, and Schaumburg, 2014).

Threshold realized variance is defined as

$$\mathsf{TRV}(X) = \sum_{i=1}^{n} |\Delta_i X|^2 \, \mathbf{1}_{\{|\Delta_i X| \leq \vartheta(\Delta_n)\}}, \tag{4.1}$$

where $\mathbf{1}_{\{.\}}$ is the indicator function and $\vartheta(\Delta_n)$ is the threshold function satisfying $\lim_{\Delta_n \to 0} \vartheta(\Delta_n) = 0$ and $\lim_{\Delta_n \to 0} \frac{\Delta_n \log(1/\Delta_n)}{\vartheta(\Delta_n)} = 0$. The difference between $\mathsf{TRV}(X)$ and the realized variance of the continuous part of X is asymptotically negligible. Consequently, $\mathsf{TRV}(X)$ remains consistent under the presence of staleness. However, it is strongly negatively biased by

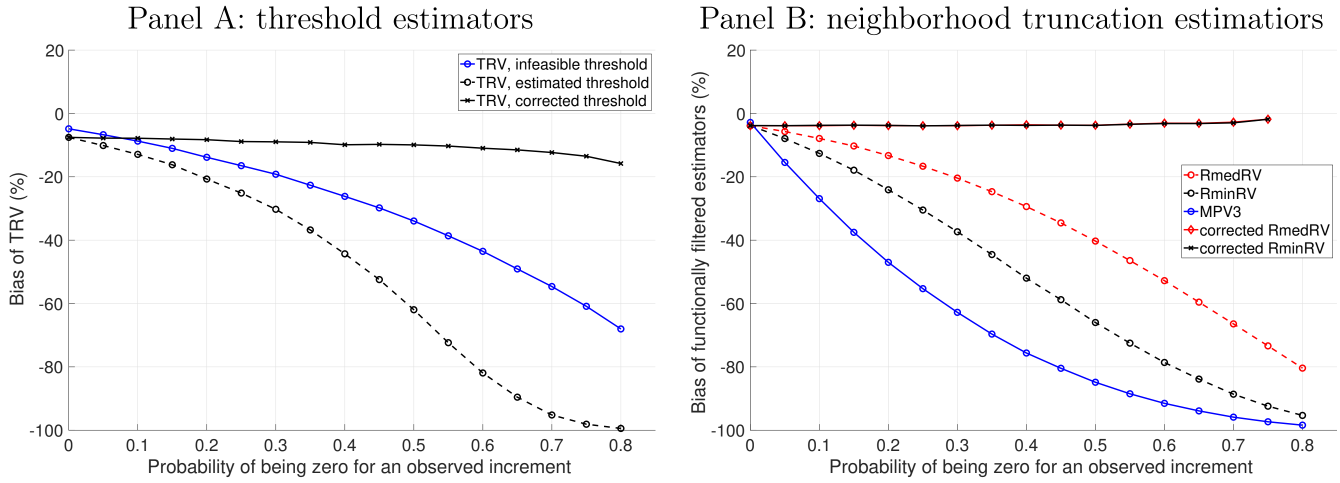

FIGURE 5: Alternative jump robust integrated variance estimators under the presence of zeros. Panel A: threshold realized variance. Infeasible $\mathsf{TRV}(X)$ is implemented with the time-varying threshold $3 \cdot \sigma_{i\Delta_n} \cdot \sqrt{\Delta_n}$. The threshold of feasible $\mathsf{TRV}(X)$ is equal to $3 \cdot \widehat{\sigma}_{i\Delta_n} \cdot \sqrt{\Delta_n}$, where $\widehat{\sigma}_{i\Delta_n}$ is estimated as in Corsi, Pirino, and Renò (2010). $\mathsf{TRV}^c(X)$ is implemented with $\widehat{\sigma}_{i\Delta_n}$ estimated by applying the algorithm of Corsi, Pirino, and Renò (2010) to rescaled returns. Panel B: robust neighborhood truncation estimators (RminRV and RmedRV) implemented as in Andersen, Dobrev, and Schaumburg (2014), tripower variation with equal powers (denoted by $MPV3$) and corrected RminRV and RmedRV computed using rescaled returns.

zeros in finite samples. The increments of X observed after a series of zeros include the previous unobserved increments of the efficient price. Consequently, they have larger variance and more likely exceed the threshold (which is fixed in practice). As a result, such increments are eliminated from the sum lowering the value of $\mathsf{TRV}(X)$. The negative bias can be even stronger when the threshold is proportional to the estimated local standard deviation (which is a common strategy of implementing truncated estimators in practice), since in this case the threshold is spuriously undervalued due to the negative bias of local volatility estimators.

Robust neighborhood truncation estimators adopt a filtering scheme truncating arbitrary functionals on return blocks. We consider a pair of such estimators (denoted by RminRV and RmedRV) applied with one-sided filtering as in the simulations in the original paper. Note that these estimators are optimized for the setup of Andersen, Dobrev, and Schaumburg (2014), which uses ultra-high frequency data and differs from our settings by the absence of staleness at moderate frequencies. The presence of price staleness generates an artificial time-change distorting the local normality of the data and producing a negative bias for both estimators.

To soften the bias of $\mathsf{TRV}(X)$ the threshold values corresponding to the increments observed after a series of zeros ought to be rescaled by the number of previous zero returns, so that the corrected threshold realized variance we propose is defined as:

$$\mathsf{TRV}^c(X) = \sum_{i=1}^{n} |\Delta_i X|^2 \, \mathbf{1}_{\left\{|\Delta_i X| \le \left(\mathbf{n}_i^Z + 1\right) \cdot \vartheta(\Delta_n)\right\}}, \tag{4.2}$$

where $\mathbf{n}_i^Z$ denotes the number of consecutive zeros observed prior to the instant $i\Delta_n$. When local volatility estimates are used to compute the threshold, they ought to be corrected for the presence of zeros as suggested in Section 2. To correct robust neighborhood truncation estimators we simply apply them to the time series of non-equispaced returns computed as in Section 2 and rescaled by the number of the preceding zeros.

Panel A of Figure 5 illustrates the bias of $\mathsf{TRV}(X)$ on the same simulated data of Section 3. The figure compares $\mathsf{TRV}(X)$ computed with the time-varying (infeasible) threshold equal to three standard deviations of the efficient returns, $\mathsf{TRV}(X)$ with estimated threshold computed as in Corsi, Pirino, and Renò (2010) and corrected $\mathsf{TRV}^c(X)$ with the threshold computed by applying the algorithm of Corsi, Pirino, and Renò (2010) to rescaled returns. All estimators are negatively biased due to truncating the largest returns. The (absolute) bias of $\mathsf{TRV}(X)$ substantially increases with the number of zeros. For example, already with 20% of stale prices, $\mathsf{TRV}(X)$ with feasible threshold underestimates integrated variance by 20%. The corrected estimator effectively eliminates the bias produced by staleness, however its variance increases with the number of zeros. Panel B shows the bias of RminRV and RmedRV. The absolute bias of the robust neighborhood truncation estimators is smaller with respect to the tripower variation. However, it is comparable with the bias of $\mathsf{TRV}(X)$. Applying the estimators to the rescaled returns allows to eliminate the bias.

