Rate-optimal Tests for Jumps in Diffusion Processes

Taesuk Lee[a] and Werner Ploberger[b]

[a]Department of Economics, University of Rochester

[b]Department of Economics, Washington University in St. Louis

December 2009

Abstract

Suppose one has given a sample of high-frequency intra-day discrete observations of a continuous-time random process (e.g. stock market data) and wants to test for the presence of jumps. We show that the power of any test of this hypothesis depends on the frequency of observation. In particular, we show that if the process is observed at intervals of length $1/n$ and the instantaneous volatility of the process is given by σ_t, at best one can detect jumps of height no smaller than $\sigma_t\sqrt{\log(n)/n}$. We construct a test which achieves this rate in the case for diffusion-type processes. With simulation experiments, we show that our tests have excellent size and power properties for realistic sample sizes and that they outperform other tests of the hypothesis that have been proposed in the recent literature. Applying our tests to high-frequency financial data (e.g. foreign exchange rates, stock market indices, and stock prices of companies included in the Dow-Jones Industrials 30 index), we detect more jumps in the data than are found by other tests.

Keywords : High Frequency Data, Jump, Likelihood Test.

[0]Further author information:

Taesuk Lee: Ph.D. candidate, E-mail: taesuk@taesuk.net

Werner Ploberger: Thomas H. Eliot Distinguished Professor, E-mail: werner@ploberger.com

1 Introduction

Continuous diffusion models have provided a simple, flexible, and powerful tool to analyze economic and financial data since the time before high frequency data become available. Because data are only observed at discrete times, we do not have full information on the trajectory of the process. Therefore we may have modeling errors due to the discreteness of the observations, but this problem can be significantly mitigated with high frequency data which is now available with technological progress.

High frequency data, however, generate their own challenges: We are not sure that the data generating process is continuous; there may be jumps. Since continuous diffusion models do not capture jumps, researchers must know whether the data contain jumps or not. Furthermore, many datasets (such as when returns are measured over short intervals - say 1 to 5 minutes) contain some contamination commonly called "market microstructure noise". Our aim in this paper is to propose an optimal test for the null hypothesis of continuous diffusion models against an alternative hypothesis of jump diffusion models while allowing for the presence of market microstructure noise in the data. Though several tests (Barndorff-Nielsen and Shephard (2006), hereafter BNS, Ait-Sahalia and J. (2009), hereafter AJ) were already introduced, their power properties were not explicitly considered. In this paper,

we derive a rate-optimal test valid under fairly general assumptions about the data generating process. Furthermore we compare power of our test and that of other competing tests.

2 Local Power Bound

As the null, we consider the usual (i.e., purely continuous) diffusion model:

$$dX_t = \mu_t dt + \sigma_t dW_t, \tag{1}$$

where $W = (W_t : t \in [0, 1))$ is a Wiener process and μ_t and σ_t are non-anticipating random processes fulfilling the usual requirements of Ito-calculus. Later on we will impose additional assumptions on μ and σ. (For example, we will assume them to be smooth to a certain extent, so that the process X_t specified in equation(1) has certain "nice" properties.)

As the alternative, we want to consider jumps present in the diffusion model :

$$dX_t = \mu_t dt + \sigma_t dW_t + J_t d\kappa_t,$$

where J_t is a non-zero random variable whose absolute value specifies the jump size and κ_t is a counting process governing whether there is a jump or not at time t.

The problem, however, is that in empirical practice, we cannot observe the entire process. Instead, it is observed only at discrete times $t = i/n$, where n is a positive integer denoting the "sample size" and

$$0 \leq i \leq n.$$

The first test for this testing problem - and still the "gold standard" for all tests was developed by Barndorff-Nielsen and Shephard (2002). Since the problem of detecting jumps is of enormous practical importance, further research has occurred in this field. An alternative test was developed by Ait-Sahalia and J. (2009), and an informal "testing procedure" was provided by Lee and Mykland (2008).

To the best of our knowledge, none of these contributions, however, discussed power of the tests. In this paper, we establish the following two results.

