Identifying the Underlying Components of High-Frequency Data: Pure vs Jump Diffusion Processes

Rodrigo Hizmeri*[a], Marwan Izzeldin[b], and Giovanni Urga[c]

[a]University of Liverpool, UK
[b]Lancaster University, UK
[c]Bayes Business School, UK

January 13, 2025

Abstract

In this paper, we examine the finite sample properties of test statistics designed to identify distinct underlying components of high-frequency financial data, specifically the Brownian component and infinite vs. finite activity jumps. We conduct a comprehensive set of Monte Carlo simulations to evaluate the tests under various types of microstructure noise, price staleness, and different levels of jump activity. We apply these tests to a dataset comprising 100 individual S&P 500 constituents from diverse business sectors and the SPY (S&P 500 ETF) to empirically assess the relative magnitude of these components. Our findings strongly support the presence of both Brownian and jump components. Furthermore, we investigate the time-varying nature of rejection rates and we find that periods with more jumps days are usually associated with an increase in infinite jumps and a decrease in finite jumps. This suggests a dynamic interplay between jump components over time.

Keywords: High-frequency data; infinite jumps; finite jumps; Brownian motion; price staleness; microstructure noise.
JEL Classification: C14, C15, C58, G01.

We are grateful to the editor (Rossen Valkanov), an associate editor, and an anonymous referee for their valuable comments. We also thank Deniz Erdemlioglu, Catherine Forbes, Aleksey kolokolov, Ignmar Nolte, Sandra Nolte, Eduardo Rossi, and participants at the 14th International Conference on Computational and Financial Econometrics (Virtual, 19-21 December 2020), and the CMAF-EMP Workshop (Lancaster, 12 March 2020), for useful discussions and comments on an earlier version of the paper. The usual disclaimer applies.

* *Corresponding Author*: Email: r.hizmeri@liverpool.ac.uk, Accounting and Finance Group, University of Liverpool Management School, Liverpool L69 7ZH, U.K.

1 Introduction

A general presumption in dealing with financial time series is that the continuous evolution of asset prices is driven by Brownian motion processes (Bachelier, 1900). However, that presumption is undermined by the failure of Brownian increments to explain heavy-tails observed in the distribution of returns. With an alternative approach, Merton (1976) proposes a finite-jump diffusion process, which successfully mimics empirical continuous and jump components. As small jumps eliminate the need of a Brownian component, the attraction of pure jump processes lies with their versatility and flexibility.[1]

The abundance of reliable high-frequency data has enabled the financial econometrics literature to shift from parametric to non-parametric approaches. As such, several non-parametric procedures to test for the presence of jumps (e.g., Aït-Sahalia and Jacod, 2009, Barndorff-Nielsen and Shephard, 2006, Jiang and Oomen, 2008, Lee and Mykland, 2008, Podolskij and Ziggel, 2010), Brownian motion (e.g., Aït-Sahalia and Jacod, 2010, Cont and Mancini, 2011, Jing et al., 2012a, Kong et al., 2015), and jump activity (e.g., Aït-Sahalia and Jacod, 2011) were proposed. As these procedures rely on in-fill asymptotics, the general assumption is that the observed price is the true price; i.e., market microstructure noise and price staleness are absent.[2]

While several studies have assessed the finite sample properties of price jump tests under more realistic scenarios (e.g., Dumitru and Urga, 2012, Huang and Tauchen, 2005, Maneesoonthorn et al., 2020), the existing literature is extremely limited regarding the finite sample properties of tests that identify the presence of a Brownian motion or jump activity. The latter tests play a crucial role in modelling financial data as they identify their different underlying components, and therefore provide useful information for risk

[1] Some infinite jump models are the variance gamma model (Madan and Seneta, 1990), the hyperbolic model of (Eberlein and Keller, 1995), the CGMY Carr et al. (2002), the COGARCH model (Klüppelberg et al., 2004), the non Gaussian Ornstein-Uhlenbeck-based models Barndorff-Nielsen and Shephard (2001), the CARMA model (Brockwell, 2001), the normal inverse Gaussian Barndorff-Nielsen (1997), among others.

[2] The few exceptions, considering the existence of microstructure noise to identify price jumps, are Podolskij and Ziggel (2010), Aït-Sahalia et al. (2012), and Lee and Mykland (2012). Recently, Kolokolov and Renò (2024) propose a staleness-adjustment that mitigate the impact of zero returns on well-known jump test statistics. However, as noted by the authors, the tests proposed by Podolskij and Ziggel (2010) afford similar performance than their adjustment.

management, portfolio allocation, and derivative pricing.

The contribution of this paper is twofold. First, we provide a comprehensive Monte Carlo simulation to assess the finite sample properties of four tests statistics developed to identify the different underlying components of high-frequency financial data in terms of a Brownian component and infinite vs. finite activity jumps. We consider the pure jump test of Kong et al. (2015) (KLJ hereafter), which evaluates the Brownian motion hypothesis, and the Aït-Sahalia and Jacod (2011) (ASJ hereafter) test able to detect the presence of finite and infinite jumps.[3] Second, as finite and infinite jumps have very different features,[4] we disentangle these jumps using the intersection between the ASJ and a price jump test at the optimal sampling frequency. This allows us to quantify the contribution of both finite and infinite jumps to the total quadratic variation. In addition, we also investigate whether the rejection rates are time-varying.

Although, in principle, any price jump test can be employed, we consider the procedure of Podolskij and Ziggel (2010) (PZ hereafter), as this test is capable of detecting jumps both of finite and infinite activity, and because it also has a noise-robust version. Evidently, in presence of microstructure noise and at very high frequencies, the performance of the noise-robust PZ test should dominate its noise-free counterpart. However, very little is known about the power and size of the noise-robust PZ test across various sampling frequencies and types of market microstructure noise. Neither Dumitru and Urga (2012) nor Maneesoonthorn et al. (2020) consider the noise-robust version of the PZ test. Thus, we also assess the finite properties of both tests as previous studies have documented that noise-robust price jump tests, such as that of Aït-Sahalia et al. (2012) and Lee and Mykland (2012), are less powerful at relatively low frequencies (e.g., Maneesoonthorn et al., 2020).

[3]Note that the tests proposed by Aït-Sahalia and Jacod (2010) and Jing et al. (2012a) are also developed to test for the presence of a Brownian motion. However, in unreported results, we confirm the findings of Kong et al. (2015) that, in a noise-free environment, their test improves upon the results of the aforementioned procedures. Thus, in this paper, we only assess the finite sample properties of the best-performing model. Further, Aït-Sahalia and Jacod (2011) also propose a test that has a null of finite jumps instead of infinite jumps; however, one version of the test is the complement of the other version, and thus their sample properties and conclusion are qualitatively similar. For brevity, in this paper, we consider the test for infinite jump activity. The unreported results are available upon request.

[4]As discussed in Aït-Sahalia et al. (2020) and Bu et al. (2023), infinite (finite) jumps are normally related to idiosyncratic (systematic) events, such as earning disappointments (FOMC meetings).

Our comprehensive simulation exercise focuses on two main features associated with market frictions: market microstructure noise and price staleness, as both these features become relevant considerations as $\Delta_n \to 0$. While sparse sampling provides a simple remedy (e.g., Ait-Sahalia et al., 2005), sampling too sparsely reduces the power of the tests. Thus, it is vital to find a trade-off between achieving sufficient statistical power and avoiding distortions that could arise from microstructure noise and price staleness, as inference about the appropriate model depends on the sampling frequency and testing technique (see, for instance, Todorov and Tauchen, 2010).

The main results of our Monte Carlo experiments can be summarized as follows. In the absence of microstructure noise, all tests exhibit very good finite sample properties, which deviate slowly from the theoretical size and power as the sampling frequency increases. Conversely, the presence of microstructure noise biases the distributions of all the tests, except for the robust PZ test, which is derived under the assumption of noisy prices.

The distributions of the tests exhibit bias only at very high frequencies when Gaussian noise is present. However, in the presence of t-distributed and Gaussian-T mixture noise, the performance of all tests is severely adversely affected. Sampling sparsely decreases the bias of the tests. Under Gaussian noise, the tests display good performance when returns are sampled every 30 seconds. When microstructure noise is t-distributed or Gaussian-T mixture, sampling every 60 seconds give satisfactory results. However, in the presence of non-Gaussian noise and degrees of freedom equal to 2.5, the standard PZ test shows severe upward bias even when returns are sampled every 90 seconds. On the other hand, contaminating the efficient price with price staleness moderately distorts the tests at higher-frequencies, although sampling returns every 1 minute seems to eliminate most of these distortions.

In our empirical illustration, we consider 100 individual stocks and the SPY (S&P 500 ETF) from January 2000 to December 2022. Guided by our Monte Carlo experiment, we sample returns every 1 minute as this sampling frequency shows sufficient statistical power with minimal distortions from microstructure noise and price staleness. Our findings provide strong evidence for both Brownian and jump components. On average, jumps

occur 19% of days across sectors and 13% of days for the SPY, with the lower SPY proportion resulting from aggregation effects. In reducing idiosyncratic risk, aggregation causes stock specific jumps to be 'lost'. While both finite and infinite jumps characterize the jump component, finite jumps contribute more to the total jump part. The estimated jump activity index, $\widehat{\beta}$, oscillates around 1.0, confirming the presence of jump types. Finally, we document significant time variation in the rejection rates, with variations ranging between 5–20%. This aligns with recent findings on time-varying jump intensity, which explain variations in news-induced realized volatility (e.g., Erdemlioglu and Yang, 2022). This suggests that the data generating process should allow for time variation, with increases in the rejection of no jumps generally accompanied by decreases in the rejection of infinite jumps. These results remain consistent even with more conservative jump detection threshold in the PZ test.

Our paper is related to the vast literature advocated to answer one of the most central empirical issues in modelling financial data: what distributional assumptions are consistent with observed stock prices. Although the practical importance and relevance of jumps in financial data are universally recognized, the relative appropriateness of large and/or small jumps, and the extent to which Brownian motion is necessary, are still open to debate. For instance, in comparing several diffusion and finite-jump diffusion models, Andersen et al. (2002) conclude that a finite-jump diffusion model is capable of catching the characteristics of the S&P 500 returns. However, using the same index, Carr et al. (2002) find that a pure jump process is the most appropriate model. The source of such discrepancies may be associated with that parametric approaches run the risk of incorrect specification for functionals in their chosen models. This is not the case with the non-parametric approaches employed in this paper.

The remainder of the paper is structured as follows. Section 2 reviews both the theoretical background and the test statistics which are respectively categorized as pure jump test (Kong et al., 2015), infinite activity jump (Aït-Sahalia and Jacod, 2011), and jump test (Podolskij and Ziggel, 2010). Section 3 describes the Monte Carlo setup, reports the simulated results. The empirical illustration including a time-variation exercise for

the rejection rates are reported in Section 4. Section 5 concludes.

2 Theoretical Background

In this section, we briefly review the test statistics considered, together with a brief description of (truncated) power variation methods which are the building blocks of most of the procedures.

2.1 Background

Let the log-price X_t follow a semimartingale defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$:

$$X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + X_t^d, \tag{1}$$

where x_0 is the initial value, b_s is the continuous and locally bounded drift term, σ_s is a strictly positive and càdlàg stochastic volatility process, W_s is a standard Brownian motion, and X_t^d is a pure-jump component. The Blumenthal-Getoor (BG) index of X_t^d measures the degree of activity of small jumps and is defined as:

$$\beta := \inf\left\{ r; \sum_{0\leq s\leq t} |\Delta X_s|^r < \infty \right\}, \tag{2}$$

where $\Delta X_s = X_s - X_{s-} \neq 0$ if jumps are present. β serves as an indicator of the activity of jumps contained in X^d. The larger the β, the more active the jumps. A finite activity jump process such as a compound Poisson process has $\beta = 0$, whereas a β-stable process has an index equal to $\beta \in (0, 2)$. Finite variation corresponds to $0 < \beta < 1$ and infinite variation to $1 < \beta < 2$.

To construct the tests, we define power and truncated power variations (see, Jacod, 2008, Mancini, 2001, 2009), as well as an estimator of the continuous part that is robust to infinite jump variation (see, Jacod and Todorov, 2014, JT hereafter).

Let denote the power variation estimator as $B(p,\infty,\Delta_n)_t$, which is outline as:

$$B(p,\infty,\Delta_n)_t = n^{p/2-1}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}|\Delta_i^n X|^p \xrightarrow{\mathbb{P}} \begin{cases}\mu_p\int_0^t|\sigma_s|^p ds, & \text{(No Jumps)},\\ \infty, & \text{(With Jumps)},\end{cases} \tag{3}$$

when $p>2$, and $\mu_p \equiv E[|U|^p] = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)$, where $U\sim\mathcal{N}(0,1)$. $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$, with $\Delta_n = 1/n$ for $0\le i\le n$. It is well-known that:

$$B(2,\infty,\Delta_n)_t = \sum_{i=1}^{\lfloor t/\Delta_n\rfloor}|\Delta_i^n X|^2 \xrightarrow{\mathbb{P}} \underbrace{\int_0^t \sigma_s^2 ds}_{\text{Integrated Variation }(IV_t)} + \underbrace{\sum_{0<s\le t}(\Delta_s X)^2}_{\text{Jump Variation}}. \tag{4}$$

We denote the truncated power variation as $B(p,\nu_n,\Delta_n)_t$:

$$B(p,\nu_n,\Delta_n)_t = n^{p/2-1}\sum_{i=1}^{\lfloor t/\Delta_n\rfloor}|\Delta_i^n X|^p \mathbb{1}_{\{|\Delta_i^n X|\le\nu_n\}} \xrightarrow{\mathbb{P}} \int_0^t \mu_p|\sigma_s|^p ds \tag{5}$$

where $\nu_n = \alpha\Delta_n^{\varpi}$ is the truncation threshold and $\alpha>0$ is expressed in units of the standard deviation of the continuous part of the process for a constant $\varpi\in(0,1/2)$. When the jump of X_t is a Lèvy process with the Blumenthal-Getoor index $\beta\in[0,2)$, as outlined in equation (2), then the required condition is given by $\varpi\ge\frac{p-2}{2(p-\beta)}$, for $p>2$.

The JT bias-corrected estimator, $C(u_n)_j^n$, is given by:

$$C_0(u_n)_j^n = 2k_n\Delta_n\sum_{j=0}^{\lfloor t/k_n\Delta_n\rfloor-1}\left(c_0(u_n)_j^n - \frac{1}{u_n^2 k_n}(\sinh(u_n^2 c_0(u)_j^n))^2\right) \xrightarrow{\mathbb{P}} \int_0^t \sigma_s^2 ds \tag{6}$$

where

$$c_0(u_n)_j^n = -\frac{1}{u_n^2}\log\left(L(u_n)_j^n \vee \frac{1}{\sqrt{k_n}}\right)$$
$$L(u_n)_j^n = \frac{1}{k_n}\sum_{l=0}^{k_n-1}\cos\left(u_n(\Delta_{2jk_n+1+2l}^n X - \Delta_{2jk_n+2+2l}^n X)/\sqrt{\Delta_n}\right)$$

where the following conditions must satisfy $k_n\Delta_n^{1/2}\to 0$, $u_n\to 0$, $\sup_n\frac{k_n\Delta_n^{1/2}}{u_n^4}<\infty$. Possible choice for k_n and u_n are $k_n\asymp 1/\sqrt{\Delta_n}(\log(1/\Delta_n))^x$ and $u_n\asymp 1/(\log(1/\Delta_n))^{x'}$ for

$0 < x' \leq x/4$, where $x \in (0, 1]$, and $x \asymp y$ means that $1/A \leq x/y \leq A$, for a constant $A \in (1, \infty)$ (for more details, see Jacod and Todorov, 2014).

2.2 Test Statistics

In what follows, we offer a short overview of the test statistics used to identify the presence of Brownian component, infinite/finite jumps, as well as testing for the presence of price jumps.