Overall, this section suggests that alternative estimators can be affected by zeros similarly to multipowers. The main reason for that is the deviation from local normality of equally spaced returns generated by price staleness. We leave further theoretical considerations (in particular on the joint effect of staleness and other microstructure frictions present at ultra-high frequencies) for future research.

5 The impact of price discretization

A relevant source of zero returns which is not considered in our model is price discretization. Including rounding in the data generating process is a notoriously challenging problem (Delattre and Jacod, 1997; Li and Mykland, 2015). Bandi, Kolokolov, Pirino, and Renò (2020) argue that the number of observed zeros due to this specific friction depends on the "Rounding Impact Ratio"

(RIR):

$$\mathsf{RIR} = \frac{d}{\sigma_P \cdot \sqrt{\Delta_n}}, \tag{5.1}$$

where d is the magnitude of the price discreteness (e.g. $d = \$0.01$ for NYSE stocks), σ_P is the volatility of price differences (price matters here since discreteness is for prices, not for log-prices). In the definition (5.1) σ_P is assumed to be constant, but when measuring RIR on data in Section 6.1 we will relax this assumption. The RIR is an index without units with a simple interpretation: it compares the width of price discretization with the volatility of the price difference, measuring the relative strength of price discretization on observed returns.

We study the impact of rounding on multipower-based quantities in Figure 6 and Figure 7. We do so with a model that rounds to the nearest cent ($d = 0.01$) the prices simulated with the model of Section 3, again at the 5-minute frequency. The unconditional probability of zeros generated by the Bernoulli variates is fixed to 10% in Figure 6 and to 30% in Figure 7. We consider different levels of rounding ranging from zero to extremely high levels. As reference values, in our individual stocks the median value of RIR is 22.46%, and RIR index is greater than 100% in 2.6% of the daily time series, as we report in Section 6.1, see, e.g. Figure 9. In order to control the strength of rounding, we use different starting values for the efficient price, which are calibrated in such a way that the average value of RIR estimated from this simulated data is equal to reported values.

Figure 6 shows that rounding has a non-negligible impact on $\mathsf{RV} = \mathsf{MV}(X;[2])$, $\mathsf{BPV} = \mathsf{MV}(X;[1\ 1])$ and $\mathsf{TriPV} = \mathsf{MV}(X;[4/3\ 4/3\ 4/3])$, and in particular that it generates a positive bias on corrected estimator which is increasing with the RIR. Moreover, corrected estimator also display a larger variance with respect to non-corrected ones. This finding complements the results of Li and Mykland (2015) who explore the effect of rounding on realized variance. One can see that corrected estimators can be less precise than traditional ones (in the relative RMSE sense) when RIR is larger than a given threshold and staleness is low. When staleness is high, instead (as shown in Figure 7), corrected estimator are more precise than traditional ones even in cases in which rounding is extremely aggressive.

However, Figure 6 also shows that corrected estimators are better than non corrected ones for realistic value of RIR and even when staleness is low for three relevant applications: jump testing, measurement of the quadratic variation due to jumps, and jump activity index estimation. For jump testing (Panel D), the size of the non-corrected BNS test is distorted (as discussed above), and the size distortion grows with RIR, while the corrected BNS test has a conservative size. This advantage for the corrected BNS test comes at the cost of power (displayed in the inset of Panel D, obtained by adding to simulated price trajectories a single jump with uniform location and size

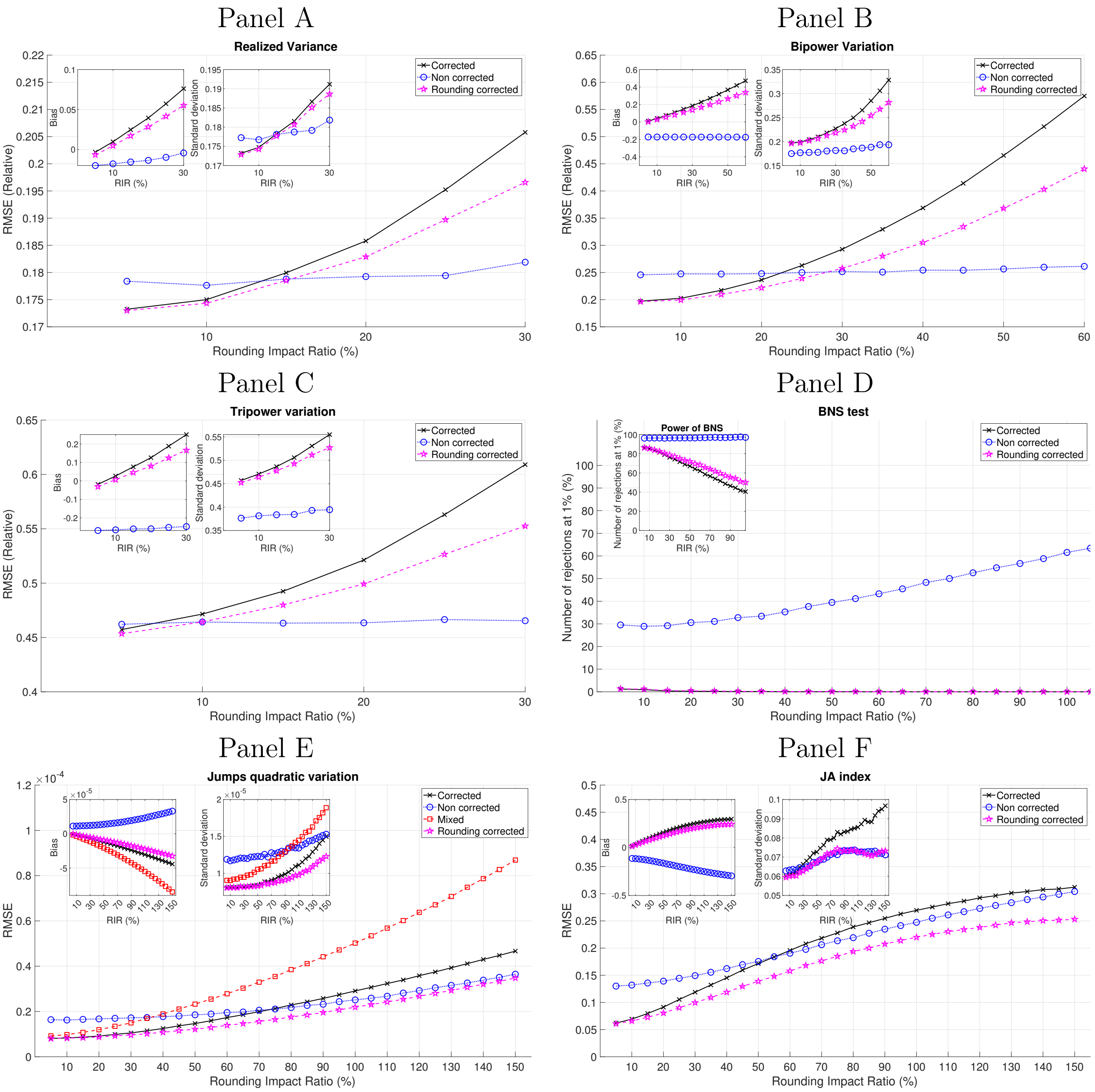

Figure 6: The joint impact of rounding and staleness on estimates based on multipower, corrected multipower and rounding corrected multipower on simulated data. The rounding corrected estimators are defined in Appendix C. Panel A: Relative RMSE, bias and standard deviation for RV. Panel B: Relative RMSE, bias and standard deviation for BPV. Panel C: Relative RMSE, bias and standard deviation for TriPV. Panel D: jump size (at 1%) and power of the BNS test. Panel E: RMSE, bias and standard deviation for the JV = RV − BPV. "Mixed" refers to the case in which we correct BPV but not RV. Panel F: RMSE, bias and standard deviation for the Kolokolov (2022) estimator of the jump acrivity index. Biases and standard deviations are in the insets. The price impact ratio is defined on simulations as RIR $= d/(<\sigma> \cdot P \cdot \sqrt{\Delta_n})$, where $d = 0.01$, $<\sigma>$ is the average generated volatility across all simulations expressed in daily units, P is the starting price of the simulation and $\Delta_n = 1/78$. The unconditional probability of staleness is 10%.

