Testing for Jumps and Jump Intensity Path Dependence*

Valentina Corradi[1], Mervyn J. Silvapulle[2] and Norman R. Swanson[3]
[1]University of Surrey, [2]Monash University and [3]Rutgers University

December 2015

Abstract

In this paper, we fill a gap in the financial econometrics literature, by developing a "jump test" for the null hypothesis that the probability of a jump is zero. The test is based on realized third moments, and uses observations over an increasing time span. The test offers an alternative to standard finite time span tests, and is designed to detect jumps in the data generating process rather than detecting realized jumps over a fixed time span. More specifically, we make two contributions. First, we introduce our largely model free jump test for the null hypothesis of zero jump intensity. Second, under the maintained assumption of strictly positive jump intensity, we introduce a "self excitement test" for the null of constant jump intensity against the alternative of path dependent intensity. The latter test has power against autocorrelation in the jump component, and is a direct test for Hawkes diffusions (see e.g., Aït-Sahalia, Cacho-Diaz and Laeven (2015)). The limiting distributions of the proposed statistics are analyzed via use of a double asymptotic scheme, wherein the time span goes to infinity and the discrete interval approaches zero; and the distributions of the tests are normal and half normal, respectively. The results from a Monte Carlo study indicate that the tests have good finite sample properties.

Keywords: diffusion model, jump intensity, jump size density, tricity.
JEL classification: C12, C22, C52, C55.

* Valentina Corradi, School of Economics, University of Surrey, Guildford, Surrey, GU2 7XH, UK, v.corradi@surrey.ac.uk. Mervyn Silvapulle, Department of Econometrics and Business Statistics, Monash University Caulfield, 26 Sir John Monash Drive, Caulfield East, Victoria 3145, Australia, mervyn.silvapulle@monash.edu. Norman Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu.
We are grateful to Brendan Beare, Francesco Bravo, Bertrand Candelon, Laura Coroneo, Prosper Dovonon, Rustam Ibragimov, Tatsushi Oka, Peter Spencer, and to seminar participants at the National University of Singapore, Notre Dame Univeristy, the University of York, the 2013 Computational and Financial Econometrics Conference in London, the 2014 CIREQ Time Series and Financial Econometrics conference in Montreal, the 2014 UK Econometric Study Group in Bristol, and the 2015 ESRC conference on Financial Contagion in Nice for useful comments and suggestions.

1 Introduction

Jump diffusions are widely used in the financial econometrics literature when analyzing returns or exchange rates, as discussed in Duffie, Pan and Singleton (2000), Singleton (2001), Anderson, Benzoni and Lund (2002), Jiang and Knight (2002), Chacko and Viceira (2003) and Eraker, Johannes and Polson (2003), among others. In this context, various estimation techniques have been developed, and the common practice is to jointly estimate the parameters of both the continuous time and the jump components of models. Thus, parameters characterizing the drift, variance, jump intensity, and jump size probability density are jointly estimated. However, an obvious non-standard feature of this class of models is that the parameters characterizing the jump size density are not identified when the jump intensity is identically zero. This is an issue both when the intensity parameter is constant, as in standard stochastic volatility models with jumps (see, e.g. Andersen, Benzoni and Lund (2002)) as well as when the intensity follows a diffusion process, as in the important case of the Hawkes diffusion models analyzed by Aït-Sahalia, Cacho-Diaz and Laeven (2015). If one estimates a jump diffusion model that contains a jump intensity parameter and if the population jump intensity happens to be zero, then a subset of the parameters in the model is not identified, which in turn precludes consistent estimation of other parameters (see Andrews and Cheng (2012)).

The above estimation problem serves to underscore the importance of pretesting for jumps. In the extant literature, there are a large variety of tests for the null of no jumps versus the alternative of jumps. Tests include those based on the comparison of two realized volatility measures, one which is robust, and the other which is not robust to the presence of jumps (see, e.g. Barndorff-Nielsen, Shephard and Winkel (2006) and Podolskji and Vetter (2009a)), tests based on a thresholding approach (see, e.g. Corsi, Pirino, and Renò (2010), Lee and Mykland (2008), and Lee, Loretan and Ploberger (2013)), and tests based on power variation, as discussed in Aït-Sahalia and Jacod (2009). Such tests are consistent against realized jumps. One feature of these tests is that they are based on observations drawn on a given finite time span, and they can thus only detect whether jumps occurred during this given time span. While this is hardly a weakness of the existing tests, there are clearly situations for which interest lies in testing for the existence of jumps in the data generating process, or within a class of models. For example, this is the case if one is interested in using (transformations of) jump diffusion processes in a variety of valuation problems, such as option pricing and default modelling (see, e.g., Duffie, Pan and Singleton (2000)). It follows that tests for jumps using an increasing time span are needed, although such tests have not previously been proposed.

In this paper, we fill the above gap in the literature by developing a "jump test" of the null hypothesis that the probability of a jump is zero. In addition, under the maintained assumption of strictly positive jump intensity, we introduce a "self excitement test" for the null of constant jump intensity against the alternative of path dependent intensity. This test is found to have power against autocorrelation in the jump component, and is a direct test for Hawkes diffusions (see Aït-Sahalia, Cacho-Diaz and Laeven 2015), in which jump intensity is modeled as a mean-reverting diffusion process. When the proposed tests are implemented prior to model specification, standard estimation of jump diffusions can be subsequently carried out, avoiding the identification problems discussed above.

Our jump test is based on realized third moments, or so-called tricity. Various realized tricity-type statistics over a finite time span have already been examined in the literature in order to: detect realized jumps, as in Jacod (2012); study the contribution of realized skewness when predicting the cross-section of equity returns, as in Amaya, Christoffersen, Jacobs and Vasquez (2015); and to test for the endogeneity of sampling times, as in Li, Mykland, Renault, Zhang and Zheng (2014). What distinguishes our tricity-type test from these is that it is analyzed using both in-fill and long-span asymptotics. The use of long-span asymptotics ensures that the suggested statistic has power against jump intensity rather than against realized jumps. Importantly, our test is also robust to the presence of leverage. The limiting behavior of the proposed statistic is readily analyzed via use of a double asymptotic scheme wherein the time span goes to infinity and the discrete interval approaches zero. Under the null hypothesis of zero intensity, the statistic has a normal limiting distribution. Under the alternative, it is necessary to distinguish between jumps with zero or non-zero third moment. In the latter case, the proposed test has a well defined Pitman drift and has power against $\sqrt{T}-$local alternatives, where T is the time span, in days. In the former case, the sample third moment approaches zero, but the probability order of the statistic is larger than that which obtains under the null, since the jump component does not contribute to the mean, while it does contribute to the variance. As the order of magnitude of the variance depends on whether the null hypothesis is true or not, we introduce a threshold estimator for the variance, which is consistent under the null of zero intensity, and bounded in probability under the alternative. Thus, inference can be performed via use of a simple t-statistic.

As alluded to above, our self excitement test is based on the autocorrelation function of the jump component. A necessary condition for test consistency is that the mean of the jump size is non-zero. We thus additionally propose a simple "zero mean jump test" for the null hypothesis of zero jump mean. The self excitement test is implemented as follows. First, carry out the jump test. If the null of zero intensity is rejected, then carry out a simple test in order to ascertain whether the jump process is zero mean. If this test is in turn rejected, carry out the self excitement test, in order to ascertain whether the jump intensity is a constant, or is path dependent.

In principle, one might consider testing for the null of zero intensity using a score, Wald or likelihood ratio test, based on discrete observations (see., e.g., Andrews (2001)). This approach requires treating jump size density parameters as nuisance parameters unidentified under the null, and requires correct specification of both the continuous and the jump components of the diffusion.

Misspecification of one or both components will invalidate the test. Additionally, the likelihood function of a jump diffusion is not generally known in closed form, and therefore estimation (which is needed for test statistic construction) is usually based on either simulated GMM (see Duffie and Singleton (1993) and Anderson, Benzoni and Lund (2002)), indirect inference (see Gourieroux and Monfort (1993) and Gallant and Tauchen (1996)), or nonparametric simulated maximum likelihood (see Fermanian and Salanié (2004) and Corradi and Swanson (2011)). However, it goes without saying that one cannot simulate a diffusion with a negative intensity parameter. This, in turn, precludes the existence of a quadratic approximation around the null parameters of the criterion function to be maximized (minimized). Given that the existence of such quadratic approximations is a necessary condition for estimation and inference about parameters on the boundary (see Andrews (1999, 2001), Beg, Silvapulle and Silvapulle (2001), and Chapter 4 in Silvapulle and Sen (2005)), we cannot rely on simulation-based estimators when testing using standard score, Wald or likelihood ratio tests.

The finite sample behavior of the tests is studied in a Monte Carlo experiment. Since the tests are not robust to microstructure noise, one needs to choose a frequency for which the noise is not too binding. For this reason, in our Monte Carlo exercise, we set the discretization interval $\Delta = 1/78$ and $\Delta = 1/156$, which correspond to moderate frequencies. The empirical size of the jump test is sensitive to the smallest values of T and Δ^{-1}, but performance is markedly better as the magnitude of these parameters is increased. Moreover, the power is quite good across all parameterizations, even in the case of jumps with zero third moment. The zero mean jump test is well sized for all parameterizations, and power is quite good, other than for parameter combinations with the smallest values of the jump intensity, the shortest time spans, and the widest discretization intervals. The self excitement test likewise has very good size properties, for all parameterizations, and has good power, which increases with the level of path dependence.

The rest of the paper is organized as follows. Section 2 describes the set-up. Section 3 and Section 4 discuss the jump intensity and self excitement tests, and derive their asymptotic properties, respectively. Section 5 reports the findings of a Monte Carlo experiment designed to examine the finite sample properties of the tests, and concluding remarks are gathered in Section 6. All proofs are collected in an appendix.

2 Set-Up

We consider stochastic volatility jump diffusions, with either constant or path dependent intensity. For $t \in \mathbb{R}^+$, consider

$$\mathrm{d}\ln X_t = \mu \mathrm{d}t + V_t^{1/2}\sqrt{1-\rho^2}\mathrm{d}W_{1,t} + V_t^{1/2}\rho \mathrm{d}W_{2,t} + Z_t \mathrm{d}N_t, \tag{1}$$

and

$$\mathrm{d}V_t = \mu(V_t, \theta)\mathrm{d}t + g\left(V_t, \theta\right)\mathrm{d}W_{2,t}, \tag{2}$$

with $-1 \leq \rho \leq 1$. Additionally, $W_{1,t}$ and $W_{2,t}$ are independent standard Brownian motions. From (1) and (2), it is immediate to see that the specification of the volatility process is rather general, as the drift and variance terms in (2) need only to ensure the existence of a strong solution, $V_t > 0$. For example, V_t can be generated by a square root process, which is the case considered in the Monte Carlo study. On the other hand, we are imposing functional structure on the variance process in the asset process, as well as a constant drift. While, the first restriction can be relaxed at the cost of more notation, and some additional assumptions, the constancy of the drift is used in the proofs, and cannot be relaxed in the test for (no) path dependence in the jump intensity.

In the sequel, we assume that the jump process, N_t, is a finite activity process. Namely, we focus on the case of a small number of large jumps. More precisely,

$$\Pr\left(N_{t+\Delta} - N_t = 1 | \mathcal{F}_t\right) = \lambda_t\Delta + o\left(\Delta\right), \tag{3}$$

$$\Pr\left(N_{t+\Delta} - N_t = 0 | \mathcal{F}_t\right) = 1 - \lambda_t\Delta + o\left(\Delta\right), \tag{4}$$

and

$$\Pr\left(N_{t+\Delta} - N_t > 1 | \mathcal{F}_t\right) = o\left(\Delta\right), \tag{5}$$

where $\mathcal{F}_t = \sigma\left(N_s,\ 0 \leq s \leq t\right)$. Additionally, in the sequel $1_{\Delta_{N_{(k+1)\Delta}}} = 1$ if a jump occurs between time $k\Delta$ and $(k+1)\Delta$, and the associated jump size, Z_k, is identically and independently distributed with density $f(z;\gamma)$.

We consider two general cases. The first is that of Poisson jumps, in which $\lambda_t = \lambda$, for all t. The second is that of a Hawkes diffusion, in which the intensity is an increasing function of past jumps (see Bowsher (2007) and Aït-Sahalia, Cacho-Diaz and Laeven (2015)). In this case:

$$\lambda_t = \lambda_\infty + \beta \int_0^t \exp\left(-a\left(t-s\right)\right)\mathrm{d}N_s,$$

with $\lambda_\infty \geq 0$, $\beta \geq 0$, $a > 0$, and $a > \beta$ (in order to ensure intensity mean reversion). Thus,

$$\mathrm{d}\lambda_t = a\left(\lambda_\infty - \lambda_t\right)\mathrm{d}t + \beta\mathrm{d}N_s \tag{6}$$

and

$$\mathrm{E}\left(\lambda_t\right) = \frac{a\lambda_\infty}{a-\beta} = \lambda.$$

If $\lambda_\infty = 0$, then $\mathrm{E}(\lambda_t) = 0$; and since λ_t can never be negative, this in turn implies that $\lambda_t = 0$ a.s., for all t (i.e., $N_t = 0$ a.s., for all t). But, if $N_t = 0$ a.s., for all t, then β cannot be identified, and consequently a is not identified. Furthermore, if $N_t = 0$ a.s., for all t, then γ cannot be identified.

In summary, if $\lambda_\infty = 0$, then β, α,and γ are not identified. By contrast, if $\lambda_\infty > 0$, then γ and β are identified. However, if $\lambda_\infty > 0$ but $\beta = 0$, then a is not identified.

These observations highlight the importance of being very clear as to which of the two assumptions, $\lambda_\infty = 0$ or $\lambda_\infty > 0$, is made for statistical inference in the foregoing Hawkes diffusion model. In practice, thus, we are concerned with the following jump test hypotheses: $H_0 : \lambda = 0$ versus $H_A : \lambda > 0$, where hereafter λ denotes expected intensity (i.e. $\lambda = \mathrm{E}(\lambda_t)$).[1] This is a nonstandard inference problem because, under H_0, some parameters are not identified and a parameter lies on the boundary of the null parameter space. Additionally, depending upon the outcome of tests of the above hypotheses, we are also interested in the following self excitement test hypotheses: $H_0 : \beta = 0$ versus $H_A : \beta > 0$.