Pure Jump Test

The pure-jump test of Kong et al. (2015, KLJ hereafter) is based on the realized characteristic function and checks whether the underlying process of a high-frequency data set can be modelled as a pure-jump process [5]. The pure jump test is of the following hypotheses:

$$H_0 : \int_0^t \sigma_s^2 ds > 0 \quad \text{v.s.} \quad H_1 : \int_0^t \sigma_s^2 ds = 0.$$

The test, under the null, limit distribution takes the following form:

$$\mathcal{T}_t = \frac{C_0(u_n)_j^n - C_1(u_n)_j^n - \gamma_n \Delta^{1/2}}{2 I_n^{1/2} \Delta_n^{1/2}} \xrightarrow{\mathcal{L}_s} \mathcal{N}(0, 1), \tag{7}$$

where

$$\begin{aligned} I_n &\equiv \frac{1}{2}(I_{n,0} + I_{n,1}) \\ I_{n,k} &= 2k_n\Delta_n \sum_{j=0}^{\lfloor t/k_n\Delta_n \rfloor - 1} \left(c_k(u_n)_j^n - \frac{\sinh(u_n^2 c_k(u_n)_j^n)}{u_n^2(k_n - 1)} \right)^2, \quad k = 0, 1 \end{aligned}$$

where γ_n is some chosen constant satisfying $\gamma_n \to 0$, and can be estimated as $\gamma_n = c^*/\log(u_n^2/\Delta_n)$, where $c^* = 0.2$, when the number of observations is moderate. $C_0(u_n)_j^n$ is estimated as in equation (6), whereas $C_1(u_n)_j^n$ can be defined as the $C_0(u_n)_j^n$, where $\Delta_{2jk_n+2l+1}^n X - \Delta_{2jk_n+2l}^n X$ is replaced by $\Delta_{2jk_n+2l}^n X - \Delta_{2jk_n+2l-1}^n X$, for $l = 1, \ldots, k_n - 1$.

[5] In finite sample terms, this test is superior to the Brownian test of Aït-Sahalia and Jacod (2010) and the modified version of the latter proposed by Jing et al. (2012a)

Finally, H_0 can be rejected if $\mathcal{T}_t < -z_\theta$ where $\mathbb{P}(\mathcal{N}(0,1) > z_\theta) = \theta$ for $\theta \in (0,1)$.

As KLJ do not provide analytical results for cases that include microstructure noise, our prior is that, in the presence of microstructure noise, the test would be unable to recognize whether the small and frequent movements are Brownian or pure jump increments.

Infinite Activity Jump Test

The infinite activity jump test proposed by (Aït-Sahalia and Jacod, 2011, ASJ hereafter) evaluates the following hypothesis:

$$H_0 : \Omega_T^{i\beta} \quad \text{v.s.} \quad H_1 : \Omega_T^f \cap \Omega_T^c,$$

where $\Omega_T^{i\beta}$ and Ω_T^f respectively refer to infinite and finite jump activity, and Ω_T^c is the diffusive part. The ASJ test is outlined as:[6]

$$S_t^{IA} = \frac{B(p', \varphi\nu_n, \Delta_n)_t B(p, \nu_n, \Delta_n)_t}{B(p', \nu_n, \Delta_n)_t B(p, \varphi\nu_n, \Delta_n)_t} \xrightarrow{\mathbb{P}} \varphi^{p'-p}. \tag{8}$$

The CLT of this test, under the null, takes the following form:[7]

$$(S_t^{IA} - \varphi^{p'-p}) \big/ \sqrt{\hat{\sigma}_t^2} \xrightarrow{\mathcal{L}_s} \mathcal{N}(0,1), \tag{9}$$

where

$$\begin{aligned}\hat{\sigma}_t^2 = \varphi^{2p'-2p} \Bigg(& \frac{B(2p, \nu_n, \Delta_n)_t}{(B(p, \nu_n, \Delta_n)_t)^2} + (1 - 2\varphi^{-p}) \frac{B(2p, \varphi\nu_n, \Delta_n)_t}{(B(p, \varphi\nu_n, \Delta_n)_t)^2} + \frac{B(2p', \nu_n, \Delta_n)_t}{(B(p', \nu_n, \Delta_n)_t)^2} \\ & + (1 - 2\varphi^{-p'}) \frac{B(2p', \varphi\nu_n, \Delta_n)_t}{(B(p', \varphi\nu_n, \Delta_n)_t)^2} - 2 \frac{B(p + p', \nu_n, \Delta_n)_t}{B(p, \nu_n, \Delta_n)_t B(p', \nu_n, \Delta_n)_t} \\ & - 2(1 - \varphi^{-p} - \varphi^{-p'}) \frac{B(p + p', \varphi\nu_n, \Delta_n)_t}{B(p, \varphi\nu_n, \Delta_n)_t B(p', \varphi\nu_n, \Delta_n)_t} \Bigg).\end{aligned}$$

[6] The convergence in probability holds only under the stated null hypothesis.
[7] The convergence in law holds only under the stated null hypothesis.

We set $p = 3$, $p' = 4$, $\varphi = 2$, $\varpi = 0.2$, and $\alpha = 8$. As shown by ASJ, this test converges to $\varphi^{p'-p}$ (1) when the underlying process has infinitely (finitely) many jumps. When microstructure noise dominates, the test also converges to $\varphi^{p'-p}$. The implication is that in the presence of microstructure noise, the test cannot distinguish whether jumps have finite or infinite activity. z_θ denotes the *upper* θ-quantile of $\mathcal{N}(0,1)$, that is, $\mathbb{P}(\mathcal{N}(0,1) > z_\theta) = \theta$, for $\theta \in (0,1)$, the test rejects H_0 when $S_t^{IA} < \varphi^{p'-p} - z_\theta \sqrt{\hat{\sigma}_t^2}$.

Jump Test

As the difference between the two capture the contribution of jumps, Podolskij and Ziggel (2010) use the difference between power and truncated power variations to construct their test statistics.[8] The test is of the following hypotheses:

$$H_0 : \Omega_T^c \quad \text{v.s.} \quad H_1 : \Omega_T^j,$$

where Ω_T^c and Ω_T^j are respectively the set of a continuous and a discontinuous price path. We outline the test as:

$$S_t^J = \frac{T(\Delta_i^n X, p)_t}{\rho_t} \xrightarrow{\mathcal{L}_s} \mathcal{N}(0,1), \tag{10}$$

where

$$T(\Delta_i^n X, p)_t = n^{(p-1)/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\Delta_i^n X|^p \left(\mathbb{1} - \eta_i \mathbb{1}_{\{|\Delta_i^n X| \le \nu_n\}} \right), \quad p \ge 2, \tag{11}$$

$$\rho_t^2 = \mathrm{Var}^*(\eta_i) B(2p, \nu_n, \Delta_n)_t. \tag{12}$$

η_i is a positive i.i.d. random variables with $\mathbb{E}[\eta_i] = 1$ and $\mathbb{E}[|\eta_i|^2] < \infty$. PZ suggest that η_i can be sampled from the distribution:

$$P^\eta = \frac{1}{2} \left(\delta_{1-\tau} + \delta_{1+\tau} \right),$$

[8] The limit distribution of the standard and noise-robust versions are all the null limit distributions.

where δ is the Dirac measure, and $\tau = 0.1$ or 0.05. ϖ is chosen such that it satisfies $\varpi \geq \frac{p-2}{2(p-\beta)}$, for $p > 2$.

PZ are amongst the few that account for microstructure noise.[9] Robust jump tests were generally ignored by subsequent research (e.g Dumitru and Urga, 2012), although more recently Maneesoonthorn et al. (2020) have shown evidence that both the Aït-Sahalia et al. (2012) and Lee and Mykland (2012) tests lose power very rapidly. Given the evidence that jump tests are very sensitive to microstructure noise, in this paper, we also examine the finite sample properties of the noise-robust PZ test.

Let $Y_t = X_t + u_t$ be a contaminated price, and u_t an additive i.i.d. process. We assume that u_t has $\mathbb{E}[u_t] = 0$ and $\mathbb{E}[u_t^2] = \omega_t^2$, and $X_t \perp u_t$ ($\perp$ means stochastic independence). We pre-filter the data using the pre-averaging method of Jacod et al. (2009), so that the additive component is eliminated. The pre-averaging returns are defined as:

$$\Delta_i^n \bar{Y} = \sum_{j=0}^{K_n - 1} g\left(\frac{j}{K_n}\right) \Delta_{i+j}^n Y,$$

where $K_n/\sqrt{n} = \Theta + o\left(n^{-1/4}\right)$, for some $\Theta > 0$, and a nonzero real-valued function $g(x) = (x \wedge 1 - x)$. The test is outlined as:

$$S_t^{J^{noise}} = \frac{T^{noise}\left(\Delta_i^n \bar{Y}\right)_t}{\sqrt{\Gamma_t}} \xrightarrow{\mathcal{L}_s} \mathcal{N}(0,1), \tag{13}$$

$$T^{noise}\left(\Delta_i^n \bar{Y}, p\right)_t = n^{(p-2)/4} \sum_{i=0}^{\lfloor 1/\Delta_n \rfloor - K_n + 1} |\Delta_i^n \bar{Y}|^p \left(\mathbb{1} - \eta_i \mathbb{1}_{\{|\Delta_i^n \bar{Y}| \leq \nu_n\}}\right), \quad p \geq 2, \tag{14}$$

$$\Gamma_t = \mathrm{Var}^*\left(\eta_i\right) n^{(p-2)/2} \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor - K_n + 1} |\Delta_i^n \bar{Y}|^{2p} \mathbb{1}_{\{|\Delta_i^n \bar{Y}| \leq \nu_n\}}, \tag{15}$$

where $\alpha > 0$, $\varpi \in (0, 1/4)$ and η_i is estimated as described above.[10] Both versions of the test reject H_0 when $S_t^J(S_t^{J^{noise}}) > z_\theta$ where $\mathbb{P}(\mathcal{N}(0,1) > z_\theta) = \theta$ for $\theta \in (0,1)$. Throughout the paper we set $\tau = 0.05$, $\alpha = 3$, and $\varpi = 0.40$ for the standard PZ test

[9] Other noise-robust tests are Aït-Sahalia et al. (2012), Jiang and Oomen (2008), and Lee and Mykland (2012).

[10] For more details on the pre-averaging methods used, see (Podolskij and Ziggel, 2010, Section 4)

and to $\varpi = 0.20$ for the noise-robust version.[11]

3 Monte Carlo Study

In general, most of the test statistics are derived under the assumption of frictionless markets. Under this assumption the observed price is the efficient price, and therefore one could implement these statistics at the highest possible frequency. However, in practice there exists market frictions that give rise to market microstructure noise (e.g., Ait-Sahalia and Yu, 2009, O'Hara, 2015) and price staleness (e.g., Bandi et al., 2020, 2017), which are very likely to impact the performance of these procedures in finite samples. In this section, we evaluate the performance of these tests under different types of microstructure noise and price staleness, the main aim being the identification of an optimal trade-off between performance and bias.

3.1 Monte Carlo design

In following Aït-Sahalia and Jacod (2010, 2011) and Jing et al. (2012a), we use a Heston stochastic volatility model that allows for both finite and infinite jumps, while for the pure-jump process $\nu_t \equiv 0$, i.e. $dX_t = dL_t$. The model is described as:

$$\begin{aligned} dX_t &= \nu_t^{1/2} dW_t^{(1)} + dL_t \\ d\nu_t &= k(\eta - \nu_t)dt + \gamma \nu_t^{1/2} dW_t^{(2)}, \end{aligned} \tag{16}$$

with $\mathbb{E}[dW_t^{(1)} dW_t^{(2)}] = \rho dt$, $\eta = 0.25^2$, $\gamma = 0.5$, $k = 5$, $\rho = -0.5$, and the pure jump process is either a finite activity compound Poisson process or a symmetric tempered β-Stable process.[12] The compound Poisson has intensity $\lambda = \{0.1,\ 0.2,\ 1.0\}$ and jumps that are uniformly distributed on $\nu_t^{1/2}\sqrt{m}([-2,-1] \cup [1,2])$, where $m = 0.7$. We set the

[11] Podolskij and Ziggel (2010) suggest setting $\alpha = 2.3$. However, in unreported simulation studies, employing this value results in an increased Type I error rate compared to more conservative values, i.e., $\alpha \geq 3$.

[12] We specifically simulate the CGMY process of Carr et al. (2002), where the trajectories could be approximately simulated by the time-changed-Brownian-motion algorithm, where the change of time is via the $\beta/2$ stable subordinator which is also a Lévy process.

following values for $\beta = \{1.00,\ 1.25,\ 1.50\}$.

We add the measurement error, so that we do not observe the true price X_t, but observe instead the contaminated price as follows:

$$Y_t = X_t + u_t, \tag{17}$$

where Y_t and X_t are respectively the contaminated and true log-price processes, and u_t is the measurement error with $\mathbb{E}[u_t] = 0$, and $\mathbb{E}[u_t^2] = \omega_t^2$. We follow Aït-Sahalia et al. (2012) and consider four settings:

$$u_t = \begin{cases} 0, & \text{(No noise)} \\ 2\nu_t^{1/2}\Delta_n^{1/2}u_t^A, & \text{(Gaussian noise)} \\ 2\nu_t^{1/2}\Delta_n^{1/2}u_t^B \Big/ \sqrt{\frac{df}{df-2}}, & \text{(T-distributed noise)} \\ 2\nu_t^{1/2}\Delta_n^{1/2}\left(u_t^A + u_t^B \Big/ \sqrt{\frac{df}{df-2}}\right), & \text{(Gaussian-T mixture noise),} \end{cases} \tag{18}$$

where u_t^A and u_t^B are mutually independent i.i.d. drawn from an $\mathcal{N}(0,1)$ distributed and a t-distribution with degree of freedom $df = \{2.5,\ 3.5, 4.5\}$, respectively. The instantaneous standard deviations of the Gaussian noise and the t-distributed noise are twice that of the diffusive increment, i.e., $(\nu_t\Delta_n)^{1/2}$, and allow for temporal heteroskedasticity and dependence in u_t. The t-distributed noise captures the large bouncebacks commonly observed in transaction data as shown in Figure 1.

Once the efficient price, X_t, is simulated and sampled to the required sampling frequency, we construct the staleness-contaminated log-price process Z_t following Kolokolov et al. (2020):

$$\begin{cases} Z_0 &= X_0, \\ Z_j &= (1 - \mathbb{B}_j)X_j + \mathbb{B}_j X_{j-1}, \end{cases} \tag{19}$$

where $\mathbb{B}_j$ are i.i.d. Bernoulli random variables with constant expected value $\mathbb{E}[\mathbb{B}_j] = \mathfrak{p}_F$. To determine the probability of price staleness, we analyze the proportion of zero returns in our dataset across the required sampling frequencies. We establish high and

medium levels of price staleness, corresponding to the 95th and 50th percentiles, respectively. For the high level, the probabilities are $\mathfrak{p}_F = \{0.70,\ 0.35,\ 0.23\ , 0.18\}$ for 5-, 30-, 60-, and 90-second frequencies, respectively. The medium level probabilities are $\mathfrak{p}_F = \{0.54,\ 0.22,\ 0.13,\ 0.10\}$ for the same frequencies. We generate data for 50 days and use 3,000 replications, which encompass 150,000 simulated days.

Figure 2 plots the distribution of the tests using a diffusion process, Equation (16)), contaminated with different types of microstructure noise. Of course, in the absence of noise effects (Figure 2a) all procedures are well-behaved, so giving good finite sample properties. When noise is added, there is a decrease in performance in all the tests, but the noise-robust PZ test. Gaussian noise (Figure 2b), the most popular type of noise in the literature, produces least distortions in all the test statistics. By contrast, t-distributed (Figure 2c) and Gaussian-T mixture (Figure 2d) noise severely affect the standard procedures.

By comparison, the standard PZ test is upward biased, which results in a high incidence of Type I error. However, the bias induced by Gaussian noise is present only at high sampling frequencies, while the bias from t-distributed and Gaussian-T mixture noise distorts the distribution of the test even when returns are sampled every 90 seconds. On the other hand, the distribution of the robust PZ test is largely unaffected whether or not the underlying process is contaminated with microstructure noise. The KLJ and ASJ tests are also more affected when the noise is non-Gaussian. Nevertheless, the distribution of both tests under the different types of microstructure noise suggest that sampling sparsely, and not at very low frequencies, can solve this issue without losing much power.

Figure 3 shows the distribution of the tests when the model is an infinite-jump diffusion process with $\beta = 1.0$. The distribution of both versions of the PZ test tends to infinity in the presence of infinite jumps, which confirms the capacity of these tests to identify jumps of infinite variation. The noise has little effect on the finite sample performance of the PZ tests. This is mainly because the noise-robust version is not affected by any type of noise as shown in Figure 2, and the standard PZ test is oversized in the presence of noise, which does not affect the power of the test but increases the spurious

detection of jumps. The distribution of the KLJ test shows similar behavior to the jump diffusion case; that is, the test is downward biased in the presence of microstructure noise. This means that the test fails to distinguish between the null and alternative hypothesis. As expected, the ASJ test is centered around 2 in the absence of noise, while in the presence of t-distributed and Gaussian-T mixture noise the distribution shifts to the left of 2, thereby indicating a rise in Type II error.