$10\sigma_J\sqrt{\Delta_n}$, where σ_J is the simulated volatility at the time of the jump), which, for the corrected BNS test, decreases when the impact of rounding increases.

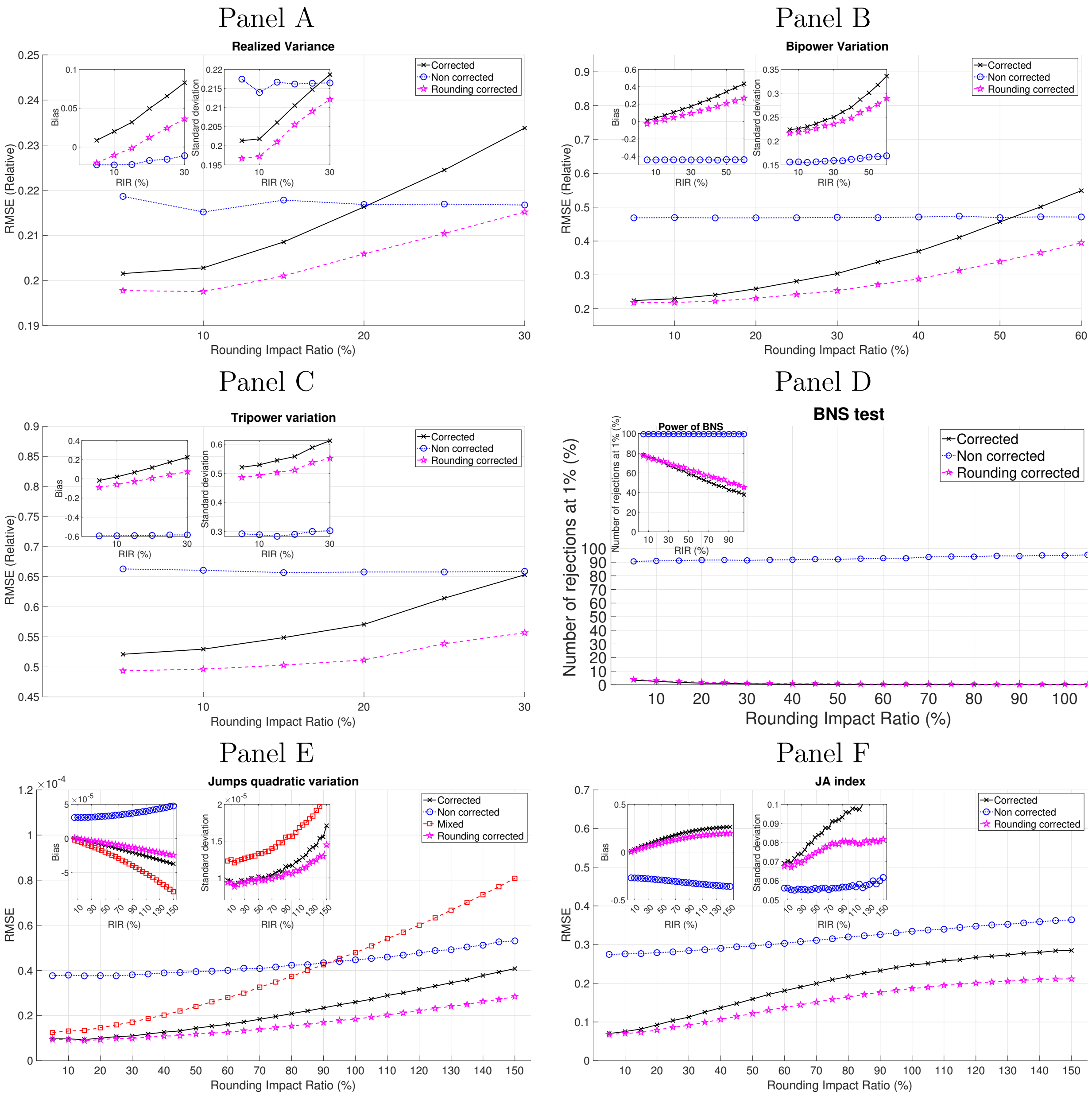

Figure 7: Same as Figure 6, but with unconditional probability of staleness set to 30%

When measuring the quadratic variation due to jumps (Panel E), using $\mathsf{JV} = \mathsf{RV} - \mathsf{BPV}$ (which, in our simulations, is zero), non-corrected estimators display a positive bias increasing with RIR, while corrected estimators display a negative bias, also increasing (negatively) with RIR. However, the bias in the difference due to rounding cancels out more efficiently for corrected estimators, which display a lower standard deviation. As a result, corrected JV is better than non-corrected JV (again, in terms of relative RMSE) for RIR up to 70% when staleness is 10%, and up to more than 150% when staleness is 30%. Given the large bias induced by rounding on corrected RV displayed in Panel A, Panel E also shows a mixed case in which we compute JV after correcting BPV but not RV. We see that this technique is worst than correcting both quantities. We need the biases in both quantities to cancel out to measure jump quadratic variation precisely.

Finally, when measuring jump activity index, the corrected estimator are better than the non-corrected ones for values of RIR up to 50% when staleness is 10%, and up to more than 150% when staleness is 30%. They display a positive bias and a standard deviation which both increase with RIR. The bias is however smaller than the negative one of the non-corrected estimator, so that the RMSE with respect to the value generated in the simulations (equal to 2) tends to be either smaller or comparable.

When the impact of rounding is large, the scaling used in corrected estimators (motivated by linearity of the variance of the Brownian motion) could be excessive. This does not have impact on relevant jump statistics, for which corrected estimators remain superior to the standard ones in realistic situations. However, one may wonder how to recover consistency of multipower estimators in the joint presence of staleness and rounding. We now show that we can improve our corrected estimators in the presence of rounding, by adopting a non-linear scaling depending on RIR and estimated staleness. This non-linear scaling is more appropriate (but more difficult to implement) in situations in which the rounding is strong. The non-linear scaling is detailed in Appendix C.

Figure 6 shows that the proposed heuristic correction, labelled "rounding corrected", improves our correction for all considered estimation targets and tests. In particular, using this correction delivers more precise estimators of the jump activity index and jump quadratic variation than the traditional multipowers for all the rounding levels, even in this case with low staleness. The comparison is even more favorable for our estimators when staleness is higher (as shown in Figure 7).

