Consistent Pretesting for Jumps*

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

June 2014

Abstract: 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. Hence, it should be standard practice to consistently pretest for jumps, prior to estimating jump diffusions. Many currently available tests have power against the presence of jumps over a finite time span (typically a day or a week); and, as already noted by various authors, jumps may not be observed over finite time spans, even if the intensity parameter is strictly positive. Such tests cannot be consistent against non-zero intensity. Moreover, sequential application of finite time span tests usually leads to sequential testing bias, which in turn leads to jump discovery with probability one, in the limit, even if the true intensity is identically zero. This paper introduces tests for jump intensity, based on both in-fill and long-span asymptotics, which solve both the test consistency and the sequential testing bias problems discussed above, in turn facilitating consistent estimation of jump diffusion models. A "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 Ait-Sahalia, Cacho-Diaz and Laeven (2013). In a series of Monte Carlo experiments, the proposed tests are evaluated, and are found to perform adequately in finite samples.

Keywords: diffusion model, jump intensity, jump size density, sequential testing bias, bootstrap.
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 Francesco Bravo, Laura Coroneo, Prosper Dovonon, Rustam Ibragimov, Peter Spencer, and to seminar participants at the University of York, the 2013 Conputational and Financial Econometrics Conference in London, and the 2014 CIREQ Time Series and Financial Econometrics conference in Montreal for useful comments and suggestions.

1 Introduction

Jump diffusions are widely used in the financial econometrics literature when analyzing returns or exchange rates, for example, 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. 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 these diffusion 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 Ait-Sahalia, Cacho-Diaz and Laeven (2013). Clearly, when one estimates a jump diffusion with jump intensity equal to zero, a subset of the parameters is not identified. This in turn precludes consistent estimation of other parameters in the model (see Andrews and Cheng (2012)).

The above estimation problem serves to underscore the importance of pretesting for jumps. In the extant literature, there is 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 Reno (2010)), and tests based on power variation, as discussed in Ait-Sahalia and Jacod (2009). All of these tests are carried out using a finite time span (typically a day or a week), and limiting distributions are found using in-fill asymptotic approximations. However, over a finite time span we may not observe jumps, even if the intensity parameter is positive. Thus, these tests are not consistent against the alternative of positive jump intensity, as pointed out by Huang and Tauchen (2005) and Ait-Sahalia and Jacod (2009), for example. Moreover, sequential application of these tests (again, typically done daily or weekly) results in a failure to control the probability of false jump discovery. This is because of the well known sequential testing size distortion problem. This paper develops new and practical methods for solving the above testing and associated estimation problems in the context of the specification of jump diffusion models.

Consider solving the above problem of pretesting and subsequent estimation in stages by first testing the null of zero versus positive intensity using a score, Wald or likelihood ratio test, as in Andrews (2001), and subsequently estimating the model using standard techniques. This would involve treating the parameters of the jump size density as nuisance parameters unidentified under the null. Furthermore, such an approach would require correct specification of both the continuous and the jump components of the diffusion; and misspecification of one or both components would in general affect the overall outcome of the test. Just as importantly, the likelihood function of a jump diffusion is not generally known in closed form, and therefore estimation (which is needed for construction of these jump tests) is usually based on either simulated GMM (see, e.g. Duffie and Singleton (1993) and Anderson, Benzoni and Lund (2002)); Indirect Inference (see, e.g. Gourieroux and Monfort (1993) and Gallant and Tauchen (1996)); or Nonparametric Simulated Maximum Likelihood (see, e.g. Fermanian and Salanie (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, e.g. Andrews (1999,2001), Beg, Silvapulle and Silvapulle (2001), and Chapter 4 in Silvapulle and Sen (2005)), we cannot rely on simulation-based estimators if attempting to pretest using standard score, Wald or likelihood ratio tests. A different variety of jump pretest is instead required. The approach taken in this paper is to propose model free jump pretests, and to subsequently estimate the jump diffusion using standard estimation techniques, depending upon the outcome of the test(s).

In particular, this paper makes two key contributions to the literature. First, we introduce model free "jump" tests for the null of zero intensity. The tests are based on both in-fill and long-span asymptotics, thus addressing the issues of consistency and sequential testing bias discussed above. Second, under the maintained assumption of strictly positive intensity, we introduce a "self excitement" test for the null of constant intensity against the alternative of path dependent intensity. The objective in this context is the provision of a direct test for Hawkes diffusions (see Ait-Sahalia, Cacho-Diaz and Laeven (2013)) in which jump intensity is modeled as a mean-reverting diffusion process. When the tests are implemented prior to model specification, standard estimation of jump diffusions can be subsequently carried out, avoiding the identification problems discussed above.

The jump tests are based on sample third moments, and are constructed using a long time span of high frequency observations. Two versions of these tests are discussed. One version does not allow for leverage, in the sense that rejection of the null may be due either to the presence of jumps or due to the presence of leverage effects in the underlying data generating process. An alternative version is robust to leverage effects, with the caveat that it is a less powerful test than its non-robust counterpart. The limiting behavior of the proposed statistics can be readily analyzed via use of a double asymptotic scheme wherein the time span goes to infinity and the discrete interval approaches zero. The tests are model free, except for a drift component, which is assumed to be constant[1]. Under the null hypothesis of zero intensity, the statistics are characterized by normal limiting distributions. Under the alternative, it is necessary to distinguish between the case in which the density of the jumps is asymmetric and the case in which the density is symmetric. In the former case, the proposed tests have a well defined Pitman drift and have power against $\sqrt{T}-$local alternatives, where T denotes the time span. In the latter case, the sample third moment approaches zero, but the probability order of the statistics 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. To ensure power under both types of alternatives, it follows that we cannot rescale the test statistics by an estimator of the variance. We instead construct bootstrap critical values, the first order validity of which is established in the sequel.

Turning now to the self excitement test, note that if the null of zero intensity is rejected, one can proceed with a second test that is carried out in order to ascertain whether jump intensity is a constant, or follows a diffusion process, as in the case of Hawkes diffusions. This test is based on the sample autocorrelation of the (log) first differences of the data, and is analyzed using asymptotic approximations closely related to those used in the analysis of the jump test statistics.

As none of the tests proposed in this paper are robust to microstructure noise, one might choose to build a dataset consisting of observations at the highest frequency for which the noise is not binding. However, the assumptions posited in order to analyze the tests herein simply require that the discrete interval approaches zero; and this does not have to occur at a minimum speed. Indeed, in our framework the time span can grow faster than the discrete interval.

The finite sample behavior of the suggested statistics is studied via Monte Carlo experimentation. The jump tests exhibit empirical size very close to nominal and empirical power close to unity, across various empirically motivated parameterizations. The self excitement test likewise has very good size and good power properties, whenever there are "enough" jumps and the degree of self excitation is not "too weak". In an examination of the finite sample behavior of the tests, when carried out in sequence (i.e., carry out the self excitement test in all instances for which the jump test rejects the null of no jumps), we also find evidence of adequate performance.

The rest of the paper is organized as follows. Section 2 describes the set-up. Section 3 provides heuristic arguments for the testing approach taken in this paper. Section 4 discusses the jump and self excitement tests, derives their asymptotic properties, and discusses asymptotically valid bootstrap based inference using an m out of n bootstrap procedure. Section 5 reports the findings of a Monte Carlo study 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.

[1] Recall that drift terms can be ignored over finite time spans, while they have a non-negligible impact on asymptotic approximations over long time spans. Hence, this "trade-off" is not surprising.

2 Set-Up

Consider the following jump diffusion model,

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

where volatility V_t is defined according to either (i), (ii), (iii), or (iv), as follows:

(i) a constant:

$$V_t = v \text{ for all } t; \tag{2}$$

(ii) a measurable function of the state variable:

$$V_t \text{ is } X_t - \text{measurable}; \tag{3}$$

(iii) a stochastic volatility process without leverage:

$$\mathrm{d}V_t = \mu_{V,t}(\theta)\mathrm{d}t + g\left(V_t, \theta\right)\mathrm{d}W_{2,t},\ \mathrm{E}\left(W_{1,t}W_{2,t}\right) = 0; \tag{4}$$

(iv) a stochastic volatility process with leverage:

$$\mathrm{d}V_t = \mu_{V,t}(\theta)\mathrm{d}t + g\left(V_t, \theta\right)\mathrm{d}W_{2,t},\ \mathrm{E}\left(W_{1,t}W_{2,t}\right) = \rho \neq 0. \tag{5}$$

Here,

$$\Pr\left(N_{t+\Delta} - N_t = 1|\mathcal{F}_t\right) = \lambda_t \Delta + o\left(\Delta\right), \tag{6}$$

$$\Pr\left(N_{t+\Delta} - N_t = 0|\mathcal{F}_t\right) = 1 - \lambda_t \Delta + o\left(\Delta\right), \tag{7}$$

and

$$\Pr\left(N_{t+\Delta} - N_t > 1|\mathcal{F}_t\right) = o\left(\Delta\right), \tag{8}$$

where $\mathcal{F}_t = \sigma\left(N_s,\ 0 \leq s \leq t\right)$, and the jump size, Z_t, 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 Hawkes diffusions, in which the intensity is an increasing function of past jumps (see Bowsher (2007) and Ait-Sahalia, Cacho-Diaz and Laeven (2013)). 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,

$$\mathrm{d}\lambda_t = a\left(\lambda_\infty - \lambda_t\right)\mathrm{d}t + \beta \mathrm{d}N_s \tag{9}$$

and $\mathrm{E}(\lambda_t) = \frac{a\lambda_\infty}{a-\beta}$. 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 $\lambda_\infty = 0$, β, α, γ are not identified. By contrast, if $\lambda_\infty > 0$, then γ and β are identified. However, if $\lambda_\infty > 0$ but $\beta = 0$, 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 hypotheses $H_0 : \lambda_\infty = 0$ versus $H_A : \lambda_\infty > 0$. 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 hypotheses (i.e., self excitement pretests): $H_0 : \beta = 0$ versus $H_A : \beta > 0$.

At this juncture, we provide further heuristic motivation by discussing various key differences between existing jump tests and tests based directly on testing the intensity proposed in the sequel.