In what follows, we focus on the finite sample properties of these procedures across different sampling frequencies using a significance level of $\theta = 0.01$. We sample the simulated data every 5, 30, 60, and 90 seconds, which corresponds to $\lfloor 1/\Delta_n \rfloor = \{4680, 780, 390, 260\}$ observations per day, respectively.

3.2 Monte Carlo Results

The full set of results of the Monte Carlo exercise is reported in Tables 1–6. In each table, the first three rows report the KLJ and PZ tests, whereas the fourth row reports the ASJ test. To facilitate a more structured analysis, we start by examining the results influenced by market microstructure noise, subsequently progressing to the implications of price staleness.

3.2.1 Finite Sample Performance under Microstructure Noise

Pure Jump Test

Tables 1, 2 and 3 report the empirical sizes of the KLJ test given the presence of a diffusive component in all underlying processes. Irrespective of whether the underlying model is a diffusion (Table 1) or a finite jump-diffusion process (Table 2),[13] the KLJ test performs well at higher frequencies when microstructure noise is absent, although it tends to be slightly upwards bias as sampling frequency decreases. When noise is present, the test is downward biased, and this bias is exacerbated when the noise is t-distributed or Gaussian-T mixture. However, we note that as the degrees of freedom increase, the

[13]Results are qualitatively similar when the jump intensity is set to $\lambda = 0.2$ (Table A.1), $\lambda = 0.1$ (Table A.2), or when only 4 large random jumps are simulated (Table A.3).

performance of the test also improves as these noises behave more like Gaussian noise. Generally, the KLJ test approaches the theoretical size when sampling returns every 30 or 60 seconds, i.e., $\lfloor 1/\Delta_n \rfloor = 780$ and 390, respectively.

Table 3 presents the results using an infinite-jump diffusion process with jump activity index equal to $\beta = \{1.00,\ 1.25,\ 1.50\}$. In the absence of microstructure noise, the KLJ test is close to its theoretical size. However, when using a $\beta = 1.50$, the test is downward biased. An explanation for this finding is that, as $\beta \to 2$, the increments are closer to those of a Brownian motion, which makes it difficult for the test to recognize true Brownian increments. When microstructure noise contaminates the true underlying process, we observe a decrease in the size of the test, which can be recovered by sampling more sparsely, i.e. every 780 to 390 observations. The results are also robust to considering the case of an infinite jump-diffusion process with residual finite jumps, as shown in Table A.4.

Table 4 reports the power of the KLJ test using a pure jump process. We simulate microstructure noise as described in equation (18), replacing ν_t with the JT estimator. This approach ensures that noise variance remains constant intraday but varies across days. In the absence of microstructure noise, the KLJ's power closely approximates its theoretical power. This power decreases as both the time interval increases and as $\beta \to 2$. However, the introduction of microstructure noise significantly impacts the test's performance at higher frequencies. For $\beta = 1$, sampling at 60-second intervals (every 390 observations) yields power values comparable to those in the noiseless scenario, regardless of the microstructure noise type. Conversely, when $\beta > 1.0$, the test exhibits very low power, improving only marginally with increased sampling intervals.

All in all, the KLJ test has high power in the absence of microstructure noise, improving further as $\Delta_n \to 0$. When microstructure noise is added, the KLJ test is undersized under the null and increases the Type II error under the alternative hypothesis. Although this is observed for all types of microstructure noise, Gaussian noise produces less severe distortions that are only observed at high sampling frequencies and as $\beta \to 2$. The distortions of t-distributed and Gaussian-T mixture noise are less obvious around

$\lfloor 1/\Delta_n \rfloor = 390$, which suggests that sampling returns every 1-min presents a good trade-off between bias and performance.[14]

Infinite Activity Jump Test

Tables 1 and 2 report the power of the ASJ test for data generating processes following a diffusion and a finite jump-diffusion, respectively. In the absence of microstructure noise, the power of the test gradually decreases as the time interval increases. For example, when the data generating process is a diffusion, the power of the test is 0.94, and when it is a finite jump-diffusion, the power is 0.936, both using 1-minute returns (i.e., $\lfloor 1/\Delta_n \rfloor = 390$). These results suggest that the loss of power due to more sparse sampling in the absence of price noise is moderate.

As shown by Aït-Sahalia and Jacod (2011) and illustrated in Figure 2, when microstructure noise dominates, the ASJ test converges to $\varphi^{p'-p}$. This convergence implies that the test fails to distinguish between its null and alternative hypotheses. Furthermore, regardless of whether the underlying process is a diffusion or a finite jump-diffusion, all types of microstructure noise increase the Type II error of the test. While the effect of Gaussian noise declines rapidly for sampling frequencies beyond $\lfloor 1/\Delta_n \rfloor = 4680$, the distortions caused by t-distributions and Gaussian-T mixture noise are more pronounced. However, these distortions are considerably reduced when sampling at frequencies of $\lfloor 1/\Delta_n \rfloor = 390$ or lower.

Table 3 reports the size of the ASJ test, presenting results when the underlying model follows an infinite jump-diffusion process. In the absence of microstructure noise effects, the size of the ASJ test closely approximates its theoretical size. As $\beta \to 2$ and the number of intraday observations decreases, Type I error increases, reflecting the increased difficulty in distinguishing between infinite jumps and Brownian increments. When the underlying process is contaminated with Gaussian noise and $\beta \in [1.0, 1.25]$, the results show minimal deviation from the noiseless case. This deviation becomes more pronounced

[14] As a robustness-check, we also implement the test of Jing et al. (2012a), which shows being oversized for time intervals lower than 4,680 (5-seconds). Note that this comparison is in line with what reported in Kong et al. (2015). Thus, we do not report these results which are available upon request.

when $\beta = 1.5$. In contrast, t-distributed and Gaussian-T mixture noise significantly increase the number of spurious rejections in the ASJ test, regardless of the value of β. As sampling decreases, for df $= \{3.5, 4.5\}$, the number of spurious rejections initially falls, but then rises again at $\lfloor 1/\Delta_n \rfloor = 260$.

In summary, the ASJ test exhibits relatively robust performance in the presence of Gaussian noise. Conversely, the Type I error increases when the noise follows a t-distribution or Gaussian-T mixture. However, for values of $\beta = 1.0$, which are on average the values observed in our empirical exercise, the test performs relatively well when returns are sampled at 1-minute intervals. Furthermore, our results remain consistent under various scenarios, including different levels of jump intensity (Tables A.1–A.3), infinite jumps with residual finite jumps (Table A.4), and under variance-gamma type-jumps (Table A.5). To conclude, sparse sampling at approximately 1-minute intervals provides a simple mitigation strategy, as we observe that sampling at higher or lower frequencies leads, on average, to an increase in spurious rejections.

Jump Test

The null hypothesis $H_0 : \Omega_T^c$ holds when the underlying model follows a diffusion process, and therefore Table 1 reports the size of the PZ tests. When prices have no noise, both versions of the PZ tests produce sizes close to the theoretical size across all sampling frequencies. While the noise-robust version of the test perform extremely well irrespective of the type of noise employed, the standard PZ test is oversized under microstructure effects. Nonetheless, the standard PZ test appears somewhat robust to Gaussian noise, with a Type I error rate close to the theoretical size when returns are sampled at 30-second intervals or longer. This demonstrates the sensitivity of the standard version to microstructure noise, which increases the spurious detection of jumps, making it difficult to identify true jump days.

Tables 2 and 3 report results when the underlying process follows a finite- and infinite-jump diffusion, respectively. Given the previous findings for size, it is unsurprising that the standard PZ test shows power values of 1 across most sampling frequencies. This is

primarily because microstructure noise causes the test to explode to infinity, rendering it unable to distinguish between the null and alternative hypotheses. To illustrate this, the size of the test when $\lfloor 1/\Delta_n \rfloor = 780$ under t-distributed noise ($df = 2.5$) is 0.744. This suggests that the test is likely to identify almost every day as a jump day, even when the true process has no discontinuities. Conversely, the noise-robust PZ test, S_t^{Jnoise}, is unaffected by microstructure noise, thereby achieving similar finite sample performance relative to the no noise case. Additionally, our results remain qualitatively similar under various levels of jump intensity (Tables A.1–A.3), infinite jumps with residual finite jumps (Table A.4), and under variance-gamma type-jumps (Table A.5).

Despite the capacity of the PZ tests to detect jumps of finite and infinite activity, it is surprising to find such high power levels when $\beta = 1.50$. As $\beta \to 2$, the infinite jumps are akin to Brownian increments, whereby one would expect the tests to struggle in disentangling the small jumps from the Brownian increments. However, as shown in Figure 3, the distribution of both tests is shifted to the right, confirming their ability to detect small jumps.

To summarise, our analysis reveals that the standard PZ test encounters minimal microstructure problems when dealing with Gaussian noise. However, t-distributed and Gaussian-T mixture noise increase the spurious detection of jumps under this standard test. In contrast, the noise-robust PZ test demonstrates consistent performance in detecting jumps of both finite and infinite activity, regardless of the presence or type of microstructure noise.[15]

3.2.2 Finite Sample Performance under Price Staleness

Pure Jump Test

Tables 5 and 6 present the empirical sizes of the KLJ test when the underlying process is contaminated with price staleness. We initially focus on Table 5, which reports the results for a diffusion process ($\lambda = 0$) and finite jump-diffusion processes ($\lambda > 0$). As

[15] As a robustness-check, we considered the tests of Aït-Sahalia and Jacod (2009) and their noise-robust version (Aït-Sahalia et al., 2012). Consistent with the findings in Maneesoonthorn et al. (2020), both tests show a sharp decrease in power after 5-seconds, with the standard test adversely affected by all types of noise. Results are available upon request.

can be seen, the presence of zero returns moderately increases the Type I error of test at higher frequencies for both the high and medium level of price staleness. However, as the sampling frequency decreases, the size of the test converges to its theoretical values. This observation holds true irrespective of the presence or absence of finite jumps. When price staleness is added to the infinite jump-diffusion process (Table 6), we observe very little distortions with values that are slightly larger than the theoretical size.

Infinite Activity Jump Test

When the diffusion or finite jump-diffusion processes (Tables 5 and 6) are contaminated with zero returns, we observe a moderate downward bias in the performance of the test. For instance, at the 5- and 60-second frequencies, the test exhibits a power of 0.999 and 0.940, respectively, under the diffusion process in Table 1. However. when a high level of price staleness is added to the efficient price, the power of the test decreases to 0.988 and 0.902, respectively. This increase in Type II error, although moderate, suggests that the ratio of power variations computed at different frequencies becomes slightly distorted in the presence of price staleness when the alternative hypothesis is true. This observation aligns with the findings reported by Kolokolov and Renò (2024) for a test based on power variations estimated at different frequencies.[16] By contrast, when the infinite jump-diffusion process is contaminated with zero returns, the test is slightly oversized, although the values remain close to the theoretical size. One possible rationale for this finding is that the heavy-tailed nature of the infinite jumps might dominate the impact of zero returns on the test statistic. This suggests that the underestimation of power variations induced by zero returns is less severe when jumps have infinite activity.

Jump Test

We start by analyzing the size of the test, specifically the scenario where $\lambda = 0$ in Table 5. The findings indicate that both tests are oversized at very high frequencies. However, as the sampling frequency decreases, the size of the tests converges towards

[16] Kolokolov and Renò (2024) study the jump activity index of Todorov and Tauchen (2010) under a diffusion process.

the theoretical values. Analogously, the power of the tests, as reported in Table 5 for finite jumps ($\lambda > 0$) and in Table 6 for infinite jumps, is also very close to the theoretical values, irrespective of whether finite or infinite jumps are considered. As expected, the power of the robust PZ test diminishes more rapidly than that of the standard PZ as the sampling frequency decreases. This is because the test loses efficiency as the sampling interval increases. The robustness of the PZ tests against price staleness has also been documented by Kolokolov and Renò (2024), who show that this test performs as good as their staleness-robust correction. The good performance is due to the reliance on power variations as opposed to multipower variations, which are known to be very sensitive to price staleness.

3.2.3 Recommendations and Practical Guidance

The Monte Carlo study evaluates test performance under realistic market conditions, focusing on microstructure noise and price staleness across sampling frequencies. Without market frictions, all tests show good finite sample properties, though performance declines at lower frequencies. The introduction of microstructure noise has varying impacts: Gaussian noise mainly affects tests at very high frequencies, while t-distributed and Gaussian-T mixture noise substantially impair all tests except the noise-robust PZ test. These effects can be mitigated by sampling less frequently – every 30-60 seconds for Gaussian noise and every 60 seconds for non-Gaussian noise. The standard PZ test shows particular sensitivity to non-Gaussian noise with low degrees of freedom. Price staleness creates moderate distortions at higher frequencies, but these largely disappear with 1-minute sampling, with results remaining robust across different jump intensities and types.

While our Monte Carlo simulations suggest that 60-seconds sampling offers an effective compromise between statistical power and market friction effects, this frequency represents an optimal choice under average market conditions. In practice, the optimal sampling frequency likely varies both across assets and over time, depending on market microstructure characteristics. During periods of high liquidity, characterized by tight

bid-ask spreads and low price staleness, sampling frequencies higher than 60-seconds improve the accuracy of the tests. Conversely, during periods of market stress or for less liquid assets, where spreads widen and price staleness increases, lower sampling frequencies may help controlling the Type I and II error of the tests. Regular monitoring of liquidity metrics can help identify when sampling frequency adjustments are needed.

4 Empirical Study

4.1 Data

Our empirical application considers 100 individual stocks from the S&P 500 basket and the SPDR S&P 500 ETF (SPY) for the period January 3, 2000 to December 30, 2022.[17] Ten representative stocks taken from each sector, vary in terms of liquidity and market capitalization, so ensuring a heterogeneous representation on each sector. Taking direction from our simulations, we sample the data every 1-minute, i.e. $\lfloor 1/\Delta_n \rfloor = 390$. The use of the noise-robust PZ test throughout our empirical analysis is motivated by its excellent performance both in the absence and presence of microstructure noise, and because the standard PZ test is sensitive to non-Gaussian noise even when the data are sampled using 1-minute intervals. To mitigate spurious detection of jumps, which is inherent to multiple testing, we identify rejection rates using FDR-adjusted p-values in the spirit of Bajgrowicz et al. (2016). The subsequent analysis employs the same parameters as those considered in the Monte Carlo section.

4.2 Empirical Rejections by Sector and Market Capitalization

Panel A of Table 7 reports proportion of rejection rates over our sample for the noise-robust PZ test S_t^{Jnoise}, the ASJ test, S_t^{IA}, and the KLJ test, $\mathcal{T}_t$. For each sector, we report the average across the 10 individual stocks. Panel B shows the contribution of the

[17] Our data is obtained from Refinitiv Tick History. We apply standard filtering procedures to eliminate outliers, following the approach outlined by Barndorff-Nielsen et al. (2008). We then match our data with the Center for Research in Security Prices (CRSP) database, which allows us to adjust for stock split and dividends. Finally, we sample down to 1-minute intervals using the previous tick interpolation method.

continuous and discontinuous component to the total variance estimated as $B\left(2,\infty,\Delta_n\right)_t$. For completeness, we also report the index of jump activity, $\widehat{\beta}$, estimated as in Jing et al. (2012b). We estimate the variables of Panel B as follows: $C_t = B\left(2,\infty,\Delta_n\right)_t \cdot \mathbb{1}(no\ jumps) + B\left(2,\nu_n,\Delta_n\right)_t \cdot \mathbb{1}(jumps)$. Therefore, $J_t = B\left(2,\infty,\Delta_n\right)_t - C_t$. Finally, $C_T = \frac{\sum_{t\in(0,T]} C_t}{\sum_{t\in(0,T]} C_t + J_t}$ and $J_T = \frac{\sum_{t\in(0,T]} J_t}{\sum_{t\in(0,T]} C_t + J_t}$. The contribution of finite and infinite jumps to the total jump component are obtained as $FJ_t = J_t \cdot \mathbb{1}(finite\ jumps)$ and $IJ_t = J_t \cdot \mathbb{1}(infinite\ jumps)$. Thus, $FJ_T = \frac{\sum_{t\in(0,T]} FJ_t}{\sum_{t\in(0,T]} J_t}$ and $IJ_T = \frac{\sum_{t\in(0,T]} IJ_t}{\sum_{t\in(0,T]} J_t}$.