However, some words of caution are needed. First, in our simulation setting we assume we know the tick size d and the probability of staleness $p^{\emptyset}$, while in practice both have to be observed or estimated. Second, our correction is heuristic in the sense that, for sake of simplicity, it is based on several approximations which are detailed in Appendix C. When RIR in the data is low, we recommend to use the estimator (2.10). Notice that RIR can be made arbitrarily low by decreasing the sampling frequency, at the cost of increasing the estimator variance, generating the usual bias-variance tradeoff similar to that generated by the presence of market microstructure noise (Bandi and Russell, 2008). The optimization of this bias-variance problem due to the joint presence of rounding and staleness is left for future research.

6 Empirical application

The data set we use here is the collection of 442 stocks belonging to the S&P500 index and quoted on the New York Stock Exchange (NYSE). The sample ranges from January 2, 2003 to March 10, 2022, for a total of $5,087$ trading days, and consists of one-minute transaction prices recorded from 9:30 to 16:00, New York GMT. Moreover, we use prices of SPY (the SPDR S&P 500 trust exchange-traded fund) and VIX (the CBOE Volatility Index) sampled at the one-minute frequency in the same daily time span (these two data sets extend to June 30, 2022). The data, coming from multiple exchanges and electronic networks, went trough a standard filtering procedure. Data are cross-checked, tested and verified so that outliers and bad ticks are removed. In our applications, we use 5-minute grids. We trim data at the beginning and at the end of the day with zero volume, so that we do not include zero returns resulting from delayed opening or anticipated market closure. We compute the quantities of interest only in days in which we have more than 20 returns and the percentage of zero returns is less than 80%. This leaves us with $N = 2,013,613$ daily time series (plus $4,908$ on SPY and $4,909$ on VIX).

Figure 8 shows the time series of the daily fraction of zero returns in our datasets at the 5-minute frequency. The Figure shows that, consistently with markets becoming more liquid, the fraction of zero returns in the data was declining over time. However, the Figure also shows that the fraction of zero in the data is relevant, and still not negligible, in current times.

6.1 Staleness versus price discreteness: empirical evidence

The discussion in Section 5 makes clear that it is important, before applying the proposed correction to multipower estimators, to quantify the amount of zeros in the data due to price discreteness. On this regard, Bandi, Kolokolov, Pirino, and Renò (2020) make two important points. The first one is of empirical nature: the amount of staleness in the data, on the top of price discreteness, is substantial at all frequencies, such that it cannot be ignored. The second is of theoretical nature: the economic interpretation of zeros due to price discreteness is similar to that due to staleness. Indeed, price discreteness bites when volume is low. If staleness is a signature of lack of trading, then price discreteness is also a signature of lack of trading, in the sense that volume is not sufficient to move the price of more than one tick. They refer to this idea as *near idleness*. This idea, not considered in our simulation setting, suggests that the impact of rounding is probably less than that discussed in Section 5.

To assess the aggressiveness of price discretization, we estimate the Rounding Impact Ratio for

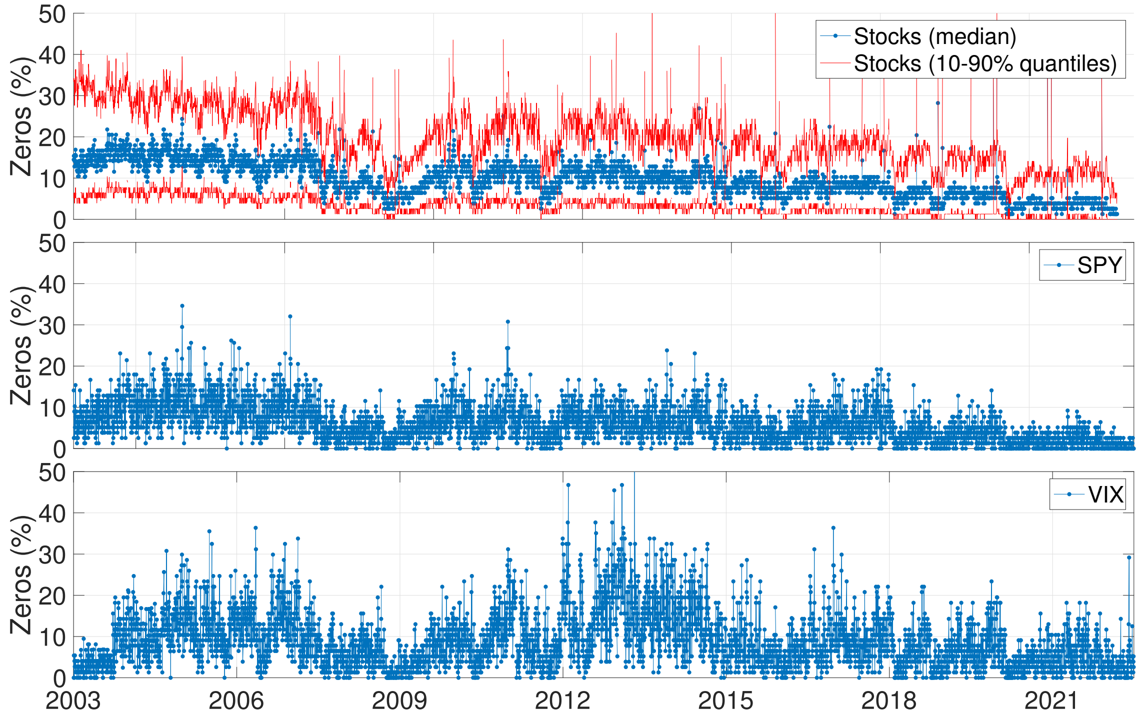

Figure 8: Time series of the daily fraction of zero returns in our datasets from January 2003 to June 2022 at the sampling frequency of 5 minutes. Top panel: stocks (we report the median daily value and the 10%-90% quantiles across stocks). Center panel: SPY. Bottom panel: VIX.

each stock and each day via

$$\widehat{\mathsf{RIR}} = \frac{\hat{d}}{\sqrt{\Delta_n \sum_{i=1}^{n} (P_{i\Delta_n} - P_{(i-1)\Delta_n})^2}}, \tag{6.1}$$

where $\Delta_n = 1/n$, n is the number of 5-minute returns used in a day, $P_{i\Delta_n}$ is the price at time $i\Delta_n$ and $\hat{d}$ is an estimate of the discreteness level. The denominator in Eq. (6.1) estimates price volatility over the day. The discreteness level $\hat{d}$ is estimated, for each day and stock, as follows. We first compute the absolute values of price differences at the highest available frequency (in our case: 1-minute) in the considered day and in the previous 5 days. We then round these absolute price differences to the sixth significant digit (to cluster returns whose difference is only due to numerical precision). The estimate $\hat{d}$ is the most frequent non-zero outcome of this procedure. While for our data-sample the NYSE minimum tick is $d = \$0.01$, we prefer to estimate it for two reasons: i) our prices are adjusted for corporate actions; ii) there could be larger rounding levels because of several reasons, including larger bid-ask spreads or psychological rounding, especially for illiquid stocks. The estimated $\widehat{d}$ is exactly 0.01 in 64.85%, below 0.01 in 22.29%, and above