Another important class of jump diffusion is the affine jump diffusion, in which intensity is an affine function of a state variable (see e.g., Duffie, Pan and Singleton (2000) and Singleton (2001)). For example, in our set-up one can define intensity to be an affine function of the volatility process,

$$\lambda_t = \lambda_0 + \lambda_1 V_t, \;\; \lambda_0, \lambda_1 \geq 0, \tag{7}$$

so that the probability of a jump is positively correlated with the volatility level. Whenever λ_t is generated as in (6) or as in (7), it displays autocorrelation. However, in the former case, the positive autocorrelation in λ_t is reflected in a positive autocorrelation in $(N_{t+\Delta} - N_t)$. This is because the realization of a jump at time t increases the probability of a jump realization at time $t + \Delta$. By contrast, in the affine case, a jump realization at time t, does not affect the probability of having another jump at time $t+\Delta$. However, given (7), jumps appear positively correlated during periods of increasing volatility and vice-versa. Thus, jumps are not independently distributed. Nevertheless we cannot observe a clear pattern in their autocorrelation structure. As a consequence, our test for zero intensity has power against affine jumps. However, the test for self-excitation has power against positive autocorrelation in jumps, and may have no power against affine jumps.

If we are willing to parametrically specify the continuous and the jump components of the model, and most importantly if the transition density is known in closed form, then it is easy to construct a consistent test for jumps, based only on a long time span of discrete observations. In particular one can easily test $H_0 : \lambda = 0$ versus $H_A : \lambda > 0$. This fact can be illustrated by considering a score test. Suppose that the skeleton of the process, $\ln X_t$ in (1) is observed. Namely, $\ln X_1, \ln X_2, ..., \ln X_T$, is observed. Now, using the notation in (1)-(5), let $\delta = (\theta, \mu, \rho, \lambda, \gamma) = (\vartheta, \gamma)$. It follows immediately that, provided the transition density is known in closed form, the likelihood

[1] Note that testing for $\lambda = 0$ (> 0) implies and is implied by $\lambda_\infty = 0$ (> 0). Also, if $\beta = 0$, $\lambda = \lambda_\infty$.

can be written as:

$$l_T(\vartheta,\gamma) = \frac{1}{T}\sum_{t=1}^{T-1} l_t(\vartheta,\gamma) = \frac{1}{T}\sum_{t=1}^{T-1} \ln f_{t+1|t}\left(Y_{t+1}|Y_t,\vartheta,\gamma\right).$$

The score statistic for testing H_0 is thus[2]:

$$K_T\left(\gamma\right) = \max\left\{0, \left(R\widehat{\mathcal{I}}_T(\gamma)^{-1}\widehat{V}_T(\gamma)\widehat{\mathcal{I}}_T(\gamma)^{-1}R'\right)^{-1/2} U_T\left(\gamma\right)\right\},$$

where R is a $1\times p$ matrix, with p denoting the dimension of ϑ, and where

$$\begin{aligned} U_T\left(\gamma\right) &= \sqrt{T}\left(R\widehat{\mathcal{I}}_T(\gamma)^{-1}\nabla_\vartheta l_T\left(\widehat{\vartheta}_T,\gamma\right)\right), \\ \widehat{\mathcal{I}}_T(\gamma) &= \frac{1}{T}\sum_{t=1}^{T}\nabla_{\vartheta\vartheta} l_t\left(\widehat{\vartheta}_T,\gamma\right), \\ \widehat{\vartheta}_T &= \arg\max_{\vartheta} l_T\left(\vartheta,\gamma\right) \text{ s.t. } R\vartheta = \lambda_\infty = 0, \end{aligned} \tag{8}$$

and

$$\widehat{V}_T(\gamma) = \frac{1}{T}\sum_{j=-\tau_T}^{\tau_T}\sum_{t=\tau_T}^{T-\tau_T}\omega_j\nabla_\vartheta l_t\left(\widehat{\vartheta}_T,\gamma\right)\nabla_\vartheta l_{t+j}\left(\widehat{\vartheta}_T,\gamma\right)',\ \omega_j = 1-\frac{j}{1+\tau_T}. \tag{9}$$

Now, given mild regularity assumptions controlling the smoothness of the likelihood, under the null of $\lambda = 0$,

$$\sup_{\gamma\in\Gamma} K\left(\gamma\right) \xrightarrow{d} \sup_{\gamma\in\Gamma}\max\left\{0, \left(R\mathcal{I}(\gamma)^{-1}V(\gamma)\mathcal{I}(\gamma)^{-1}R'\right)^{-1/2} Z\left(\gamma\right)\right\},$$

where $\sup_{\gamma\in\Gamma}\left|\widehat{\mathcal{I}}_T(\gamma)-\mathcal{I}(\gamma)\right| = o_p(1)$, $\sup_{\gamma\in\Gamma}\left|\widehat{V}_T(\gamma)-V(\gamma)\right| = o_p(1)$, and $Z(\cdot)$ is a Gaussian process with covariance kernel,

$$C(\gamma_1,\gamma_2) = \begin{pmatrix} R\mathcal{I}(\gamma_1)^{-1}V(\gamma_1,\gamma_1)\mathcal{I}(\gamma_1)^{-1}R' & R\mathcal{I}(\gamma_1)^{-1}V(\gamma_1,\gamma_2)\mathcal{I}(\gamma_2)^{-1}R' \\ R\mathcal{I}(\gamma_2)^{-1}V(\gamma_1,\gamma_2)\mathcal{I}(\gamma_1)^{-1}R' & R\mathcal{I}(\gamma_2)^{-1}V(\gamma_2,\gamma_2)\mathcal{I}(\gamma_2)^{-1}R' \end{pmatrix},$$

where $V(\gamma_1,\gamma_2) = \mathrm{p}\lim_{\mathrm{T}\to\infty}\widehat{V}_T(\gamma_1,\gamma_2)$.

Note also that $\sup_{\gamma\in\Gamma} K\left(\gamma\right)$ diverges to infinity under the alternative. This test has power against $\sqrt{T}-$local alternatives. Additionally, the limiting behavior of the test depends on the quadratic approximation of the likelihood around $\lambda = 0$ (see Andrews (2001)). Hence, if the

[2] If λ is not scalar (for example, consider allowing for different up and down jump intensities, as in Chacko and Viceira (2003)), then the score statistic can be written as:

$$\begin{aligned} K\left(\gamma\right) &= U_T\left(\gamma\right)'\left(R\widehat{\mathcal{I}}_T(\gamma)^{-1}\widehat{V}_T(\gamma)\widehat{\mathcal{I}}_T(\gamma)^{-1}R'\right)^{-1}U_T\left(\gamma\right) \\ &\quad - \inf_{\lambda\geq 0}\left(U_T\left(\gamma\right)-\lambda\right)'\left(R\widehat{\mathcal{I}}_T(\gamma)^{-1}\widehat{V}_T(\gamma)\widehat{\mathcal{I}}_T(\gamma)^{-1}R'\right)^{-1}\left(U_T\left(\gamma\right)-\lambda\right) \end{aligned}$$

likelihood is known in closed form, and if both the continuous and the jump components of the model, including the density of the jumps size, are correctly specified, then inference can be easily carried out using this score test, or using analogous Wald or likelihood ratio tests. However, it is well known that for most empirically relevant models the likelihood is usually not known in closed form. In such cases, as discussed in the introduction, one often relies on simulation based estimation techniques such as simulated GMM, indirect inference, or nonparametric simulated maximum likelihood. However, as one cannot simulate observations with negative intensity, a quadratic approximation of the criterion function cannot be constructed, and these sorts of tests are not applicable. It is for this reason that we instead focus on simple moment based jump and self-excitement tests.

3 Test of $\lambda = 0$ (Jump Test)

As mentioned in the introduction, tests based on high frequency observations over a finite time span are model free, but have power only against realized jumps, and thus cannot be consistent against the alternative $\lambda > 0$. On the other hand, tests based on discrete observations over a long time span are consistent against $\lambda > 0$, but require correct specification of both the continuous and jump components, as well as knowledge of the transition density. In order to have tests that are consistent against $\lambda > 0$, but are still to a large extent model free, we use functions of sample moments and rely on double in-fill and long-time span asymptotic approximations.

In the sequel, assume the existence of a sample of n^+ observations over an increasing time span T^+ and a shrinking discrete interval Δ, so that $n^+ = \frac{T^+}{\Delta}$, with $T^+ \to \infty$ and $\Delta \to 0$. Define $n = \frac{T}{\Delta} = n^+ - \frac{T^+ - T}{\Delta}$, with $T^+ > T$, and $T^+/T \to \infty$. We first test for zero jump intensity ($\lambda = 0$). The hypotheses of interest are:

$$H_0^\lambda : \lambda = 0$$

versus

$$H_A^\lambda : \lambda > 0,$$

where

$$H_A^\lambda = H_A^{\lambda(1)} \cup H_A^{\lambda(2)} : \left(\lambda > 0 \text{ and } \mathrm{E}\left(Z_k^3\right) \neq 0\right) \cup \left(\lambda > 0 \text{ and } \mathrm{E}\left(Z_k^3\right) = 0\right).$$

Notice that the alternative hypothesis is the union of two different alternatives, depending on whether $\mathrm{E}\big(Z_k^3\big) \neq 0$ or $\mathrm{E}\big(Z_k^3\big) = 0$. Let:

$$\widehat{\mu}_{3,T,\Delta} = \frac{1}{T} \sum_{k=1}^{n} \left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n} \right)^3$$

$$
-\frac{1}{T^{+}}\sum_{k=1}^{n^{+}}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n^{+}\Delta}-\ln X_{\Delta}}{n^{+}}\right)^{3}1\left\{\left|\ln X_{k\Delta}-\ln X_{(k-1)\Delta}\right|\leq\tau\left(\Delta\right)\right\}, \tag{10}
$$

with $T^{+}/T\rightarrow\infty$, $\tau\left(\Delta\right)\rightarrow 0$, and $\tau\left(\Delta\right)/\Delta^{1/2}\rightarrow\infty$, and define the statistic:[3]

$$
S_{T,\Delta}=\frac{T^{1/2}}{\Delta}\widehat{\mu}_{3,T,\Delta}. \tag{11}
$$

The logic underlying the suggested statistic is the following. As outlined in the proof of Theorem 1,

$$
\begin{aligned}
&\frac{1}{T}\sum_{k=1}^{n}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n\Delta}-\ln X_{\Delta}}{n}\right)^{3}\\
&=\mathrm{E}\left(Z_{k}1_{\Delta_{N_{(k+1)\Delta}}}\right)^{3}+\frac{3}{2}\Delta\rho^{3}\mathrm{E}\left(V_{k\Delta}^{1/2}g\left(V_{k\Delta},\theta\right)\right)+o_{p}\left(\frac{\Delta}{T^{1/2}}\right),
\end{aligned}
$$

so that $\left|\frac{1}{\sqrt{T\Delta}}\sum_{k=1}^{n}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n\Delta}-\ln X_{\Delta}}{n}\right)^{3}\right|$ may diverge to infinity in probability either because of the presence of jumps or because of the presence of leverage, or both. Hence, we need to correct for the skewness due to the presence of leverage. This is the role played by the second term on the RHS of (10). In fact, regardless of the presence of jumps, provided that $\tau\left(\Delta\right)\rightarrow 0$ at an appropriate rate,

$$
\begin{aligned}
&\frac{1}{T^{+}}\sum_{k=1}^{n^{+}}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n^{+}\Delta}-\ln X_{\Delta}}{n^{+}}\right)^{3}1\left\{\left|\ln X_{k\Delta}-\ln X_{(k-1)\Delta}\right|\leq\tau\left(\Delta\right)\right\}\\
&=\frac{3}{2}\Delta\rho^{3}\mathrm{E}\left(V_{k\Delta}^{1/2}g\left(V_{k\Delta},\theta\right)\right)+o_{p}\left(\frac{\Delta}{T^{+1/2}}\right).
\end{aligned}
$$

The requirement of $T^{+}/T\rightarrow\infty$ ensures that the contribution of leverage estimation error is negligible.

Furthermore, as shown in a number of lemmata in the Appendix, the thresholding effect is asymptotically negligible under the null, and thus for $T=T^{+}$ the statistic is degenerate under H_{0}^{λ}. On the other hand, if $T^{+}/T\rightarrow 0$, the limiting behaviour is driven by the threshold term, and the statistic lacks power against jumps.

With regard to the $\tau\left(\Delta\right)$ term in (10), note that thresholding is a well established technique for disentangling the jump component from the continuous component in various estimation and testing frameworks (see Mancini (2009) and Mancini and Renò (2011)). In these papers the threshold

[3] The first term on the RHS of $\widehat{\mu}_{3,T,\Delta}$ can also be expressed as $\frac{1}{T}\sum_{k=n^{+}-n+1}^{n^{+}}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n\Delta}-\ln X_{\Delta}}{n}\right)^{3}$. However, it is necessary to use a longer sample size for the second term.

sequence, $\tau(\Delta)$, is selected so that $\tau(\Delta) \to 0$ and $\frac{\tau(\Delta)}{\sqrt{\Delta \log(1/\Delta)}} \to 0$, which is dictated by the law of the iterated logarithm of the Brownian component of the model. In the theorems below we would require mildly stronger conditions on $\tau(\Delta)$ because the time span is growing.

Before establishing the asymptotic properties of $S_{T,\Delta}$, it should be pointed out that a central limit theorem for realized third moments has been proven in Li, Mykland, Renault, Zhang and Zheng (2014). Their Theorem 2 establishes asymptotic mixed normality for tricity in the case of unequal random times and finite time spans. Namely, they show that $\frac{1}{\Delta}\sum_{t_k \leq 1}\left(\ln X_{t_{k+1}} - \ln X_{t_k}\right)^3$ has a mixed normal limiting distribution under the null of exogenous sampling time, and diverges otherwise.

In the sequel, we need the following assumption.

Assumption A: (i) $\ln X_t$ and V_t are generated as in (1) and (2), with $\mu(v,\theta)$ and $g(v,\theta)$ twice continuously differentiable, satisfying local Lipschitz and growth conditions for all $\theta \in \Theta$, **(ii)** V_t is geometrically ergodic, **(iii)** $\mathrm{E}\left(V_t^{\frac{m}{2}}\right) < \infty$ and $\mathrm{E}\left(g(V_t)^2\right) < \infty$ for even integer $m > 6$, (**iv**) N_t satisfies (3)-(5), and λ_t is either constant or it satisfies (6), with $\lambda_\infty \geq 0$, $\beta \geq 0$, $a > 0$, and $a > \beta$, and (**v**) the jump size, Z_k, is independently and identically distributed, with density $f(z;\gamma)$, and $\mathrm{E}\left(|Z_k|^\kappa\right) < \infty$, for $\kappa \geq 6$.