1. Clearly, when n, the number of observations, increases, we should expect our any "reasonable" test to have "better" power. In particular, we want to consider the power against local alternatives of jump components in the process. Specifically, we consider for each n the alternative hypotheses characterized by

$$\left| J_t^{(n)} \right| = c_n$$

where we assume that c_n is a positive sequence converging to zero and the process κ_t remains (uniformly) bounded as n to ∞. (The latter assumption implies that there is a maximum number of jumps.) Let $\varepsilon > 0$ be arbitrary. Then we show that - even if $\sigma_t = \sigma$ for all t - it is impossible to construct tests that have nontrivial power against local alternatives characterized by

$$c_n = (1 - \varepsilon)\sigma \frac{\sqrt{\ln n}}{\sqrt{n}} \tag{2}$$

2. As our main result, we show that one can devise a test so that (in the general case of time-varying σ_t) if

$$c_n = (\sqrt{2} + \varepsilon)\sigma_t \frac{\sqrt{\ln n}}{\sqrt{n}}$$

the power of this test converges to one. So in a certain sense, our test attains the "optimal" rate. This is an improvement over the BNS and AJ tests: their local alternatives shrink with the order at the faster rate of $n^{-1/4}$ or - in the case of AJ - with the order of $n^{-1/2+1/p}$, where p is a positive number determining the test statistic.

Let us first deal with our first assertion. Let's consider the very simplest case, namely $\mu_t = 0$ and $\sigma_t = 1$, so that the underlying process X_t is a Wiener processes under the null. Then let us assume that under the alternative we only have one jump, and the timing of the jump is distributed uniformly on the interval $[0, 1]$. We first show that even under this rather ideal set of conditions one cannot construct tests with nontrivial power if the sequence c_n follows (2).

Theorem 1 *We want to test the null of X_t being a Wiener process W_t of known variance against the alternative that*

$$X_t = W_t + c_n I\left\{\tau \geq t\right\},$$

where τ is an independent random variable following a uniform distribution. Suppose we observe the process X_t only at the time points $0, 1/n, 2/n, ...1$. If c_n follows (2) (or is smaller than this bound). Then, it is impossible to construct nontrivial tests.

Proof. We assume that the variance of the Wiener process W is known. Without loss of generality, we can assume this variance to be 1. Let P_n be the probability measure of $(X_0, X_{1/n}, X_{2/n}, ..., X_1)$ under the null, and Q_n be the measure under the alternative. Let the z_i be defined as

$$z_i = \left(X_{i/n} - X_{(i-1)/n}\right)\sqrt{n}$$

Then we can easily see that z_i are i.i.d. from standard normal and that

$$\frac{dQ_n}{dP_n} = \frac{1}{n}\sum_{i=1}^{n} \exp((c_n\sqrt{n})z_i - \frac{1}{2}(c_n\sqrt{n})^2)$$

Since each of the z_i is standard normal, the expectation of each $\exp((c_n\sqrt{n})z_i - \frac{1}{2}(c_n\sqrt{n})^2)$ equals one.

Moreover, $E\left(\exp((c_n\sqrt{n})z_i - \frac{1}{2}(c_n\sqrt{n})^2)\right)^2 = E(\exp(2(c_n\sqrt{n})z_i - (c_n\sqrt{n})^2)) = \exp((c_n\sqrt{n})^2)$. Hence the variance of $\frac{dQ_n}{dP_n}$ is smaller than $\exp((c_n\sqrt{n})^2)/n$, which converges to 0 if c_n follows (2). Therefore

$$\frac{dQ_n}{dP_n} \to 1$$

in probability. Hence, for an arbitrary $\eta > 0$

$$P_n\left(\left[\left|\frac{dQ_n}{dP_n}-1\right|>\eta\right]\right)\to 0.$$

Now let A_n be a sequence of events. Then we have

$$\begin{aligned}(1-\eta)P_n(A_n)-P_n\left(\left[\left|\frac{dQ_n}{dP_n}-1\right|>\eta\right]\right) &< Q_n(A_n)\\ &< (1+\eta)P_n(A_n)+P_n\left(\left[\left|\frac{dQ_n}{dP_n}-1\right|>\eta\right]\right).\end{aligned}$$

Since η was arbitrary, we can conclude that

$$P_n(A_n)-Q_n(A_n)\to 0.$$

Since A_n is an arbitrary sequence of events, we can conclude that the total variation between P_n and Q_n converges to 0 and hence that for all measurable functions φ_n with $0 \leq \varphi_n \leq 1$ we have

$$\int \varphi_n dP_n - \int \varphi_n dQ_n \to 0.$$

But this is exactly what we want to show: For every sequence of tests, the power under the null (P_n) is the same as under the alternative (Q_n). ■