3 Testing for Jumps or Jump Intensity - Heuristic Arguments

In recent years, a large variety of tests for jumps have been developed. One common feature of these tests is that they are all performed using high frequency observations over a finite time span. We thus argue that none of these tests is consistent against the alternative $\lambda_\infty > 0$. Many of the extant tests can be broadly classified as belonging in one of three groups: (i) Hausman type tests (ii); threshold type tests; and (iii) higher order power variation tests.

Hausman type tests are based on the comparison of non-robust and robust realized volatility measures (see, e.g. Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen, Shephard and Winkel (2006), and Huang and Tauchen (2005)). More recently, related tests have been proposed that are based on comparisons using pre-averaged volatility measures, in order to obtain tests that are robust to microstructure noise (see, e.g. Podolskij and Vetter (2009a)). Hausman type tests are able to detect whether $\sum_{j=N_t}^{N_{t+1}} c_j^2 = 0$ or $\sum_{j=N_t}^{N_{t+1}} c_j^2 > 0$, where N_t denotes the number of jumps up to time t, and c_j is the (random) size of the jumps. However, $\lambda_\infty > 0$ does not imply that $\sum_{j=N_t}^{N_{t+1}} c_j^2 > 0$, given that $\Pr\left(N_{t+1} - N_t > 0\right) < 1$.

Threshold type tests are based on the difference between "standard" volatility measures and trimmed realized measures, where the trimming is implemented at a threshold level which allows for the separation of jump and continuous components. Such tests have power against jump size, but not necessarily against jump intensity (see, e.g. Corsi, Pirino and Reno (2010) and Lee and Mykland (2008)).

A more recent class of tests is based on higher order power variation, and is motivated by the fact that for $p > 2$, $\sum_{i=1}^{n-1} \left|X_{(t+(i+1)\Delta} - X_{t+i\Delta}\right|^p$ converges to $\sum_{t\leq s\leq t+1} \left|X_s - X_{s-}\right|^p$, where

$\sum_{t\leq s\leq t+1}|X_s - X_{s-}|^p$ is strictly positive if there are jumps and zero otherwise (see, e.g. Ait-Sahalia and Jacod (2009) and Ait-Sahalia, Jacod and Li (2012)). Even in this case, power obtains because of jump size, and not because of jump probability.

More generally, for tests performed on a finite time span, are able to distinguish between[2]:

$$\Omega_t^c = \{\omega : s \to X_s \text{ is continuous on } [t, t+1)\}$$

and

$$\Omega_t^j = \{\omega : s \to X_s \text{ has jumps on } [t, t+1)\}.$$

Hence, all of the tests discussed above are dependent upon pathwise behavior. Clearly, one might decide in favor of Ω_t^c, even if $\lambda_\infty > 0$, simply because jumps are not observed over the interval $[t, t+1)$. Lee, Loretan and Ploberger (2013) discuss the optimality properties of jump tests against local alternative defined in terms of jump sizes. It follows that in order to carry out a consistent jump test, one must test the composite hypothesis:

$$\Omega_T^c = \cap_{t=0}^{T-1}\Omega_t^c,$$

versus it negation. Broadly speaking, one must test the composite null hypothesis that none of the daily (or weekly, say) paths contain jumps. In fact, under mild conditions on the degree of heterogeneity of the process, failure to reject $\Omega_\infty^c = \lim_{T\to\infty}\cap_{t=1}^{T-1}\Omega_t^c$ implies failure to reject $\lambda_\infty = 0$. The difficulty herein lies in how to implement a test for Ω_T^c, when T gets large. Needless to say, sequential application of finite time span jump tests leads to sequential test bias, and for T large Ω_T^c is rejected with probability going to unity. At issue here is the control of overall size when testing composite hypotheses. One common approach to this problem is based on controlling the overall Family-Wise Error-Rate (FWER), which ensures that no single hypothesis is rejected at a level larger than a fixed value, say α. This is typically accomplished by sorting individual $p-$values, and using a rejection rule which depends on the overall number of hypotheses. For further discussion, see Holm (1979), who develops modified Bonferroni bounds, White (2000), who develops the so-called "reality check", and Romano and Wolff (2005), who provide a refinement of the reality check. However, when the number of hypotheses in the composite grows with the sample size, the null will (almost) never be rejected. In other words, approaches based on the FWER are far too conservative for our purpose.

An alternative approach, which allows for the number of hypotheses in the composite to grow to infinity, is based on the Expected False Discovery Rate (E-FDR). When using this approach, one controls the expected number of false discoveries (rejections). For further discussion, see Benjamini and Hochberg (1995) and Storey (2003). Although the E-FDR approach applies to the case of a

[2] Jump test inconsistency has been pointed out by Huang and Tauchen (2005) and Ait-Sahalia and Jacod (2009), among others.

growing number of hypotheses, it is very hard to implement in the presence of generic dependences across p-values, as it is in our context.

In summary, a key advantage of jump tests based on high frequency observations over finite time spans is that they are virtually model free, as minimal regularity assumptions on the underlying process are required. A key disadvantage is that they are not consistent against the alternative of positive jump intensity. On the other hand, if more structure is imposed, 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 $H_0 : \lambda_\infty = 0$ against $H_A : \lambda_\infty > 0$. This fact can be illustrated by considering a score test. Suppose that the skeleton of the process $\ln X_t$ in *Eq.* (1) is observed. Namely, $\ln X_1, \ln X_2, ..., \ln X_T$, is observed, with V_t defined as in *Eq.* (4) or *Eq.* (5); and for sake of simplicity suppose that $\lambda_t = \lambda_\infty$. Now, using the notation in *Eqs.* (1)-(8), let $\delta = (\theta, \mu, \rho, \lambda_\infty, \gamma) = (\vartheta, \gamma)$. It immediately follows that, provided the transition density is known in closed form, the likelihood can be written as:

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

The score statistic for testing H_0 is thus[3]:

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

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

$$\begin{aligned} U_T(\gamma) &= \sqrt{T}\left(R\widehat{\mathcal{I}}_T(\gamma)^{-1}\nabla_\vartheta l_T\left(\widehat{\vartheta}_T,\gamma\right)\right), \\ \widehat{\mathcal{I}}_T(\gamma) &= \frac{1}{T}\sum_{t=1}^{T}\nabla_{\vartheta\vartheta} l_t\left(\widehat{\vartheta}_T,\gamma\right), \\ \widehat{\vartheta}_T &= \arg\min_{\vartheta} l_T(\vartheta,\gamma) \text{ s.t. } R\vartheta = \lambda_\infty = 0, \end{aligned} \tag{10}$$

and

$$\widehat{V}_T(\gamma) = \frac{1}{T}\sum_{j=-\tau_T}^{\tau_T}\sum_{t=\tau_T}^{T-\tau_T}\omega_j\nabla_\vartheta l_t\left(\widehat{\vartheta}_T,\gamma\right)\nabla_\vartheta l_{t+j}\left(\widehat{\vartheta}_T,\gamma\right)',\ \omega_j = 1 - \frac{j}{1+\tau_T}. \tag{11}$$

Now, given mild regularity assumptions controlling the smoothness of the likelihood, under the null

[3] 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:

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

of $\lambda_\infty = 0$,

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

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 $\sup_{\gamma\in\Gamma} K(\gamma)$ 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 $\lambda_\infty = 0$ (see Andrews (2001)). Hence, if the likelihood is known in closed form, and if both the continuous and the jump components of the model 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 interesting 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.

4 Long Time Span Jump Tests

4.1 Test of $\lambda_\infty = 0$ (no leverage effects)

From the above discussion, recall that tests based on high frequency observations over a finite time span are model free, but are not consistent against the alternative $\lambda_\infty > 0$. On the other hand, tests based on discrete observations over a long time span are consistent against $\lambda_\infty > 0$, but require correct specification of both the continuous and jump components, as well as knowledge of the transition density. This is because long time spans ensure test consistency against non-zero intensities (note that the intensity parameter cannot be identified on a given finite time span), while the use of high frequency observations does not require knowledge of the parametric specification of the diffusion. In order to have tests that are consistent against $\lambda_\infty > 0$, but still (almost) model free, we use functions of sample moments and rely on double in-fill and long-time span asymptotic approximations. The only (small) price to pay is that the drift component is assumed to be constant. This follows because the drift term can be ignored over a finite time span, while it has a non-negligible impact on asymptotic approximations over a long time spans.

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 $\Delta \to 0$. Proceed in two steps. First test for zero jump intensity ($\lambda_\infty = 0$). Then, if the null is rejected, test for path dependence ($\beta = 0$). In the first step, thus, interest lies in the following hypotheses:

$$H_0 : \lambda_\infty = 0$$

$$\begin{aligned} H_A = H_A^{(1)} \cup H_A^{(2)} &: \left(\lambda_\infty > 0 \text{ and } \mathrm{E}\left((Z_t - \mathrm{E}(Z_t))^3\right) \neq 0\right) \\ &\cup \left(\lambda_\infty > 0 \text{ and } \mathrm{E}\left((Z_t - \mathrm{E}(Z_t))^3\right) = 0\right). \end{aligned}$$

Notice that the alternative hypothesis is the union of two different alternatives, designed to allow for both symmetric and asymmetric jump size density. Let $Y_{k\Delta} = \ln X_{k\Delta} - \frac{\Delta}{T}\sum_{k=1}^{n} \ln X_{k\Delta}$, and $Y_{(k-1)\Delta} = \ln X_{(k-1)\Delta} - \frac{\Delta}{T}\sum_{k=2}^{n} \ln X_{(k-1)\Delta}$. Also, let:

$$\widehat{\lambda}_{T,\Delta} = \frac{1}{T}\sum_{k=2}^{n}\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3,$$

and define the statistic:

$$S_{T,\Delta} = \frac{T^{1/2}}{\Delta}\widehat{\lambda}_{T,\Delta}. \tag{12}$$

The asymptotic behavior of $S_{T,\Delta}$ is analyzed under the following set of assumptions.

Assumption A: (i) $\ln X_t$ is generated by *Eq.* (1) and V_t is defined in *Eq.* (2), (3), or (4). **(ii)** $\ln X_t$ is generated by *Eq.* (1) and V_t is defined in *Eq.* (5). For C a generic constant, (**iii**) $\mathrm{E}\left(|V_t|^k\right) \leq C$, $k \geq 3$, (**iv**) N_t satisfies *Eqs*. (6)-(8), and λ_t is either constant or it satisfies *Eq.* (9). (**v**) The jump size, Z_t, is independently and identically distributed, and $\mathrm{E}\left(|Z_t|^k\right) \leq C$, for $k \geq 6$.