Theorem 1: *Let Assumption* **A** *hold. Also, assume that as* $T \to \infty$, $\Delta \to 0$, $T\Delta \to \infty$, $\sqrt{T}\Delta \to 0$, *and* $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)} \to 0$, *for even* $m > 6$, *and* $T^+/T \to \infty$. *Then,*

(i) Under H_0^λ :

$$S_{T,\Delta} \xrightarrow{d} N\left(0, \omega_0\right),$$

where

$$\omega_0 = \left(15\left(1-\rho^2\right)^3 + 15\rho^6 + 45\left(1-\rho^2\right)^2\rho^2 + 45\left(1-\rho^2\right)\rho^4\right)\mathrm{E}\left(V_{k\Delta}^3\right).$$

(ii) Under $H_A^{\lambda(1)}$, *there exists an* $\varepsilon > 0$, *such that:*

$$\lim_{T\to\infty,\Delta\to 0} \Pr\left(\frac{\Delta}{\sqrt{T}}\left|S_{T,\Delta}\right| > \varepsilon\right) = 1.$$

(iii) Under $H_A^{\lambda(2)}$, *there exists an* $\varepsilon > 0$, *such that:*

$$\lim_{T\to\infty,\Delta\to 0} \Pr\left(\Delta\left|S_{T,\Delta}\right| > \varepsilon\right) = 1.$$

Remark 1: It follows immediately that $S_{T,\Delta}$ converges to a normal random variable under the null hypothesis, diverges at rate $\frac{\sqrt{T}}{\Delta}$ under the alternative of jumps with non-zero third moment, and diverges at the slower rate of $\frac{1}{\Delta}$ under the alternative of jumps symmetric around zero. As shown in the appendix (see equation (23)), as $\Delta \to 0$ and $T \to \infty$, we have that $\widehat{\mu}_{3,T,\Delta} = \lambda\mathrm{E}\left(Z^3\right) + 2\lambda^2\Delta\mathrm{E}\left(Z\right)\mathrm{E}\left(Z^2\right) + o_p\left(\Delta\right)$. Now, if $\mathrm{E}\left(Z_k^3\right) \neq 0$, (i.e., under $H_A^{\lambda(1)}$), the test has a well defined Pitman drift against $\sqrt{T}-$alternatives. On the other hand, if jumps are symmetric around zero (i.e., $\mathrm{E}\left(Z_k^3\right) = \mathrm{E}\left(Z_k\right) = 0$), then $\lambda \mathrm{E}\left(Z_k^3\right)$ is not identified, and under $H_A^{\lambda(2)}$ the Pitman drift is zero. Indeed, $\widehat{\mu}_{3,T,\Delta} \xrightarrow{p} 0$ regardless of whether $\lambda = 0$ or $\lambda > 0$ in this case. Although it is not possible to distinguish between H_0^λ and $H_A^{\lambda(2)}$ on the basis of the different locations of the limiting distribution (i.e., the Pitman drift), it is possible to distinguish between them on the basis of different scales of the limiting distribution of $\frac{T^{1/2}}{\Delta}\widehat{\mu}_{3,T,\Delta}$. This is because the order of magnitude of the variance of $\frac{T^{1/2}}{\Delta}\widehat{\mu}_{3,T,\Delta}$ is larger when $\lambda > 0$ and $\mathrm{E}\left(Z_k^3\right) = \mathrm{E}\left(Z_k\right) = 0$ than when $\lambda = 0$. Broadly speaking, under H_0^λ, $S_{T,\Delta} \xrightarrow{d} N\left(0, \omega_0\right)$, while under $H_A^{\lambda(2)}$, $\Delta S_{T,\Delta} \xrightarrow{d} N\left(0, \omega_1\right)$, with $\omega_1 \neq \omega_0$. This is what allows one to distinguish between H_0^λ and $H_A^{\lambda(2)}$.

Remark 2: The test has power not only against constant intensity and self-exciting intensity, as in (6), but also against affine jump diffusions where intensity is an affine function of volatility, for example.

Remark 3: As the variance of the statistic is of larger order under the alternative of positive jump intensity, we cannot construct a variance estimator which is consistent under all hypotheses. Thus, our aim is to construct an estimator for the variance of $S_{T,\Delta}$ which is consistent under the null and bounded in probability under the (union of) alternatives. This is done by using a threshold variance estimator, which filters out the contribution of the jump component. In particular, define:

$$\begin{aligned} &\widehat{\sigma}^2_{\lambda,T,\Delta} \\ &= \frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)^6 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \leq \tau\left(\Delta\right)\right\}. \quad (12) \end{aligned}$$

It follows that the t-statistic version of the jump test is,

$$t_{\lambda,T,\Delta} = \frac{S_{T,\Delta}}{\widehat{\sigma}_{\lambda,T,\Delta}}.$$

The following corollary summarizes the limiting behavior of $t_{\lambda,T,\Delta}$.

Corollary 2: *Let Assumption **A** hold. Also, assume that as* $T \to \infty$, $\Delta \to 0$, $T\Delta \to \infty$, $\sqrt{T}\Delta \to 0$, $T^+/T \to \infty$, $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)} \to 0$ *for even* $m > 6$, *and* $\tau^7\left(\Delta\right)\Delta^{-2} \to 0$. *Then,*

(i) Under H_0^λ :

$$t_{\lambda,T,\Delta} \xrightarrow{d} N\left(0,1\right).$$

(ii) Under $H_A^{\lambda(1)}$, *there exists an* $\varepsilon > 0$, *such that:*

$$\lim_{T\to\infty,\Delta\to 0} \Pr\left(\frac{\Delta}{\sqrt{T}}\left|t_{\lambda,T,\Delta}\right| > \varepsilon\right) = 1.$$

(iii) Under $H_A^{\lambda(2)}$ *there exists an* $\varepsilon > 0$, *such that:*

$$\lim_{T\to\infty,\Delta\to 0} \Pr\left(\Delta\left|t_{\lambda,T,\Delta}\right| > \varepsilon\right) = 1.$$

Remark 4: Note that the variance estimator can be constructed using the entire time span of T^+ observations, and the statement in Corollary 2 still holds, provided that we replace $\sqrt{T}\Delta \to 0$ and $\sqrt{T}\tau^2(\Delta) \to 0$ with $\sqrt{T^+}\Delta \to 0$ and $\sqrt{T^+}\tau^2(\Delta) \to 0$. In general, the price of having a statistic which allows for possible leverage effects is that we need to use a longer time span for estimating the leverage contribution. A possible alternative approach would be to pretest for $H_0^\rho : \rho = 0$ vs. $H_A^\rho : \rho \neq 0$. Here, under H_0^ρ,

$$\frac{\Delta}{\sqrt{T^+}} \sum_{k=1}^{n^+} \left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n^+\Delta} - \ln X_\Delta}{n^+} \right)^3 1\left\{ \left| \ln X_{k\Delta} - \ln X_{(k-1)\Delta} \right| \leq \tau(\Delta) \right\}$$

is asymptotically normal, while it diverges to (likely minus) infinity under H_A^ρ. This means that if we do not reject the null of no leverage, we do not need to recenter the statistic in (10) and we can use all T^+ observations for tricity estimation.

Remark 5: Consider selection of the threshold sequence. From Corollary 2, it follows that $\tau(\Delta)$ should approach zero faster than $T^{-1/4}$ and faster than $\Delta^{2/7}$, but slower than $\Delta^{\frac{1}{2}-\frac{3}{m}}$, where m denotes the number of finite moments of $V_t^{1/2}$, with $m > 6$ by Assumption A(iii). For example, if V_t follows a square root process, so that all finite moments exist, we can set $\tau(\Delta) = c\Delta^\eta$, with $\frac{2}{7} < \eta < \frac{1}{2}$ and $\Delta = T^{-\delta}$, where $\frac{7}{8} < \delta < 1$. These conditions ensure that $T\Delta \to \infty$, $\sqrt{T}\Delta \to 0$, and $\sqrt{T}\tau^2(\Delta) \to 0$. Additionally, in order to implement the statistic, we choose the constant c in a data driven manner. A natural solution is to use $\widehat{\sigma}_{\mu_Z}$ as defined in (15) below.

Remark 6: Since the suggested statistic is not robust to the presence of microstructure noise, the optimal discrete interval, Δ, is the highest frequency at which microstructure noise doesn't bind. Visual inspection of the signature plots of Andersen, Bollerslev and Diebold (2000) provides a useful tool for the choice of interval. It should also be noted that the statistic is constructed over an increasing time span; and hence it is not straightforward to ascertain whether simple pre-averaging will yield a statistic that is robust to microstructure noise (as in the case of the realized pre-average power variation discussed in Podolskji and Vetter (2009b)). In our Monte Carlo experiment, we implement values of Δ for which noise is either not binding or moderately binding. Careful exploration of this issue is left to future research.

Remark 7: In this paper, we only derive tests for the null of zero jump intensity in asset returns. However, the same approach can be used for testing equivalent hypotheses for volatility. Such tests would require estimators of the spot volatility, say $V_{k\Delta}^2$, which can be constructed using a finer grid of observations than that used in the above tests, such as if there were M observations over each interval of order Δ. The order of magnitude of the error due to the estimation of the spot volatility is derived in Bandi and Renò (2012), under various settings.

4 Testing for Self-Exciting Jumps

If the null hypothesis of zero intensity is rejected, one can proceed to test the null of no self-excitation or path dependence. If the drift in (1) and the intensity are constant, and if the jump size is independently distributed, then $\left(\ln X_{(k+1)\Delta} - \ln X_{k\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)$ is a martingale difference process and so it is uncorrelated. In fact, if λ_t is generated as in (6), then it follows from Aït-Sahalia, Cacho-Diaz and Laeven (2015), that

$$\begin{aligned} &\mathrm{E}\left(\left(\ln X_{(k+\tau)\Delta} - \ln X_{(k+\tau-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)\left(X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)\right) \\ &= \frac{\beta\lambda_{\infty}\left(2a-\beta\right)}{2\left(a-\beta\right)} \exp\left(-\left(a-\beta\right)\tau\right)\left(\mathrm{E}\left(Z\right)\right)^2 \Delta^2 + o\left(\Delta^2\right). \end{aligned} \tag{13}$$

In this case, it is natural to test null hypothesis of $\beta = 0$, against the alternative that $\beta > 0$, in order to test for path dependence. In turn, it is immediate to see that in order to identify β, we require not only $\lambda_{\infty} > 0$, but also $\mathrm{E}\left(Z_k\right) \neq 0$. In fact, failure to reject the null may be simply due to the fact that $\mathrm{E}\left(Z_k\right) = 0$. Hence, before testing for jump self-excitation, it remains to pretest for the null of $\mathrm{E}\left(Z_k\right) = 0$ versus $\mathrm{E}\left(Z_k\right) \neq 0$.

4.1 Test of $(Z_k) = 0$ (Zero Mean Jump Test)

We test the null of zero mean jumps, against its negation. The hypotheses of interest are:

$$H_0^{\mu_Z} : \mathrm{E}\left(Z_k\right) = 0$$

and

$$H_A^{\mu_Z} : \mathrm{E}\left(Z_k\right) \neq 0.$$

Let

$$\begin{aligned} \widehat{\mu}_{T,\Delta}^{Z} = &\frac{1}{T}\sum_{k=0}^{n}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right) \\ &- \frac{1}{T^{+}}\sum_{k=1}^{n^{+}}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right) 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \leq \tau\left(\Delta\right)\right\} \end{aligned} \tag{14}$$

and

$$\widehat{\sigma}_{\mu_Z}^2 = \frac{1}{T}\sum_{k=0}^{n}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)^2 \tag{15}$$

and define

$$t_{\mu_Z,T,\Delta} = \sqrt{T}\frac{\widehat{\mu}_{T,\Delta}^{Z}}{\widehat{\sigma}_{\mu_Z}}. \tag{16}$$

Since $\mathrm{E}\big(\frac{1}{T}\sum_{k=0}^{n}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}\right)\big)=\mu+\lambda\mathrm{E}\left(Z_k\right)+O\left(\Delta\right)$, from the first term on the RHS of (14), we cannot disentangle the contribution to the mean of the continuous and jump components. However, thanks to the thresholding, the second term on the RHS of (14) provides a $\sqrt{T^+}-$consistent estimator of the drift, μ; and as $T^+/T\rightarrow\infty$, estimation error is negligible. As in the case of Theorem 1, if $T=T^+$, the statistic is degenerate under the null. Note also that the first term on the RHS of (14) is not recentered using $\frac{\ln X_{n\Delta}-\ln X_{\Delta}}{n}$. This is because otherwise we would recenter both the continuous as well as the jump components, and the statistic would not have any power against non-zero jump mean.

Theorem 3: *Let Assumption* ***A*** *hold. Also, assume that as* $T\rightarrow\infty$, $\Delta\rightarrow 0$, $T\Delta\rightarrow\infty$, $\sqrt{T}\Delta\rightarrow 0$, $T^+/T\rightarrow\infty$, $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)}\rightarrow 0$ *for even* $m>6$, *and* $\sqrt{T}\tau^2\left(\Delta\right)\rightarrow 0$.[4] *Then,*

(i) Under $H_0^{\mu_Z}$:

$$t_{\mu_Z,T,\Delta}\xrightarrow{d}N\left(0,1\right).$$

(ii) Under $H_A^{\mu_Z}$, *there exists an* $\varepsilon>0$ *such that:*

$$\lim_{T\rightarrow\infty,\Delta\rightarrow 0}\Pr\left(\left|\frac{t_{\mu_Z,T,\Delta}}{\sqrt{T}}\right|>\varepsilon\right)=1,$$

where $\hat{\mu}_{T,\Delta}^{Z},\hat{\sigma}_{\mu_Z}$, and $t_{\mu_Z,T,\Delta}$ are defined as in (14), (15), and (16), respectively.

Remark 1: Note that $\widehat{\sigma}_{\mu_Z}^2$ in (15) can be constructed using the entire time span, T^+, provided that we replace $\sqrt{T}\Delta\rightarrow 0$ and $\sqrt{T}\tau^2\left(\Delta\right)\rightarrow 0$ with $\sqrt{T^+}\Delta\rightarrow 0$ and $\sqrt{T^+}\tau^2\left(\Delta\right)\rightarrow 0$.

4.2 Test of $\beta=0$ (Self Excitement Test)

Note that for this test, we can use the entire time span, T^+, as leverage plays no role in autocorrelation calculations. Our objective is to test the following hypotheses:

$$H_0^{\beta}:\beta=0$$

and

$$H_A^{\beta}:\beta>0,$$

under the maintained assumption that $\mathrm{E}\left(Z_k\right)\neq 0$. Define the statistic:

$$S_{T^+,\Delta}^{\beta}=\max\left\{0,t_{\beta,T^+,\Delta}\right\},$$

[4] Note that $\sqrt{T}\Delta\rightarrow 0$ is implied by $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)}\rightarrow 0$ for even $m>6$, and $\sqrt{T}\tau^2\left(\Delta\right)\rightarrow 0$, so that the conditions of the theorem can be restated without the requirement that $\sqrt{T}\Delta\rightarrow 0$.

where

$$t_{\beta,T^+,\Delta} = \frac{\sqrt{\frac{T^+}{\Delta}}\widehat{\beta}_{T^+,\Delta}}{\widehat{\sigma}_{\beta,T^+,\Delta}}, \tag{17}$$

with

$$\widehat{\beta}_{T^+,\Delta} = \frac{1}{T^+}\sum_{k=2}^{n^+-1}\left(\ln X_{(k+1)\Delta} - \ln X_{k\Delta} - \frac{\ln X_{n^+\Delta} - \ln X_{\Delta}}{n^+}\right)\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n^+\Delta} - \ln X_{\Delta}}{n^+}\right) \tag{18}$$

and

$$\widehat{\sigma}^2_{\beta,T^+,\Delta} = \frac{1}{T^+\Delta}\sum_{k=2}^{n^+-1}\left(\ln X_{(k+1)\Delta} - \ln X_{k\Delta} - \frac{\ln X_{n^+\Delta} - \ln X_{\Delta}}{n^+}\right)^2\left(\ln X_{(k+1)\Delta} - \ln X_{k\Delta} - \frac{\ln X_{n^+\Delta} - \ln X_{\Delta}}{n^+}\right)^2. \tag{19}$$

From (13), and recalling that $a > 0$, $\beta \geq 0$, and $a > \beta$, it follows immediately that the autocorrelation can never be negative. This is why the test is one-sided. Additionally, recall that $T\Delta \to \infty$ implies $T^+\Delta \to \infty$. The following result thus holds.