Now we want to present a test statistic, for which we will show that the local power of the test statistic attains this bound. The preceding result indicates that for fixed c_n our best statistic is an exponential sum of the z_i. We should, however, keep in mind that our z_i are increments over shorter and shorter time intervals. In order to consider relevant alternatives, we might be interested in alternatives where c becomes "large". In this case, the test statistic gives more and more increasing influence to bigger values. Continuing with this line of reasoning, it may be a good idea to look mainly at the "largest" absolute values of the increments of X_t: The standard theory of diffusion processes guarantees that, when divided by σ_t, these increments are approximately normal. Next, because we do not know σ_t, we have to estimate it. As σ_t is time-varying, a moving average of the squares of the increments seems to be a natural candidate estimator. This leads us to propose the following test statistic: Define (for an arbitrary n) the return $r_i = r_{i,n}$ by

$$r_i = r_{i,n} = \left(X_{i/n} - X_{(i-1)/n}\right)$$

Then choose an integer ℓ (the "length of the window") and reject the null whenever

$$\tau_n = \sup_i \frac{r_i^2}{(r_{i-1}^2 + r_{i-2}^2 + ..r_{i-\ell}^2)/\ell}$$

is "too large". This is quite analogous to the test statistics of Lee and Mykland (2008): We standardize the returns by an estimator for $\sigma^2(t, X_t)$. We use the usual quadratic estimator instead of the bipower estimator. One might argue that jumps could unduly affect the properties of our estimator. We think, however, that the much simpler form is justified, essentially in part for the following two reasons:

1. We assume that the jumps are separated events: *Before* the first jump occurs, our estimator for $\sigma^2(t, X_t)$ is not influenced by it.

2. We only use a window of length ℓ to estimate $\sigma^2(t, X_t)$. So a jump will only influence a small number of estimated values of $\sigma^2(t, X_t)$. Our test would get only distorted if we had two (or more) jumps within an interval of length ℓ/n, an event whose probability we assume converges to zero.

The main reason, however, for using this specific estimator is convenience. Specifically, only Lemma 6 is essential to prove the rate-optimality of our proposed test. We think analogous results will hold for more general classes of estimators.

3 Critical Values and Local Power of Our Test

Let $z_i, i = 1, .., n$ be a sequence of independent, identically distributed standard normal random variables. Assume that for each n we have given an $\ell = \ell(n)$, and let us denote by $\mathcal{F}_i$ the $\sigma-$algebra generated by $z_i, z_{i-1}, ..$ Then let us define

$$w_i = \sum_{j=1}^{\ell} z_{i-j}^2,$$

$$\widehat{\sigma}_i^2 = w_i/\ell$$

and

$$\tau_i = z_i^2 / \widehat{\sigma}_i^2.$$

For the computation of the critical values the following lemma is very helpful.

Lemma 2 *Suppose that*

$$\ell = o\left(n\right),$$

but also

$$\ell \geq 2 \log n.$$

Define for each $c > 0$, a $K_n^* = K_n^*(c)$ such that

$$2E\left(\exp\left(-K_n^{*2}\widehat{\sigma}_i^2/2\right)/\sqrt{2\pi K_n^{*2}\widehat{\sigma}_i^2}\right) = c/n \tag{3}$$

Then,

$$P\left(\max_{i=\ell+1,..n} \tau_i > K_n^*\right) \to 1 - \exp\left(-c\right) \text{ as } n \to \infty.$$

Proof. First of all we observe that $P\left(\max_{i=\ell+1,..n} \tau_i > K_n^*\right) = 1 - P\left(\max_{i=\ell+1,..n} \tau_i \leq K_n^*\right)$ and

$$P\left(\max_{i=\ell+1,..n} \tau_i \leq K_n^*\right) = E\left(\prod_{i=\ell+1,..n} I\left\{\tau_i \leq K_n^*\right\}\right).$$

It can immediately be seen that the τ_i are $\mathcal{F}_i-$measurable. We will repeatedly apply the optional sampling theorem for various stopping times. Let $\varepsilon > 0$ be arbitrary, and let $M(\varepsilon)$ be defined as in the Appendix A (14), (15), (16).