Table 7 presents the main results that provide strong evidence for the presence of a Brownian component, with less than 0.1% of days in our sample exhibiting a pure jump process.[18] This finding effectively rules out the hypothesis of a pure jump process and confirms the relevance of the Brownian component for modelling the diffusive part of the process. In addition to pronounce presence of the Brownian component, Panel A also indicates a substantial jump component in both the SPY and individual stocks across all sectors. Jumps are observed on an average of 19% of the days across the sectors, with the Energy sector displaying the lowest proportion of jumps (16% of days) and the Communication Services sector showing the highest proportion (22% of days). The SPY exhibits jumps on only 13% of days within the sample period. The lower frequency of jumps in the SPY can be attributed to the fact that aggregation reduces idiosyncratic risk, thereby diversifying away stock specific jumps.

Having established the presence of a Brownian and a jump component, we now evaluate whether the jump part exhibits finite and/or infinite activity. Table 7 (second column) reports the rejection rates of the ASJ test for both the SPY and the individual stocks classified by sector. The ASJ test provides evidence for the presence of both types of jumps, with rejection rates of 97% for the SPY and an average of 83% across sectors. The index of jump activity, $\widehat{\beta}$, oscillates around 1.0 for all the sectors and SPY. This finding corroborates that both finite and infinite jumps activity characterize the jump component of our 1-minute data.

Given that the alternative hypothesis of the ASJ test could indicate either finite

[18]The proportion of rejection rates without employing the FDR-adjusted p-values ranges between 1.5% and 3.5%, further corroborating the strong presence of a Brownian component.

jumps or Brownian motion, in Panel B (Table 7) we present the contribution of these variables to total variance. The continuous part contributes to 97% and 92% of the total variance for SPY and the sector average, respectively. Consequently, only 3% and 8% of the total variance for SPY and the sector average, respectively, can be attributed to jumps. Of these jumps components, 82% (for SPY) and 77% (for the sector average) are contributions from finite jumps. Thus, evidence supporting both types of jumps is observed. Finite jumps are typically associated with macroeconomic announcements and stock-specific news, which are more likely to generate spillover effects. Conversely, infinite jumps can be linked to continuously adjusting market microstructure dynamics, which may be attributable, but not limited to, inventory allocation, price-contingent trading, stop-loss and limit-profit orders, among other factors.

As shown by Jiang and Yao (2013), stocks with different size and liquidity levels exhibit display different levels of jump returns, and exploiting these cross-sectional differences in jumps fully account for the size and illiquidity effects. Motivated by their findings, we present rejection rates and contribution to total variance classified by sector and market capitalization. This analysis aim to shed light on the different underlying components of stocks with varying size and liquidity within the same sector.[19] We report these results in Table 8. As expected, stocks with smaller market capitalization tend to jump more frequently than those with larger size. Specifically, we find that stocks with smaller size in our sample have, on average, 2% more jump days than their largest counterparts. The most pronounced variation is observed in the Communication Services sector, where 24% of days exhibit jumps for companies with small market capitalization, compared to 19% for those with large market capitalization. This increase in the number of days with jumps, directly translate into a higher contribution to the jump component to the total variance. Again, using the Communication Services as an example, the contribution of the jump component accounts for 8% of the total variance for large-cap stocks, as opposed to the 13% observed for small-cap companies. The larger contribu-

[19] A company is classified as "Big Market Cap" if its average market capitalization exceeds the median market capitalization, where the median is computed using the 10 stocks within each sector. Conversely, companies with average market capitalization below this median are classified as "Small Market Cap".

tion of jumps to total variance has important implications for asset allocation and risk management. For instance, a risk averse investor might be expected to avoid investments with large unforeseeable movements.

Based on that the PZ test can be sensitive to the choice of the threshold, we repeat our empirical exercise by setting $\alpha = 4$. The increase of the threshold makes the test, and therefore, the identification of jumps, more conservative. The results are reported in Tables A.6 and A.7. As can be seen, increasing the threshold reduces the proportion of days with jumps from 13% to 8% for SPY, and from 19% to 11% for the sector average. While we observe a reduction in the proportion of days with jumps, the results remain qualitatively similar and supportive of the presence of jumps of both finite and infinite activity.

These findings complement those previously reported in the literature by Huang and Tauchen (2005) and Dumitru and Urga (2012). For instance, Huang and Tauchen (2005) document that approximately 10% of the days in the S&P 500 index contain jumps, with the jump component contributing to 8% of the total variance. Similarly, Dumitru and Urga (2012) find that jumps comprise between 10% to 14% of the total variance for a small sample of individual stocks. Our study extends these finding by providing evidence for the existence of both finite and infinite jumps. We document that finite jumps contribute, on average, 77% to the total jump component, while infinite jumps account for the remaining 23%. Moreover, we find that the jump component as a whole comprises 8% of the total variance, aligning with previous studies.

4.3 Time-varying Rejections

Over the sample period, covering the years from 2000 to 2022, financial markets experienced several crises periods (dot-com, sub-prime, European sovereign debt, Covid-19 pandemic) as well as the Brexit referendum. These along with the development of efficient trading systems, including electronic platforms and algorithmic trades, contributed to more frequent and faster reactions to changes in the market, which could easily generate time variation in statistical properties. To further investigate this issue, we follow

Erdemlioglu et al. (2015) and fit a probit model to the daily series of the rejection indicators $(0, 1)$ for the noise-robust PZ and ASJ test. We omit the KLJ test, as the rejection rates are almost undistinguishable from zero. We use a sixth order polynomial of time in a probit model, [20] where regressors are a constant, and time trends up to a sixth power. We have orthogonalized and standardized them to have unit variance.

Figure 4 plots, for SPY, the fitted values of FDR-adjusted rejection rates for the robust PZ and ASJ tests. We find significant time variation in the PZ and ASJ tests, with at least 4 significant polynomial coefficients per regression. In general, there is a negative correlation in the evolution of the time-varying rejections. This outcome implies that an increase in the number of jump days usually results in an increase of infinite jumps vis-à-vis finite jumps. The most notable feature of these series is the systematic increase in the rejection of the null of no jumps around 2002, 2008, 2016 and 2020.

To illustrate, these periods experienced significant market events. The dot-com bubble burst, with the NASDAQ falling by 78% in October 2002; the Global Financial Crisis (GFC) in 2007–2009, during which the S&P 500 index fell by 9% in October, 2008; the 2016 US Presidential election, which triggered a fall in the S&P 500 index of more than 5%, creating a circuit breaker to halt trading activity; and the Covid-19 pandemic, which created one of the most turbulent periods, with the S&P 500 index falling as much as 12% in March, 2020. These are examples of very large jumps. However, the increase in rejection rates of the noise-robust PZ test and the decrease in rejection rates of the ASJ test suggest that these periods are also accompanied by a large increment in small jumps, associated to electronic price-contingent trading to control for losses during these periods of high uncertainty.

To underscore the time variation and the negative relationship implying that an increase in the number of days with jumps results in an increase of infinite jumps, Figure 5 plots the rejection rates for one representative stock from each sector. All stocks display very strong time variation in rejection rates with a pronounced negative relation between these rates. As a robustness exercise, we assess whether time-variation still persists in

[20] Although the sixth order polynomial provides the best fit for the data, other order polynomials provide qualitatively similar conclusions.

the rejection rates of the noise-robust PZ test when we increase the threshold parameter α from 3 to 4. To facilitate comparison, we repeat the exercise by plotting the rejection rates of both the noise-robust PZ (with $\alpha = 4$) and the ASJ test. Figures A.1 and A.2 depict the time-varying rejection rates for SPY and one representative stock per sector, respectively. It is evident that both SPY and individual stocks exhibit strong time variation in rejection rates with a pronounced negative relationship between them.

4.3.1 Optimal Sampling Frequency Under Time-Varying Market Frictions

Our Monte Carlo simulations demonstrate that 60-second return sampling achieves an optimal balance between statistical power and market friction effects. However, these market frictions, including liquidity levels and price staleness, generally exhibit temporal variation (e.g., Zhu and Liu, 2024). Given this time-varying nature of market frictions, a uniform sampling frequency may not be optimal across different assets and time periods. We therefore recommend conducting pre-implementation analyses that focus on three key metrics: the noise variance estimator, price staleness indicators, and liquidity measures. When these analyses reveal moderate levels of noise variance and price staleness, sampling frequencies between 30 and 60 seconds may be appropriate. Conversely, in periods characterized by elevated noise and price staleness, longer sampling intervals ranging from 90 to 300 seconds may yield more reliable results. These findings suggest implementing an adaptive sampling strategy that adjusts to varying levels of market microstructure noise and price staleness, with our Monte Carlo simulations serving as calibration benchmarks.

5 Conclusions

In this paper, we studied the finite sample performance of high-frequency test statistics developed to identify the different underlying components of high-frequency financial data in terms of Brownian component (Kong et al., 2015), jump component (Podolskij and Ziggel, 2010), and the activity of the jumps (i.e., finite vs infinite jumps) (Aït-Sahalia and Jacod, 2011), under the presence of various types of market microstructure noise,

different levels of jump activity and price staleness. The performance is gauged under the assumption of Gaussian, t-distributed, and Gaussian-T mixture noise for the microstructure component. For price staleness, we set a high and medium level of staleness, selected empirically from our data. The exact knowledge of these features is crucial for correct inference in modelling stock price data in the context of optimal option pricing, risk management, and portfolio allocation.

The results of the Monte Carlo simulation exercise showed that the presence of microstructure noise generally skewed the distribution of all the test statistics, with the exception of the noise-robust PZ test, as their asymptotic distributions are derived under the assumption of noiseless prices. This skew increased the Type I or Type II errors. Specifically, we found that Gaussian noise only affected the distribution of the tests at very high frequencies, while t-distribution and Gaussian-T mixture noise completely biased the performance of the tests. These findings did not apply for the noise-robust PZ test, which performed extremely well under any type of microstructure noise. Sampling returns every 30 seconds when the noise is Gaussian, and every 60 seconds when the noise is t-distributed or Gaussian-T mixture, considerably reduced the microstructure noise effects. On the other hand, contaminating the efficient price with price staleness moderately distorted the tests at higher-frequencies, although sampling sparsely seems to eliminate most of these distortions.

We applied these tests to 100 individual stocks and SPY using 23 years of 1-minute interval data. The results strongly support the presence of both Brownian motion and jumps, rejecting the pure jump process hypothesis. Jumps occurred on 13% of days for SPY and 19% across sectors. The continuous component accounted for 97% of total variance in SPY and 92% in sector averages. Of the jump component, finite jumps comprised 97% of total variance in SPY and 92% in sector averages. From the total variance attributed to jumps (3% for SPY, 8% for sector average), finite jumps contributed 82% and 77%, respectively.

Our results also indicated significant time variation in rejection rates. These variations range from 5% to 20%, where generally an increase in the rejection of the null of no jumps

is associated with a decrease in the rejection of infinite jumps. These findings suggest that the most appropriate specification for modelling stock price data should allow for a Brownian motion, jumps of finite and infinite activity, with the jump component having both time-varying activity and intensity.

Finally, it is important to note that recent literature has identified the presence of drift bursts (e.g., Christensen et al., 2022), pockets of extreme return persistence (e.g., Andersen et al., 2023), and time-varying price staleness (e.g., Zhu and Liu, 2024). These phenomena may significantly impact the finite sample properties of test statistics. While a detailed investigation of these effects is beyond the scope of this paper, it presents an important avenue that we leave for future research.

References

Aït-Sahalia, Y. and Jacod, J. (2009). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1):184–222.

Aït-Sahalia, Y. and Jacod, J. (2010). Is Brownian motion necessary to model high-frequency data? *The Annals of Statistics*, 38(5):3093–3128.

Aït-Sahalia, Y. and Jacod, J. (2011). Testing whether jumps have finite or infinite activity. *The Annals of Statistics*, 39(3):1689–1719.

Aït-Sahalia, Y., Jacod, J., and Li, J. (2012). Testing for jumps in noisy high frequency data. *Journal of Econometrics*, 168(2):207–222.

Aït-Sahalia, Y., Kalnina, I., and Xiu, D. (2020). High-frequency factor models and regressions. *Journal of Econometrics*, 216(1):86–105.

Ait-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies*, 18(2):351–416.

Ait-Sahalia, Y. and Yu, J. (2009). High frequency market microstructure noise estimates and liquidity measures. *Annals of Applied Statistics*, 3(1):422–457.

Andersen, T. G., Benzoni, L., and Lund, J. (2002). An empirical investigation of continuous-time equity return models. *The Journal of Finance*, 57(3):1239–1284.

Andersen, T. G., Li, Y., Todorov, V., and Zhou, B. (2023). Volatility measurement with pockets of extreme return persistence. *Journal of Econometrics*, 237(2):105048.

Bachelier, L. (1900). Théorie de la spéculation. In *Annales Scientifiques de l'École Normale Supérieure*, volume 17, pages 21–86.

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

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

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

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N. (2008). Realized kernels in practice: trades and quotes. *Econometrics Journal*, 4:1–32.

Barndorff-Nielsen, O. E. (1997). Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68.

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.

Brockwell, P. J. (2001). Continuous time ARMA model. In Shanbhag, D. N. and Rao, C. R., editors, *Stochastic Processes: Theory and Methods*, volume 19, pages 249–276. North-Holland, Amsterdam.

Bu, R., Hizmeri, R., Izzeldin, M., Murphy, A., and Tsionas, M. (2023). The contribution of jump signs and activity to forecasting stock price volatility. *Journal of Empirical Finance*, 70(1):144–164.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business*, 75(2):305–332.

Christensen, K., Oomen, R., and Renò, R. (2022). The drift burst hypothesis. *Journal of Econometrics*, 227(2):461–497.

Cont, R. and Mancini, C. (2011). Nonparametric tests for pathwise properties of semimartingales. *Bernoulli*, 17(2):781–813.

Dumitru, A.-M. and Urga, G. (2012). Identifying jumps in financial assets: a comparison between nonparametric jump tests. *Journal of Business & Economic Statistics*, 30(2):242–255.

Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, 1(3):281–299.

Erdemlioglu, D., Laurent, S., and Neely, C. J. (2015). Which continuous-time model is most appropriate for exchange rates? *Journal of Banking & Finance*, 61:S256–S268.

Erdemlioglu, D. and Yang, X. (2022). News arrival, time-varying jump intensity, and realized volatility: Conditional testing approach. *Journal of Financial Econometrics*, pages 1–38.

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

Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Processes and their Applications*, 118(4):517–559.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Applications*, 119(7):2249–2276.

Jacod, J. and Todorov, V. (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *The Annals of Statistics*, 42(3):1029–1069.

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

Jiang, G. J. and Yao, T. (2013). Stock price jumps and cross-sectional return predictability. *Journal of Financial and Quantitative Analysis*, 48(5):1519–1544.

Jing, B.-Y., Kong, X.-B., and Liu, Z. (2012a). Modeling high-frequency financial data by pure jump processes. *The Annals of Statistics*, 40(2):759–784.

Jing, B.-Y., Kong, X.-B., Liu, Z., and Mykland, P. (2012b). On the jump activity index for semimartingales. *Journal of Econometrics*, 166(2):213–223.

Klüppelberg, C., Lindner, A., and Maller, R. (2004). A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. *Journal of Applied Probability*, 41(3):601–622.

Kolokolov, A., Livieri, G., and Pirino, D. (2020). Statistical inferences for price staleness. *Journal of Econometrics*, 218(1):32–81.

Kolokolov, A. and Renò, R. (2024). Jumps or staleness? *Journal of Business & Economic Statistics*, 42(2):516–532.

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

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

Lee, S. S. and Mykland, P. A. (2012). Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics*, 168(2):396–406.

Madan, D. B. and Seneta, E. (1990). The variance gamma (VG) model for share market returns. *Journal of Business*, pages 511–524.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, 64(19-47):44.

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

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, 116(2):257–270.