Theorem 4: *Let Assumption **A** hold. Also, assume that $E(Z) \neq 0$, $\lambda_\infty > 0$, and as $n \to \infty$, $T \to \infty$, $\Delta \to 0$ and $T\Delta \to \infty$. Then,*
(i) Under H_0 :

$$S^{\beta}_{T^+,\Delta} \xrightarrow{d} \max\{0, \mathcal{Z}\},$$

where $\mathcal{Z}$ is a standard normal random variable.
(ii) Under H_A, there exists an $\varepsilon > 0$ such that:

$$\lim_{T^+\to\infty,\Delta\to 0} \Pr\left(\frac{1}{\sqrt{T^+\Delta}}S^{\beta}_{T^+,\Delta} > \varepsilon\right) = 1.$$

It follows that $S^{\beta}_{T^+,\Delta}$ converges to an half-normal random variable under the null, and diverges at rate $\sqrt{T^+\Delta}$ under the alternative.

Remark 1: The test statistic is only a function of the first autocovariance term. It follows immediately that one can construct a test based on an increasing number of autocovariance terms, with the number of terms chosen adaptively (see, e.g. Escanciano and Lobato (2009)).

Remark 2: If the nulls of zero intensity, zero jump mean and no self-excitation are all rejected, then one can proceed to estimate the full Hawkes diffusion using GMM, as in Aït-Sahalia, Cacho-Diaz and Laeven (2015).

Remark 3: In this section, we consider self-exciting intensity. However, from an empirical point of view, an interesting case is that of financial contagion, where the contagion is due to "common" jumps. In this case, the jump intensity is an increasing function not only of its own past jumps but also of past jumps in other assets. In order to test for (no) cross-excitation, it suffices to construct a statistic based on cross correlations instead of autocorrelations (see Theorem 4 in Aït-Sahalia, Cacho-Diaz and Laeven (2015)). For example, let:

$$\widehat{\beta}_{T^+,\Delta}^{(I,II)}$$
$$= \frac{1}{T^+}\sum_{k=2}^{n^+-1}\left(\ln X_{(k+1)\Delta}^{(I)} - \ln X_{k\Delta}^{(I)} - \frac{\ln X_{n^+\Delta}^{(I)} - \ln X_{\Delta}^{(I)}}{n^+}\right)\left(\ln X_{k\Delta}^{(II)} - \ln X_{(k-1)\Delta}^{(II)} - \frac{\ln X_{n^+\Delta}^{(II)} - \ln X_{\Delta}^{(II)}}{n^+}\right),$$

and note that if the jump intensity in asset II does not depend on past jumps in asset I, then $\widehat{\beta}_{T,\Delta}^{(I,II)} \xrightarrow{p} 0$. On the other hand, if the intensity in asset II increases when there is a jump in asset I, then $\left|\widehat{\beta}_{T^+,\Delta}^{(I,II)}\right|$ has a strictly positive probability limit.

5 Monte Carlo Experiment

In this section we present the findings of a Monte Carlo experiment designed to evaluate the finite sample properties of: (i) the "jump test" for the null of zero jump intensity, based on $t_{\lambda,T,\Delta} = \frac{S_{T,\Delta}}{\widehat{\sigma}_{\lambda,T,\Delta}}$, where $S_{T,\Delta} = \frac{T^{1/2}}{\Delta}\widehat{\mu}_{3,T,\Delta}$, $\widehat{\mu}_{3,T,\Delta}$ is defined in (10), and $\widehat{\sigma}^2_{\lambda,T,\Delta}$ is defined in (12); (ii) the "zero mean jump test" for the null of zero mean jumps based on $t_{\mu_Z,T,\Delta} = \sqrt{T}\frac{\widehat{\mu}_{T,\Delta}^Z}{\widehat{\sigma}_{\mu_Z}}$, where $\widehat{\mu}_{T,\Delta}^Z$ is defined in (14) and $\widehat{\sigma}^2_{\mu_Z}$ is defined in (15); and the "self excitement test" for the null of no jump-path dependence, based on $S_{T^+,\Delta}^{\beta} = \max\left\{0, t_{\beta,T^+,\Delta}\right\}$, where $t_{\beta,T^+,\Delta} = \frac{\sqrt{\frac{T^+}{\Delta}}\widehat{\beta}_{T^+,\Delta}}{\widehat{\sigma}_{\beta,T^+,\Delta}}$, $\widehat{\beta}_{T^+,\Delta}$ is defined in (18), and $\widehat{\sigma}^2_{\beta,T^+,\Delta}$ is defined in (19).

Data used in the experiment are generated according to the following data generating processes (DGPs):

$$\mathrm{d}\ln X_t = \mu \mathrm{d}t + \sqrt{V_t}\mathrm{d}W_{1,t} + Z_t \mathrm{d}N_t,$$

where volatility is modeled as a square-root process:

$$\mathrm{d}V_t = \kappa_v(\theta_v - V_t)\mathrm{d}t + \zeta\sqrt{V_t}\mathrm{d}W_{2,t},$$

with $\mathrm{E}(W_{1,t}W_{2,t}) = \rho$. We set $\mu = 0.1$, $\rho = \{0, -0.25\}$, $\kappa_v = 5$, $\theta_v = 0.16$, and $\zeta = 0.5$. Additionally, N_t satisfies the conditions in (3)-(5). The jump size, Z_k, is identically and independently distributed with density $f(z;\gamma)$. We consider three jump densities: $f(z;\gamma) = N(0.0,\sigma)$, $f(z;\gamma) = N(0.5,\sigma)$, and $f(z;\gamma) = \varsigma e^{-\varsigma z}$. For the cases where Z_k is a normal random variable, $\sigma = \{0.1, 0.2, 0.3, 0.4, 0.7\}$; and for the case where Z_t is characterized by the exponential density, $\varsigma = 5$. The jump intensity evolves according to:

$$\lambda_t = \lambda_\infty + \beta\int_0^t \exp\left(-a\left(t-s\right)\right)\mathrm{d}N_s, \tag{20}$$

where $\lambda_\infty = \{0.3, 0.5, 0.7, 0.9\}$ and $(a, \beta) = \{(0,0), (3,2),\ (5,4),\ (7,5)\}$. Note that the case where $(a, \beta) = (0,0)$ is consistent with both the case of no jumps (i.e., $\lambda_\infty = 0$) and with the case of constant jump intensity (i.e., $\lambda_t = \lambda_\infty > 0$, for all t). In the constant jump intensity case, we consider Poisson jumps, with parameter λ_∞.

We simulate observations using a Milstein discretization scheme, with discrete interval $h = 1/312$, and consider two intra-daily sampling frequencies: $\Delta = 1/78$ and $\Delta = 1/156$. In an empirical context, these values are consistent with 5-minute and 2.5-minute sampling frequencies, assuming a 6.5 hour trading day (see e.g., Aït-Sahalia and Jacod (2009)). Loosely speaking, we view our values of Δ as associated with noise which is either not binding or moderately binding. Recall also that for the $t_{\lambda,T,\Delta}$ and $t_{\mu_Z,T,\Delta}$ tests, a key assumption is that $T\Delta \to \infty$ and $\sqrt{T}\Delta \to 0$, leading to a restriction that $1/\Delta < T < 1/\Delta^2$. In particular, when $\Delta = 1/78$ we set $T = \{60, 70, 80, 90, 100, 110, 120, 130\}$ and when $\Delta = 1/156$ we set $T = \{160, 180, 200, 220, 240, 260, 280, 300\}$. Notice that all values of T satisfy the condition, with the exception of $T = 60$ and $T = 70$. These sample sizes are included in order to provide some evidence on the performance of the tests when the condition is broken. In all experiments, we perform 1000 Monte Carlo replications.

When implementing the test for no jumps $t_{\lambda,T,\Delta}$, in order to disentangle the contribution of jumps from that of leverage, we also need to select the thresholding sequence $\tau(\Delta)$ and the "longer" time span T^+. We set $T^+ = 10T$, to satisfy the requirement that $T^+/T \to \infty$, which ensures that the contribution of leverage estimation error is negligible. We then set $\tau(\Delta) = c\Delta^\eta$, with $\frac{2}{7} < \eta < \frac{1}{2}$, and given that for most choices of T, $\Delta = T^{-\delta}$, $\frac{7}{8} < \delta < 1$, we have $\tau(\Delta) = cT^{-\delta\eta}$, and for $\delta\eta \in (\frac{1}{4}, \frac{1}{2})$, the rate conditions $\frac{\Delta^{\frac{1}{2}}}{\tau(\Delta)} \to 0$, $\sqrt{\Delta}\tau(\Delta)^{-1}$ and $\sqrt{T}\tau^2(\Delta) \to 0$ are satisfied.[5] Finally, we set $c = \widehat{\sigma}_{\mu_Z}$, where $\widehat{\sigma}^2_{\mu_Z}$ is defined in (**??**). The choice of $\delta\eta$ is also important in finite sample applications. Consider the case where there are no jumps, so that $\widehat{\sigma}^2_{\lambda,T,\Delta}$, which is used in the construction of $t_{\lambda,T,\Delta}$, is a consistent variance estimator. Recalling (10)-(12), it is immediate to see that $\tau(\Delta)$ plays a role both in the numerator and the denominator of $t_{\lambda,T,\Delta}$. In our experimental setup, small thresholds (e.g. $\delta\eta = 0.4$) result in a too small variance estimator, leading to an oversized test. Not surprisingly, the empirical power is not affected by the choice of the threshold parameter. Below, we only report results for the case where $\delta\eta = 0.251$.

The findings from our Monte Carlo experiment are reported in Tables 1-5. In these tables, rejection frequencies based on tests implemented at a 10% nominal level are reported for various values of λ_∞, $(a, \beta), \sigma$, and Δ.

Tables 1 and 2 contain results for the jump test. Empirical rejection frequencies under $H_0^\lambda : \lambda_\infty = 0$ are given in the first 4 rows of entries in Table 1. When $\Delta = 1/78$, rejection frequencies are very near nominal levels only for $T = 80$ and $T = 90$. Indeed, when T is increased to 130, the empirical size deteriorates substantively and is over 20%. However, rejection frequencies are

[5] Note tht the condition $\sqrt{T}\tau(\Delta) \to 0$ s required only for Theorem 3.

close to the nominal 10% level for $T = 240, 260, 280$, and 300 when $\Delta^{-1} = 1/156$. This finding is quite interesting, and it suggests that the range of permissible T and Δ permutations for which the size properties of our jump test are adequate is wide. Of course, for extremely large values of T, performance should be again expected to deteriorate, since we require that $T < 1/\Delta^2$. Also, the empirical size is not affected by the presence of leverage.

Empirical rejection frequencies under $H_A^{\lambda(1)} : \lambda_\infty > 0$ and $\mathrm{E}\left(Z_k^3\right) \neq 0$ are given in the last 16 rows of entries in Table 1. Recall, that in this case all jump densities have non-zero third moment, and the test has well defined Pitman drift against $\sqrt{T}-$ alternatives. Not surprisingly, rejection frequencies are thus near unit, regardless of jump density specification.

The more challenging alternative is $H_A^{\lambda(2)} : \lambda_\infty > 0$ and $\mathrm{E}\left(Z_k^3\right) = 0$. In this case, the test has zero Pitman drift, and that the ability of the test to distinguish between H_0^λ and $H_A^{\lambda(2)}$ derives solely from the different order of magnitude of the variance under the two hypotheses. Results for DGPs generated under $H_A^{\lambda(2)}$ are gathered in Table 2, for $\sigma = 0.1, 0.2, 0.4$, and $\lambda_\infty = 0.3, 0.5, 0.7$. As might be expected, the power increases as σ and λ_∞ increase. However, the value of σ plays a much bigger role than that of λ_∞. This is not surprising, given that what drives the power is the order of magnitude of the variance.

Table 3 summarizes experimental findings for the zero mean jump test, based on $t_{\mu_Z,T,\Delta} = \sqrt{T}\frac{\widehat{\mu}^Z_{T,\Delta}}{\widehat{\sigma}_{\mu_Z}}$. This test can be thought of as a pre-test, prior to testing for self excitement, and after testing for jumps, say. The reason for this is that in order to identify β when testing $H_0^\beta : \beta = 0$ vs. $H_A^\beta : \beta > 0$, we require not only that $\lambda_\infty > 0$ but also that $E(Z_k) \neq 0$. Inspection of the rejection frequencies reported in the table indicates that the test is well sized, for all values of T, when $\lambda_\infty = 0.3$ and $\sigma = 0.2$. Similar results obtain for other values of λ_∞ and σ, and are hence not reported. Overall, the power is quite good, except for the cases in which T is small, λ_∞ is small, $\Delta^{-1} = 78$, and there is no self-excitation.

Finally, Tables 4 (empirical size) and 5 (empirical power) summarize experimental findings for our self excitement test, based on $S^\beta_{T^+,\Delta} = \max\left\{0, t_{\beta,T^+,\Delta}\right\}$, where $t_{\beta,T^+,\Delta} = \frac{\sqrt{\frac{T^+}{\Delta}}\widehat{\beta}_{T^+,\Delta}}{\widehat{\sigma}_{\beta,T^+,\Delta}}$. Although 192 $(\lambda_\infty,\sigma,\Delta^{-1},T)$ permutations are reported in Table 4, even cursory examination of the table indicates that the test is very well sized, with rejection frequencies very close to the nominal 10% level, in all cases.

It remains only to examine the performance of the self excitement test, under $H_A^\beta : \beta > 0$. The findings are reported in Table 5, where empirical power is summarized for various values of λ_∞,(a,β), Δ^{-1}, and T.[6] For the case where Z_k is a normal random variable, we report rejection frequencies for only one value of σ (i.e., $\sigma = 0.2$). However, it should be noted that empirical power generally declines as σ increases from 0.1 to 0.7, for fixed values of the other parameters.