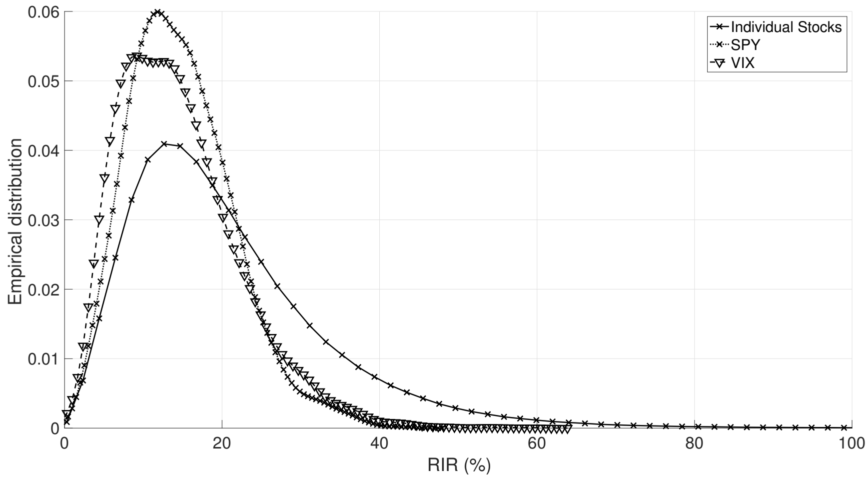

FIGURE 9: Pooled distribution, for all the considered S&P500 stocks, the SPY ETF and the VIX Index from January 2003 to June 2022, of the daily estimated Rounding Impact Ratio ($\widehat{RIR}$) as in formula (6.1), with a sampling frequency of 5 minutes.

0.01 in 12.87% of the days.

Figure 9 shows the distribution of the daily $\widehat{\mathsf{RIR}}$ at the 5-minute frequency in our three samples. We can see that price discreteness is different from zero, especially for individual stocks, but not large. The median value for individual stocks is 17.97%, and for this sample $\widehat{\mathsf{RIR}}$ is greater than 100% in 0.17% of the daily time series only. On the two other datasets, the median $\widehat{\mathsf{RIR}}$ is 14.14% for SPY and 13.38% for VIX. With these estimates of the RIR values, our simulation experiments suggest that corrected estimators are virtually always more precise than traditional ones.

Next, we estimate the amount of price staleness in the data using the $\widehat{p}_n$ estimator proposed in Bandi, Kolokolov, Pirino, and Renò (2020), which is based on matching discrete counting at each tick with zeros generated by the simple rounding model used in our simulations in which rounding is independent from staleness. We relate the number of observed zeros (which are due to both staleness and price discreteness) to estimated staleness and stock liquidity, proxied by the daily dollar traded volume.

Panel A of Figure 10 shows the scatter plot of the average of percentage of zeros for each stock against the logarithm of the average trading dollar volume. As expected, less intensely traded stocks display a higher percentage of zero returns. Panel B of Figure 10 shows the scatter plot of the average of estimate of the probability of staleness for each stock against the logarithm of the average trading dollar volume. It shows that staleness is clearly present in the data, more prominently, as expected, for stocks that are less liquid. The ratio between the average probability of staleness and the average percentage of zero returns, again against the logarithm

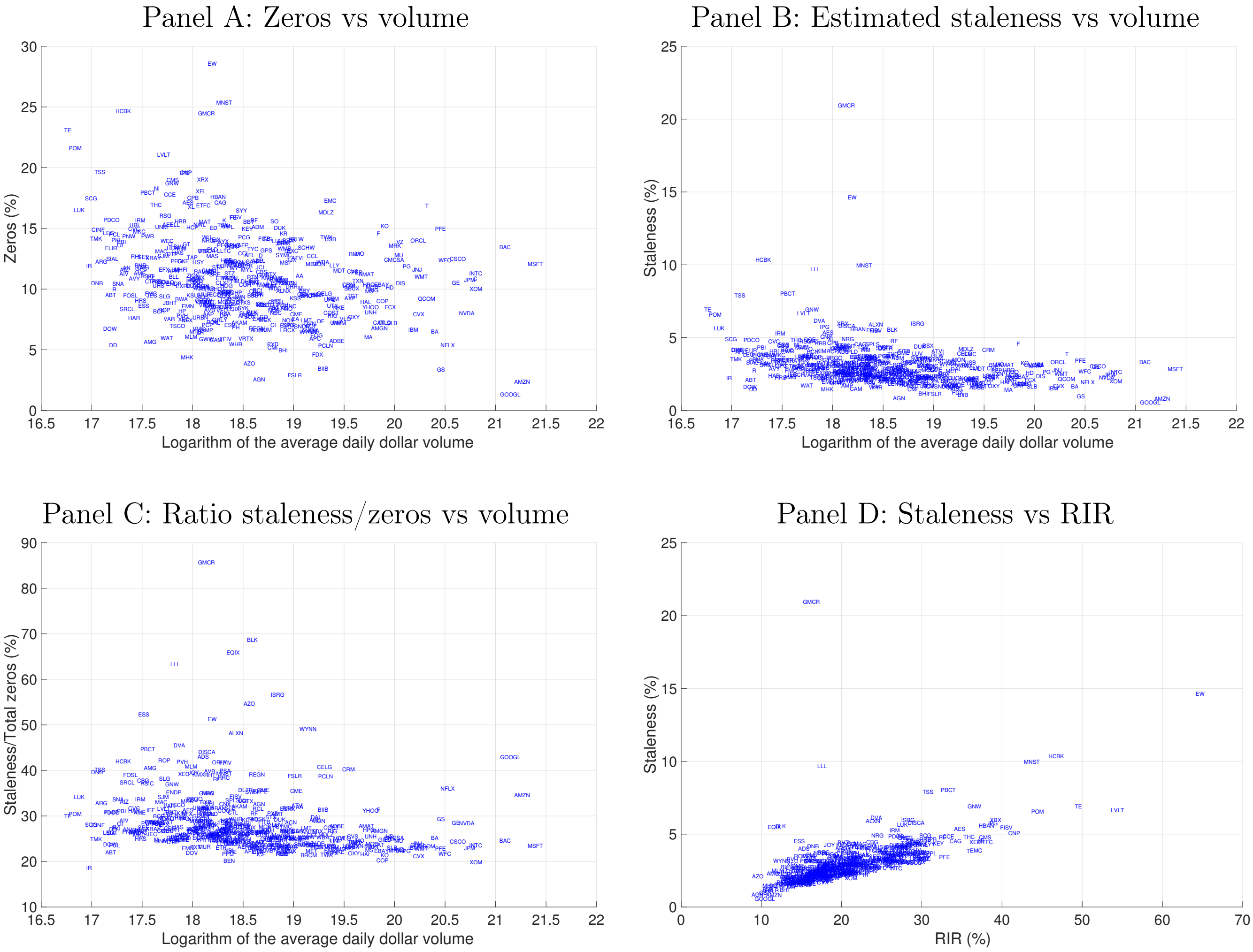

Figure 10: Panel A: average daily number of zero five-minute returns (per cent) as a function of logarithm of the average daily trading volume. Panel B: average daily estimates of the unconditional probability of observing a zero return due to price staleness, $\widehat{p}_n$, estimated as proposed in Bandi, Kolokolov, Pirino, and Renò (2020) using five-minute returns as a function of logarithm of the average daily trading volume. Panel C: ratio of the average estimate $\widehat{p}_n$ (estimated as in Panel B) and average daily number of zero five-minute returns as a function of logarithm of the average daily trading volume. Panel D: average daily estimates $\widehat{p}_n$ (computed as in Panel B) as a function of the average estimates of the $\widehat{\mathsf{RIR}}$ index, as computed by formula (6.1).

of the average trading dollar volume, is shown in Panel C. Staleness accounts for more than 20% of the observed zeros, with this percentage increasing as the stocks become less liquid. Moreover, Panel D, which displays the scatter plot of average estimated staleness with the average RIR, shows that the rounding and zeros are strongly correlated in the data, which is a phenomenon that can be explained, as argued, by near idleness. The reported estimates for the probability of staleness in Figure 10 thus need to be considered a lower bound on the economic phenomenon of absence of trading, since rounding is stronger when staleness is stronger, so many of zeros attributed to price discreteness would be attributed to staleness in a model that internalizes this correlation.