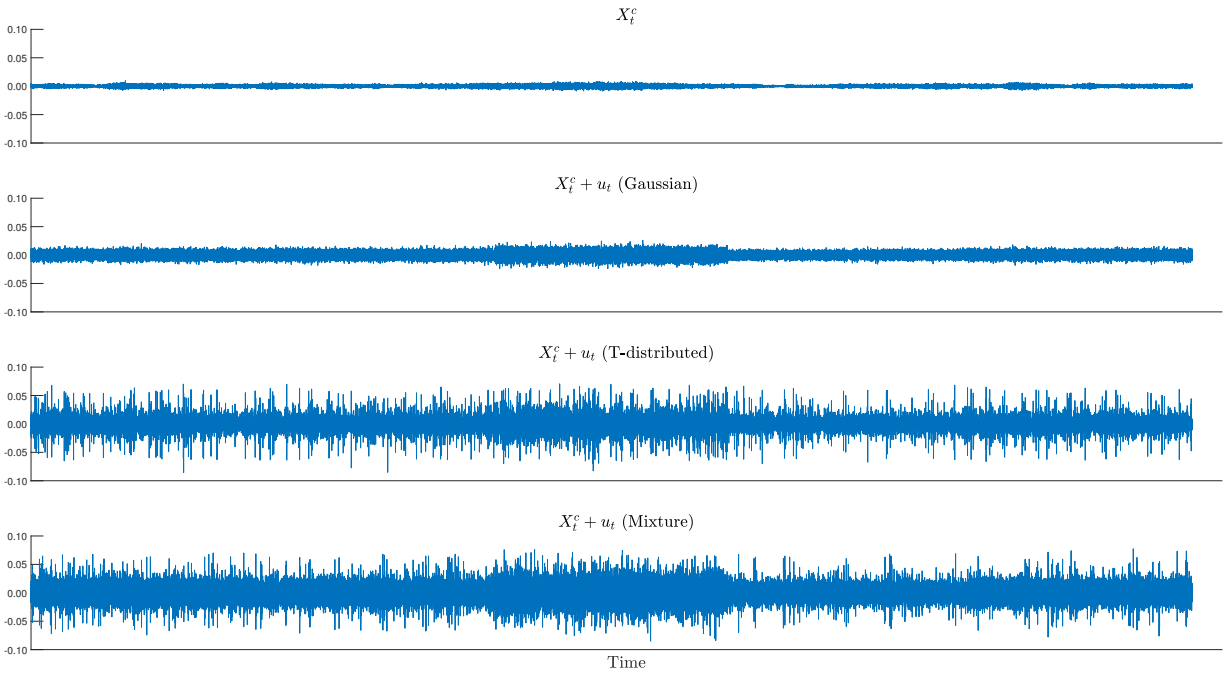
Podolskij, M. and Ziggel, D. (2010). New tests for jumps in semimartingale models. *Statistical Inference for Stochastic Processes*, 13(1):15–41.

Todorov, V. and Tauchen, G. (2010). Activity signature functions for high-frequency data analysis. *Journal of Econometrics*, 154(2):125–138.

Zhu, H. and Liu, Z. (2024). On bivariate time-varying price staleness. *Journal of Business & Economic Statistics*, 42(1):229–242.

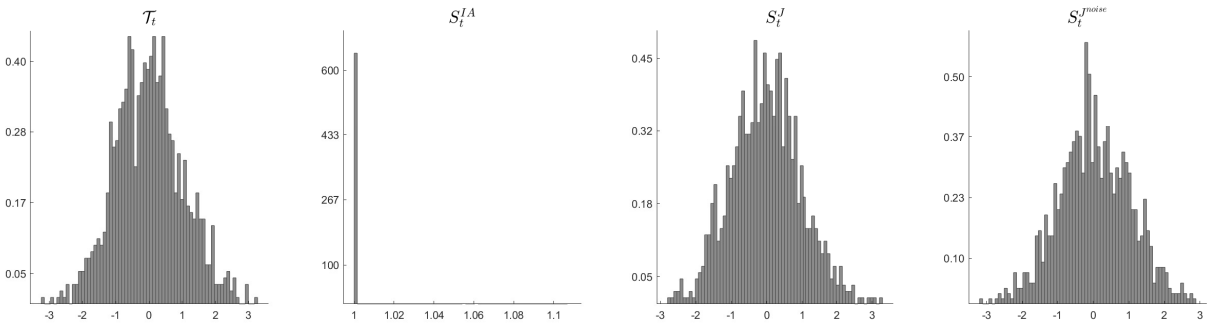
Tables and Figures

Figure 1: A Realization of the True and Contaminated Continuous Part of the log-price Increments

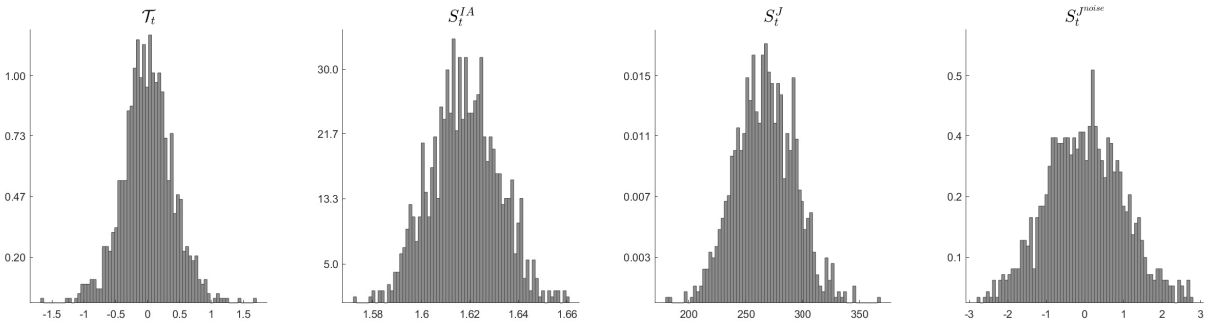


Note: The figures depicts a realization of the noisy continuous part of the log price increments. From top to bottom, the continuous part is contaminated with no noise, Gaussian noise, t-distributed noise, and Gaussian-T mixture noise.

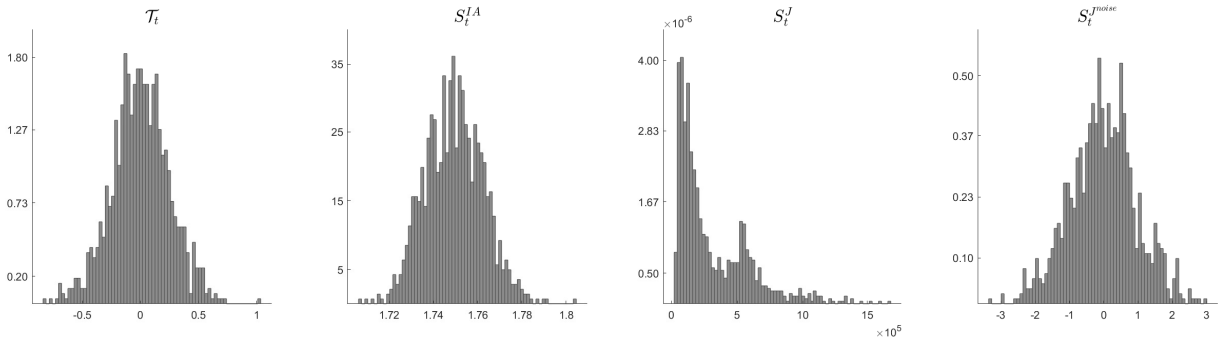
Figure 2: Distribution of the Tests Statistics using a Diffusion Process with Different Types of Noise



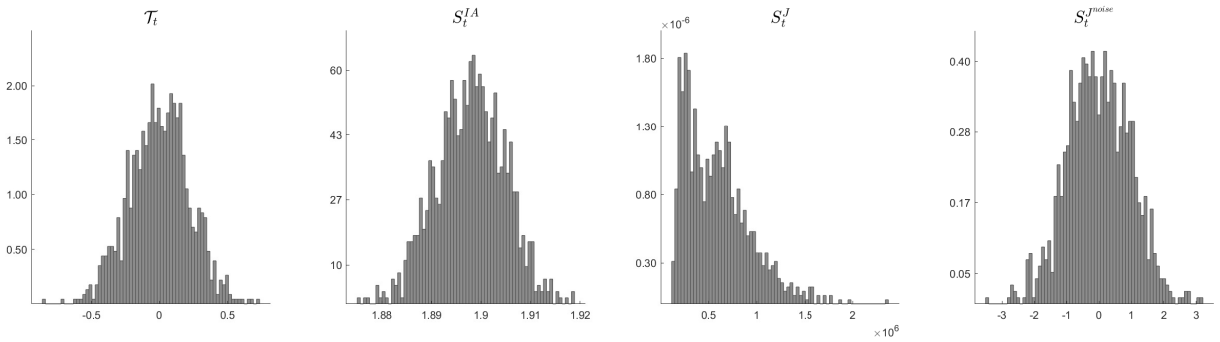
(a) X_t^c $(u_t = 0)$



(b) $X_t^c + u_t$ (Gaussian)



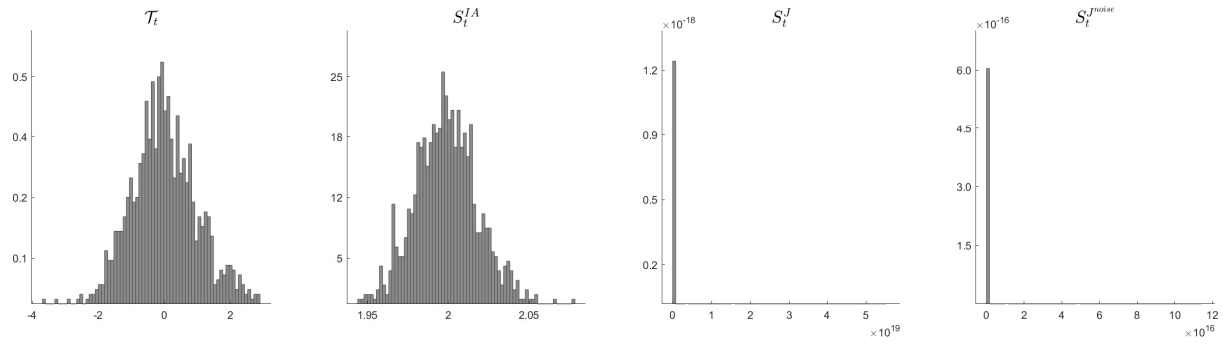
(c) $X_t^c + u_t$ (t-distributed, $df = 2.5$)



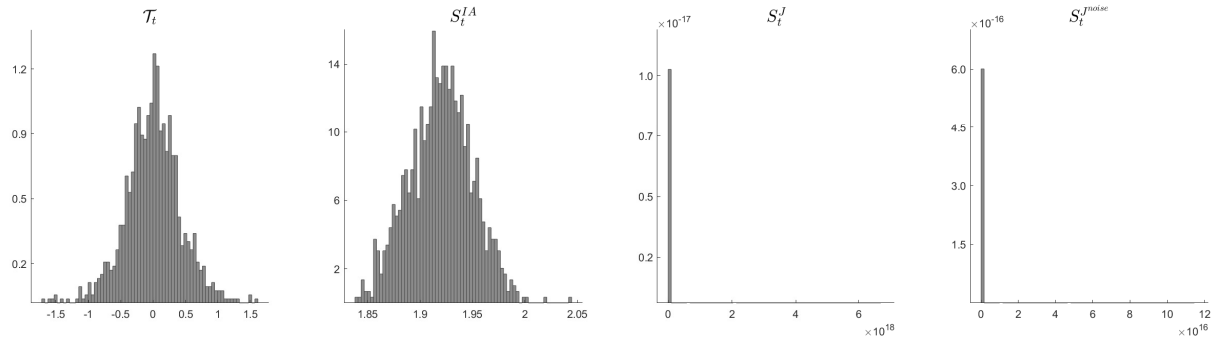
(d) $X_t^c + u_t$ (Gaussian-T Mixture)

Note: The figure plots the simulated distribution of the different tests under a diffusion process (equation (16)) with (a) no noise, (b) Gaussian noise, (c) t-distributed noise with 2.5 degrees of freedom, and (d) Gaussian-T mixture noise. X_t^c denotes the diffusive component.

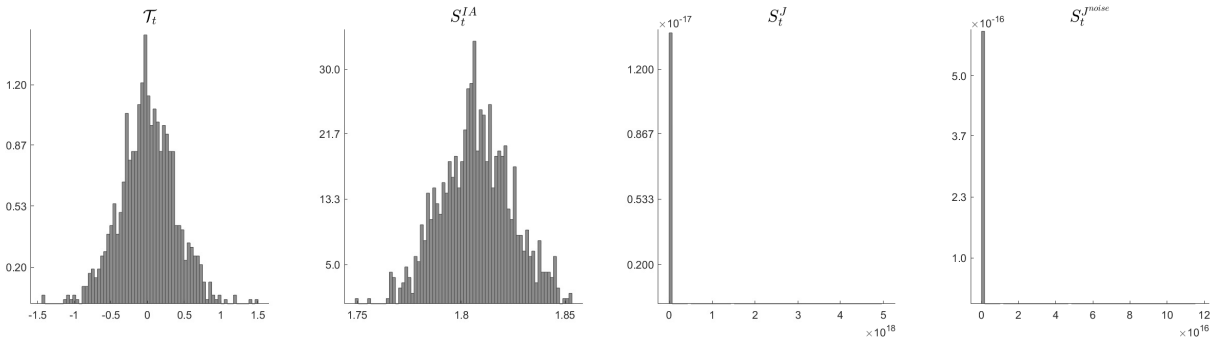
Figure 3: Distribution of the Test Statistics using an Infinite Jump Diffusion Process with Different Types of Noise



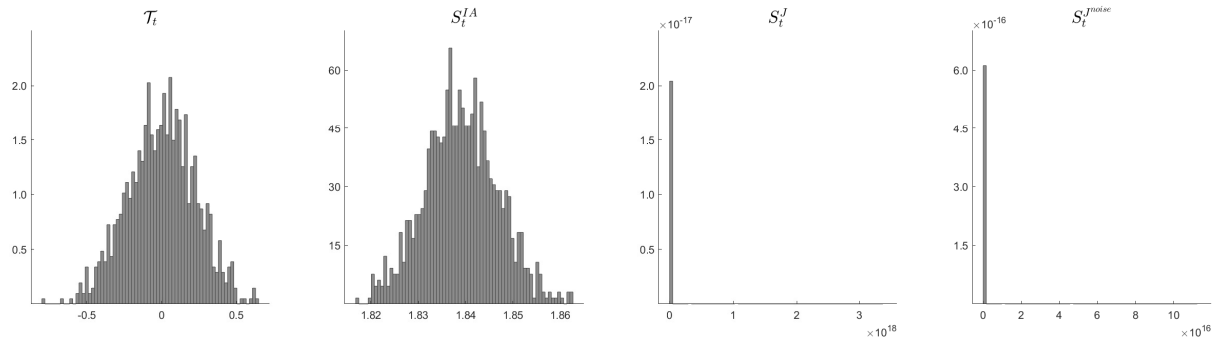
(a) $X_t^c + X_t^d$ $(u_t = 0)$



(b) $X_t^c + X_t^d + u_t$ (Gaussian)