[6]As the test is based on the first autocovariance term, leverage plays no role in the test statistic. Not surprisingly, then, including leverage (or not) has no qualitatively noteworthy impact on our findings, and only results for the non-zero leverage case are reported.

One possible explanation for this finding is that "noisiness" is induced when estimating the first autocovariance term, when jumps are extremely large.

Empirical power also declines when the value of a is increased, with $(a - \beta)$ fixed (compare rejection frequencies for $(a, \beta) = (3, 2)$ with those for $(a, \beta) = (5, 4)$). The reason for this follows from (13)-(20), where it is immediate to see that the smaller is a and the smaller is $(a-\beta)$, the higher is the degree of self-excitation. This means that our lowest degree of self-excitation is associated with the case where $(a, \beta) = (7, 5)$. Empirical power is correspondingly the lowest in this case, bottoming out at around 0.30. When the degree of self-excitation is strongest (i.e., $(a, \beta) = (3, 2)$) and there are "enough" jumps (i.e., $\lambda_\infty = 0.7$), rejection frequencies are (roughly) in the range 0.70 to 0.80.

In summary, all of our tests perform as expected, given the asymptotic theory describing their large sample behavior. Moreover, the finite sample performance of the tests is found to be good, in all cases, with the important obvious caveat that jump process characteristics and sampling frequencies affect the ability of the tests to perform adequately.

6 Concluding Remarks

If the intensity parameter in a jump diffusion model is identically zero, then parameters characterizing the jump size density cannot be identified. In general, this lack of identification precludes consistent estimation of identified parameters. In the extant literature, there are a large variety of tests for the null of no jumps versus the alternative of jumps, including tests based on the comparison of two realized volatility measures, one which is robust, and the other which is not robust to the presence of jumps (see, e.g. Barndorff-Nielsen, Shephard and Winkel (2006) and Podolskji and Vetter (2009a)), tests based on a thresholding approach (see, e.g. Corsi, Pirino, and Renò (2010), Lee and Mykland (2008), and Lee, Loretan and Ploberger (2013)), and tests based on power variation (see e.g. Aït-Sahalia and Jacod (2009)). One feature of these tests is that they are based on observations drawn on a given finite time span, and thus they can only detect realized jumps. This paper introduces a test which is instead able to detect jumps in the data generating process. Our test is based on realized tricity and make use of high frequency observations measured over a long time-span. Importantly, the test is robust to the presence of leverage. It has a normal limiting distribution, and so inference is straightforward. A so-called "self-excitement" test is also introduced, which is designed to have power against path dependent intensity, thus providing a direct test for the Hawkes diffusion model of Aït-Sahalia, Cacho-Diaz and Laeven (2015). The finite sample behavior of the suggested statistics is studied via Monte Carlo experimentation, and is found to be adequate under a variety of realistic data generating processes.

7 Appendix

Lemma 1: Let Assumptions **A(i)-(iii)** hold. Also, as $T \to \infty$, $T\Delta^2 \to 0$, $\tau(\Delta) \to 0$ and $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)} \to 0$, with $m > 6$ and even. Then if $\lambda = 0$,

$$P\left(\max_{k \le n} \left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| > \varepsilon\tau(\Delta)\right) \to 0.$$

Proof of Lemma 1: Recalling that $T\Delta^2 \to 0$,

$$\begin{aligned}
&P\left(\max_{k \le n} \left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| > \varepsilon\tau(\Delta)\right) \\
&\le \sum_{k=1}^{n} P\left(\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| > \varepsilon\tau(\Delta)\right) \\
&= \sum_{k=1}^{n} P\left(\left|\mu\Delta + V_{k\Delta}^{1/2}\sqrt{1-\rho^2}\left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right)\right| > \varepsilon\tau(\Delta)\right) \\
&\le \frac{T}{\Delta}\frac{1}{\varepsilon^m\tau(\Delta)^m}\mathrm{E}\left(\left|\mu\Delta + V_{k\Delta}^{1/2}\sqrt{1-\rho^2}\left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right)\right|^m\right) \\
&\le C\mathrm{E}\left(V_{k\Delta}^{m/2}\right)\Delta^{\frac{m}{2}-3}\tau(\Delta)^{-m} \to 0,
\end{aligned}$$

for $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)} \to 0$ and $m > 6$, as $\mathrm{E}\left(V_{k\Delta}^{m/2}\right)$ is finite, by A(iii).

Lemma 2: Let Assumptions **A(iv)-(v)** hold. Then, for all $l \ge 2$ even,

$$\mathrm{E}\left(\frac{1}{T}\sum_{k=1}^{n}\left(Z_k 1_{\Delta_{N_{(k+1)\Delta}}} - \Delta\mathrm{E}(Z_k)\lambda\right)^l 1\left\{\left|Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right| \le \tau(\Delta)\right\}\right) \le C\tau(\Delta)^{l+1}.$$

Proof of Lemma 2: For l even, and for $c_{l-k,k} > 0$, whenever l and k are even,

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{T}\sum_{k=1}^{n}\left(Z_k 1_{\Delta_{N_{(k+1)\Delta}}} - \Delta\mathrm{E}(Z_k)\lambda\right)^l 1\left\{\left|Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right| \le \tau(\Delta)\right\}\right) \\
&= \frac{1}{\Delta}\int_{-\tau(\Delta)}^{\tau(\Delta)}\left((z - \Delta\mathrm{E}(Z)\lambda)\right)^l \mathrm{E}\left(1_{\Delta_{N_{(k+1)\Delta}}}\right) f_Z(z)\mathrm{d}z \\
&= \lambda\int_{-\tau(\Delta)}^{\tau(\Delta)}\left(z - \Delta\mathrm{E}(Z_k)\lambda\right)^l f_Z(z)\mathrm{d}z \\
&= \lambda\sum_{k=0}^{l} c_{l-k,k}\int_{-\tau(\Delta)}^{\tau(\Delta)} Z^{l-k}\left(\Delta\mathrm{E}(Z_k)\lambda\right)^k f_Z(z)\mathrm{d}z\left(1 + o(\Delta)\right) \\
&= \lambda c_{l,0}\int_{-\tau(\Delta)}^{\tau(\Delta)} Z^l f_Z(z)\mathrm{d}z\left(1 + o(\Delta)\right) \\
&\le C\tau(\Delta)^{l+1}.
\end{aligned}$$

Lemma 3: Let Assumptions **A(i)-(v)** hold. If as $T, \Delta^{-1} \to \infty$, $T^{+}/T \to \infty$, $\frac{\Delta^{\frac{1}{2}-\frac{3}{m}}}{\tau(\Delta)} \to 0$, with $m > 6$ and $\sqrt{T}\tau^{2}(\Delta) \to 0$, then,

$$\sqrt{T}\left(\frac{1}{T^{+}}\sum_{k=1}^{n^{+}}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right)1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \leq \tau(\Delta)\right\} - \mu\right) = o_{p}(1).$$

Proof of Lemma 3: Note that $\sqrt{T}\tau(\Delta) \to 0$ implies that $T\Delta^{2} \to 0$, and so by Lemma 1,

$$\begin{aligned}
&\frac{\sqrt{T}}{T^{+}}\sum_{k=1}^{n^{+}}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right)1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \leq \tau(\Delta)\right\} \\
&= \frac{\sqrt{T}}{T^{+}}\sum_{k=1}^{n^{+}}\left(\mu\Delta + V_{k\Delta}^{1/2}\sqrt{1-\rho^{2}}\left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right)\right)\left(1 - 1_{\Delta_{N_{(k+1)\Delta}}}\right) \\
&+ \frac{\sqrt{T}}{T^{+}}\sum_{k=1}^{n^{+}}\left(\mu\Delta + V_{k\Delta}^{1/2}\sqrt{1-\rho^{2}}\left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right) + Z_{k}\right) \\
&\times 1\left\{\left|Z_{k}1_{\Delta_{N_{(k+1)\Delta}}}\right| \leq \tau(\Delta)\right\} + o_{p}(1) \\
&= I_{T,\Delta} + II_{T,\Delta}.
\end{aligned}$$

We need to show that: (i) $I_{T,\Delta} = \sqrt{T}\mu + o_{p}(1)$; and (ii) $II_{T,\Delta} = o_{p}(1)$.

For $T^{+}/T \to \infty$,

$$\frac{\sqrt{T}}{T^{+}}\sum_{k=1}^{n^{+}}\left(V_{k\Delta}^{1/2}\sqrt{1-\rho^{2}}\left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right)\right)\left(1 - 1_{\Delta_{N_{(k+1)\Delta}}}\right) = o_{p}(1)$$

and

$$\frac{\sqrt{T}}{T^{+}}\sum_{k=1}^{n^{+}}1_{\Delta_{N_{(k+1)\Delta}}} = o_{p}(1)$$

as $\operatorname{var}\left(1_{\Delta_{N_{(k+1)\Delta}}}\right) = O(\Delta)$ and $T/T^{+} \to 0$. Thus, (i) is established. With regard to (ii), as

$$\begin{aligned}
\mathrm{E}\left(1\left\{\left|Z_{k}1_{\Delta_{N_{(k+1)\Delta}}}\right| \leq \tau(\Delta)\right\}\right) &= \mathrm{E}\left(1\left\{\left|Z_{k}\right| \leq \tau(\Delta)\right\}\right)\mathrm{E}\left(1_{\Delta_{N_{(k+1)\Delta}}}\right) \\
&\leq O\left(\tau(\Delta)\Delta\right)
\end{aligned}$$

it follows that

$$\begin{aligned}
&\frac{\sqrt{T}}{T^{+}}\sum_{k=1}^{n^{+}}\left(\mu\Delta + V_{k\Delta}^{1/2}\sqrt{1-\rho^{2}}\left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right)\right) \\
&\times 1\left\{\left|Z_{k}1_{\Delta_{N_{(k+1)\Delta}}}\right| \leq \tau(\Delta)\right\} = o_{p}(1).
\end{aligned}$$

Finally,

$$
\begin{aligned}
&\frac{\sqrt{T}}{T^+}\sum_{k=1}^{n^+}\mathrm{E}\left(|Z_k|\right)1\left\{\left|Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right|\leq\tau\left(\Delta\right)\right\}\\
&=\frac{\sqrt{T}}{\Delta}\int_{-\tau(\Delta)}^{\tau(\Delta)}|z|f_Z(z)\mathrm{d}z\mathrm{E}\left(1_{\Delta_{N_{(k+1)\Delta}}}\right)\\
&\leq C\sqrt{T}\tau^2\left(\Delta\right)=o(1).
\end{aligned}
$$

Then (ii) follows by a straightforward application of the Markov inequality.

Proof of Theorem 1:

Part (i): From the multivariate Milstein formula (see Kloeden and Platen (1999), Section 10.3), we have that:

$$
\begin{aligned}
&\ln X_{(k+1)\Delta}-\ln X_{k\Delta}\\
&=\Big(\mu\Delta+V_{k\Delta}^{1/2}\sqrt{1-\rho^2}\left(W_{1,(k+1)\Delta}-W_{1,k\Delta}\right)+V_{k\Delta}^{1/2}\rho\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)\\
&+\frac{1}{4}\rho V_{(k-1)\Delta}^{-1/2}g\left(V_{(k-1)\Delta},\theta\right)\left(\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)^2-\Delta\right)\\
&+\frac{1}{4}\sqrt{1-\rho^2}V_{(k-1)\Delta}^{-1/2}g\left(V_{k\Delta},\theta\right)\int_{k\Delta}^{(k+1)\Delta}\int_{k\Delta}^{s_2}\mathrm{d}W_{2,s_1}\mathrm{d}W_{1,s_2}\Big)\left(1+o_p(1)\right)
\end{aligned}
\tag{21}
$$

Also,

$$
\frac{\ln X_T}{n}=\left(\mu\Delta+\sqrt{1-\rho^2}\frac{\Delta}{T}\int_0^T V_s^{1/2}\mathrm{d}W_{1,s}+\rho\frac{\Delta}{T}\int_0^T V_s^{1/2}\mathrm{d}W_{2,s}\right)\left(1+o_p(1)\right)\tag{22}
$$

and

$$
\frac{\ln X_\Delta}{n}=\left(\mu\frac{\Delta^2}{T}+\sqrt{1-\rho^2}\frac{\Delta}{T}\int_0^\Delta V_s^{1/2}\mathrm{d}W_{1,s}+\rho\frac{\Delta}{T}\int_0^\Delta V_s^{1/2}\mathrm{d}W_{2,s}\right)\left(1+o_p(1)\right).
$$

Thus, $\frac{\ln X_\Delta}{n}=O_p\left(\frac{\Delta^{3/2}}{T}\right)$ and $\frac{\ln X_\Delta}{n^+}=O_p\left(\frac{\Delta^{3/2}}{T^+}\right)$. These terms can thus be ignored given that they are $o_p\left(\Delta\right)$. Now,

$$
\begin{aligned}
&\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_T}{n}\right)^3\\
&=\Big(\sqrt{1-\rho^2}V_{k\Delta}^{1/2}\left(W_{1,(k+1)\Delta}-W_{1,k\Delta}\right)+\rho V_{k\Delta}^{1/2}\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)\\
&+\frac{1}{4}\rho V_{k\Delta}^{-1/2}g\left(V_{k\Delta},\theta\right)\left(\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)^2-\Delta\right)+\\
&+\frac{1}{4}\sqrt{1-\rho^2}V_{(k-1)\Delta}^{-1/2}g\left(V_{k\Delta},\theta\right)\int_{k\Delta}^{(k+1)\Delta}\int_{k\Delta}^{s_2}\mathrm{d}W_{2,s_1}\mathrm{d}W_{1,s_2}\Big)^3\left(1+o_p(1)\right).
\end{aligned}
$$

Straightforward but tedious algebra shows that

$$\begin{aligned}
&\mathrm{E}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_T}{n}\right)^3\\
&=\frac{3}{2}\rho^3\mathrm{E}\left(V_{k\Delta}^{1/2}g\left(V_{k\Delta},\theta\right)\right)\Delta^2.
\end{aligned}$$

Because of Lemma 1,

$$\begin{aligned}
\widehat{\mu}_{3,T,\Delta}=&\frac{1}{T}\sum_{k=1}^{n}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_T}{n}\right)^3\\
&-\frac{1}{T^+}\sum_{k=1}^{n^+}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_T}{n^+}\right)^3\left(1+o_p(1)\right).
\end{aligned}$$