Theorem 1: *Let Assumptions A(i) and A(iii)-(v) hold. Also, assume that as $n \to \infty$, $T \to \infty$ and $\Delta \to 0$.*
(i) Under H_0 :

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

with $\omega_0 = 15E\left(V_{k\Delta}^3\right) + 4\left(\mathrm{E}(V_{k,\Delta})\right)^3 - 12E(V_{k,\Delta})\,E\left(V_{k,\Delta}^2\right)$ and $S_{T,\Delta}$ defined as in Eq. (12).
(ii) Under $H_A^{(1)}$, there exists an $\varepsilon > 0$, such that:

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

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

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

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 asymmetric jumps, and diverges at the slower rate of $\frac{1}{\Delta}$ under the alternative of symmetric jumps. As shown in the Appendix, as $\Delta \to 0$ and $T \to \infty$, we have that $\widehat{\lambda}_{T,\Delta} \xrightarrow{p} \lambda_\infty \mathrm{E}\left(\left(Z_t - \mathrm{E}\left(Z_t\right)\right)^3\right)$. Now, if $\mathrm{E}\left(\left(Z_t - \mathrm{E}\left(Z_t\right)\right)^3\right) \neq 0$, then under $H_A^{(1)}$, the test has power against $\frac{\sqrt{T}}{\Delta}-$alternatives. Nevertheless, under H_0 the statistic has Pitman drift "only" against $\sqrt{T}-$alternatives. This is because the limiting distribution of the statistic under $H_A^{(1)}$ differs from that under H_0 both in terms of the location and scale. In fact, under $H_A^{(1)}$ the statistic has a mean (Pitman drift) of order $\sqrt{T}$, and has a standard deviation of order Δ^{-1}, while under H_0 the limiting distribution has mean zero and finite variance. On the other hand, if $\mathrm{E}\left(\left(Z_t - \mathrm{E}\left(Z_t\right)\right)^3\right) = 0$, then λ_∞ is not identified, and so under $H_A^{(2)}$ the Pitman drift is zero. Indeed, $\widehat{\lambda}_{T,\Delta} \xrightarrow{p} 0$ regardless of whether $\lambda_\infty = 0$ or $\lambda_\infty > 0$. Although it is not possible to distinguish between H_0 and $H_A^{(2)}$ based on the different locations of the limiting distribution (Pitman drift), it is possible to distinguish between H_0 and $H_A^{(2)}$ based on the different scales of the limiting distribution of $\frac{T^{1/2}}{\Delta}\widehat{\lambda}_{T,\Delta}$. This is because the order of magnitude of the variance of $\frac{T^{1/2}}{\Delta}\widehat{\lambda}_{T,\Delta}$ is larger when $\lambda_\infty > 0$ and $\mathrm{E}\left(\left(Z_t - \mathrm{E}\left(Z_t\right)\right)^3\right) = 0$ than when $\lambda_\infty = 0$. Broadly speaking, under H_0, $S_{T,\Delta} \xrightarrow{d} N\left(0, \omega_0\right)$, while under $H_A^{(2)}$, $\Delta S_{T,\Delta} \xrightarrow{d} N\left(0, \omega_1\right)$, with $\omega_1 \neq \omega_0$. This is what allows one to distinguish between H_0 and $H_A^{(2)}$.

If the moments of V_t were known, then an estimator of the variance which is consistent for the true variance under the null and bounded in probability under the alternative could be constructed; and consequently one could carry out inference on a simple $t-$statistic. However, spot volatilities are not generally observed, and hence the moments are not generally known. Heuristically, one may think of using $\widehat{\sigma}^2_{\lambda,T,\Delta} = \frac{1}{T\Delta^2}\sum_{k=1}^n \left(Y_{k\Delta} - \overline{Y}_\Delta\right)^6$ as an estimator of $\mathrm{var}\left(\frac{\sqrt{T}}{\Delta}\widehat{\lambda}_{T,\Delta}\right)$. However, while $\widehat{\sigma}_{\lambda,T,\Delta}$ is consistent for the "true"standard deviation under the null, it is order $O_p\left(\Delta^{-1}\right)$ under either of the alternatives. As a consequence, under $H_A^{(2)}$, $t_{\lambda,T,\Delta} = \frac{\frac{\sqrt{T}}{\Delta}\widehat{\lambda}_{T,\Delta}}{\widehat{\sigma}_{\lambda,T,\Delta}}$ would remain bounded in probability. Hence, $t_{\lambda,T,\Delta}$ does not have power against the alternative of jumps characterized by a symmetric distribution. Needless to say, if one rules out the possibility of symmetric jumps, then one could simply compare $t_{\lambda,T,\Delta}$ with standard normal critical values. However, in order to allow for the possibility of symmetric jumps, the statistic should not be rescaled, and hence inference should be based on the use of the bootstrap.

Finally, not that $\frac{1}{T}\sum_{k=2}^n \left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^4 \xrightarrow{p} \lambda_\infty \mathrm{E}\left(\left(Z_t - \mathrm{E}\left(Z_t\right)\right)^4\right)$, and hence a statistic based on the sample fourth moment has a well defined Pitman drift against $\sqrt{T}-$alternatives, regardless whether the jump size density is symmetric or not. We did not attempt to construct a statistic based on the sample fourth moment as, under the null hypothesis it could not have a limiting Gaussian distribution, because of the boundary issue and the impossibility of having a quadratic approximation around $\lambda_\infty = 0$.

4.2 Bootstrap Critical Values

Given that the variance is of a different order of magnitude under the null and under each alternative, the "standard" nonparametric bootstrap is not asymptotically valid. This issue arises because the variance of the bootstrap statistic mimics the sample variance. This implies that the bootstrap statistic is of order Δ^{-1} under the alternative. This is not be a problem under $H_A^{(1)}$, since the statistic is of order $\sqrt{T}\Delta^{-1}$, but is a problem under $H_A^{(2)}$, since the actual and bootstrap statistics would be of the same order. To ensure power against $H_A^{(2)}$, it suffices to ensure that the bootstrap statistic is of a smaller order than the actual statistic. This can be accomplished by resampling observations over a rougher grid, $\widetilde{\Delta}$, using the same time span, T.

Set the new discrete interval to be $\widetilde{\Delta}$, such that $\Delta/\widetilde{\Delta} \to 0$, and resample, with replacement, $\left(Y^*_{k\widetilde{\Delta}} - Y^*_{(k-1)\widetilde{\Delta}}, ..., Y^*_{\widetilde{n}\widetilde{\Delta}} - Y^*_{(\widetilde{n}-1)\widetilde{\Delta}}\right)$ from $\left(Y_{k\widetilde{\Delta}} - Y_{(k-1)\widetilde{\Delta}}, ..., Y_{\widetilde{n}\widetilde{\Delta}} - Y_{(\widetilde{n}-1)\widetilde{\Delta}}\right)$, where $\widetilde{n} = \frac{T}{\widetilde{\Delta}}$. Now, let:

$$\widetilde{\lambda}_{T,\widetilde{\Delta}} = \frac{1}{T}\sum_{k=2}^{\widetilde{n}} \left(Y_{k\widetilde{\Delta}} - Y_{(k-1)\widetilde{\Delta}}\right)^3,$$

and

$$\widetilde{\lambda}^*_{T,\widetilde{\Delta}} = \frac{1}{T}\sum_{k=2}^{\widetilde{n}} \left(Y^*_{k\widetilde{\Delta}} - Y^*_{(k-1)\widetilde{\Delta}}\right)^3.$$

Further, define the bootstrap statistic:

$$S^*_{T,\widetilde{\Delta}} = \frac{\sqrt{T}}{\widetilde{\Delta}}\left(\widetilde{\lambda}^*_{T,\widetilde{\Delta}} - \widetilde{\lambda}_{T,\widetilde{\Delta}}\right).$$

Finally, let $c^*_{\alpha,B,\Delta,\widetilde{\Delta}}$ and $c^*_{(1-\alpha),B,\Delta,\widetilde{\Delta}}$ be the $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ critical values of the empirical distribution of $S^*_{T,\widetilde{\Delta}}$, constructed using B bootstrap replications.

Theorem 2: *Let Assumptions A(i) and A(iii)-(v) hold. Also, assume that as* $n \to \infty$, $B \to \infty$, $T \to \infty$, $\Delta \to 0$, $\widetilde{\Delta} \to 0$ *and* $\Delta/\widetilde{\Delta} \to 0$.

(i) Under H_0 :

$$\lim_{T,B\to\infty,\Delta,\widetilde{\Delta}\to 0} \Pr\left(c^*_{\alpha/2,B,\Delta,\widetilde{\Delta}} \leq S_{T,\Delta} \leq c^*_{(1-\alpha/2),B,\Delta,\widetilde{\Delta}}\right) = 1-\alpha.$$

(ii) Under $H_A^{(1)} \cup H_A^{(2)}$:

$$\lim_{T,B\to\infty,\Delta,\widetilde{\Delta}\to 0} \Pr\left(c^*_{\alpha/2,B,\Delta,\widetilde{\Delta}} \leq S_{T,\Delta} \leq c^*_{(1-\alpha/2),B,\Delta,\widetilde{\Delta}}\right) = 0.$$

It is immediate to see that rejecting the null whenever $\frac{\sqrt{T}}{\Delta}\widehat{\lambda}_{T,\Delta} < c^*_{\alpha/2,B,\Delta,\widetilde{\Delta}}$ or $\frac{\sqrt{T}}{\Delta}\widehat{\lambda}_{T,\Delta} > c^*_{(1-\alpha/2),B,\Delta,\widetilde{\Delta}}$, and otherwise failing to reject, delivers a test with asymptotic size equal to α and asymptotic power equal to unity. Note that the bootstrap statistic is of P^*-probability order $\frac{1}{\widetilde{\Delta}}$ under both $H_A^{(1)}$ and $H_A^{(2)}$, while the actual statistic is of $P-$probability order $\frac{\sqrt{T}}{\Delta}$ under $H_A^{(1)}$ and $\frac{1}{\Delta}$ under $H_A^{(2)}$. Hence, the condition that $\frac{\Delta}{\widetilde{\Delta}} \to 0$ ensures unit asymptotic power under $H_A^{(2)}$. As the suggested statistics are 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 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 make the statistic robust to microstructure noise (as in the case of the realized pre-average power variation discussed in Podolskji and Vetter (2009b)). Future exploration of this issue is left to future research.