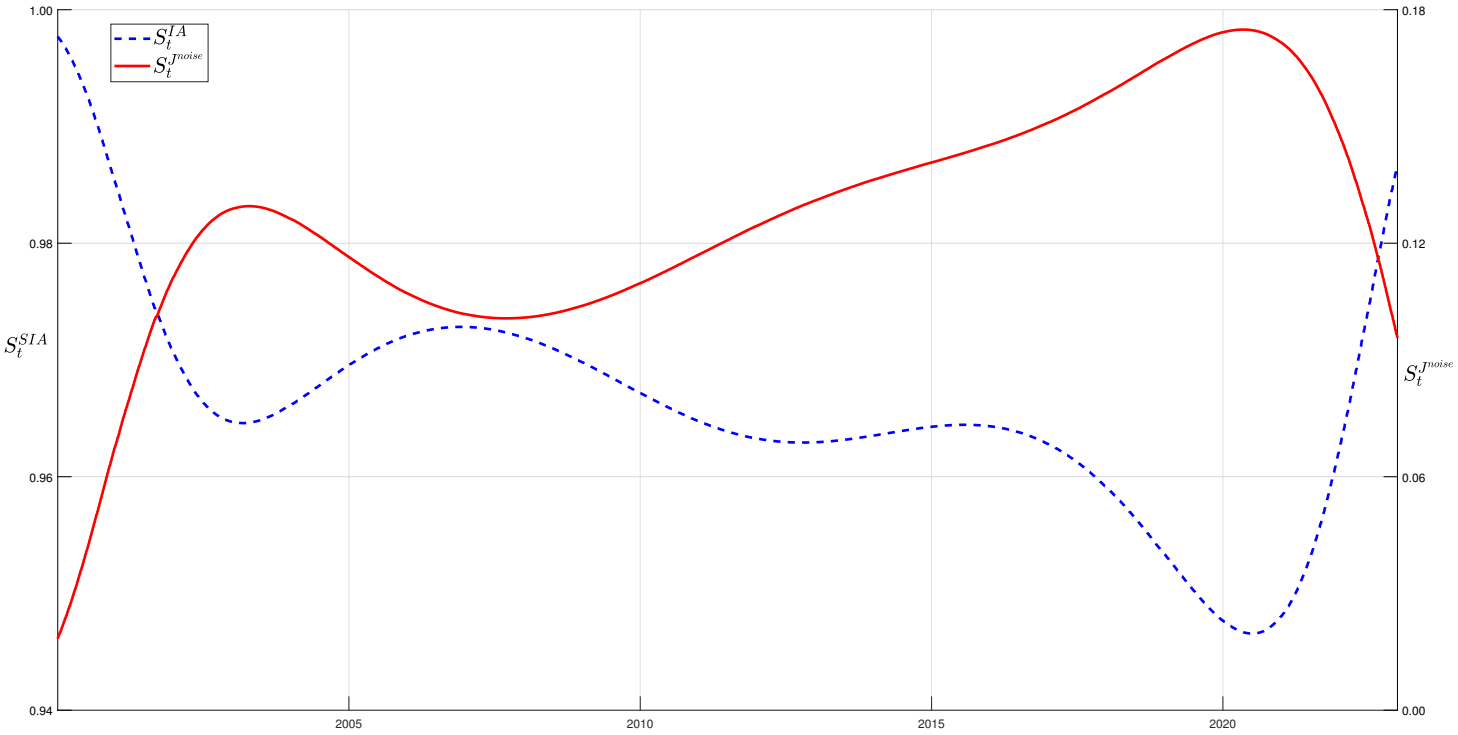
(c) $X_t^c + X_t^d + u_t$ (t-distributed, $df = 2.5$)



(d) $X_t^c + X_t^d + u_t$ (Gaussian-T Mixture)

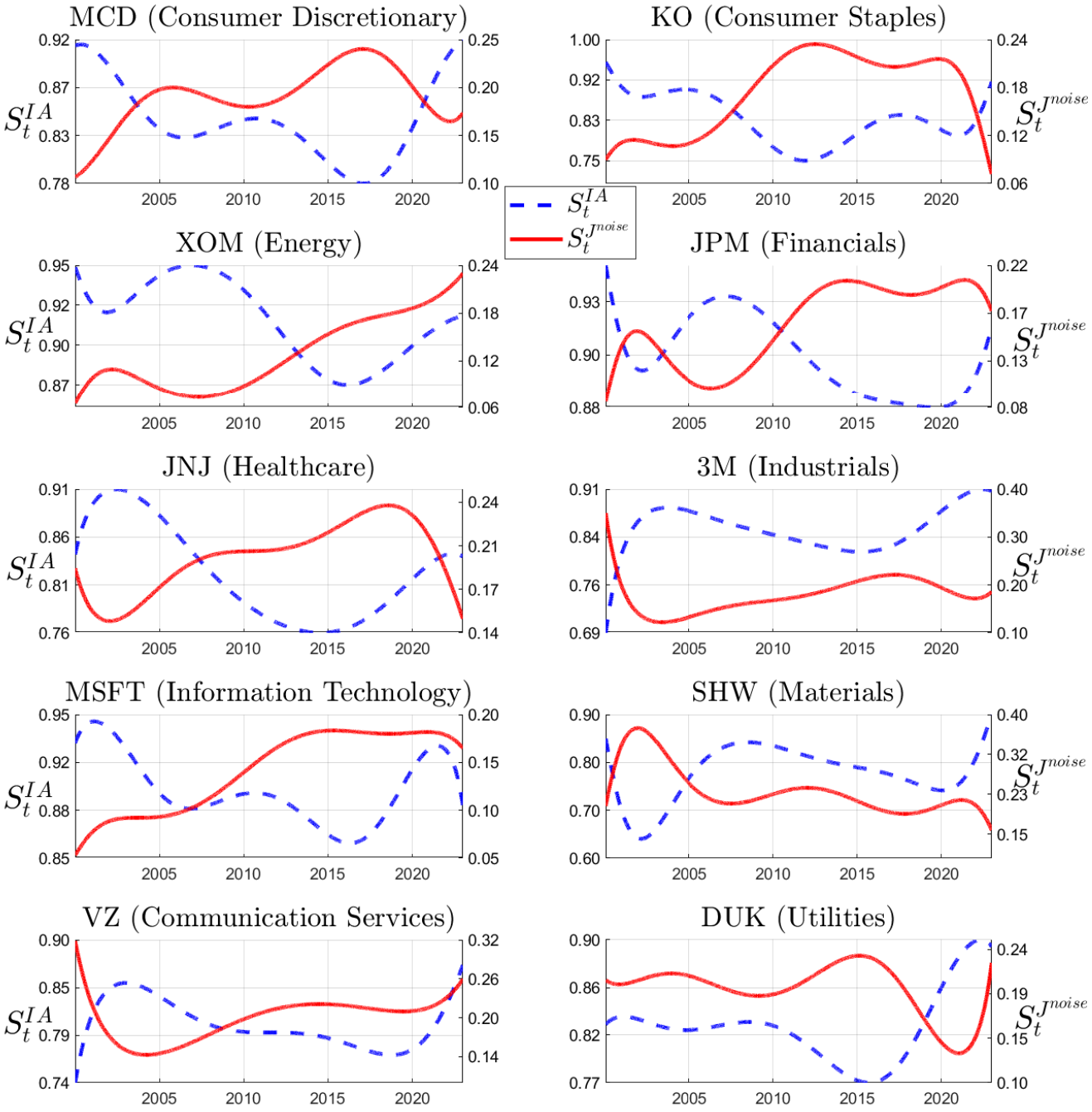
Note: The figure plots the simulated distribution of the different tests under a infinite jump diffusion process (equation (16) and $\beta = 1.0$) with (a) no noise, (b) Gaussian noise, (c) t-distributed noise with 2.5 degrees of freedom, and (d) Gaussian-T mixture noise. X_t^c and X_t^d denote the diffusive and pure jump components.

Figure 4: SPY – Time Variation in Rejection Rates



Note: The figure depicts the time variation in rejection rates predicted by a probit model with a 6th order polynomial in time. The left (right) y-axis denotes the probability of rejection over time for the ASJ (noise-robust PZ test).

Figure 5: Individual Stocks – Time Variation in Rejection Rates



Note: The figure depicts the time variation in rejection rates predicted by a probit model with a 6th order polynomial in time. The left (right) y-axis of each subplot denotes the probability of rejection over time for the ASJ (noise-robust PZ test).

Table 1: Monte Carlo Rejection Rates under a Diffusion Process

$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
	No Noise				Gaussian Noise			
$\mathcal{T}_t$	0.012	0.015	0.018	0.021	0.003	0.017	0.025	0.030
S_t^J	0.012	0.014	0.016	0.015	0.204	0.023	0.019	0.019
S_t^{Jnoise}	0.010	0.009	0.007	0.008	0.010	0.008	0.007	0.009
S_t^{IA}	0.999	0.971	0.940	0.922	0.850	0.977	0.947	0.926
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\mathcal{T}_t$	0.001	0.003	0.006	0.014	0.000	0.004	0.008	0.012
S_t^J	1.000	0.744	0.307	0.163	1.000	0.878	0.388	0.204
S_t^{Jnoise}	0.010	0.009	0.007	0.009	0.009	0.008	0.007	0.009
S_t^{IA}	0.285	0.687	0.912	0.925	0.662	0.817	0.917	0.927
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\mathcal{T}_t$	0.002	0.008	0.012	0.017	0.003	0.010	0.014	0.018
S_t^J	1.000	0.235	0.071	0.041	1.000	0.502	0.137	0.071
S_t^{Jnoise}	0.010	0.009	0.009	0.009	0.010	0.009	0.009	0.009
S_t^{IA}	0.435	0.866	0.936	0.923	0.683	0.921	0.944	0.930
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\mathcal{T}_t$	0.006	0.009	0.018	0.021	0.003	0.010	0.019	0.025
S_t^J	0.999	0.080	0.031	0.024	1.000	0.331	0.084	0.050
S_t^{Jnoise}	0.010	0.010	0.009	0.009	0.010	0.009	0.009	0.009
S_t^{IA}	0.578	0.946	0.962	0.924	0.769	0.928	0.964	0.935

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under a diffusion process the KLJ ($\mathcal{T}_t$) and PZ (S_t^{Jnoise}, S_t^J) tests report their empirical size, while the ASJ test (S_t^{SIA}) reports its empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table 2: Monte Carlo Rejection Rates under a Finite Jump-Diffusion Process, $\lambda = 1.0$

$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
	No Noise				Gaussian Noise			
$\mathcal{T}_t$	0.012	0.013	0.015	0.019	0.003	0.016	0.023	0.027
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.995	0.971	0.956	0.928	0.998	0.988	0.976	0.956
S_t^{IA}	0.999	0.964	0.936	0.916	0.795	0.954	0.937	0.920
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\mathcal{T}_t$	0.004	0.004	0.009	0.013	0.000	0.004	0.008	0.011
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.998	0.988	0.976	0.956	0.998	0.985	0.965	0.935
S_t^{IA}	0.286	0.684	0.891	0.910	0.560	0.762	0.899	0.908
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\mathcal{T}_t$	0.006	0.005	0.016	0.018	0.001	0.006	0.012	0.017
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.998	0.989	0.978	0.965	0.998	0.983	0.976	0.946
S_t^{IA}	0.533	0.860	0.908	0.853	0.763	0.815	0.919	0.914
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\mathcal{T}_t$	0.009	0.008	0.018	0.022	0.002	0.009	0.011	0.018
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.998	0.988	0.974	0.955	0.998	0.981	0.971	0.954
S_t^{IA}	0.783	0.942	0.913	0.863	0.876	0.907	0.932	0.949

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under a finite-jump diffusion process the KLJ test ($\mathcal{T}_t$) reports its empirical size, while the PZ (S_t^{Jnoise}, S_t^J) and the ASJ (S_t^{SIA}) tests report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table 3: Monte Carlo Rejection Rates under an Infinite Jump-Diffusion Process

	$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
		No Noise				Gaussian Noise			
$\beta = 1.00$	$\mathcal{T}_t$	0.016	0.017	0.016	0.020	0.006	0.011	0.012	0.013
	S_t^J	0.999	0.993	0.993	0.990	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	0.996	0.993	0.991	0.979	0.996	0.992	0.990	0.978
	S_t^{IA}	0.010	0.008	0.014	0.020	0.039	0.024	0.026	0.031
$\beta = 1.25$	$\mathcal{T}_t$	0.015	0.014	0.012	0.016	0.001	0.011	0.011	0.013
	S_t^J	1.000	0.999	0.996	0.993	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	0.999	0.996	0.989	0.969	0.999	0.999	0.986	0.972
	S_t^{IA}	0.010	0.012	0.017	0.025	0.048	0.031	0.035	0.039
$\beta = 1.50$	$\mathcal{T}_t$	0.001	0.002	0.002	0.002	0.000	0.001	0.001	0.001
	S_t^J	0.999	0.993	0.993	0.991	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	0.997	0.954	0.954	0.909	0.997	0.993	0.956	0.908
	S_t^{IA}	0.024	0.031	0.037	0.045	0.105	0.047	0.041	0.044
		t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\beta = 1.00$	$\mathcal{T}_t$	0.003	0.002	0.002	0.003	0.002	0.009	0.007	0.008
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	1.000	0.983	0.969	0.929	1.000	0.981	0.962	0.919
	S_t^{IA}	0.237	0.114	0.047	0.042	0.186	0.100	0.049	0.041
$\beta = 1.25$	$\mathcal{T}_t$	0.003	0.002	0.002	0.002	0.002	0.006	0.006	0.009
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	0.997	0.958	0.926	0.908	0.997	0.957	0.926	0.904
	S_t^{IA}	0.336	0.144	0.064	0.048	0.299	0.128	0.069	0.052
$\beta = 1.50$	$\mathcal{T}_t$	0.003	0.002	0.002	0.002	0.002	0.001	0.001	0.001
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	0.996	0.927	0.905	0.879	0.996	0.927	0.909	0.880
	S_t^{IA}	0.470	0.193	0.144	0.062	0.307	0.163	0.142	0.079

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under an infinite-jump diffusion process the KLJ ($\mathcal{T}_t$) and ASJ (S_t^{SIA}) tests report their empirical size, while the PZ tests ($S_t^{J^{noise}}$, S_t^J) report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day, β is the jump activity index, and the significance level is $\theta = 0.01$.

Table 3: Monte Carlo Rejection Rates under an Infinite Jump-Diffusion Process ***(Continued)***

	$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
		t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\beta = 1.00$	$\mathcal{T}_t$	0.002	0.002	0.006	0.008	0.002	0.008	0.010	0.014
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	S_t^{Jnoise}	0.999	0.989	0.970	0.940	0.999	0.989	0.967	0.939
	S_t^{IA}	0.229	0.085	0.043	0.052	0.174	0.096	0.039	0.047
$\beta = 1.25$	$\mathcal{T}_t$	0.002	0.001	0.002	0.002	0.003	0.012	0.010	0.012
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	S_t^{Jnoise}	0.997	0.970	0.939	0.918	0.996	0.966	0.933	0.917
	S_t^{IA}	0.261	0.105	0.050	0.071	0.224	0.111	0.045	0.071
$\beta = 1.50$	$\mathcal{T}_t$	0.000	0.001	0.001	0.002	0.003	0.001	0.009	0.005
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	S_t^{Jnoise}	0.996	0.959	0.918	0.897	0.995	0.960	0.919	0.891
	S_t^{IA}	0.397	0.123	0.076	0.092	0.296	0.138	0.065	0.085
		t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\beta = 1.00$	$\mathcal{T}_t$	0.002	0.006	0.010	0.014	0.003	0.011	0.010	0.014
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	S_t^{Jnoise}	0.999	0.992	0.980	0.949	0.999	0.994	0.982	0.946
	S_t^{IA}	0.167	0.044	0.025	0.032	0.105	0.039	0.016	0.027
$\beta = 1.25$	$\mathcal{T}_t$	0.002	0.007	0.009	0.011	0.003	0.011	0.009	0.013
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	S_t^{Jnoise}	0.997	0.989	0.948	0.928	0.997	0.982	0.942	0.922
	S_t^{IA}	0.229	0.050	0.031	0.054	0.142	0.054	0.028	0.039
$\beta = 1.50$	$\mathcal{T}_t$	0.002	0.002	0.002	0.003	0.000	0.001	0.001	0.001
	S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	S_t^{Jnoise}	0.996	0.979	0.908	0.907	0.996	0.979	0.907	0.901
	S_t^{IA}	0.281	0.104	0.091	0.109	0.168	0.095	0.071	0.087

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under an infinite-jump diffusion process the KLJ ($\mathcal{T}_t$) and ASJ (S_t^{SIA}) tests report their empirical size, while the PZ tests (S_t^{Jnoise}, S_t^J) report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day, β is the jump activity index, and the significance level is $\theta = 0.01$.

Table 4: Monte Carlo Rejection Rates under a Pure Jump Process

$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
	No Noise				Gaussian Noise			
$\beta = 1.00$	0.999	0.978	0.942	0.941	0.987	0.974	0.939	0.937
$\beta = 1.25$	0.997	0.946	0.909	0.877	0.852	0.927	0.886	0.871
$\beta = 1.50$	0.954	0.903	0.798	0.810	0.307	0.836	0.762	0.794
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\beta = 1.00$	0.579	0.939	0.920	0.928	0.209	0.883	0.909	0.913
$\beta = 1.25$	0.284	0.824	0.813	0.852	0.037	0.647	0.745	0.815
$\beta = 1.50$	0.106	0.596	0.662	0.746	0.007	0.267	0.512	0.683
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\beta = 1.00$	0.843	0.968	0.930	0.935	0.260	0.924	0.919	0.922
$\beta = 1.25$	0.492	0.903	0.847	0.866	0.027	0.731	0.780	0.834
$\beta = 1.50$	0.125	0.751	0.728	0.775	0.003	0.344	0.567	0.709
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\beta = 1.00$	0.930	0.974	0.937	0.938	0.290	0.931	0.921	0.924
$\beta = 1.25$	0.644	0.915	0.854	0.868	0.027	0.749	0.785	0.834
$\beta = 1.50$	0.171	0.798	0.744	0.785	0.002	0.377	0.580	0.716

Note: The table reports rejection rates across sampling frequencies for the KLJ test. Under a pure jump process the KLJ test reports its empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations, β is the jump activity index per day, and the significance level is $\theta = 0.01$.