Now, write

$$\begin{aligned}
&\frac{\sqrt{T}}{\Delta}\widehat{\mu}_{3,T,\Delta}\\
&=\frac{\Delta}{\sqrt{T}}\sum_{k=1}^{n}\left(\frac{1}{\Delta^2}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n\Delta}-\ln X_{\Delta}}{n}\right)^3-\frac{3}{2}\rho^3\mathrm{E}\left(V_{k\Delta}^{1/2}g\left(V_{k\Delta},\theta\right)\right)\right)\\
&-\frac{\sqrt{T}}{T^+}\Delta\sum_{k=1}^{n^+}\left(\frac{1}{\Delta^2}\left(\ln X_{k\Delta}-\ln X_{(k-1)\Delta}-\frac{\ln X_{n^+\Delta}-\ln X_{\Delta}}{n^+}\right)^3-\frac{3}{2}\rho^3\mathrm{E}\left(V_{k\Delta}^{1/2}g\left(V_{k\Delta},\theta\right)\right)\right)\left(1+o_p(1)\right)\\
&=\left(I_{T,\Delta}+II_{T,T^+,\Delta}\right)\left(1+o_p(1)\right).
\end{aligned}$$

It is immediate to see that $\mathrm{E}(I_{T,\Delta})=0$. Also, recalling that for m even, the $m-$th central moment of a standard normal is equal to $\frac{m!}{2^{m/2}(m/2)!}$,

$$\begin{aligned}
&\omega_0=\mathrm{var}\left(I_{T,\Delta}\right)\\
&=\mathrm{var}\left(\frac{1}{\sqrt{T}\Delta}\sum_{k=1}^{n}\left(\left(1-\rho^2\right)^{1/2}V_{k\Delta}^{1/2}\left(W_{1,(k+1)\Delta}-W_{1,k\Delta}\right)+\rho V_{k\Delta}^{1/2}\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)\right)^3\right)+o(1)\\
&=\mathrm{E}\left(\left(\left(1-\rho^2\right)^{1/2}V_{k\Delta}^{1/2}\left(W_{1,(k+1)\Delta}-W_{1,k\Delta}\right)+\rho V_{k\Delta}^{1/2}\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)\right)^6\right)\\
&=15\left(1-\rho^2\right)^3\mathrm{E}\left(V_{k\Delta}^3\right)+15\rho^6\mathrm{E}\left(V_{k\Delta}^3\right)+45\left(1-\rho^2\right)^2\rho^2\mathrm{E}\left(V_{k\Delta}^3\right)+45\left(1-\rho^2\right)\rho^4\mathrm{E}\left(V_{k\Delta}^3\right).
\end{aligned}$$

Hence, by the central limit theorem for martingale differences,

$$I_{T,\Delta}\xrightarrow{d}N\left(0,\omega_0\right).$$

Since $T^+/T\to\infty$, $II_{T,T^+,\Delta}$ is of smaller probability order than $I_{T,\Delta}$, and thus is $o_p(1)$. The statement in Part (i) then follows.

Part (ii): Let

$$\ln X_{(k+1)\Delta} - \ln X_{k\Delta} = \left(\ln X^c_{(k+1)\Delta} - \ln X^c_{k\Delta}\right) + \left(\ln X^d_{(k+1)\Delta} - \ln X^d_{k\Delta}\right),$$

where $\left(\ln X^c_{(k+1)\Delta} - \ln X^c_{k\Delta}\right)$ is defined as in the RHS of (21), and

$$\left(\ln X^d_{(k+1)\Delta} - \ln X^d_{k\Delta}\right) = Z_k 1_{\Delta N_{(k+1)\Delta}},$$

where Z_k denotes a draw from the jump size density, say f_Z, and $1_{\Delta N_{(k+1)\Delta}} = 1$, if $\Delta N_{(k+1)\Delta} = 1$, and equals zero otherwise. Also

$$\frac{\ln X_T}{n} = \frac{\ln X^c_T}{n} + \frac{1}{n}\sum_{i=0}^{N_T} Z_i = \frac{\ln X^c_T}{n} + \Delta\lambda \mathrm{E}\left(Z_k\right) + o_p(1),$$

with $\frac{\ln X^c_T}{n}$ defined as in the RHS of (22). Write,

$$\begin{aligned}
\frac{\sqrt{T}}{\Delta}\widehat{\mu}_{3,T,\Delta} &= \Bigg(\frac{1}{\Delta\sqrt{T}}\sum_{k=1}^{n}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)^3 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \le \tau\left(\Delta\right)\right\} \\
&\quad - \frac{\sqrt{T}}{\Delta T^+}\sum_{k=1}^{n^+}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n^+\Delta} - \ln X_{\Delta}}{n^+}\right)^3 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \le \tau\left(\Delta\right)\right\}\Bigg) \\
&\quad + \frac{1}{\Delta\sqrt{T}}\sum_{k=1}^{n}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)^3 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| > \tau\left(\Delta\right)\right\} \\
&= A_{T,T^+,\Delta} + B_{T,\Delta}.
\end{aligned}$$

By Lemma 1,

$$\begin{aligned}
A_{T,T^+,\Delta} &= \frac{1}{\Delta\sqrt{T}}\sum_{k=1}^{n}\left(\left(\ln X^c_{k\Delta} - \ln X^c_{(k-1)\Delta} - \frac{\ln X^c_{n\Delta}}{n}\right)^3 - \frac{3}{2}\rho^3 \mathrm{E}\left(V^{1/2}_{k\Delta} g\left(V_{k\Delta},\theta\right)\right)\right) \\
&\quad + \Bigg(\frac{1}{\Delta\sqrt{T}}\sum_{k=1}^{n}\left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}\left(Z_k\right)\right)^3 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \le \tau\left(\Delta\right)\right\} \\
&\quad - \frac{\sqrt{T}}{\Delta T^+}\sum_{k=1}^{n^+}\left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}\left(Z_k\right)\right)^3 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \le \tau\left(\Delta\right)\right\}\Bigg) \\
&\quad + \text{ cross terms} \\
&= A_{1,T,T^+,\Delta} + A_{2,T,T^+,\Delta} + \text{ cross terms.}
\end{aligned}$$

By the same argument as that used in Part (i), $A_{1,T,T^+,\Delta} \xrightarrow{d} N(0,\omega_0)$, while $A_{2,T,T^+,\Delta} +$ cross terms is of a smaller probability order than $B_{T,\Delta}$.

Because of Lemma 1,

$$B_{T,\Delta} = \frac{1}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} \left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}\left(Z_k\right)_i \right)^3 + o_p(1).$$

Now,

$$\begin{aligned}
&\frac{1}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} \left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}\left(Z_k\right) \right)^3 \\
&= \frac{1}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} Z_k^3 1_{\Delta N_{(k+1)\Delta}} - 2\lambda \mathrm{E}\left(Z_k\right) \frac{\Delta}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} Z_i^2 1_{\Delta N_{(k+1)\Delta}} \\
&+ 2\lambda^2 \mathrm{E}\left(Z_k\right)^2 \frac{\Delta^2}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} Z_k 1_{\Delta N_{(k+1)\Delta}} - \sqrt{T}\Delta\lambda^3 \mathrm{E}\left(Z_k\right)^3 \\
&= \frac{1}{\sqrt{T}} \sum_{k=1}^{n} \left(\frac{1}{\Delta} Z_k^3 1_{\Delta N_{(k+1)\Delta}} - \lambda \mathrm{E}\left(Z_k^3\right) \right) + \frac{\sqrt{T}}{\Delta} \lambda \mathrm{E}\left(Z_k^3\right) \\
&- 2\lambda \mathrm{E}\left(Z_k\right) \frac{\Delta}{\sqrt{T}} \sum_{k=1}^{n} \left(\frac{1}{\Delta} Z_k^2 1_{\Delta N_{(k+1)\Delta}} - \lambda \mathrm{E}\left(Z_k^2\right) \right) - \sqrt{T} 2\lambda^2 \mathrm{E}\left(Z_k\right) \mathrm{E}\left(Z_k^2\right) + o_p(1),
\end{aligned} \tag{23}$$

since $\sqrt{T}\Delta\lambda^3 \mathrm{E}\left(Z_k\right)^3 = o(1)$ and $\frac{\Delta^2}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} Z_k 1_{\Delta N_{(k+1)\Delta}} = o_p(1)$, for $\sqrt{T}\Delta \to 0$.
From (23), we see that the statistic has $\frac{\sqrt{T}}{\Delta}$ Pitman drift, whenever $\mathrm{E}\left(Z_k^3\right) \neq 0$. The statement in Part (ii) then follows.

Part (iii): When $\mathrm{E}\left(Z_k^3\right) = \mathrm{E}\left(Z_k\right) = 0$,

$$\begin{aligned}
&\frac{1}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} \left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}\left(Z_k\right) \right)^3 \\
&= \frac{1}{\sqrt{T}} \sum_{k=1}^{n} \frac{1}{\Delta} Z_k^3 1_{\Delta N_{(k+1)\Delta}} + o_p(1).
\end{aligned}$$

We now show that $\mathrm{var}\left(S_{T,\Delta}\right) = O\left(\frac{1}{\Delta^2}\right)$, regardless of whether the jump intensity is constant or path dependent.
If $\beta = 0$ (no path dependent intensity), then:

$$\begin{aligned}
&\mathrm{var}\left(S_{T,\Delta}\right) \\
&= \mathrm{var}\left(\frac{1}{\sqrt{T}\Delta} \sum_{k=1}^{n} Z_k^3 1_{\Delta N_{(k+1)\Delta}} \right) (1 + o(1)) \\
&= \frac{1}{T\Delta^2} \sum_{k=1}^{n} \mathrm{var}\left(Z_k^3 1_{\Delta N_{(k+1)\Delta}} \right) (1 + o(1))
\end{aligned}$$

$$= \frac{1}{\Delta^3}\mathrm{var}\left(Z_k^3 1_{\Delta N_{(k+1)\Delta}}\right)(1+o(1)) = O\left(\frac{1}{\Delta^2}\right).$$

Alternatively, if $\beta > 0$, one must take autocovariance terms into account when carrying out similar calculations. However, given A(iv), the order of magnitude of the variance is still $O\left(\frac{1}{\Delta^2}\right)$. Given that $\sqrt{T}\Delta \to 0$, $S_{T,\Delta}$ is of probability order Δ^{-1}, and the statement in Part (iii) follows.

Proof of Corollary 2:

Part (i). We need to show that $\widehat{\sigma}^2_{\lambda,T,\Delta} - \omega_0 = o_p(1)$, with ω_0 defined as in the statement of Theorem 1. By Lemma 1,

$$\widehat{\sigma}^2_{\lambda,T,\Delta} = \frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(\ln X_{k\Delta} - \ln X_{(k-1)\Delta} - \frac{\ln X_{n\Delta} - \ln X_{\Delta}}{n}\right)^6 + o_p(1).$$

The statement of Part (i) follows directly by the law of large numbers (for iid processes if $\beta = 0$ and for ergodic mixing processes if $\beta > 0$).

Parts (ii)-(iii): We need to show that $\widehat{\sigma}^2_{\lambda,T,\Delta} = O_p(1)$.Now, note that:

$$\begin{aligned}
\widehat{\sigma}^2_{\lambda,T,\Delta} &= \frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(\ln X^c_{k\Delta} - \ln X^c_{(k-1)\Delta} - \frac{\ln X^c_{n\Delta} - \ln X^c_{\Delta}}{n}\right)^6 \\
&+ \frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}(Z)\right)^6 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \le \tau(\Delta)\right\} \\
&+ \text{ cross terms } + o_p(1). \qquad (24)
\end{aligned}$$

The first term on the RHS of (24) is a consistent estimator of ω_0. It suffices to show that the second term on the RHS of (24) is $O_p(1)$. This follows because the cross term cannot be of a larger order than the second term. Given Lemma 1,

$$\begin{aligned}
&\frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}(Z_k)\right)^6 1\left\{\left|\ln X_{k\Delta} - \ln X_{(k-1)\Delta}\right| \le \tau(\Delta)\right\} \\
&= O_p(1)\left(\frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}(Z_k)\right)^6 1\left\{\left|Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right| \le \tau(\Delta)\right\}\right),
\end{aligned}$$

and by Lemma 2,

$$\begin{aligned}
&P\left(\frac{1}{T\Delta^2}\sum_{k=1}^{n}\left(Z_k 1_{\Delta_{N_{(k+1)\Delta}}} - \mathrm{E}\left(Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right)\right)^6 1\left\{\left|Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right| \le \tau(\Delta)\right\} > \varepsilon\right) \\
&\le \frac{1}{\Delta^2\varepsilon}\mathrm{E}\left(\frac{1}{T}\sum_{k=1}^{n}\left(Z_k 1_{\Delta_{N_{(k+1)\Delta}}} - \mathrm{E}\left(Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right)\right)^6 1\left\{\left|Z_k 1_{\Delta_{N_{(k+1)\Delta}}}\right| \le \tau(\Delta)\right\}\right)
\end{aligned}$$

$$\to 0,$$

provided that $\tau(\Delta)^7 \Delta^{-2} \to 0$.

Proof of Theorem 3:

Part (i): By Lemma 3,

$$\begin{aligned}
&\sqrt{T}\widehat{\mu}^{Z}_{T,\Delta} \\
&= \frac{1}{\sqrt{T}} \sum_{k=0}^{n} \Big(\sqrt{1-\rho^2} V_{k\Delta}^{1/2} \left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + \rho V_{k\Delta}^{1/2} \left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right) \\
&+\mu\Delta + Z_k 1_{\Delta N_{(k+1)\Delta}} \Big) - \sqrt{T}\mu + o_p(1).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\sqrt{T}\widehat{\mu}^{Z}_{T,\Delta} \\
&= \frac{1}{\sqrt{T}} \sum_{k=0}^{n} \Big(\sqrt{1-\rho^2} V_{k\Delta}^{1/2} \left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + \rho V_{k\Delta}^{1/2} \left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right) + Z_k 1_{\Delta N_{(k+1)\Delta}} \Big) \\
&\xrightarrow{d} N\left(0, \sigma^2_{\mu_Z}\right),
\end{aligned}$$

with $\sigma^2_{\mu_Z} = \mathrm{E}\left(V_{k\Delta}\right) + \lambda \mathrm{E}\left(Z_k^2\right)$. As $\widehat{\sigma}^2_{\mu_Z} = \sigma^2_{\mu_Z} + o_p(1)$, the statement in Part (i) follows directly.