4.3 Test of $\lambda_\infty = 0$ (leverage effects)

The statements in Theorems 1 and 2 require absence of leverage effects. In particular, the results presented in these theorems rely on the fact that under the null of no jumps, returns are symmetrically distributed. More precisely, all results are derived under the assumption that $\mathrm{E}\Big(\big(Y_{k\Delta} - Y_{(k-1)\Delta}\big)^3\Big) = 0$, whenever there are no jumps. However, in the presence of leverage, if V_t is generated as in *Eq.* (5), $\mathrm{E}\left(\left(\int_{(k-1)\Delta}^{k\Delta} V_s^{1/2}\mathrm{d}W_{1,s}\right)^3\right) \neq 0$, and is instead of order Δ^2. For example, if V_t is generated by a square root process (i.e., $\mathrm{d}V_t = \kappa\left(\theta - V_t\right)\mathrm{d}t + \eta V_t^{1/2}\mathrm{d}W_{2,t}$), then $\mathrm{E}\Big(\big(Y_{k\Delta} - Y_{(k-1)\Delta}\big)^3\Big) = \lambda_\infty \mathrm{E}\left(Z_t - \mathrm{E}\left(Z_t\right)\right)^3\Delta + \frac{\eta\theta\rho}{2\kappa}\Delta^2$ (see Ait-Sahalia, Cacho-Diaz and Laeven (2013)). Although, the contribution to the third moment of the asymmetric jump component is of a larger order than that of the leverage component, inference based on the comparison of $S_{T,\Delta}$ with the bootstrap critical values $c^*_{\alpha,B,\Delta,\widetilde{\Delta}}$ and $c^*_{(1-\alpha),B,\Delta,\widetilde{\Delta}}$ will lead to the rejection of the null of no jumps, even if the null is true. This is established in the theorem below.

Theorem 3: *Let Assumptions A(ii)-(v) hold. Also, assume that as* $n \to \infty$, $T \to \infty$, $\Delta \to 0$, $\widetilde{\Delta} \to 0$ *and* $\Delta/\widetilde{\Delta} \to 0$. *Then, under both* H_0 *and* $H_A^{(1)} \cup H_A^{(2)}$:

$$\lim_{T,B\to\infty,\Delta,\widetilde{\Delta}\to 0} \Pr\left(c^*_{\alpha/2,B,\Delta,\widetilde{\Delta}} \le S_{T,\Delta} \le c^*_{(1-\alpha/2),B,\Delta,\widetilde{\Delta}}\right) = 0.$$

It follows that, in the presence of leverage, we always reject the null of no jumps, regardless as to whether it is true or false. To avoid spurious rejection due to the presence of leverage, use the following modified statistic:

$$\widetilde{S}_{T,\Delta} = \frac{1}{T^{1/2+\varepsilon}} S_{T,\Delta}, \tag{13}$$

with $\varepsilon > 0$, arbitrarily small.

Theorem 4: *Let Assumption A(ii)-(v) hold. Also, assume that as* $n \to \infty$, $T \to \infty$, $\Delta \to 0$, $\widetilde{\Delta} \to 0$, *and* $(T^{1/2+\varepsilon}\Delta)/\widetilde{\Delta} \to 0$.
(i) Under H_0 :

$$\lim_{T,B\to\infty,\Delta,\widetilde{\Delta}\to 0} \Pr\left(c^*_{\alpha/2,B,\Delta,\widetilde{\Delta}} \leq \widetilde{S}_{T,\Delta} \leq c^*_{(1-\alpha/2),B,\Delta,\widetilde{\Delta}}\right) = 1.$$

(ii) Under $H_A^{(1)} \cup H_A^{(2)}$:

$$\lim_{T,B\to\infty,\Delta,\widetilde{\Delta}\to 0} \Pr\left(c^*_{\alpha/2,B,\Delta,\widetilde{\Delta}} \leq \widetilde{S}_{T,\Delta} \leq c^*_{(1-\alpha/2),B,\Delta,\widetilde{\Delta}}\right) = 0.$$

It follows that inference based on the comparison of $\widetilde{S}_{T,\Delta}$ with the bootstrap critical values $c^*_{\alpha,B,\Delta,\widetilde{\Delta}}$ and $c^*_{(1-\alpha),B,\Delta,\widetilde{\Delta}}$ delivers a test with zero asymptotic size and unit asymptotic power. Needless to say, the statements in Theorem 4 are also valid when there is no leverage. However, in the this case tests should be based on $S_{T,\Delta}$, in order to maximize power.

4.4 Test of $\beta = 0$

If the null hypothesis of zero intensity is rejected, one can proceed to test the null of no self-excitation or path dependence. The null in this case is $\beta = 0$, and the alternative is $\beta > 0$, with β defined as in *Eq*. (9). As shown by Ait-Sahalia, Cacho-Diaz and Laeven (2013), β can be identified from the autocorrelation function. In particular, they show that given *Eq*. (9),

$$\begin{aligned} &\mathrm{E}\left(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)\left(Y_{(k+\tau)\Delta} - Y_{(k+\tau-1)\Delta}\right)\right) \\ &= \frac{\beta\lambda_\infty\left(2a-\beta\right)}{2\left(a-\beta\right)} \exp\left(-\left(a-\beta\right)\tau\right)\left(\mathrm{E}\left(Z\right)\right)^2 \Delta^2 + o\left(\Delta^2\right). \end{aligned} \tag{14}$$

Given that $\lambda_\infty > 0$, it follows that $\mathrm{E}\left(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)\left(Y_{(k+\tau)\Delta} - Y_{(k+\tau-1)\Delta}\right)\right) = 0$ if and only if $\beta = 0$. Our objective is to test the following hypotheses:

$$H_0 : \beta = 0$$
$$H_A : \beta > 0.$$

Define the statistic:

$$Z_{T,\Delta} = \max\left\{0, t_{\beta,T,\Delta}\right\},$$

where

$$t_{\beta,T,\Delta} = \frac{\sqrt{\frac{T}{\Delta}}\widehat{\tau}_{T,\Delta}}{\widehat{\sigma}_{\beta,T,\Delta}}, \tag{15}$$

with

$$\widehat{\tau}_{T,\Delta} = \frac{1}{T} \sum_{k=2}^{n-1} \left(Y_{k\Delta} - Y_{(k-1)\Delta} \right) \left(Y_{(k+1)\Delta} - Y_{k\Delta} \right) \tag{16}$$

and

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

From *Eq.* (14), 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.

Theorem 5: *Let Assumption A(i) or A(ii) and A(iii)-(v) hold, and let* λ_t *be the solution to Eq. (9). Also, assume that* $E(Z) \neq 0$, *and as* $n \to \infty$, $T \to \infty$ *and* $\Delta \to 0$.
(i) Under H_0 :

$$Z_{T,\Delta} \xrightarrow{d} \max\{0, Z\},$$

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

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

It follows that $Z_{T,\Delta}$ converges to an half-normal random variable under the null, and diverges at rate $\sqrt{\frac{\Delta}{T}}$ 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 and no self-excitation are both rejected, then one can proceed to estimate the full Hawkes diffusion using GMM, as in Ait-Sahalia, Cacho-Diaz and Laeven (2013). Of course, there is still a positive probability that the nulls have been falsely rejected. Because of this, one should consider carrying out somewhat conservative inference, using a “small” significance level.

Remark 3: In this paper, we only derive model free 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 $V^2_{k\Delta}$, which can be constructed using a finer grid of observations than used in the above tests, such as if there are 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 Reno (2012), under various settings.

Remark 4: In this section, we consider the case of 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 Ait-Sahalia, Cacho-Diaz and Laeven (2013)). For example, let:

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

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

5 Monte Carlo Results

In this section we carry out a set of experiments designed to evaluate the finite sample properties of (i) the test for the null of zero intensity, based on $S_{T,\Delta}$, as defined in *Eq.* (12) and, for the case of leverage, based on $\widetilde{S}_{T,\Delta}$, as defined in *Eq.* (13); (ii) the test for the null of no jump path dependence based on $Z_{T,\Delta}$, as discussed in the previous section; and (iii) the overall procedure according to which, if we reject the null of no jumps, according to either $S_{T,\Delta}$ or $\widetilde{S}_{T,\Delta}$, we then proceed to test the null of no path dependence, using $Z_{T,\Delta}$. We now outline the data generating processes (DGPs) used in the simulation experiments, namely:

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

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 have set $\mu = 0.5$, $\rho = \{0, -0.5\}$, $\kappa_v = 5$, $\theta_v = 0.04$, and $\zeta = 0.5$. Additionally, N_t satisfies the conditions in *Eqs.* (6)-(8) and for the jump size density we consider two cases, (a) Z_t an *iid* $N(0.5, 0.01)$ random variable, and (b) Z_t an exponential random variable with parameter equal to 5. The jump intensity evolves according to:

$$\lambda_t = \lambda_\infty + \beta \int_0^t \exp\left(-a\left(t-s\right)\right) \mathrm{d}N_s, \tag{17}$$

where $\lambda_\infty = \{1/20, 1/10, 1/5, 3/10, 2/5\}$, and $(a, \beta) = \{(0,0), (0.2, 0.1), (2,1), \text{ and } (5,4)\}$. Note that the case where $(a,\beta) = \{(0,0)\}$ is consistent with both the case of no jumps (i.e., $\lambda_\infty = 0$) and constant jump intensity (i.e., $\lambda_t = \lambda_\infty$).

We simulate observations using a Milstein discretization scheme, with discrete interval $h = 1/100$. For DGPs with $\rho = 0$, we sample the simulated observations using $\Delta = 1/60$ when constructing the test statistics and $\widetilde{\Delta} = 1/20$ when constructing bootstrap statistics. For DGPs with $\rho = -0.5$, we set $\Delta = 1/100$ and $\widetilde{\Delta} = 1/10$ for test statistics and bootstrap statistics, respectively. In all experiments, we perform 1000 Monte Carlo replications. Finally, recall that only the jump intensity test uses bootstrap critical values (see the statements in Theorems 2 and 4).