Table 5: Monte Carlo Rejection Rates under (Finite Jump-)Diffusion Process and Price Staleness

	$[1/\Delta_n]$	4680	780	390	260	4680	780	390	260
		High Level				Medium Level			
$\lambda = 0$	$\mathcal{T}_t$	0.067	0.042	0.029	0.018	0.061	0.035	0.020	0.013
	S_t^J	0.085	0.030	0.018	0.015	0.099	0.022	0.013	0.012
	$S_t^{J^{noise}}$	0.075	0.021	0.014	0.011	0.075	0.021	0.014	0.011
	S_t^{IA}	0.988	0.941	0.902	0.890	0.990	0.939	0.911	0.912
$\lambda = 1.0$	$\mathcal{T}_t$	0.059	0.037	0.027	0.017	0.053	0.032	0.018	0.014
	S_t^J	1.000	1.000	0.999	0.999	1.000	1.000	0.999	0.999
	$S_t^{J^{noise}}$	0.998	0.988	0.975	0.955	0.998	0.988	0.975	0.955
	S_t^{IA}	0.988	0.936	0.901	0.849	0.989	0.934	0.913	0.865
$\lambda = 0.2$	$\mathcal{T}_t$	0.065	0.041	0.029	0.016	0.059	0.036	0.017	0.013
	S_t^J	1.000	1.000	0.999	0.999	1.000	1.000	0.999	0.999
	$S_t^{J^{noise}}$	0.997	0.983	0.965	0.934	0.997	0.983	0.965	0.934
	S_t^{IA}	0.988	0.941	0.918	0.899	0.990	0.942	0.926	0.909

Note: The table reports rejection rates, for the four test statistics outlined in Section 2.2, across various levels of price staleness. Results are reported for three DGPs: diffusion process ($\lambda = 0$), jump-diffusion processes with jump intensity of $\lambda = 1.0$ and $\lambda = 0.2$, respectively. The probability of staleness is set to $\mathfrak{p}_F = \{0.70, 0.35, 0.23, 0.18\}$ for the high level, and to $\mathfrak{p}_F = \{0.54,\ 0.22,\ 0.13,\ 0.10\}$ for the medium level. Under a finite-jump diffusion process the KLJ test ($\mathcal{T}_t$) reports its empirical size, while the PZ ($S_t^{J^{noise}}$, S_t^J) and the ASJ (S_t^{SIA}) tests report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table 6: Monte Carlo Rejection Rates under Infinite Jump-Diffusion Process and Price Staleness

	$[1/\Delta_n]$	4680	780	390	260	4680	780	390	260
		High Level				Medium Level			
$\beta = 1.00$	$\mathcal{T}_t$	0.016	0.017	0.019	0.030	0.014	0.020	0.022	0.034
	S_t^J	0.999	0.993	0.982	0.966	0.999	0.993	0.982	0.966
	$S_t^{J^{noise}}$	0.999	0.978	0.953	0.905	0.999	0.978	0.952	0.913
	S_t^{IA}	0.013	0.011	0.016	0.022	0.014	0.010	0.015	0.021
$\beta = 1.25$	$\mathcal{T}_t$	0.016	0.014	0.016	0.025	0.013	0.016	0.018	0.027
	S_t^J	1.000	0.999	0.996	0.993	1.000	0.999	0.996	0.993
	$S_t^{J^{noise}}$	0.996	0.952	0.937	0.910	0.996	0.953	0.939	0.912
	S_t^{IA}	0.023	0.013	0.016	0.020	0.024	0.013	0.015	0.019
$\beta = 1.50$	$\mathcal{T}_t$	0.015	0.016	0.021	0.030	0.012	0.022	0.026	0.035
	S_t^J	0.999	0.993	0.986	0.978	0.999	0.993	0.986	0.978
	$S_t^{J^{noise}}$	0.995	0.929	0.902	0.864	0.995	0.929	0.907	0.872
	S_t^{IA}	0.056	0.042	0.012	0.027	0.055	0.046	0.012	0.023

Note: The table reports rejection rates, for the four test statistics outlined in Section 2.2, across various levels of price staleness. The probability of staleness is set to $\mathfrak{p}_F = \{0.70, 0.35, 0.23, 0.18\}$ for the high level, and to $\mathfrak{p}_F = \{0.54, 0.22, 0.13, 0.10\}$ for the medium level. Under an infinite-jump diffusion process the KLJ ($\mathcal{T}_t$) and ASJ (S_t^{SIA}) tests report their empirical size, while the PZ tests ($S_t^{J^{noise}}$, S_t^J) report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day, β is the jump activity index, and the significance level is $\theta = 0.01$.

Table 7: Empirical Test Rejections and Contribution to Total Variance by Sector

	Panel A: Test Rejections			Panel B: Components				
	S_t^{Jnoise}	S_t^{IA}	$\mathcal{T}_t$	C_T	J_T	FJ_T	IJ_T	$\widehat{\beta}$
SPY	0.126	0.966	0.000	0.966	0.034	0.817	0.183	1.018
Consumer Discretionary	0.204	0.844	0.000	0.919	0.081	0.798	0.202	0.919
Consumer Staples	0.205	0.811	0.001	0.908	0.092	0.738	0.262	1.225
Energy	0.161	0.876	0.000	0.939	0.061	0.811	0.189	0.871
Financials	0.175	0.881	0.000	0.939	0.061	0.849	0.151	0.945
Healthcare	0.201	0.816	0.000	0.898	0.102	0.757	0.243	1.211
Industrials	0.190	0.848	0.000	0.919	0.081	0.765	0.235	1.072
Information Technology	0.168	0.873	0.000	0.950	0.050	0.789	0.211	0.872
Materials	0.217	0.800	0.000	0.898	0.102	0.754	0.246	0.898
Communication services	0.219	0.776	0.000	0.898	0.102	0.767	0.233	0.856
Utilities	0.201	0.793	0.000	0.915	0.085	0.697	0.303	1.161
Sector Average	0.194	0.832	0.000	0.918	0.082	0.773	0.227	1.003

Note: The table reports in two panels the rejection rates and the contribution of the continuous and discontinuous part to total variance estimated as $B\left(2,\infty,\Delta_n\right)_t$. The identify rejections rates using FDR-adjusted p-value. Panel A presents the number of rejections for each test, which is standardized by the total number of days in the sample data. The rejection rate is the average across the 10 stocks of each sector. Panel B depicts the contribution of the continuous and discontinuous part to total variance, as well as the contribution of finite and infinite activity jumps to the total jump component, J_T. $\widehat{\beta}$ is an estimate of the Blumenthal-Getoor index as in Jing et al. (2012b). $C_t = B\left(2,\infty,\Delta_n\right)_t \cdot \mathbb{1}(no\ jumps) + B\left(2,\nu_n,\Delta_n\right)_t \cdot \mathbb{1}(jumps)$. $J_t = B\left(2,\infty,\Delta_n\right)_t - C_t$. Hence, $C_T = \frac{\sum_{t\in(0,T]} C_t}{\sum_{t\in(0,T]} C_t + J_t}$ and $J_T = \frac{\sum_{t\in(0,T]} J_t}{\sum_{t\in(0,T]} C_t + J_t}$. The contribution of finite and infinite jumps to the total jump component are obtained as $FJ_t = J_t \cdot \mathbb{1}(finite\,jumps)$ and $IJ_t = J_t \cdot \mathbb{1}(infinite\,jumps)$. Thus, $FJ_T = \frac{\sum_{t\in(0,T]} FJ_t}{\sum_{t\in(0,T]} J_t}$ and $IJ_T = \frac{\sum_{t\in(0,T]} IJ_t}{\sum_{t\in(0,T]} J_t}$

Table 8: Empirical Rejection Rates and Contribution to Total Variance Classified by Market Capitalization and Sector

	Panel A: Rejections			Panel B: Components			
	S_t^{Jnoise}	S_t^{IA}	$\mathcal{T}_t$	C_T	J_T	FJ_T	IJ_T
	Big Market Cap Companies						
Consumer Discretionary	0.190	0.864	0.000	0.927	0.073	0.786	0.214
Consumer Staples	0.184	0.831	0.001	0.917	0.083	0.764	0.236
Energy	0.146	0.890	0.000	0.952	0.048	0.776	0.224
Financials	0.165	0.907	0.000	0.945	0.055	0.868	0.132
Healthcare	0.193	0.832	0.000	0.903	0.097	0.777	0.223
Industrials	0.184	0.867	0.000	0.928	0.072	0.754	0.246
Information Technology	0.153	0.896	0.000	0.956	0.044	0.784	0.216
Materials	0.210	0.811	0.000	0.903	0.097	0.759	0.241
Communication services	0.194	0.811	0.000	0.921	0.079	0.760	0.240
Utilities	0.190	0.817	0.000	0.922	0.078	0.694	0.306
Sector Average	0.181	0.853	0.000	0.927	0.073	0.772	0.228
	Small Market Cap Companies						
Consumer Discretionary	0.218	0.824	0.000	0.911	0.089	0.811	0.189
Consumer Staples	0.225	0.792	0.001	0.899	0.101	0.712	0.288
Energy	0.175	0.862	0.000	0.925	0.075	0.846	0.154
Financials	0.185	0.854	0.000	0.933	0.067	0.830	0.170
Healthcare	0.209	0.799	0.000	0.893	0.107	0.738	0.262
Industrials	0.196	0.828	0.000	0.909	0.091	0.775	0.225
Information Technology	0.183	0.850	0.000	0.944	0.056	0.794	0.206
Materials	0.224	0.790	0.000	0.892	0.108	0.748	0.252
Communication Services	0.244	0.741	0.000	0.875	0.125	0.774	0.226
Utilities	0.211	0.770	0.000	0.908	0.092	0.701	0.299
Sector Average	0.207	0.811	0.000	0.909	0.091	0.773	0.227

Note: See Notes to Table 7. The top (bottom) panel reports the results for the biggest (smallest) 5 companies of each sector selected by market capitalization.

INTERNET APPENDIX

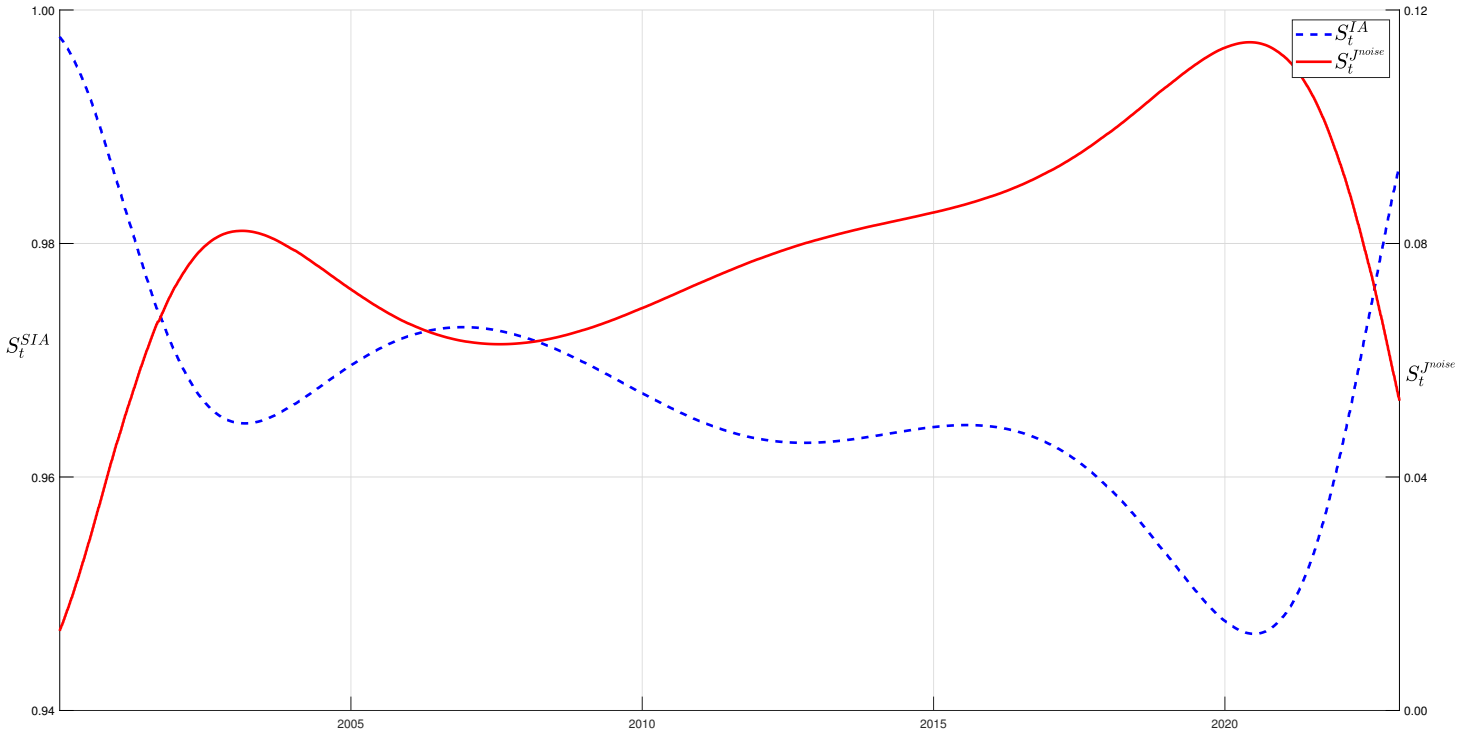
Identifying the Underlying Components of High-Frequency Data: Pure vs Jump Diffusion Processes

Abstract

This appendix collects additional simulation and empirical results supporting the main paper.

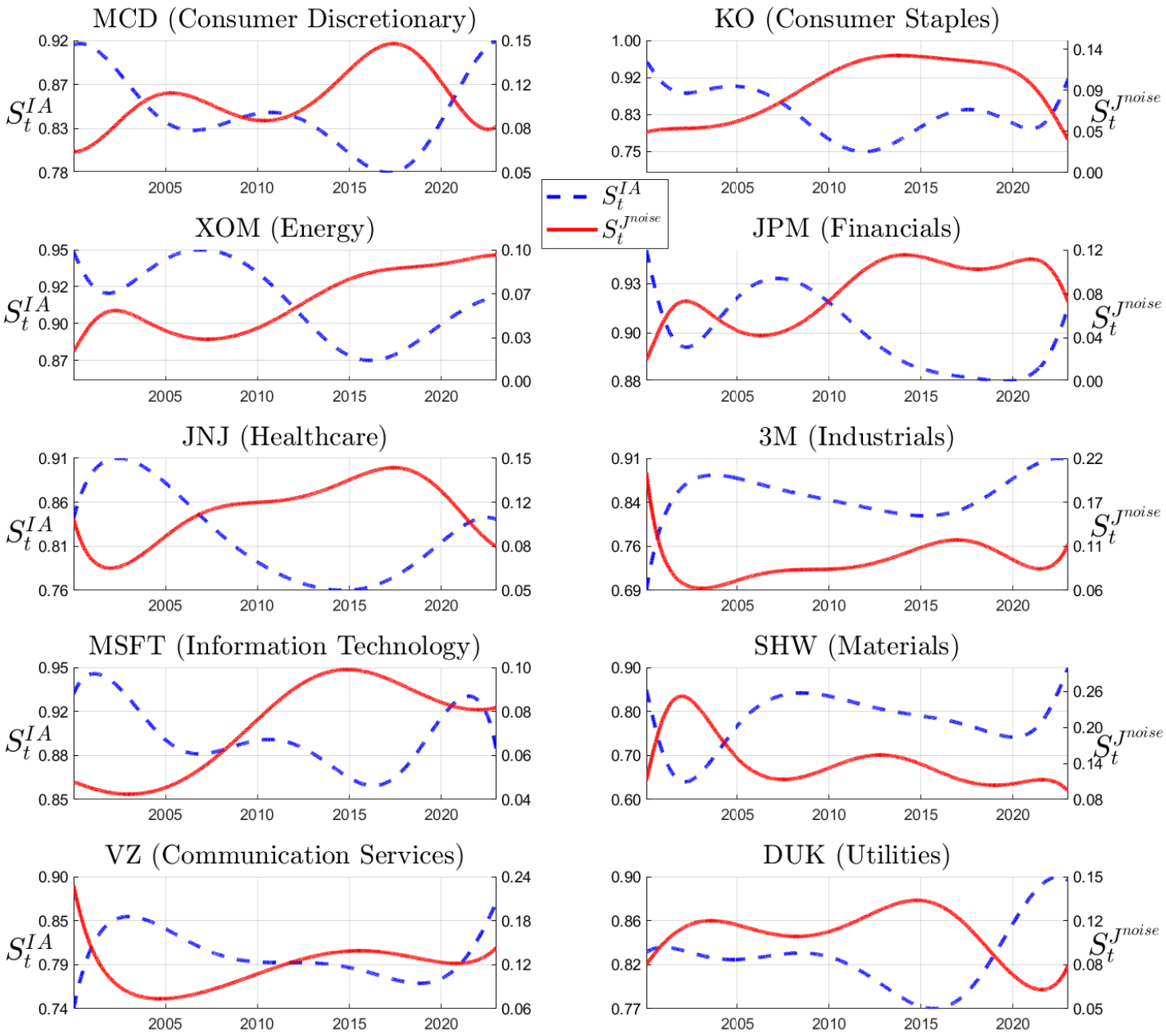
Supplementary Tables and Figures

Figure A.1: SPY – Time Variation in Rejection Rates



Note: The figure depicts the time variation in rejection rates predicted by a probit model with a 6th order polynomial in time. The left (right) y-axis denotes the probability of rejection over time for the ASJ (noise-robust PZ test). Rejection rates are computed using the FDR-adjusted p-values.