(ii) The proof is immediate, as

$$\begin{aligned}
&\sqrt{T}\widehat{\mu}^{Z}_{T,\Delta} \\
&= \frac{1}{\sqrt{T}} \sum_{k=0}^{n} \Big(\sqrt{1-\rho^2} V_{k\Delta}^{1/2} \left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + \rho V_{k\Delta}^{1/2} \left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right) + Z_k 1_{\Delta N_{(k+1)\Delta}} \Big) + \lambda\sqrt{T}\mathrm{E}\left(Z_k\right).
\end{aligned}$$

Proof of Theorem 4:

Part (i): Note that,

$$\begin{aligned}
&\sqrt{\frac{T^+}{\Delta}}\widehat{\beta}_{T,\Delta} \\
&= \frac{1}{\sqrt{T^+\Delta}} \sum_{k=1}^{n^+-1} \Big(\Big(\sqrt{1-\rho^2} V_{k\Delta}^{1/2} \left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) + \rho V_{k\Delta}^{1/2} \left(W_{2,(k+1)\Delta} - W_{2,k\Delta}\right) \\
&+Z_k 1_{\Delta N_{(k+1)\Delta}} - \Delta\lambda \mathrm{E}\left(Z\right) \Big) \Big(\sqrt{1-\rho^2} V_{k\Delta}^{1/2} \left(W_{1,k\Delta} - W_{1,(k-1)\Delta}\right) \\
&\rho V_{k\Delta}^{1/2} \left(W_{2,k\Delta} - W_{2,(k-1)\Delta}\right) + Z_{k-1} 1_{\Delta N_{k\Delta}} - \Delta\lambda \mathrm{E}\left(Z\right) \Big) \Big) + o_p\left(1\right) \\
&= \frac{1}{\sqrt{T^+\Delta}} \sum_{k=1}^{n^+-1} \left(1-\rho^2\right) V_{k\Delta} \left(W_{1,(k+1)\Delta} - W_{1,k\Delta}\right) \left(W_{1,k\Delta} - W_{1,(k-1)\Delta}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{2}{\sqrt{T^{+}\Delta}}\sum_{k=1}^{n^{+}-1}\sqrt{1-\rho^{2}}\rho V_{k\Delta}\left(W_{1,(k+1)\Delta}-W_{1,k\Delta}\right)\left(W_{2,k\Delta}-W_{2,(k-1)\Delta}\right)\\
&+\frac{1}{\sqrt{T^{+}\Delta}}\sum_{k=1}^{n^{+}-1}\rho^{2}V_{k\Delta}\left(W_{2,(k+1)\Delta}-W_{2,k\Delta}\right)\left(W_{2,k\Delta}-W_{2,(k-1)\Delta}\right)\\
&+\frac{1}{\sqrt{T^{+}\Delta}}\sum_{k=1}^{n^{+}-1}\left(Z_{k}Z_{k-1}1_{\Delta N_{(k+1)\Delta}}1_{\Delta N_{k\Delta}}-\Delta^{2}\lambda^{2}\mathrm{E}\left(Z\right)^{2}\right)+o_{p}\left(1\right). \qquad (25)
\end{aligned}
$$

Under the null of $\beta=0$,

$$
\begin{aligned}
&\mathrm{E}\left(Z_{k}Z_{k-1}1_{\Delta N_{(k+1)\Delta}}1_{\Delta N_{(k+1)\Delta}}\right)\\
&=\mathrm{E}\left(Z_{k}\right)^{2}\mathrm{E}\left(1_{\Delta N_{(k+1)\Delta}}\right)\mathrm{E}\left(1_{\Delta N_{k\Delta}}\right)=\Delta^{2}\lambda^{2}\mathrm{E}\left(Z_{k}\right)^{2}.
\end{aligned}
$$

Thus, under H_0, all of the terms on the RHS of (25) have zero mean. Also,

$$
\begin{aligned}
\sigma_{\beta}^{2}&=\mathrm{var}\left(\sqrt{\frac{T^{+}}{\Delta}}\widehat{\tau}_{T,\Delta}\right)\\
&=\left(\left(1-\rho^{2}\right)^{2}+4\rho\left(1-\rho^{2}\right)+\rho^{4}\right)\mathrm{E}\left(V_{k\Delta}^{2}\right)+\lambda^{2}\left(\mathrm{E}\left(Z_{k}^{2}\right)\right)^{2},
\end{aligned}
$$

and since $\widehat{\sigma}_{\beta,T,\Delta}^{2}=\sigma_{\beta,T,\Delta}^{2}+o_p(1)$, by the central limit theorem for martingale differences,

$$
\sqrt{\frac{T^{+}}{\Delta}}t_{\beta,T^{+},\Delta}\rightarrow N\left(0,\sigma_{\beta}^{2}\right).
$$

The statement in Part (i) follows from the continuous mapping theorem.

Part (ii): The first three terms on the RHS of (25) are asymptotically normal, under both hypotheses. With regard to the fourth term, note that, under the alternative, from Hawkes (1971),

$$
\begin{aligned}
&\mathrm{E}\left(Z_{k}Z_{k-1}1_{\Delta N_{(k+1)\Delta}}1_{\Delta N_{k\Delta}}-\Delta^{2}\lambda^{2}\mathrm{E}\left(Z_{k}\right)^{2}\right)\\
&=\Delta^{2}\frac{\beta\lambda\left(2a-\beta\right)}{2\left(a-\beta\right)}\exp\left(-\left(a-\beta\right)\right)\mathrm{E}\left(Z_{k}\right)^{2},
\end{aligned}
$$

and so the fourth term on the RHS of (25) can be written as:

$$
\begin{aligned}
&\frac{1}{\sqrt{T^{+}\Delta}}\sum_{k=1}^{n^{+}-1}\left(\left(Z_{k}Z_{k-1}1_{\Delta N_{(k+1)\Delta}}1_{\Delta N_{k\Delta}}-\Delta^{2}\lambda^{2}\mathrm{E}\left(Z_{k}\right)^{2}\right)\right.\\
&\left.-\Delta^{2}\frac{\beta\lambda\left(2a-\beta\right)}{2\left(a-\beta\right)}\exp\left(-\left(a-\beta\right)\right)\mathrm{E}\left(Z_{k}\right)^{2}\right)\\
&+\sqrt{T^{+}\Delta}\frac{\beta\lambda\left(2a-\beta\right)}{2\left(a-\beta\right)}\exp\left(-\left(a-\beta\right)\right)\mathrm{E}\left(Z_{k}\right)^{2}
\end{aligned}
$$

$$= O_p(1) + \sqrt{T^+ \Delta} \frac{\beta\lambda\left(2a-\beta\right)}{2\left(a-\beta\right)} \exp\left(-\left(a-\beta\right)\right) \mathrm{E}\left(Z_k\right)^2 .$$

8 Reference

Aït-Sahalia, Y., J. Cacho-Diaz and R. Laeven (2015). Modeling Financial Contagion Using Mutually Exciting Jump Processes. *Journal of Financial Economics,* 117, 585-606.

Aït-Sahalia, Y., J. Jacod (2009). Testing for Jumps in a Discretely Observed Process. *Annals of Statistics,* 37, 2202-2244.

Aït-Sahalia, Y., J. Jacod and J. Li (2012). Testing for Jumps in Noisy High Frequency Data. *Journal of Econometrics,* 168, 207-222.

Andersen, T.G., T. Bollerslev and F.X. Diebold (2000). Great Realizations. *Risk,* 105-108.

Amaya, D., P. Christoffersen, K. Jacobs and A. Vasquez (2015). Does Realized Skewness Predict the Cross-Section of Equity Returns? *Journal of Financial Economics,* 118, 135-167.

Andersen, Benzoni and Lund (2002). An Empirical Investigation of Continuous Time Equity Return Models. *Journal of Finance,* 62, 1239-1283.

Andrews, D.W.K. (1999). Estimation When a Parameter is on the Boundary. *Econometrica,* 67, 1341-1383.

Andrews, D.W.K. (2001). Testing When a Parameter is on the Boundary of the Maintained Hypothesis. *Econometrica,* 69, 683-734.

Andrews, D.W.K. and X. Cheng (2012). Estimation and Inference with Weak, Semi-strong and Strong Identification. *Econometrica,* 80, 2153-2211.

Bandi, F.M. and R. Renò (2012). Time Varying Leverage Effects. *Journal of Econometrics,* 169, 94-113.

Barndorff-Nielsen, O.E. and N. Shephard (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics,* 2, 1-48.

Barndorff-Nielsen, O.E., N. Shephard and M. Winkel (2006). Limit Theorem for Multipower Variation in the Presence of Jumps. *Stochastic Processes and Their Applications,* 116, 796-806.

Beg, A.B.M.R.A., M.J. Silvapulle and P. Silvapulle (2001). Testing Against Inequality Constraints When Some Nuisance Parameters are Present Only Under the Alternative: Test of ARCH in ARCH-M Models. *Journal of Business and Economic Statistics,* 19, 245-253.

Benjamini, Y., and Y. Hochberg (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society, B*, 57, 289-300.

Bowsher, C.G. (2007). Modeling Security Market Event in Continuous Time: Intensity-Based Multivariate, Point Process Models. *Journal of Econometrics,* 141, 876-912.

Chacko, G., and L.M. Viceira (2003). Spectral GMM estimation of continuous-time Processes. *Journal of Econometrics,* 116, 259-292.

Corradi, V. and N.R. Swanson (2011). Predictive Density Construction and Testing with Multiple Possibly Misspecified Diffusion Models. *Journal of Econometrics,* 161, 304-324.

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

Duffie, D. and K. Singleton (1993). Simulated Moment Estimation of Markov Models of Asset Prices. *Econometrica,* 61, 929-952.

Duffie, D., J. Pan and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump Diffusion. *Econometrica,* 68, 1343-1376.

Fermanian, J.-D. and B. Salanié (2004). A Nonparametric Simulated Maximum Likelihood Estimation Method. *Econometric Theory,* 20, 701-734.

Eraker, B., M. Johannes, and N. Polson (2003). The Impact of Jumps in Volatility and Returns. *Journal of Finance* 58, 1269-1300.

Escanciano, J.C. and I.N. Lobato (2009). An Automatic Data-Driven Portmanteau Test for Testing Serial Autocorrelation. *Journal of Econometrics,* 151, 140-149.

Gallant, A.R. and G. Tauchen (1996). Which Moments to Match. *Econometric Theory,* 12, 657-681.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect Inference. *Journal of Applied Econometrics,* 8, 203-227.

Hawkes, A.G. (1971). Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika,* 58, 83-90.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6, 65–70.

Huang, X. and G.E. Tauchen (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics,* 3, 456-499.

Jacod, J. (2012). Statistics and High Frequency Data, in *Statistical Methods for Stochastic Differential Equations,* M. Kessler, A. Limdner and M. Sorensen Eds, Chapman & Hall, Monographs on Statistics and Probability, v. 124.

Jiang, G.J. and J.L. Knight (2002). Estimation of Continuous-time Processes via the Empirical Characteristic Function. *Journal of Business and Economic Statistics,* 20, 198-212.

Kloeden, P.E. and E. Platen (1999). *Numerical Solutions of Stochastic Differential Equations.* Springer Verlag.

Li, Y., P.A. Mykland, E. Renault, L. Zhang, Z. Zheng (2014). Realized Volatility when Sampling Times are Possibly Endogenous. *Econometric Theory,* 30, 580-05.

Lee, S. and P.A. Mykland (2008). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. *Review of Financial Studies,* 21, 2535-2563.

Lee, T., M. Loretan and W. Ploberger (2013). Rate-Optimal Tests for Jumps in Diffusion Processes. *Statistical Papers.* 54, 1009-1041.

Mancini, C. (2009). Nonparametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics,* 36, 270-296.

Mancini, C. and R. Renò (2011). Threshold Estimation of Markov Models with Jumps and Interest Rate Modelling. *Journal of Econometrics,* 160, 77-92.

Podolskij, M. and M. Vetter (2009a). Bipower Type Estimation in a Noisy Diffusion Setting. *Stochastic Processes and Their Applications.* 119, 2803-2832.

Podolskij, M. and M. Vetter (2009b). Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps. *Bernoulli,* 15, 634-658.

Romano, J.P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, 73, 1237-1282.

Silvapulle, M.J. and P.K. Sen (2005). *Constrained Statistical Inference: Inequality, Order and Shape Restrictions.* Wiley.

Singleton, K.J. (2001). Estimation of Affine Asset Pricing Models Using Empirical Characteristic Function. *Journal of Econometrics,* 102, 111-141.

Storey, J.D. (2003). The Positive False Discovery Rate: a Bayesian Interpretation and the q-value. *Annals of Statistics*, 31, 2013-2035.

White, H. (2000). A Reality Check For Data Snooping. *Econometrica*, 68, 1097-1127.

Table 1: **Monte Carlo Experiments - Jump Test (Empirical Size and Empirical Power, $E(Z_k) \neq 0$)** *

λ_∞	(a,β)	σ	Δ	$T1$	$T2$	$T3$	$T4$	$T5$	$T6$	$T7$	$T8$
						EMPIRICAL SIZE					
						No Leverage					
–	–	–	1/78	0.072	0.086	0.106	0.148	0.142	0.166	0.188	0.210
–	–	–	1/156	0.032	0.040	0.052	0.056	0.080	0.090	0.098	0.108
						Leverage					
–	–	–	1/78	0.060	0.102	0.114	0.132	0.154	0.182	0.204	0.222
–	–	–	1/156	0.036	0.040	0.044	0.064	0.070	0.088	0.096	0.104
						EMPIRICAL POWER					
						Constant Intensity, Z_k is Exp(5)					
0.3	(0,0)	–	1/78	0.992	0.996	0.998	0.998	1.000	1.000	1.000	1.000
0.3	(0,0)	–	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
						Constant Intensity, Z_k is $N(0.5,\sigma^2)$					
0.3	(0,0)	0.2	1/78	0.976	0.986	0.998	0.998	1.000	1.000	1.000	1.000
0.3	(0,0)	0.2	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
						Self Excitement, Z_k is Exp(5)					
0.3	(3,2)	–	1/78	0.992	0.996	0.996	1.000	1.000	1.000	1.000	1.000
0.3	(3,2)	–	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	(5,4)	–	1/78	0.992	0.994	0.998	0.998	0.998	1.000	1.000	1.000
0.3	(5,4)	–	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	(7,5)	–	1/78	0.988	0.992	0.996	1.000	1.000	1.000	1.000	1.000
0.3	(7,5)	–	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
						Self Excitement, Z_k is $N(0.5,\sigma^2)$					
0.3	(3,2)	0.2	1/78	0.992	0.994	0.996	1.000	1.000	1.000	1.000	1.000
0.3	(3,2)	0.2	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	(5,4)	0.2	1/78	0.982	0.988	0.996	0.996	0.996	1.000	1.000	1.000
0.3	(5,4)	0.2	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	(7,5)	0.2	1/78	0.978	0.988	0.992	0.996	0.996	1.000	1.000	1.000
0.3	(7,5)	0.2	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

* Entries in the table are rejection frequencies for the "jump test" based on $t_{\lambda,T,\Delta}$. Results are tabulated for the following sample size (T) and discretization (Δ) permutations: $\Delta = 1/78$ – $T1$:T=60, $T2$:T=70, $T3$:T=80, $T4$:T=90, $T5$:T=100, $T6$:T=110, $T7$:T=120, $T8$:T=130. For $\Delta = 1/156$ – $T1$:T=160, $T2$:T=180, $T3$:T=200, $T4$:T=220, $T5$:T=240, $T6$:T=260, $T7$:T=280, $T8$:T=300. In all experiments, we perform 1000 Monte Carlo replications. For complete details, refer to Section 5 of the paper.