Conducting Monte Carlo experiments involving bootstrap estimators are always quite computationally demanding. In our experiments the computational burden is potentially even higher than usual, since we rely on a joint in-fill and long-span asymptotics, and since we need to control for the discretization error when simulating according to diffusion processes. To cope with this computational cost, we construct bootstrap critical values using the Warp-Speed approach of Giacomini, Politis and White (2013). This approach involves carrying out only one bootstrap replication for each simulated sample, and then averaging the bootstrap statistics over all Monte Carlo replications. Hence, the overall number of bootstrap replications is equal to the number of Monte Carlo replications, which is 1000 in our case. Under mild regularity conditions, the accuracy of the Warp-Speed bootstrap approaches is the same as that of the usual bootstrap, as both the sample size and the number of Monte Carlo replications go to infinity (see Corollary 5 in Giacomini, Politis and White (2013)).

The findings from our simulation studies are reported in Tables 1-4. Table 1 reports the empirical size, with the nominal size set at 10%. In Table 1 the first three columns consider DGPs with no leverage, while the last three columns consider DGPs with leverage. The first row displays the rejection frequencies for the test of $H_0 : \lambda_\infty = 0$ vs $H_A : \lambda_\infty > 0$, with the first three entries reporting results for the $S_{T,\Delta}$ and the last three corresponding to $\widetilde{S}_{T,\Delta}$. For $S_{T,\Delta}$ the empirical size is very accurate, and is effectively equal to the nominal size, for $T \geq 500$, while, as stated in Theorem 4, the rejection frequencies under the null for $\widetilde{S}_{T,\Delta}$ are identically equal to zero. The first and second quadrants in Table 1 contain rejection frequencies for $H_0 : \beta = 0$ vs $H_A : \beta > 0$, when data are generated with jumps characterized by constant intensity and jump sizes normally or exponentially distributed. Here, each row denotes a different intensity, from the lowest (1/20) in the top row to the highest (2/5) in the bottom row. The third and fourth quadrants report rejection frequencies for the full sequential procedure. In particular, they report how many times $H_0 : \lambda_\infty = 0$ vs $H_A : \lambda_\infty > 0$ has been rejected and how many times $H_0 : \beta = 0$ vs $H_A : \beta > 0$ has not been rejected. Overall, the empirical size is quite close to the nominal, though slightly smaller for the rows characterized by a large intensity parameter.

Table 2 contains the results of power experiments using $S_{T,\Delta}$ (non-leverage case), and using $\widetilde{S}_{T,\Delta}$ (leverage case). Panels A and B report rejection frequencies for DGPs generated with Poisson jumps characterized by constant intensity ranging from 1/20 (top row in each quadrant) to 2/5 (bottom row in each quadrant). The empirical power is essentially unity, even for the lowest intensity, and hence the test is powerful even in the presence of relatively few jumps. This is true also for the case of normal (symmetric) jumps, despite of the fact that the Pitman drift is zero, so that rejections are due only to the different order of magnitude of the variance. Interestingly, when $\Delta/\widetilde{\Delta}$ is small enough, the power in the leverage cases is as high as in the non-leverage case. Panels C and D report rejection frequencies in the case of no-leverage and self-exciting jumps (i.e., when λ_t is generated as in *Eq.* (17)), with exponential and normal jump densities, respectively. Panels E and F report analogous results for the leverage case. For cases where the mean intensity is low, rejection frequencies in the leverage case (i.e., tests based on $\widetilde{S}_{T,\Delta}$) are now slightly lower than corresponding rejection frequencies in the non-leverage case (i.e., tests based on $\widetilde{S}_{T,\Delta}$). However, for $\lambda_\infty \geq 1/5$, rejection frequencies are rather close to unity.

Table 3 contains results from power experiments (i.e., $H_0 : \beta = 0$ vs $H_A : \beta > 0$), under the maintained assumption that $\lambda_\infty > 0$. Note that for the test for no-path dependent intensity, we use the same statistic $\max\{0, t_{\beta,T,\Delta}\}$, as defined in *Eq.* (15), regardless of the presence of leverage or not. However, for the sake of completeness, we still report result for the no-leverage case in Panels A and B, and for the leverage case in Panels C and D. From *Eqs.* (14) and (17) is immediate to see that the smaller is a and the smaller is $(a - \beta)$, the higher is the level of self-excitation. For example, when $a = 0.1$ and $\beta = 0.2$, rejection frequencies are above 0.9, regardless of mean intensity. However, the case where $a = 0.1$ and $\beta = 0.2$ imply a somewhat implausibly high level of path dependence. For an intermediate degree of self-excitation, we consider $a = 2$ and $\beta = 1$. In this case, the rejection frequencies are reasonably high, from 0.65 to 0.80, when there are "enough" jumps (i.e., when $\lambda_\infty \geq 3/10$). Finally, in the case of low self-excitation (i.e., $a = 5, \beta = 4$) the power is slightly below 0.50, even for the highest mean intensity.

Table 4 summarizes experimental findings based on implementation of the full (sequential) procedure. Namely, whenever we reject $H_0 : \lambda_\infty = 0$ vs $\lambda_\infty > 0$, we proceed to test $H_0 : \beta = 0$ vs $H_A : \beta > 0$. Entries in the table denote that rejection frequencies indicating that *both* null hypotheses are rejected (sequentially). As the power of the jump intensity test against path-dependent jumps is very close to 1 (see Panels C-F in Table 2), it is not surprising to note that the entries in Table 4 are very close to those in Table 3.

In summary, the test for zero jump intensity has excellent empirical size and power across all cases. On the other hand, the test for no path dependence and the sequential procedure have the very good empirical size but good power only when jumps are frequent enough and the degree of self-excitation is not too low.

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. Hence, consistent estimation of jump diffusions requires consistent pretesting for the null of zero jump intensity. Currently available tests, which are based on high frequency observations over a finite time-span, are model free but are not consistent. On the other hand, tests based on discrete observations over a long time-span are consistent, but require full specification of the model as well as knowledge of a closed form expression for the transition density. This paper introduces novel (almost) model free tests which are consistent against the alternative of positive intensity. They are based on sample third moments, and make use of high frequency observations over a long time-span. Inference is based on m out n type bootstrap critical values, whose first order validity is established. A "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 Ait-Sahalia, Cacho-Diaz and Laeven (2013). The finite sample behavior of the suggested statistics is studied via Monte Carlo experimentation. The jump tests exhibit empirical size very close to nominal and empirical power close to unity. The self excitement test likewise has very good size and good power properties, whenever there are “enough” jumps and the degree of self excitation is not “too weak”.

7 Appendix

Proof of Theorem 1:

(i) Under H_0 $N_t = 0$, since the drift term is constant,

$$
\begin{aligned}
& S_{T,\Delta} \\
&= \frac{1}{\sqrt{T}\Delta}\sum_{k=2}^{n}\left(\left(\ln X_{k\Delta} - \frac{\Delta}{T}\sum_{k=1}^{n}\ln X_{k\Delta}\right) - \left(\ln X_{(k-1)\Delta} - \frac{\Delta}{T}\sum_{k=2}^{n}\ln X_{(k-1)\Delta}\right)\right)^3 \\
&= \frac{1}{\sqrt{T}\Delta}\sum_{k=2}^{n}\left(\int_{(k-1)\Delta}^{k\Delta}\sqrt{V_s}\mathrm{d}W_{1,s} - \frac{\Delta}{T}\sum_{k=1}^{n}\int_{(k-1)\Delta}^{k\Delta}\sqrt{V_s}\mathrm{d}W_{1,s}\right)^3 (1+o_p(1)) \\
&= \frac{1}{\sqrt{T}\Delta}\sum_{k=2}^{n}\left(\Delta^{1/2}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)^3 (1+o_p(1)) \\
&- \frac{\sqrt{T}}{\Delta^2}\left(\frac{\Delta}{T}\sum_{k=2}^{n}\Delta^{1/2}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)^3 (1+o_p(1)) \\
&- \frac{2}{\sqrt{T}\Delta}\sum_{k=2}^{n}\Delta V_{(k-1)\Delta}\epsilon_{k\Delta}^2\left(\frac{\Delta}{T}\sum_{k=1}^{n}\Delta^{1/2}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)(1+o_p(1)) \\
&+ \frac{2}{\sqrt{T}\Delta}\sum_{k=2}^{n}\Delta^{1/2}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\left(\frac{\Delta}{T}\sum_{k=2}^{n}\Delta^{1/2}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)^2 (1+o_p(1)) \\
&= I_{T,\Delta} + II_{T,\Delta} + III_{T,\Delta} + IV_{T,\Delta},
\end{aligned} \tag{18}
$$

where $\epsilon_{k\Delta}$ is an *iid* $\mathrm{N}(0,1)$ random variable, and where the $o_p(1)$ terms denote terms approaching zero as $\Delta \to 0$, uniformly in T.