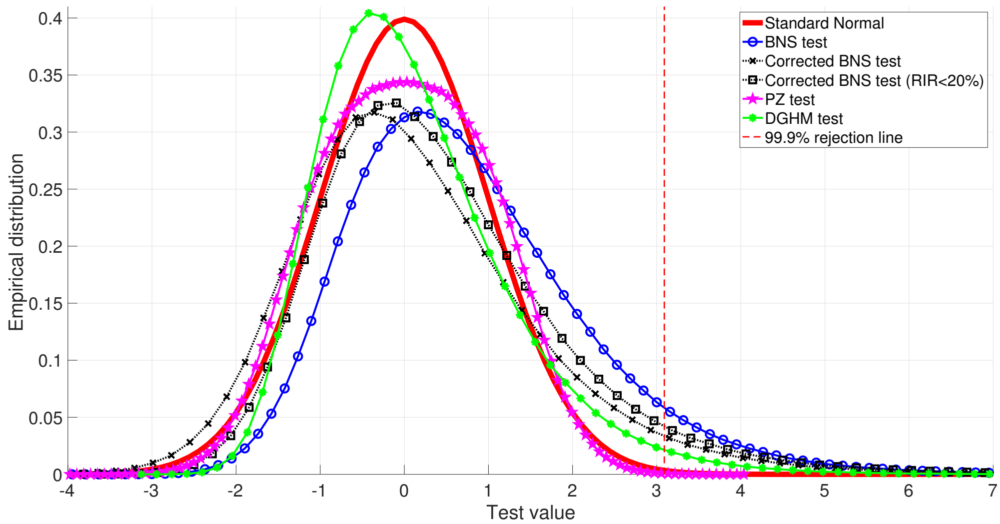

FIGURE 11: Pooled distribution, for all the considered S&P500 stocks from January 2003, to June 2022, of the daily BNS tests and corrected BNS^c tests using five-minutes returns. In total, we compute $2,013,613$ daily tests.

In a nutshell, this Section illustrated three important facts: the first is that price discreteness (as measured by the Rounding Impact Ratio) is not as large in the data as to compromise the proposed correction in the vast majority of cases. The second, already documented in Bandi et al. (2020), is that staleness is an important empirical component of the data, especially for illiquid stocks. The third, also already suggested in Bandi et al. (2020) and empirically documented here, is that zeros due to price discreteness are correlated with zeros due to staleness, suggesting that the two phenomena are driven by similar economic forces. From an econometric perspective, this explains why the proposed correction (devised in a model which accounts for staleness only) is still useful in the presence of price discreteness.

6.2 Revising the evidence on jumps in financial markets

Figure 11 shows the distribution of the daily BNS test and of its version corrected for staleness pooled across days and stocks. It also shows the DGHM and the PZ test. BNS test distribution is centered around a positive value and largely skewed to the right. As a consequence, it reveals jumps in 6.14% of the cases at the 99.9% confidence interval. This is the classic puzzle of the BNS test (Huang and Tauchen, 2005), that is that jumps appear to be too many in the data. Our Monte Carlo experiments suggest that this happens not because prices are too active, but because they are too stale. The distribution of the staleness-corrected BNS test, also shown in Figure 11, is indeed more centered around zero and still skewed to the right, but with a less pronounced right tail. This distribution also presents a left tail which depends on rounding; indeed, the distribution of the corrected test on stock/days with RIR less than 20% does not display the left tail, consistently

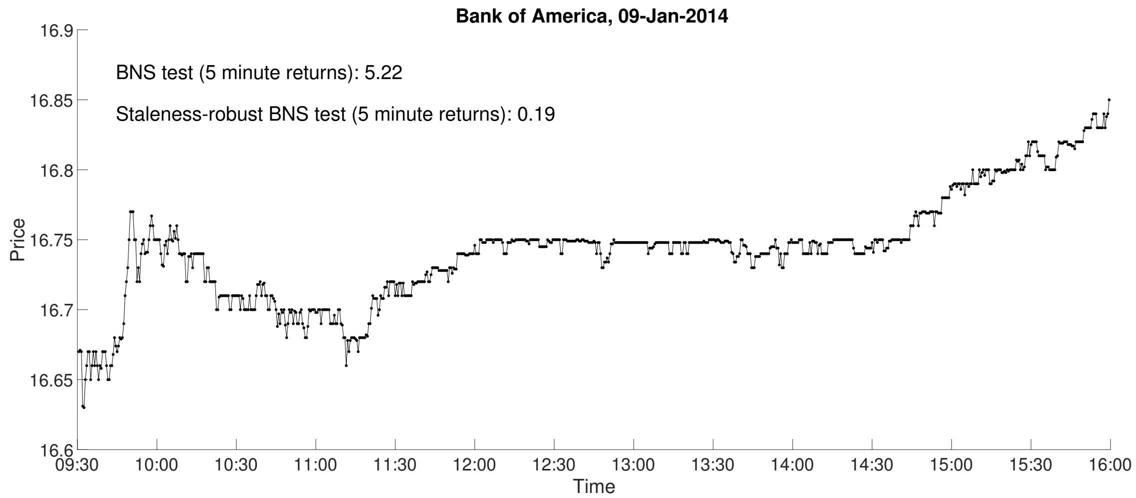

Figure 12: Time series of traded prices of the BAC stock on January 9, 2014. The standard BNS test would detect a jump at any reasonable confidence interval. The staleness-robust test we propose would not.

with our simulations. The corrected BNS test would detect jumps in just 3.70% of the days at the 99.9% confidence interval.[4] The DGHM test reveals even less jumps on average (1.85% at 99.9%). The PZ test is very sensitive to the choice of the threshold (described in footnote 3). At 99.9%, it detects jumps in 6.81% of the days with $c_\theta = 5$, 2.26% of the days with $c_\theta = 6$, and 0.58% of the days with $c_\theta = 7$. These estimates for the number of jumps in the data should be considered an upper bound because of the multiple testing problem (Bajgrowicz, Scaillet, and Treccani, 2016). If we use their universal threshold $\sqrt{2 \log N} = 5.3977$, the percentage of detected jumps by the corrected BNS test would be 0.34%. Thus, jumps are a non-negligible feature of stock data; but they appear to occur much more rarely than what previously found.

Figure 12 shows an iconic example of the distortion of traditional tests. It shows a day in which the Bank of America (BAC) stock price moves extremely smoothly, the largest 5-minute price change being \$0.10 and 21.79% of the 5-minute returns being zero. No jump is visible, but the BNS test value is 5.22, with a p-value of 10^{-8}. As argued, the BNS test is tricked by staleness in the data. When we use corrected BNS, the value of the test is just 0.19, consistent with what we see in the Figure.