Table 2: **Monte Carlo Experiments - Jump Test (Empirical Power, $E(Z_k) = 0$) ***

λ_∞	(a, β)	σ	Δ	$T1$	$T2$	$T3$	$T4$	$T5$	$T6$	$T7$	$T8$
				Constant Intensity, Z_k ***is*** $N(0.0, \sigma^2)$							
0.3	(0,0)	0.1	1/78	0.078	0.132	0.136	0.142	0.190	0.202	0.224	0.240
0.3	(0,0)	0.1	1/156	0.166	0.176	0.200	0.216	0.222	0.262	0.274	0.268
0.3	(0,0)	0.2	1/78	0.382	0.464	0.482	0.498	0.524	0.540	0.574	0.578
0.3	(0,0)	0.2	1/156	0.768	0.794	0.794	0.798	0.802	0.818	0.834	0.848
0.3	(0,0)	0.4	1/78	0.796	0.844	0.874	0.890	0.908	0.904	0.916	0.928
0.3	(0,0)	0.4	1/156	0.972	0.970	0.966	0.970	0.968	0.988	0.982	0.984
0.5	(0,0)	0.1	1/78	0.086	0.130	0.132	0.178	0.194	0.218	0.244	0.258
0.5	(0,0)	0.1	1/156	0.224	0.274	0.288	0.296	0.306	0.304	0.344	0.334
0.5	(0,0)	0.2	1/78	0.462	0.520	0.576	0.588	0.596	0.636	0.636	0.648
0.5	(0,0)	0.2	1/156	0.832	0.838	0.856	0.850	0.860	0.882	0.892	0.886
0.5	(0,0)	0.4	1/78	0.888	0.894	0.918	0.936	0.936	0.950	0.956	0.946
0.5	(0,0)	0.4	1/156	0.978	0.982	0.962	0.982	0.978	0.974	0.978	0.976
0.7	(0,0)	0.1	1/78	0.122	0.170	0.192	0.188	0.214	0.224	0.262	0.286
0.7	(0,0)	0.1	1/156	0.294	0.300	0.326	0.348	0.362	0.388	0.364	0.370
0.7	(0,0)	0.2	1/78	0.574	0.604	0.644	0.654	0.660	0.676	0.682	0.706
0.7	(0,0)	0.2	1/156	0.856	0.864	0.896	0.886	0.866	0.874	0.886	0.872
0.7	(0,0)	0.4	1/78	0.918	0.914	0.934	0.938	0.940	0.956	0.958	0.944
0.7	(0,0)	0.4	1/156	0.972	0.972	0.984	0.976	0.986	0.980	0.980	0.984

* See notes to Table 1.

Table 3: **Monte Carlo Experiments - Test of** $E(Z_k) = 0$ *****

λ_∞	(a,β)	σ	Δ	$T1$	$T2$	$T3$	$T4$	$T5$	$T6$	$T7$	$T8$
EMPIRICAL SIZE											
Constant Intensity, Z_k ***is*** $N(0.0,\sigma^2)$											
0.3	(0,0)	0.2	1/78	0.070	0.062	0.062	0.078	0.082	0.082	0.086	0.100
0.3	(0,0)	0.2	1/156	0.082	0.086	0.076	0.066	0.070	0.068	0.092	0.090
Self Excitement, Z_k ***is*** $N(0.0,\sigma^2)$											
0.3	(5,4)	0.2	1/78	0.060	0.056	0.068	0.080	0.060	0.078	0.080	0.082
0.3	(5,4)	0.2	1/156	0.088	0.088	0.090	0.094	0.094	0.078	0.094	0.082
EMPIRICAL POWER											
Constant Intensity, Z_k ***is*** Exp***(5)***											
0.3	(0,0)	–	1/78	0.182	0.236	0.264	0.272	0.320	0.324	0.334	0.366
0.3	(0,0)	–	1/156	0.858	0.880	0.928	0.934	0.964	0.972	0.970	0.982
0.5	(0,0)	–	1/78	0.414	0.458	0.480	0.508	0.562	0.584	0.616	0.650
0.5	(0,0)	–	1/156	0.998	0.998	0.998	1.000	1.000	1.000	1.000	1.000
0.7	(0,0)	–	1/78	0.580	0.652	0.716	0.742	0.784	0.814	0.836	0.870
0.7	(0,0)	–	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Constant Intensity, Z_k ***is*** $N(0.5,\sigma^2)$											
0.3	(0,0)	0.2	1/78	0.138	0.174	0.170	0.190	0.208	0.220	0.208	0.226
0.3	(0,0)	0.2	1/156	0.676	0.718	0.728	0.778	0.810	0.868	0.872	0.886
0.5	(0,0)	0.2	1/78	0.274	0.310	0.324	0.352	0.394	0.428	0.428	0.454
0.5	(0,0)	0.2	1/156	0.954	0.966	0.978	0.986	0.996	0.998	0.998	1.000
0.7	(0,0)	0.2	1/78	0.396	0.450	0.502	0.552	0.596	0.624	0.642	0.694
0.7	(0,0)	0.2	1/156	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Self Excitement, Z_k ***is*** Exp***(5)***											
0.3	(5,4)	–	1/78	0.650	0.664	0.674	0.690	0.694	0.704	0.708	0.716
0.3	(5,4)	–	1/156	0.928	0.946	0.968	0.972	0.982	0.990	0.994	0.994
0.5	(5,4)	–	1/78	0.836	0.846	0.854	0.872	0.872	0.874	0.894	0.904
0.5	(5,4)	–	1/156	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	(5,4)	–	1/78	0.950	0.954	0.964	0.970	0.972	0.978	0.976	0.980
0.7	(5,4)	–	1/156	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Self Excitement, Z_k ***is*** $N(0.5,\sigma^2)$											
0.3	(5,4)	0.2	1/78	0.602	0.624	0.628	0.624	0.644	0.638	0.642	0.658
0.3	(5,4)	0.2	1/156	0.852	0.870	0.878	0.886	0.906	0.940	0.944	0.958
0.5	(5,4)	0.2	1/78	0.784	0.804	0.812	0.820	0.810	0.828	0.828	0.844
0.5	(5,4)	0.2	1/156	0.980	0.988	0.992	0.998	1.000	1.000	1.000	1.000
0.7	(5,4)	0.2	1/78	0.916	0.928	0.922	0.936	0.942	0.946	0.960	0.956
0.7	(5,4)	0.2	1/156	0.996	0.994	0.998	1.000	1.000	1.000	1.000	1.000

* See notes to Table 1. Entries in the table are rejection frequencies for the zero mean jump test, based on $t_{\mu_Z,T,\Delta} = \sqrt{T}\frac{\widehat{\mu}^Z_{T,\Delta}}{\widehat{\sigma}_{\mu_Z}}$.

Table 4: **Monte Carlo Experiments - Self Excitement Test (Empirical Size) ***

λ_∞	(a,β)	σ	Δ	$T1$	$T2$	$T3$	$T4$	$T5$	$T6$	$T7$	$T8$
					Z_k ***is*** *Exp(**5**)*						
0.3	(0,0)	–	1/78	0.114	0.108	0.120	0.138	0.124	0.146	0.126	0.130
0.3	(0,0)	–	1/156	0.092	0.094	0.080	0.090	0.082	0.092	0.080	0.076
0.5	(0,0)	–	1/78	0.088	0.088	0.096	0.098	0.108	0.118	0.122	0.114
0.5	(0,0)	–	1/156	0.070	0.076	0.068	0.070	0.088	0.078	0.076	0.068
0.7	(0,0)	–	1/78	0.088	0.098	0.122	0.128	0.108	0.124	0.118	0.122
0.7	(0,0)	–	1/156	0.076	0.076	0.084	0.074	0.090	0.078	0.082	0.086
					Z_k ***is*** $N(0.5,\sigma^2)$						
0.3	(0,0)	0.1	1/78	0.116	0.124	0.126	0.124	0.136	0.144	0.138	0.146
0.3	(0,0)	0.1	1/156	0.084	0.106	0.094	0.092	0.094	0.086	0.078	0.070
0.3	(0,0)	0.2	1/78	0.118	0.122	0.136	0.142	0.136	0.150	0.146	0.134
0.3	(0,0)	0.2	1/156	0.102	0.098	0.086	0.088	0.094	0.092	0.084	0.080
0.3	(0,0)	0.4	1/78	0.120	0.116	0.132	0.140	0.130	0.146	0.150	0.132
0.3	(0,0)	0.4	1/156	0.082	0.098	0.086	0.094	0.088	0.088	0.076	0.084
0.5	(0,0)	0.1	1/78	0.094	0.098	0.116	0.132	0.134	0.134	0.142	0.136
0.5	(0,0)	0.1	1/156	0.082	0.084	0.078	0.090	0.100	0.086	0.088	0.074
0.5	(0,0)	0.2	1/78	0.094	0.094	0.102	0.122	0.116	0.126	0.122	0.124
0.5	(0,0)	0.2	1/156	0.086	0.086	0.088	0.088	0.092	0.080	0.076	0.066
0.5	(0,0)	0.4	1/78	0.110	0.096	0.120	0.116	0.118	0.134	0.124	0.122
0.5	(0,0)	0.4	1/156	0.096	0.090	0.078	0.092	0.092	0.088	0.090	0.084
0.7	(0,0)	0.1	1/78	0.104	0.112	0.128	0.132	0.130	0.132	0.104	0.112
0.7	(0,0)	0.1	1/156	0.086	0.086	0.078	0.082	0.094	0.084	0.094	0.098
0.7	(0,0)	0.2	1/78	0.092	0.102	0.116	0.132	0.118	0.126	0.120	0.126
0.7	(0,0)	0.2	1/156	0.072	0.078	0.074	0.084	0.084	0.082	0.092	0.086
0.7	(0,0)	0.4	1/78	0.076	0.094	0.092	0.104	0.110	0.122	0.104	0.098
0.7	(0,0)	0.4	1/156	0.078	0.068	0.086	0.072	0.092	0.092	0.098	0.084

* See notes to Table 1. Entries in the table are rejection frequencies for the "self excitement test" of the null of no jump path dependence, based on $S^{\beta}_{T+,\Delta} = \max\{0, t_{\beta,T+,\Delta}\}$.

Table 5: **Monte Carlo Experiments - Self Excitement Test (Empirical Power) ***

λ_∞	(a,β)	σ	Δ	$T1$	$T2$	$T3$	$T4$	$T5$	$T6$	$T7$	$T8$
						Z_k ***is*** *Exp(**5**)*					
0.3	(3,2)	–	1/78	0.570	0.568	0.564	0.560	0.558	0.572	0.560	0.554
0.3	(3,2)	–	1/156	0.566	0.558	0.556	0.552	0.554	0.558	0.548	0.554
0.3	(5,4)	–	1/78	0.418	0.414	0.426	0.420	0.424	0.430	0.420	0.422
0.3	(5,4)	–	1/156	0.430	0.432	0.422	0.428	0.436	0.430	0.428	0.424
0.3	(7,5)	–	1/78	0.332	0.324	0.338	0.342	0.322	0.336	0.316	0.324
0.3	(7,5)	-	1/156	0.318	0.320	0.324	0.314	0.324	0.318	0.300	0.310
0.5	(3,2)	–	1/78	0.704	0.708	0.704	0.706	0.706	0.696	0.700	0.700
0.5	(3,2)	–	1/156	0.704	0.704	0.706	0.708	0.700	0.710	0.710	0.704
0.5	(5,4)	–	1/78	0.550	0.548	0.544	0.534	0.540	0.536	0.520	0.524
0.5	(5,4)	–	1/156	0.520	0.520	0.516	0.516	0.530	0.530	0.526	0.524
0.5	(7,5)	–	1/78	0.446	0.424	0.438	0.444	0.434	0.434	0.430	0.436
0.5	(7,5)	–	1/156	0.430	0.432	0.452	0.436	0.432	0.432	0.412	0.410
0.7	(3,2)	–	1/78	0.788	0.782	0.780	0.770	0.776	0.772	0.768	0.772
0.7	(3,2)	–	1/156	0.780	0.780	0.772	0.770	0.762	0.752	0.760	0.750
0.7	(5,4)	–	1/78	0.644	0.636	0.648	0.636	0.620	0.616	0.622	0.604
0.7	(5,4)	–	1/156	0.618	0.612	0.614	0.626	0.616	0.606	0.608	0.620
0.7	(7,5)	–	1/78	0.510	0.522	0.500	0.498	0.492	0.500	0.486	0.492
0.7	(7,5)	–	1/156	0.486	0.474	0.468	0.456	0.446	0.438	0.446	0.444
						Z_k ***is*** $N(0.5, \sigma^2)$					
0.3	(3,2)	0.2	1/78	0.560	0.568	0.566	0.560	0.552	0.562	0.548	0.548
0.3	(3,2)	0.2	1/156	0.534	0.540	0.534	0.542	0.540	0.544	0.540	0.542
0.3	(5,4)	0.2	1/78	0.388	0.394	0.398	0.382	0.388	0.392	0.378	0.372
0.3	(5,4)	0.2	1/156	0.396	0.378	0.382	0.384	0.402	0.384	0.380	0.380
0.3	(7,5)	0.2	1/78	0.292	0.308	0.302	0.304	0.286	0.292	0.284	0.284
0.3	(7,5)	0.2	1/156	0.284	0.278	0.272	0.278	0.268	0.274	0.268	0.262
0.5	(3,2)	0.2	1/78	0.676	0.686	0.666	0.674	0.668	0.658	0.642	0.640
0.5	(3,2)	0.2	1/156	0.654	0.652	0.638	0.632	0.630	0.624	0.634	0.626
0.5	(5,4)	0.2	1/78	0.506	0.510	0.502	0.480	0.464	0.462	0.460	0.442
0.5	(5,4)	0.2	1/156	0.476	0.464	0.458	0.450	0.450	0.452	0.454	0.442
0.5	(7,5)	0.2	1/78	0.386	0.388	0.380	0.372	0.362	0.356	0.346	0.344
0.5	(7,5)	0.2	1/156	0.380	0.390	0.384	0.386	0.390	0.364	0.358	0.356
0.7	(3,2)	0.2	1/78	0.706	0.692	0.692	0.686	0.688	0.694	0.682	0.668
0.7	(3,2)	0.2	1/156	0.708	0.696	0.694	0.668	0.674	0.662	0.654	0.654
0.7	(5,4)	0.2	1/78	0.586	0.576	0.566	0.572	0.560	0.556	0.538	0.530
0.7	(5,4)	0.2	1/156	0.560	0.562	0.562	0.556	0.544	0.532	0.530	0.522
0.7	(7,5)	0.2	1/78	0.428	0.432	0.428	0.416	0.406	0.402	0.402	0.396
0.7	(7,5)	0.2	1/156	0.420	0.418	0.416	0.420	0.410	0.404	0.390	0.384

* See notes to Table 4.