Let $\mathcal{F}_{(k-1)\Delta} = \sigma\left(V_\Delta, ..., V_{(k-1)\Delta}\right)$ and note that under A(i), $\mathrm{E}\Big(V_{(k-1)\Delta}^{3/2}\epsilon_{k\Delta}^3 \left|\mathcal{F}_{(k-1)\Delta}\right.\Big) = V_{(k-1)\Delta}^{3/2}\mathrm{E}\big(\epsilon_{k\Delta}^3 \left|\mathcal{F}_{(k-1)\Delta}\right.\big) = 0$, and hence $V_{(k-1)\Delta}^{3/2}\epsilon_{k\Delta}^3$ is a martingale difference sequence. It follows that,

$$
\begin{aligned}
& \mathrm{var}\left(\frac{1}{\sqrt{T}\Delta}\sum_{k=1}^{n}\left(\Delta^{1/2}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)^3\right) \\
&= \frac{1}{T\Delta^2}\frac{T}{\Delta}\Delta^3\mathrm{E}\left(V_{(k-1)\Delta}^3\epsilon_{k\Delta}^6\right) = 15\mathrm{E}\left(V_{(k-1)\Delta}^3\right).
\end{aligned} \tag{19}
$$

We first show that $II_{T,\Delta}$ and $IV_{T,\Delta}$ are $o_p(1)$.

$$
\begin{aligned}
II_{T,\Delta} &= \frac{2\sqrt{T}}{\Delta^2}\frac{\Delta^3}{T^{3/2}}\left(\sqrt{\frac{\Delta}{T}}\sum_{k=1}^{n}\left(\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)^3\right)(1+o_p(1)) \\
&= \frac{\Delta}{T}O_p(1) = o_p(1),
\end{aligned}
$$

since, recalling *Eq.* (19), $\sqrt{\frac{\Delta}{T}}\sum_{k=1}^{n}\left(\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)^{3}$ satisfies a CLT for martingale difference sequences. Additionally,

$$\begin{aligned}
IV_{T,\Delta} &= 2\frac{\Delta}{T}\sqrt{\frac{\Delta}{T}}\sum_{k=1}^{n}\left(\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)\left(\sqrt{\frac{\Delta}{T}}\sum_{k=1}^{n}\left(\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right)\right)^{2}(1+o_p(1)) \\
&= \frac{\Delta}{T}O_p\left(1\right) = o_p(1).
\end{aligned}$$

Now, $\operatorname{var}\left(I_{T,\Delta}\right) = 15\mathrm{E}\left(V_{(k-1)\Delta}^{3}\right)$, given *Eq.* (19), and

$$\begin{aligned}
\operatorname{var}\left(III_{T,\Delta}\right) &= 4\left(\mathrm{E}\left(V_{k,\Delta}\right)\right)^{2}\operatorname{var}\left(\sqrt{\frac{\Delta}{T}}\sum_{k=1}^{n}\sqrt{V_{(k-1)\Delta}}\epsilon_{k\Delta}\right) \\
&= 4\left(\mathrm{E}\left(V_{k,\Delta}\right)\right)^{2}\mathrm{E}\left(V_{k,\Delta}\right) = 4\left(\mathrm{E}\left(V_{k,\Delta}\right)\right)^{3},
\end{aligned}$$

with

$$\operatorname{cov}\left(I_{T,\Delta}, III_{T,\Delta}\right) = -4\mathrm{E}\left(V_{k,\Delta}\right)\mathrm{E}\left(V_{k,\Delta}^{2}\right)\mathrm{E}\left(\epsilon_{k\Delta}^{4}\right).$$

Thus,

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

with $\omega_0 = 15\mathrm{E}\left(V_{(k-1)\Delta}^{3}\right) + 4\left(\mathrm{E}\left(V_{k,\Delta}\right)\right)^{3} - 12\mathrm{E}\left(V_{k,\Delta}\right)\mathrm{E}\left(V_{k,\Delta}^{2}\right)$. The statement in (i) then follows.

(ii) Under $H_A^{(1)}$, there are additional jump components, including:

$$\frac{1}{\sqrt{T}}\sum_{k=1}^{n}\frac{1}{\Delta}\left(Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right) - \frac{\Delta}{T}\sum_{k=1}^{n}Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right)\right)^{3},$$

plus related cross-terms. Now,

$$\begin{aligned}
&\frac{\Delta}{T}\sum_{k=1}^{n}\frac{1}{\Delta}\left(Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right) - \frac{\Delta}{T}\sum_{k=1}^{n}Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right)\right)^{3} \\
&\xrightarrow{pr} \lambda\mathrm{E}\left(Z - \mathrm{E}\left(Z\right)\right)^{3}.
\end{aligned}$$

Thus, whenever $\mathrm{E}\left(Z - \mathrm{E}\left(Z\right)\right)^{3} \neq 0$, $S_{T,\Delta}$ is of probability order $\frac{\sqrt{T}}{\Delta}$, and $\lim_{T\to\infty,\Delta\to 0}\Pr\left(\frac{\Delta}{\sqrt{T}}\left|S_{T,\Delta}\right| > \varepsilon\right) = 1$. The statement in (ii) then follows.

(iii) Under $H_A^{(2)}$, by the law of large numbers,

$$\begin{aligned}
&\frac{\Delta}{T}\sum_{k=1}^{n}\frac{1}{\Delta}\left(Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right) - \frac{\Delta}{T}\sum_{k=1}^{n}Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right)\right)^{3} \\
&\xrightarrow{pr} 0.
\end{aligned}$$

Since $\mathrm{E}\left(Z - \mathrm{E}\left(Z\right)\right)^3 = 0$, and given the central limit theorem,

$$\sqrt{\frac{\Delta}{T}}\sum_{k=1}^{n}\frac{1}{\Delta}\left(Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right) - \frac{\Delta}{T}\sum_{k=1}^{n}Z_{(k-1)\Delta}\left(N_{k\Delta} - N_{(k-1)\Delta}\right)\right)^3$$
$$= O_p(1).$$

Moreover, if $\beta = 0$ (no path dependent intensity), then:

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

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 $O\left(\frac{1}{\Delta^2}\right)$. Hence, $S_{T,\Delta}$ is of probability order Δ^{-1} and the statement in (iii) follows.

Proof of Theorem 2: Hereafter, let E^* and var^* denote the mean and variance operators under the bootstrap probability measure P^*, conditional on the sample, and let d^* denote convergence in distribution under P^*.

(i) Note that:

$$\mathrm{E}^*\left(S^*_{T,\widetilde{\Delta}}\right) = \frac{1}{\sqrt{T}\widetilde{\Delta}}\sum_{k=1}^{\widetilde{n}}\mathrm{E}^*\left(\left(Y^*_{k\widetilde{\Delta}} - Y^*_{(k-1)\widetilde{\Delta}}\right)^3 - \left(Y_{k\widetilde{\Delta}} - Y_{(k-1)\widetilde{\Delta}}\right)^3\right).$$

Now, since:

$$\mathrm{E}^*\left(Y^*_{k\widetilde{\Delta}} - Y^*_{(k-1)\widetilde{\Delta}}\right)^3 = \frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}\left(Y_{k\widetilde{\Delta}} - Y_{(k-1)\widetilde{\Delta}}\right)^3,$$

it follows immediately that $\mathrm{E}^*\left(S^*_{T,\widetilde{\Delta}}\right) = 0$. Now, consider $\mathrm{var}^*\left(S^*_{T,\widetilde{\Delta}}\right)$. First note that, by the same argument as that used in the proof of Theorem 1, part (i),

$$\begin{aligned}
&S^*_{T,\widetilde{\Delta}}\\
&= \frac{1}{\sqrt{T}\widetilde{\Delta}}\sum_{k=1}^{\widetilde{n}}\left(V^{*3/2}_{(k-1)\widetilde{\Delta}}\widetilde{\Delta}^{3/2}\epsilon^{*3}_{k\widetilde{\Delta}} - 2V^*_{(k-1)\widetilde{\Delta}}\widetilde{\Delta}\epsilon^{*2}_{k\widetilde{\Delta}}\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V^{*1/2}_{(k-1)\widetilde{\Delta}}\widetilde{\Delta}^{1/2}\epsilon^*_{k\widetilde{\Delta}}\right)
\end{aligned}$$

$$
-\frac{1}{\sqrt{T}\widetilde{\Delta}}\sum_{k=1}^{\widetilde{n}}\left(V_{(k-1)\widetilde{\Delta}}^{3/2}\widetilde{\Delta}^{3/2}\epsilon_{k\widetilde{\Delta}}^{3}-2V_{(k-1)\widetilde{\Delta}}\widetilde{\Delta}\epsilon_{k\widetilde{\Delta}}^{2}\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{1/2}\widetilde{\Delta}^{1/2}\epsilon_{k\widetilde{\Delta}}\right)+o_p^*(1)+o_p(1)
$$

$$
=\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}\left(V_{(k-1)\widetilde{\Delta}}^{*3/2}\epsilon_{k\widetilde{\Delta}}^{*3}-V_{(k-1)\widetilde{\Delta}}^{3/2}\epsilon_{k\widetilde{\Delta}}^{3}\right)-2\left(\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*}\epsilon_{k\widetilde{\Delta}}^{*2}\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*1/2}\epsilon_{k\widetilde{\Delta}}^{*}\right.
$$

$$
\left.-\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{1/2}\epsilon_{k\widetilde{\Delta}}\right)+o_p^*(1)+o_p(1)
$$

$$
=\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}\left(V_{(k-1)\widetilde{\Delta}}^{*3/2}\epsilon_{k\widetilde{\Delta}}^{*3}-V_{(k-1)\widetilde{\Delta}}^{3/2}\epsilon_{k\widetilde{\Delta}}^{3}\right)-2\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\left(\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*1/2}\epsilon_{k\widetilde{\Delta}}^{*}\right.
$$

$$
\left.-\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{1/2}\epsilon_{k\widetilde{\Delta}}\right)+o_p^*(1)+o_p(1).
$$

Hence,

$$
\begin{aligned}
&\operatorname{var}^*\left(S_{T,\widetilde{\Delta}}^*\right)\\
&=\operatorname{var}^*\left(\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*3/2}\epsilon_{k\widetilde{\Delta}}^{*3}\right)\\
&+9\left(\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\right)^2\operatorname{var}^*\left(\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*1/2}\epsilon_{k\widetilde{\Delta}}^{*}\right)\\
&-6\left(\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\right)\operatorname{cov}^*\left(\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*3/2}\epsilon_{k\widetilde{\Delta}}^{*3},\sqrt{\frac{\widetilde{\Delta}}{T}}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{*1/2}\epsilon_{k\widetilde{\Delta}}^{*}\right)
\end{aligned}
$$

and so,

$$
\begin{aligned}
&\operatorname{var}^*\left(S_{T,\widetilde{\Delta}}^*\right)\\
&=\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{3}\epsilon_{k\widetilde{\Delta}}^{6}+4\left(\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\right)^2\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\\
&-4\left(\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}\epsilon_{k\widetilde{\Delta}}^{2}\right)\frac{\widetilde{\Delta}}{T}\sum_{k=1}^{\widetilde{n}}V_{(k-1)\widetilde{\Delta}}^{2}\epsilon_{k\widetilde{\Delta}}^{4}\\
&=\operatorname{var}\left(S_{T,\widetilde{\Delta}}\right)+o_p(1).
\end{aligned}
$$

As $S_{T,\widetilde{\Delta}}^*\xrightarrow{d^*}N\left(0,\operatorname{var}^*\left(S_{T,\widetilde{\Delta}}^*\right)\right)$, the statement in (i) follows.