Figure A.2: Individual Stocks – Time Variation in Rejection Rates



Note: The figure depicts the time variation in rejection rates predicted by a probit model with a 6th order polynomial in time. The left (right) y-axis of each subplot denotes the probability of rejection over time for the ASJ (noise-robust PZ test). Rejection rates are computed using the FDR-adjusted p-values.

Table A.1: Monte Carlo Rejection Rates under a Finite Jump-Diffusion Process, $\lambda = 0.2$

$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
	No Noise				Gaussian Noise			
$\mathcal{T}_t$	0.013	0.014	0.015	0.018	0.001	0.009	0.013	0.023
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.995	0.972	0.942	0.919	0.997	0.983	0.965	0.935
S_t^{SIA}	0.999	0.968	0.934	0.908	0.806	0.912	0.941	0.932
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\mathcal{T}_t$	0.003	0.007	0.015	0.017	0.002	0.004	0.012	0.018
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.997	0.984	0.965	0.934	0.997	0.984	0.965	0.936
S_t^{SIA}	0.284	0.584	0.840	0.886	0.759	0.870	0.913	0.929
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\mathcal{T}_t$	0.002	0.008	0.013	0.024	0.002	0.009	0.011	0.017
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.997	0.984	0.963	0.937	0.997	0.983	0.961	0.927
S_t^{SIA}	0.534	0.863	0.912	0.902	0.829	0.899	0.917	0.935
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\mathcal{T}_t$	0.002	0.009	0.013	0.025	0.003	0.010	0.012	0.018
S_t^J	1.000	1.000	1.000	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.998	0.983	0.965	0.936	0.997	0.982	0.966	0.935
S_t^{SIA}	0.784	0.945	0.935	0.911	0.895	0.913	0.944	0.948

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under a finite-jump diffusion process the KLJ test ($\mathcal{T}_t$) reports its empirical size, while the PZ (S_t^{Jnoise}, S_t^J) and the ASJ (S_t^{SIA}) tests report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table A.2: Monte Carlo Rejection Rates under a Finite Jump-Diffusion Process, $\lambda = 0.1$

$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
	No Noise				Gaussian Noise			
$\mathcal{T}_t$	0.011	0.015	0.012	0.017	0.003	0.007	0.011	0.015
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.994	0.972	0.941	0.913	0.997	0.984	0.963	0.934
S_t^{SIA}	0.999	0.969	0.938	0.916	0.817	0.919	0.944	0.938
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\mathcal{T}_t$	0.001	0.007	0.013	0.018	0.002	0.006	0.012	0.017
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.997	0.982	0.966	0.936	0.997	0.983	0.957	0.935
S_t^{SIA}	0.284	0.586	0.845	0.890	0.763	0.855	0.917	0.929
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\mathcal{T}_t$	0.002	0.009	0.018	0.022	0.002	0.008	0.011	0.018
S_t^J	1.000	1.000	0.999	0.999	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.997	0.984	0.964	0.932	0.998	0.983	0.963	0.935
S_t^{SIA}	0.534	0.864	0.914	0.907	0.860	0.892	0.925	0.937
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\mathcal{T}_t$	0.002	0.009	0.012	0.022	0.003	0.010	0.012	0.019
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.997	0.980	0.964	0.933	0.997	0.981	0.964	0.934
S_t^{SIA}	0.785	0.946	0.937	0.917	0.901	0.923	0.958	0.956

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under a finite-jump diffusion process the KLJ test ($\mathcal{T}_t$) reports its empirical size, while the PZ (S_t^{Jnoise}, S_t^J) and the ASJ (S_t^{SIA}) tests report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table A.3: Monte Carlo Rejection Rates under a Finite Jump-Diffusion Process with 4 Large Jumps

$\lfloor 1/\Delta_n \rfloor$	4,680	780	390	260	4,680	780	390	260
	No Noise				Gaussian Noise			
$\mathcal{T}_t$	0.013	0.014	0.015	0.017	0.002	0.009	0.013	0.015
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S_t^{Jnoise}	0.998	0.990	0.979	0.907	0.993	0.978	0.948	0.908
S_t^{SIA}	0.999	0.971	0.939	0.915	1.000	0.978	0.945	0.920
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\mathcal{T}_t$	0.000	0.009	0.012	0.014	0.002	0.005	0.012	0.018
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S_t^{Jnoise}	0.996	0.980	0.957	0.915	0.996	0.977	0.945	0.918
S_t^{SIA}	0.283	0.684	0.891	0.848	0.665	0.865	0.887	0.863
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\mathcal{T}_t$	0.002	0.008	0.013	0.020	0.002	0.009	0.011	0.019
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	0.999	1.000
S_t^{Jnoise}	0.995	0.981	0.953	0.907	0.996	0.983	0.952	0.909
S_t^{SIA}	0.534	0.866	0.915	0.908	0.855	0.920	0.937	0.927
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\mathcal{T}_t$	0.002	0.009	0.013	0.025	0.002	0.010	0.011	0.019
S_t^J	1.000	1.000	0.999	1.000	1.000	1.000	1.000	0.999
S_t^{Jnoise}	0.997	0.968	0.948	0.905	0.909	0.980	0.949	0.902
S_t^{SIA}	0.785	0.946	0.937	0.920	0.873	0.938	0.957	0.936

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under a finite-jump diffusion process with only 4 large jumps, the KLJ test ($\mathcal{T}_t$) reports its empirical size, while the PZ (S_t^{Jnoise}, S_t^J) and the ASJ (S_t^{SIA}) tests report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table A.4: Monte Carlo Rejection Rates under a Infinite Jump-Diffusion Process with Residual Finite Jumps

	$[1/\Delta_n]$	4680	780	390	260
$\beta = 1.00$	$\mathcal{T}_t$	0.010	0.012	0.016	0.024
	S_t^J	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	1.000	0.998	0.989	0.981
	S_t^{SIA}	0.036	0.022	0.015	0.019
$\beta = 1.25$	$\mathcal{T}_t$	0.013	0.017	0.018	0.027
	S_t^J	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	1.000	0.998	0.989	0.979
	S_t^{SIA}	0.047	0.027	0.012	0.016
$\beta = 1.50$	$\mathcal{T}_t$	0.003	0.003	0.005	0.007
	S_t^J	1.000	1.000	1.000	1.000
	$S_t^{J^{noise}}$	1.000	0.996	0.985	0.968
	S_t^{SIA}	0.062	0.053	0.047	0.055

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. Under an infinite-jump diffusion process with residual finite jumps the KLJ test ($\mathcal{T}_t$) reports its empirical size, while the PZ ($S_t^{J^{noise}}$, S_t^J) and the ASJ (S_t^{SIA}) tests report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day and the significance level is $\theta = 0.01$.

Table A.5: Monte Carlo Rejection Rates under Infinite Jump-Diffusion with Variance-Gamma Jumps

$[1/\Delta_n]$	4680	780	390	260	4680	780	390	260
	No Noise				Gaussian Noise			
$\mathcal{T}_t$	0.017	0.021	0.023	0.030	0.001	0.008	0.012	0.021
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S_t^{Jnoise}	0.998	0.986	0.980	0.963	0.999	0.991	0.988	0.966
S_t^{SIA}	0.012	0.013	0.106	0.018	0.053	0.034	0.010	0.015
	t-distributed Noise, $df = 2.5$				Mixture Noise, $df = 2.5$			
$\mathcal{T}_t$	0.029	0.018	0.018	0.023	0.003	0.015	0.014	0.015
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S_t^{Jnoise}	0.991	0.980	0.972	0.923	0.993	0.988	0.979	0.920
S_t^{SIA}	0.167	0.082	0.051	0.032	0.183	0.079	0.058	0.036
	t-distributed Noise, $df = 3.5$				Mixture Noise, $df = 3.5$			
$\mathcal{T}_t$	0.020	0.013	0.017	0.023	0.002	0.012	0.013	0.014
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S_t^{Jnoise}	0.999	0.990	0.979	0.964	0.999	0.992	0.981	0.962
S_t^{SIA}	0.134	0.073	0.042	0.026	0.139	0.069	0.047	0.029
	t-distributed Noise, $df = 4.5$				Mixture Noise, $df = 4.5$			
$\mathcal{T}_t$	0.015	0.013	0.019	0.023	0.002	0.013	0.012	0.014
S_t^J	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
S_t^{Jnoise}	1.000	0.990	0.987	0.967	1.000	0.992	0.984	0.964
S_t^{SIA}	0.107	0.062	0.033	0.017	0.119	0.051	0.032	0.020

Note: The table reports rejection rates across sampling frequencies for the four test statistics outlined in Section 2.2. The variance-gamma jump process is $VG(t/\eta, \phi, \nu_t^{1/2}, 0)$, where $\eta = 0.1$ and $\phi = -0.05$. Under an infinite-jump diffusion process the KLJ ($\mathcal{T}_t$) and ASJ (S_t^{SIA}) tests report their empirical size, while the PZ tests (S_t^{Jnoise}, S_t^J) report their empirical power. $\lfloor 1/\Delta_n \rfloor$ represents the number of intraday observations per day, β is the jump activity index, and the significance level is $\theta = 0.01$.

Table A.6: Empirical Test Rejections and Contribution to Total Variance by Sector

	Panel A: Test Rejections			Panel B: Components				
	S_t^{Jnoise}	S_t^{IA}	$\mathcal{T}_t$	C_T	J_T	FJ_T	IJ_T	$\widehat{\beta}$
SPY	0.079	0.966	0.000	0.973	0.027	0.811	0.189	1.018
Consumer Discretionary	0.115	0.844	0.000	0.948	0.052	0.912	0.088	0.919
Consumer Staples	0.117	0.811	0.001	0.935	0.065	0.851	0.149	1.225
Energy	0.084	0.876	0.000	0.960	0.040	0.913	0.087	0.871
Financials	0.095	0.881	0.000	0.963	0.037	0.924	0.076	0.945
Healthcare	0.113	0.816	0.000	0.929	0.071	0.883	0.117	1.211
Industrials	0.103	0.848	0.000	0.945	0.055	0.892	0.108	1.072
Information Technology	0.089	0.873	0.000	0.970	0.030	0.890	0.110	0.872
Materials	0.124	0.800	0.000	0.932	0.068	0.899	0.101	0.898
Communication Services	0.129	0.776	0.000	0.930	0.070	0.897	0.103	0.856
Utilities	0.113	0.794	0.000	0.945	0.055	0.819	0.181	1.161
Sector Average	0.108	0.832	0.000	0.946	0.054	0.888	0.112	1.003

Note: The table reports in two panels the rejection rates and the contribution of the continuous and discontinuous part to total variance estimated as $B\left(2,\infty,\Delta_n\right)_t$. The identify rejections rates using FDR-adjusted p-values. Panel A presents the number of rejections for each test, which is standardized by the total number of days in the sample data. The rejection rate is the average across the 10 stocks of each sector. Panel B depicts the contribution of the continuous and discontinuous part to total variance, as well as the contribution of finite and infinite activity jumps to the total jump component, J_T. $\widehat{\beta}$ is an estimate of the Blumenthal-Getoor index as in Jing et al. (2012b). $C_t = B\left(2,\infty,\Delta_n\right)_t \cdot \mathbb{1}(no\ jumps) + B\left(2,\nu_n,\Delta_n\right)_t \cdot \mathbb{1}(jumps)$. $J_t = B\left(2,\infty,\Delta_n\right)_t - C_t$. Hence, $C_T = \frac{\sum_{t\in(0,T]} C_t}{\sum_{t\in(0,T]} C_t+J_t}$ and $J_T = \frac{\sum_{t\in(0,T]} J_t}{\sum_{t\in(0,T]} C_t+J_t}$. The contribution of finite and infinite jumps to the total jump component are obtained as $FJ_t = J_t \cdot \mathbb{1}(finite\,jumps)$ and $IJ_t = J_t \cdot \mathbb{1}(infinite\,jumps)$. Thus, $FJ_T = \frac{\sum_{t\in(0,T]} FJ_t}{\sum_{t\in(0,T]} J_t}$ and $IJ_T = \frac{\sum_{t\in(0,T]} IJ_t}{\sum_{t\in(0,T]} J_t}$

Table A.7: Empirical Rejection Rates and Contribution to Total Variance Classified by Market Capitalization and Sector

	Panel A: Rejections			Panel B: Components			
	S_t^{Jnoise}	S_t^{IA}	$\mathcal{T}_t$	C_T	J_T	FJ_T	IJ_T
	Big Market Cap Companies						
Consumer Discretionary	0.106	0.864	0.000	0.953	0.047	0.900	0.100
Consumer Staples	0.104	0.831	0.001	0.939	0.061	0.854	0.146
Energy	0.074	0.890	0.000	0.970	0.030	0.869	0.131
Financials	0.086	0.907	0.000	0.968	0.032	0.931	0.069
Healthcare	0.106	0.832	0.000	0.930	0.070	0.881	0.119
Industrials	0.099	0.867	0.000	0.953	0.047	0.878	0.122
Information Technology	0.079	0.896	0.000	0.973	0.027	0.881	0.119
Materials	0.119	0.811	0.000	0.937	0.063	0.902	0.098
Communication Services	0.112	0.811	0.000	0.946	0.054	0.880	0.120
Utilities	0.105	0.817	0.000	0.951	0.049	0.817	0.183
Sector Average	0.099	0.853	0.000	0.952	0.048	0.879	0.121
	Small Market Cap Companies						
Consumer Discretionary	0.125	0.824	0.000	0.942	0.058	0.924	0.076
Consumer Staples	0.131	0.792	0.001	0.932	0.068	0.847	0.153
Energy	0.093	0.862	0.000	0.950	0.050	0.956	0.044
Financials	0.103	0.854	0.000	0.958	0.042	0.916	0.084
Healthcare	0.120	0.799	0.000	0.927	0.073	0.885	0.115
Industrials	0.107	0.828	0.000	0.938	0.062	0.907	0.093
Information Technology	0.099	0.850	0.000	0.966	0.034	0.899	0.101
Materials	0.129	0.790	0.000	0.927	0.073	0.897	0.103
Communication Services	0.146	0.742	0.000	0.915	0.085	0.914	0.086
Utilities	0.121	0.770	0.000	0.939	0.061	0.821	0.179
Sector Average	0.118	0.811	0.000	0.939	0.061	0.897	0.103

Note: See Notes to Table A.6. The top (bottom) panel reports the results for the biggest (smallest) 5 companies of each sector selected by market capitalization.