The relative contribution of jumps to total quadratic variation is overestimated by traditional bipower variation too. If we define $\mathsf{JV} = \mathsf{RV} - \mathsf{BPV}$ where RV is realized volatility and BPV is bipower variation, the average relative contribution JV / RV at the five minutes frequency would be estimated to be 8.63%. However, this is again an artificially inflated result due to staleness: when

[4] If we test using 1-minute data, the BNS test would reveal massive jumps (in 33.76% of the days), while the corrected BNS would detect jumps in 4.30% of the days, fairly consistently with the 5-minute results. However, the 1-minute data are also contaminated by market microstructure noise, which is absent from our theoretical analysis.

estimating the average relative contribution of jumps to total quadratic variation using staleness-robust multipowers, the outcome is just 1.16% unconditionally, and 4.35% if we consider only days with RIR below median. This result is in agreement with the fact that the jump contribution to quadratic variation is much smaller than what implied by traditional estimates, as well as with our simulation results. Using ultra-high frequency data, Christensen, Oomen, and Podolskij (2014) uncover the puzzling discrepancy between the amount of jump variation measured at different frequencies: they report a jump variation at the tick frequency which is roughly one sixth than what found at five minutes (1.3% versus 7.3% for DJIA constituents). They show that we need ultra-high frequency data to avoid a lot of sequential small moves being mistakenly found to indicate one large jump at sparser sampling frequencies. We show that a similar reduction in the contribution of jumps to price variation can be obtained when correcting for the presence of staleness.

We now turn to the question: is Brownian motion necessary to model high frequency data? The empirical literature is ambiguous on this point. Todorov and Tauchen (2011) support the presence of Brownian motion in the S&P500 futures, while Jing, Kong, and Liu (2012) reject the presence of Brownian motion on one year of Microsoft high frequency prices. On the other hand, Todorov and Tauchen (2011) find that the JAI of the VIX index is significantly smaller than 2, advocating that the volatility process is a pure jump process without Brownian motion. Indeed, as they show in their paper, the activity of the VIX is the same as that of the volatility of the stock index.

Our results strongly support the presence of Brownian motion in stock prices.[5] Figure 13 (panel A) shows the empirical distribution of JAI estimates on individual stocks and SPY for every day in our sample with RIR below median (roughly 1 million estimates). We use the Kolokolov's estimator with $r_+ = 1.5$, which is more precise than the two-scale estimator of Andersen, Bondarenko, Todorov, and Tauchen (2015), as shown in Figure 4. Figure 13 shows very clearly, and pervasively, that the JAI of stock prices, as well as that of the index, does not depart significantly from the value of 2, supporting the omnipresence of the Brownian motion in the price dynamics.

Do we really need to dispense with the Brownian motion when modeling volatility? To answer this question, we now consider five-minute returns on the VIX index. We only consider days with more than 60 observed daily 5-minute returns. The data for VIX, as in the case of individual stocks, are heavily contaminated by the presence of staleness: on average, the daily fraction of zero returns at the five-minute frequency is 8.80%, a figure which is consistent with the amount of

[5]The test of Jing, Kong, and Liu (2012) boils down to counting the excessive number of returns below a predetermined threshold. Of course, the presence of zeros impacts severely their test toward rejection of the Brownian motion null, as discussed in the introduction.

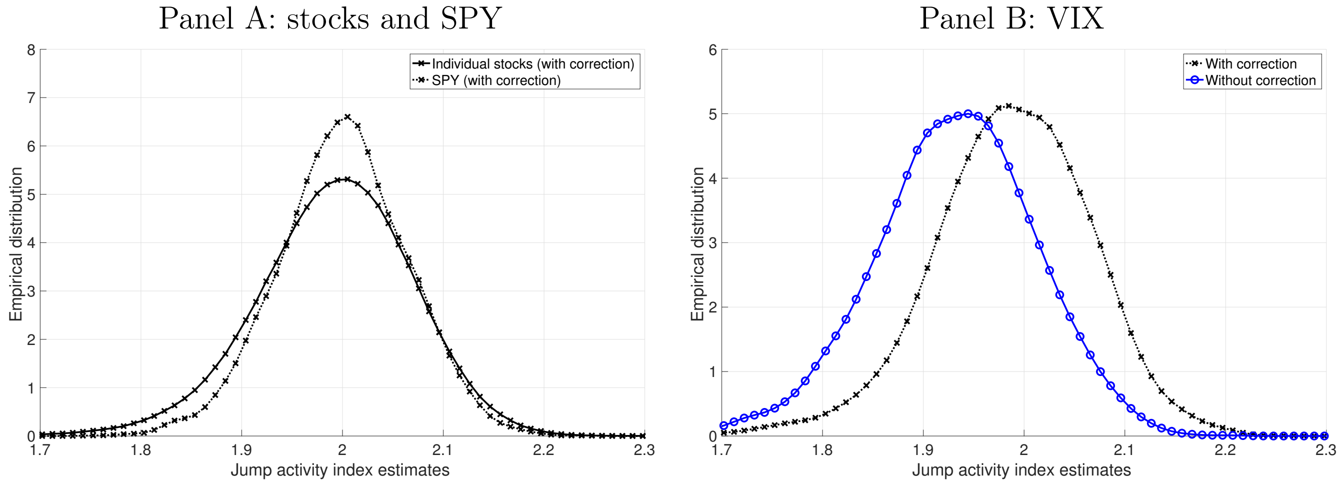

FIGURE 13: Empirical distribution of daily JAI estimates using the Kolokolov (2022) estimator on 5-minute returns in the period January 2003-June 2022. Panel A: Individual stocks and SPY using the staleness-corrected estimator for days with RIR below median. Panel B: VIX using the original versus staleness-corrected estimator.

staleness observed in stock prices. Thus, we reappraise the estimation of the JAI of the VIX after robustifying the estimators for the presence of staleness.

Figure 13 (panel B) shows the empirical densities of the daily jump activity estimates using the original Kolokolov (2022) estimator and with the version corrected for zeros. If we do not correct for staleness, we again get an estimated JAI which is significantly lower than 2. When instead we correct, the distribution of the estimates is symmetrically centered around the value of 2. Figure 14 clarifies that the bias of the non-corrected jump activity estimator is strongly correlated with the frequency of zero returns, as predicted by our theory. The panels show the scatter plot of daily jump activity estimates with the percentage of zero returns in that day. The left panel considers the estimator based on traditional multipowers, and shows that the higher the percentage of zeros, the lower the estimate of the JAI, as predicted by the theory (while, of course, staleness should have nothing to do with jump activity). The right panel consider the estimators obtained with corrected multipowers, and shows no relation between the two quantities, with the JAI estimates centered around 2. The distribution of daily point estimates using corrected estimators is consistent with the normal distribution predicted by Kolokolov (2022), and does not present any significant time variation.