(ii) $S^*_{T,\Delta}$ now contains the following additional term:

$$J^*_{T,\widetilde{\Delta}} = \frac{1}{\sqrt{T}\widetilde{\Delta}} \sum_{k=1}^{\widetilde{n}} \left(Z^{*3}_{(k-1)\widetilde{\Delta}} \left(N^*_{(k-1)\widetilde{\Delta}} - N^*_{(k-1)\widetilde{\Delta}} \right)^3 - Z^3_{(k-1)\widetilde{\Delta}} \left(N_{(k-1)\widetilde{\Delta}} - N_{(k-1)\widetilde{\Delta}} \right)^3 \right),$$

plus additional cross product terms, which cannot be of larger P^*-order than $J^*_{T,\widetilde{\Delta}}$. Now, $\mathrm{E}^*\left(J^*_{T,\widetilde{\Delta}}\right) = 0$, and

$$\begin{aligned}
&\mathrm{var}^*\left(J^*_{T,\widetilde{\Delta}}\right) \\
&= \frac{1}{\widetilde{\Delta}^2} \frac{\widetilde{\Delta}}{T} \sum_{k=1}^{\widetilde{n}} \frac{1}{\widetilde{\Delta}} \left(Z^6_{(k-1)\widetilde{\Delta}} \left(N_{(k-1)\widetilde{\Delta}} - N_{(k-1)\widetilde{\Delta}} \right)^6 \right) \\
&= \frac{1}{\widetilde{\Delta}^2} O_p(1),
\end{aligned}$$

and so $S^*_{T,\widetilde{\Delta}}$ is of P^*-order $\frac{1}{\widetilde{\Delta}}$. Recalling that $S_{T,\widetilde{\Delta}}$ is of $P-$order $\frac{1}{\Delta}$, with $\Delta/\widetilde{\Delta} \to 0$, the statement in (ii) follows.

Proof of Theorem 3: Suppose that H_0 is true, and so $\lambda_\infty = 0$. In this case: $\mathrm{E}\Big(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3\Big) = O\left(\Delta^2\right)$, and by the same argument as in the proof of Theorem 1,

$$\begin{aligned}
&\frac{1}{T} \sum_{k=1}^{n} \left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3 \\
&= \frac{1}{T} \sum_{k=1}^{n} \left(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3 - \mathrm{E}\left(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3 \right) \right) + \mathrm{E}\left(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3 \right) \\
&= O_p\left(\frac{\Delta}{\sqrt{T}}\right) + \mathrm{E}\left(\left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3 \right) \Delta^{-1} = O_p\left(\frac{\Delta}{\sqrt{T}}\right) + O(\Delta).
\end{aligned}$$

Hence,

$$S_{T,\Delta} = \frac{1}{\sqrt{T}\Delta} \sum_{k=1}^{n} \left(Y_{k\Delta} - Y_{(k-1)\Delta}\right)^3 = O_p(1) + O(\sqrt{T}).$$

On the other hand, it is immediate to see from the proof of Theorem 2, part (i), that $\mathrm{E}\left(S^*_{T,\widetilde{\Delta}}\right) = 0$, regardless of the presence of leverage. This is because the mean of the bootstrap statistic is always zero. Hence, the comparison of $S_{T,\Delta}$ with the critical values of $S^*_{T,\widetilde{\Delta}}$ will lead to a rejection of the null, with probability approaching one.

Proof of Theorem 4:

(i) Under H_0, $S_{T,\Delta} = O_p(\sqrt{T})$, so that $\widetilde{S}_{T,\Delta} = O_p(T^{-\varepsilon}) = o_p(1)$. As $S^*_{T,\widetilde{\Delta}}$ has a well defined, zero mean, normal limiting distribution (regardless the presence of leverage), the statement in (i) follows.

(ii) Under $H_A^{(1)}$, $S_{T,\Delta}$ is of probability order $\frac{\sqrt{T}}{\Delta}$, and so $\widetilde{S}_{T,\Delta}$ is of probability order $\frac{1}{T^\varepsilon \Delta}$. Under

$H_A^{(2)}$, $S_{T,\Delta}$ is of probability order $\frac{1}{\Delta}$, and so $\widetilde{S}_{T,\Delta}$ is of probability order $\frac{1}{T^{1/2+\varepsilon}\Delta}$. Now, from the proof of Theorem 2, we have that under both $H_A^{(1)}$ and $H_A^{(2)}$, $S^*_{T,\widetilde{\Delta}}$ diverges at rate $\frac{1}{\widetilde{\Delta}}$. As $\frac{T^{1/2+\varepsilon}\Delta}{\widetilde{\Delta}} \to 0$, the statement follows.

Proof of Theorem 5:

(i) Recall *Eq.* (16), which can be written as follows:

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

Let $\overline{Z}_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right) = Z_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right) - \mathrm{E}\left(Z\right)\mathrm{E}\left(\lambda\right)\Delta$, and note that under the null of constant intensity,

$$\begin{aligned}
&\mathrm{var}\left(\frac{1}{\sqrt{T\Delta}}\sum_{k=1}^{n-1}\overline{Z}_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right)\overline{Z}_{(k-1)\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)\right)\\
&= \Delta^{-2}\mathrm{E}\left(\left(\overline{Z}_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right)\right)^2\left(\overline{Z}_{(k-1)\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)\right)^2\right)\\
&= \Delta^{-2}\mathrm{E}\left(\left(\overline{Z}_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right)\right)^2\right)\mathrm{E}\left(\left(\overline{Z}_{(k-1)\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)\right)^2\right)\\
&= \left(\mathrm{E}\left(\lambda\right)\mathrm{E}\left(\left(Z-\mathrm{E}\left(Z\right)\right)^2\right)\right)^2,
\end{aligned}$$

$$\begin{aligned}
&\mathrm{var}\left(\frac{1}{\sqrt{T\Delta}}\sum_{k=1}^{n-1}\Delta V_{(k-1)\Delta}^{1/2}V_{k\Delta}^{1/2}\epsilon_{k\Delta}\epsilon_{(k+1)\Delta}\right)\\
&= \mathrm{E}\left(V_{(k-1)\Delta}V_{k\Delta}\right),
\end{aligned}$$

and

$$\mathrm{cov}\left(\frac{1}{\sqrt{T\Delta}}\sum_{k=1}^{n-1}\overline{Z}_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right), \frac{1}{\sqrt{T\Delta}}\sum_{k=1}^{n-1}\Delta^{1/2}V_{(k-1)\Delta}^{1/2}\right) = 0.$$

Now, consider $\widehat{\sigma}^2_{\beta,T,\Delta}$, the expression for which can be written as follows:

$$\begin{aligned}
\widehat{\sigma}^2_{\beta,T,\Delta} &= \frac{\Delta}{T}\sum_{k=1}^{n-1} V^2_{(k-1)\Delta} V^2_{k\Delta}\epsilon^2_{k\Delta}\epsilon^2_{(k+1)\Delta} \\
&+ \frac{\Delta}{T}\sum_{k=1}^{n-1} \Delta^{-2} Z^2_{(k-1)\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)^2 Z^2_{k\Delta}\left(N_{(k-1)\Delta}-N_{k\Delta}\right)^2 + o_p(1) \\
&= \left(\mathrm{E}\left(V^2_{(k-1)\Delta}\epsilon^2_{k\Delta}\right)\right)^2 + \left(\mathrm{E}\left(\lambda\right)\mathrm{E}\left(\left(Z-\mathrm{E}\left(Z\right)\right)^2\right)\right)^2 + o_p(1).
\end{aligned}$$

The statement then follows immediately from the central limit theorem for *iid* random variables and from the continuous mapping theorem.
(ii) Note also that $\widehat{\tau}_{T,\Delta}$ can be written as follows:

$$\begin{aligned}
&\widehat{\tau}_{T,\Delta} \\
&= \Delta\frac{\Delta}{T}\sum_{k=1}^{n-1}\Delta^{-2}\overline{Z}_{(k-1)\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)\overline{Z}_{k\Delta}\left(N_{(k+1)\Delta}-N_{k\Delta}\right) + o_p(1),
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\Delta}{T}\sum_{k=1}^{n-1}\Delta^{-2}\overline{Z}_{(k-1)\Delta}\overline{Z}_{k\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)\left(N_{(k+1)\Delta}-N_{k\Delta}\right) \\
&\xrightarrow{p} \frac{\beta\lambda\left(2a-\beta\right)}{2\left(a-\beta\right)}\exp\left(-\left(a-\beta\right)\tau\right)\mathrm{E}\left(Z\right)^2 \\
&> 0.
\end{aligned}$$

The statement in the theorem follows by noting that:

$$\frac{\Delta}{T}\sum_{k=1}^{n-1}\Delta^{-2}Z^2_{(k-1)\Delta}\left(N_{k\Delta}-N_{(k-1)\Delta}\right)^2 Z^2_{k\Delta}\left(N_{(k-1)\Delta}-N_{k\Delta}\right)^2 = O_p(1)$$

under both hypotheses, given that for $\beta < a$, the time dependence of jumps declines at an exponential rate.

8 Reference

Ait-Sahalia, Y., J. Cacho-Diaz and R. Laeven (2013). Modeling Financial Contagion Using Mutually Exciting Jump Processes. Working Paper, Princeton University.

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

Ait-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.

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. Reno (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, "Predictive Density Construction and Testing with Multiple Possibly Misspecified Diffusion Models", *Journal of Econometrics,* 161, 304-324, 2011.

Corsi, F., D. Pirino and R. Reno (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.

Giacomini, R., D. Politis and H. White (2013). A Warp-Speed Method for Conducting Monte Carlo Experiments Involving Bootstrap Estimators. *Econometric Theory,* 29, 567-589.

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

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.

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.

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.

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: Experimental Results - Empirical Size *
Data Generated According to a Stochastic Volatility Process

No Leverage in DGP			Leverage in DGP		
T=300	T=500	T=1000	T=300	T=500	T=1000
Jump Test					
0.110	0.104	0.102	0.000	0.000	0.000
Self Excitement Test *Jumps Generated Using Exponential Dist.*					
0.119	0.102	0.103	0.111	0.108	0.117
0.090	0.100	0.088	0.111	0.116	0.118
0.080	0.076	0.081	0.088	0.075	0.083
0.080	0.081	0.070	0.083	0.072	0.074
0.058	0.067	0.067	0.072	0.064	0.066
Self Excitement Test *Jumps Generated Using Normal Dist.*					
0.117	0.101	0.101	0.111	0.111	0.110
0.086	0.110	0.100	0.109	0.116	0.124
0.090	0.094	0.102	0.096	0.092	0.097
0.097	0.102	0.086	0.088	0.094	0.100
0.078	0.080	0.088	0.081	0.083	0.080
Sequential Jump and Self Excitement Test *Jumps Generated Using Exponential Dist.*					
0.111	0.102	0.103	0.107	0.099	0.114
0.090	0.100	0.088	0.110	0.116	0.118
0.080	0.076	0.081	0.088	0.075	0.083
0.080	0.081	0.070	0.083	0.072	0.074
0.058	0.067	0.067	0.072	0.064	0.066
Sequential Jump and Self Excitement Test *Jumps Generated Using Normal Dist.*					
0.114	0.101	0.101	0.100	0.105	0.107
0.086	0.110	0.100	0.109	0.116	0.124
0.090	0.094	0.102	0.096	0.092	0.097
0.097	0.102	0.086	0.088	0.094	0.100
0.078	0.080	0.088	0.081	0.083	0.080

* Notes: Entries denote rejection frequencies based on comparing $S_{T,\Delta}$ with 10% critical values calculated using the bootstrap (*Jump Test*) and based on comparing $Z_{T,\Delta}$ with 10% critical values from the normal distribution (*Self Excitement Test*). Additionally, results are reported for the sequential test whereby $Z_{T,\Delta}$ is constructed whenever implementation of $S_{T,\Delta}$ results in a rejection of the null of no jumps. Results are reported in "blocks" of 5 rows. Each of these 5 rows of entries corresponds to a different jump intensity. As discussed above, "average jump arrival times" are assumed to be every *{20 days, 10 days, 5 days, 10/3 days, 5/2 days}*, and the rows correspond to these arrivals, in order from least frequent to most frequent. All experiments are based on 1,000 Monte Carlo iterations. For complete details see above.

Table 2: Experimental Results - Jump Test Empirical Power *
Data Generated According to a Stochastic Volatility Process

T=300	T=500	T=1000	T=300	T=500	T=1000
Panel A: No Leverage, No Self-Excitement in DGP					
Exponential Dist.			Normal Dist.		
0.938	0.991	1.000	0.973	1.000	1.000
0.999	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
Panel B: Leverage, No Self-Excitement in DGP					
Exponential Dist.			Normal Dist.		
0.933	0.960	0.976	0.903	0.931	0.985
0.981	0.999	1.000	0.994	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000
1.000	1.000	1.000	1.000	1.000	1.000

$a = 0.1$,$\beta = 0.2$			$a = 2$,$\beta = 1$			$a = 5$,$\beta = 4$		
T=300	T=500	T=1000	T=300	T=500	T=1000	T=300	T=500	T=1000
Panel C: No Leverage, Self-Excitement in DGP, Jumps Generated Using Exponential Dist.								
0.917	0.939	0.958	0.545	0.511	0.797	0.811	0.917	1.000
0.975	0.995	1.000	0.682	0.865	0.996	0.921	0.978	1.000
0.995	1.000	1.000	0.993	0.999	1.000	0.998	1.000	1.000
1.000	1.000	1.000	0.999	1.000	0.999	1.000	1.000	1.000
1.000	1.000	1.000	1.000	0.999	1.000	0.999	1.000	0.999
Panel D: No Leverage, Self-Excitement in DGP, Jumps Generated Using Normal Dist.								
0.952	0.978	0.983	0.919	0.986	0.995	0.974	0.998	0.995
0.984	0.983	0.972	0.986	0.996	0.985	0.988	0.989	0.992
0.986	0.969	0.966	0.980	0.983	0.955	0.969	0.974	0.955
0.973	0.968	0.960	0.964	0.960	0.929	0.956	0.952	0.941
0.952	0.952	0.968	0.952	0.936	0.916	0.963	0.940	0.903
Panel E: Leverage, Self-Excitement in DGP, Jumps Generated Using Exponential Dist.								
0.906	0.917	0.944	0.345	0.261	0.288	0.370	0.812	0.729
0.974	0.987	0.999	0.458	0.471	0.471	0.725	0.699	0.509
0.995	0.999	1.000	0.767	0.757	0.720	0.802	0.834	0.956
1.000	1.000	1.000	0.941	0.953	0.964	0.976	0.943	1.000
1.000	1.000	1.000	0.998	0.997	0.998	0.996	0.999	1.000
Panel F: Leverage, Self-Excitement in DGP, Jumps Generated Using Normal Dist.								
0.948	0.939	0.929	0.592	0.668	0.697	0.784	0.921	0.949
0.974	0.963	0.957	0.915	0.939	0.967	0.985	0.988	0.978
0.966	0.940	0.899	0.964	0.963	0.932	0.981	0.965	0.915
0.957	0.930	0.908	0.950	0.937	0.903	0.960	0.951	0.887
0.951	0.932	0.870	0.942	0.922	0.880	0.952	0.934	0.851

* See notes to Table 1. In results reported in Panels A-B, jump intensity follows a Poisson distribution, with constant intensity, while in Panels C-F, intensity follows a Hawkes diffusion, and (a, β) parameterize path dependence strength. For complete details see above.

Table 3: Experimental Results - Self-Excitement Test Empirical Power *

Data Generated According to a Stochastic Volatility Process

$a=0.1,\beta=0.2$			$a=2,\beta=1$			$a=5,\beta=4$		
T=300	T=500	T=1000	T=300	T=500	T=1000	T=300	T=500	T=1000
Panel A: No Leverage, Self-Excitement in DGP, Jumps Generated Using Exponential Dist.								
0.889	0.885	0.861	0.203	0.224	0.215	0.151	0.124	0.130
0.965	0.963	0.937	0.316	0.343	0.282	0.172	0.167	0.164
0.952	0.932	0.919	0.456	0.457	0.448	0.223	0.200	0.181
0.916	0.906	0.874	0.576	0.555	0.530	0.283	0.298	0.250
0.882	0.867	0.839	0.610	0.624	0.567	0.360	0.311	0.264
Panel B: No Leverage, Self-Excitement in DGP, Jumps Generated Using Normal Dist.								
0.898	0.894	0.875	0.204	0.246	0.220	0.169	0.144	0.139
0.969	0.973	0.970	0.322	0.368	0.327	0.194	0.186	0.176
0.973	0.959	0.945	0.480	0.496	0.482	0.258	0.236	0.236
0.962	0.941	0.932	0.618	0.614	0.595	0.331	0.339	0.322
0.935	0.924	0.911	0.672	0.675	0.642	0.402	0.408	0.322
Panel C: Leverage, Self-Excitement in DGP, Jumps Generated Using Exponential Dist.								
0.921	0.905	0.900	0.242	0.248	0.241	0.185	0.153	0.144
0.968	0.960	0.965	0.372	0.373	0.344	0.164	0.184	0.166
0.949	0.950	0.905	0.530	0.527	0.517	0.294	0.288	0.231
0.943	0.925	0.910	0.663	0.624	0.629	0.326	0.321	0.295
0.932	0.912	0.860	0.738	0.693	0.654	0.428	0.384	0.353
Panel D: Leverage, Self-Excitement in DGP, Jumps Generated Using Normal Dist.								
0.920	0.909	0.906	0.256	0.250	0.262	0.183	0.161	0.135
0.979	0.974	0.980	0.385	0.386	0.380	0.176	0.207	0.191
0.976	0.966	0.949	0.560	0.550	0.557	0.314	0.304	0.282
0.970	0.958	0.951	0.680	0.654	0.666	0.360	0.369	0.342
0.966	0.952	0.925	0.766	0.745	0.718	0.470	0.431	0.411

* See notes to Table 2.

Table 4: Experimental Results - Sequential Jump and Self Excitement Test Empirical Power *
Data Generated According to a Stochastic Volatility Process

$a = 0.1$,$\beta = 0.2$			$a = 2$,$\beta = 1$			$a = 5$,$\beta = 4$		
T=300	T=500	T=1000	T=300	T=500	T=1000	T=300	T=500	T=1000
Panel A: No Leverage, Self-Excitement in DGP, Jumps Generated Using Exponential Dist.								
0.855	0.869	0.859	0.167	0.167	0.194	0.127	0.118	0.130
0.941	0.960	0.937	0.283	0.330	0.282	0.164	0.164	0.164
0.947	0.932	0.919	0.455	0.457	0.448	0.223	0.200	0.181
0.916	0.906	0.874	0.575	0.555	0.529	0.283	0.298	0.250
0.882	0.867	0.839	0.610	0.623	0.567	0.359	0.311	0.263
Panel B: No Leverage, Self-Excitement in DGP, Jumps Generated Using Normal Dist.								
0.887	0.888	0.868	0.197	0.244	0.219	0.164	0.144	0.139
0.967	0.964	0.963	0.320	0.366	0.321	0.189	0.185	0.175
0.968	0.954	0.939	0.477	0.491	0.474	0.252	0.230	0.226
0.955	0.930	0.928	0.614	0.608	0.569	0.323	0.329	0.307
0.922	0.910	0.905	0.663	0.659	0.622	0.397	0.388	0.297
Panel C: Leverage, Self-Excitement in DGP, Jumps Generated Using Exponential Dist.								
0.869	0.875	0.892	0.170	0.165	0.167	0.101	0.133	0.114
0.945	0.951	0.965	0.306	0.318	0.291	0.137	0.154	0.121
0.944	0.949	0.905	0.513	0.501	0.487	0.274	0.270	0.227
0.943	0.925	0.910	0.655	0.619	0.623	0.325	0.314	0.295
0.932	0.912	0.860	0.737	0.693	0.653	0.427	0.384	0.353
Panel D: Leverage, Self-Excitement in DGP, Jumps Generated Using Normal Dist.								
0.913	0.897	0.881	0.205	0.204	0.224	0.158	0.154	0.132
0.967	0.955	0.949	0.372	0.373	0.367	0.175	0.205	0.185
0.962	0.938	0.899	0.543	0.539	0.514	0.309	0.292	0.241
0.951	0.927	0.905	0.665	0.627	0.623	0.346	0.351	0.293
0.948	0.923	0.864	0.746	0.717	0.662	0.453	0.405	0.336

* See notes to Table 2.