Our conclusion, based on the proposed evidence, is that we cannot reject Brownian motion as a driving factor not only of stock prices (and, of course, of the market index), but also of of the stock index volatility, in agreement with the assumptions made by traditional stochastic volatility models in continuous-time finance. We also conclude that the rejection of the presence of the

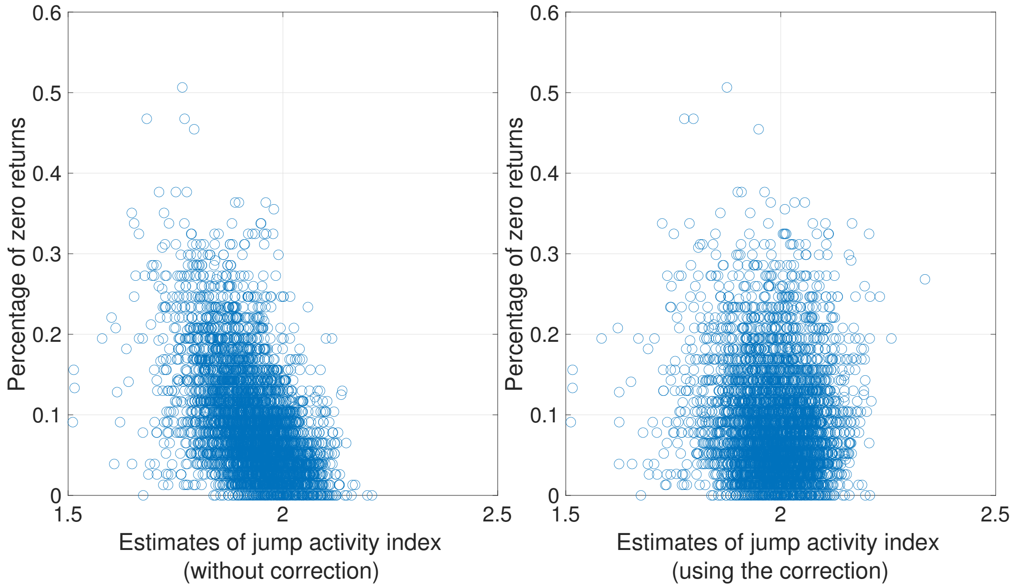

FIGURE 14: Scatter plot of daily estimates of the JAI on VIX data using the Kolokolov (2022) method versus the percentage of zero returns in each day. Left panel: original estimator. Right panel: staleness-corrected estimator. For staleness corrected estimator, we estimate (using the feasible estimator of Kolokolov, 2022) an average standard deviation of the point estimates of 0.0636.

Brownian motion found in the recent empirical literature is actually an artifact due to ignoring the presence of staleness in the data.

7 Conclusions

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

When applying the robustified estimator to stock and VIX data, we find results which are very different from those obtained in the literature so far. Jumps are much less frequent, contribute much less to jump variation and are much less vibrant than what suggested by the existing empirical literature. For the volatility of the stock index, our findings indicate that its activity is compatible with that assumed by standard stochastic volatility models, whose shocks are driven by Brownian motion, undermining the empirical relevance of pure jump processes in finance.

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

Finally, our paper also shows that additional work is required on modeling rounding theoretically, as for example in Li and Mykland (2015). The estimators studied in this paper work reasonably well unless rounding is too aggressive, and it would be desirable to devise a correction that works in this situation. While we propose a heuristic correction which soften the impact of rounding when it is too aggressive, the problem of the estimation of multipower variation in the joint presence of rounding and staleness remains open and appears challenging. However, our results show that a proper understanding of rounding on high-frequency prices can help in deliver more accurate measurements and more powerful tests about the nature of the price dynamics.

References

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

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

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

Aït-Sahalia, Y., and P. A. Mykland, 2003, "The effects of random and discrete sampling when estimating continuous–time diffusions," *Econometrica*, 71(2), 483–549.

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

Andersen, T., T. Bollerslev, and D. Dobrev, 2007, "No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications," *Journal of Econometrics*, 138(1), 125–180.

Andersen, T., O. Bondarenko, V. Todorov, and G. Tauchen, 2015, "The fine structure of equity-index option dynamics," *Journal of Econometrics*, 187(2), 532–546.

Andersen, T., D. Dobrev, and E. Schaumburg, 2009, "Duration-based volatility estimation," *Working paper*.

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

——— , 2014, "A Robust Neighborhood Truncation Approach To estimation of integrated quarticity," *Econometric Theory*, 30, 3–59.

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

Bandi, F., and J. Russell, 2008, "Microstructure noise, realized variance, and optimal sampling," *Review of Economic Studies*, 75(2), 339–369.

Bandi, F. M., A. Kolokolov, D. Pirino, and R. Renò, 2020, "Zeros," *Management Science*, 66(8), 3466–3479.

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

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

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

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

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

Catania, L., R. Di Mari, and P. Santucci de Magistris, 2020, “Dynamic Discrete Mixtures for High-Frequency Prices,” *Journal of Business & Economic Statistics*, 40(2), 559–577.

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

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

Delattre, S., and J. Jacod, 1997, “A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors,” *Bernoulli*, 3(1), 1–28.

Dovonon, P., S. Gonçalves, U. Hounyo, and N. Meddahi, 2019, “Bootstrapping High-Frequency Jump Tests,” *Journal of the American Statistical Association*, 114(526), 793–803.

Francq, C., and G. Sucarrat, 2022, “Volatility Estimation when the Zero-Process is Nonstationary,” *Journal of Business and Economic Statistics*, Forthcoming.

Griffin, J. E., and R. C. Oomen, 2008, “Sampling returns for realized variance calculations: tick time or transaction time?,” *Econometric Reviews*, 27(1-3), 230–253.

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

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

Hong, S. Y., I. Nolte, S. J. Taylor, and X. Zhao, 2021, “Volatility estimation and forecasts based on price durations,” *Journal of Financial Econometrics*, Forthcoming.

Hounyo, U., and R. T. Varneskov, 2017, “A local stable bootstrap for power variations of pure-jump semimartingales and activity index estimation,” *Journal of Econometrics*, 198, 10–28.

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

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

Jing, B.-Y., X.-B. Kong, and Z. Liu, 2012, “Modeling high-frequency financial data by pure jump processes,” *The Annals of Statistics*, 40, 759–784.

Kolokolov, A., 2022, “Estimating jump activity using multipower variation,” *Journal of Business and Economic Statistics*, 1(40), 128–140.

Kolokolov, A., G. Livieri, and D. Pirino, 2020, “Statistical Inferences for Price Staleness,” *Journal of Econometrics*, 218(1), 32–81.

Kong, X.-B., B.-Y. Jing, and Z. Liu, 2015, “Testing for pure-jump processes for high-frequency data,” *The Annals of Statistics*, 43(2), 847–877.

Lee, S., and P. Mykland, 2008, “Jumps in financial markets: A new nonparametric test and jump dynamics,” *Review of Financial Studies*, 21(6), 2535–2563.

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

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

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

Li, Y., P. A. Mykland, E. Renault, L. Zhang, and X. Zheng, 2014, "Realized volatility when sampling times are possibly endogenous," *Econometric Theory*, 30(3), 580–605.

Li, Y., I. Nolte, S. Nolte, and S. Yu, 2022, "Testing for Jumps in a Discretely Observed Price Process with Endogenous Sampling Times," Working Paper.

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

Maneesoonthorn, W., G. M. Martin, and C. S. Forbes, 2020, "High-frequency jump tests: Which test should we use?," *Journal of Econometrics*, 219(2), 478–487.

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

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

Sucarrat, G., and S. Grønneberg, 2022, "Risk Estimation with a Time-Varying Probability of Zero Returns," *Journal of Financial Econometrics*, 20(2), 278–309.

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

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

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

Todorov, V., and G. Tauchen, 2011, "Volatility jumps," *Journal of Business and Economic Statistics*, 3(29), 356–371.

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