Let us define the stopping time ν as follows: Define ν to be the first index $m \leq n-1$ so that

$$\begin{aligned}
\sum_{j=\ell+1}^{m+1} \log E\left(I\left\{\tau_i \leq K_n^*\right\}|\mathcal{F}_{i-1}\right) &< -c(1+\varepsilon)^3 \text{ or } \widehat{\sigma}_{m+1}^2 < M\left(\varepsilon\right)^2/K_n \text{ or} \\
\sum_{j=\ell+1}^{m+1} \log E\left(I\left\{\tau_i \leq K_n^*\right\}|\mathcal{F}_{i-1}\right) &> -c(1-\varepsilon)
\end{aligned}$$

and

$$n \text{ if no such } m \text{ exists.}$$

Observe that ν is indeed a stopping time adapted to $\mathcal{F}_i$: Since for $i \leq m+1$ $E\left((\tau_i \leq K_n^*)/\mathcal{F}_{i-1}\right)$ as well as $\widehat{\sigma}_{m+1}^2$ are $\mathcal{F}_m$-measurable, the event

$$[\nu = n] \in \mathcal{F}_m.$$

We contend that

$$\lim_{n\to\infty} P([\nu = n]) = 1. \tag{4}$$

To demonstrate (4), it is sufficient to first show that

$$P(\left[\inf \widehat{\sigma}_i^2 > M(\varepsilon)^2 / K_n\right]) \to 1 \tag{5}$$

and then, because $\log E\left(I\left\{\tau_i \leq K_n^*\right\} | \mathcal{F}_{i-1}\right) \leq 0$, then that

$$P\left[\left\{\sum_{j=\ell+1}^{n} \log E\left(I\left\{\tau_i \leq K_n^*\right\} | \mathcal{F}_{i-1}\right) \geq -c(1+\varepsilon)^3\right\} \cap \left\{\inf \widehat{\sigma}_i^2 > M(\varepsilon)^2 / K_n\right\}\right] \to 1. \tag{6}$$

Equation (5) is an immediate consequence of Lemma 6, which shows that

$$P\left[\inf \widehat{\sigma}_i^2 \leq M(\varepsilon)^2 / K_n\right] \leq nP\left[\widehat{\sigma}_i^2 \leq M(\varepsilon)^2 / K_n\right] \to 0.$$

To prove that (6) is valid, first observe that

$$E\left(I\left\{\tau_i \leq K_n^*\right\} | \mathcal{F}_{i-1}\right) = 2\Phi(\sqrt{K_n^* \widehat{\sigma}_i^2}) - 1.$$

If $\widehat{\sigma}_i^2 > M(\varepsilon)^2 / K_n$, we can use inequality (16) to conclude that

$$\log\left(2\Phi(\sqrt{K_n^* \widehat{\sigma}_i^2}) - 1\right) \geq -2(1+\varepsilon)^2 \exp\left(-K_n^* \widehat{\sigma}_i^2/2\right) / \sqrt{2\pi K_n^* \widehat{\sigma}_i^2}$$

Hence

$$\begin{aligned} &\left[\sum_{j=\ell+1}^{n} \log E\left(I\left\{\tau_i \leq K_n^*\right\} | \mathcal{F}_{i-1}\right) \geq -c(1+\varepsilon)^3\right] \cap \left[\inf \widehat{\sigma}_i^2 > M(\varepsilon)^2 / K_n\right] \\ \subseteq \quad &\left[-2(1+\varepsilon)^2 \sum_{j=\ell+1}^{n} \exp\left(-K_n^* \widehat{\sigma}_i^2/2\right) / \sqrt{2\pi K_n^* \widehat{\sigma}_i^2} \geq -c(1+\varepsilon)^3\right] \cap \left[\inf \widehat{\sigma}_i^2 > M(\varepsilon)^2 / K_n\right] \end{aligned}$$

Since we already know that $P\left(\left[\inf \widehat{\sigma}_i^2 > M(\varepsilon)^2 / K_n\right]\right) \to 1$, it is sufficient to show that

$$P\left(\left[2 \sum_{j=\ell+1}^{n} \exp\left(-K_n^* \widehat{\sigma}_i^2/2\right) / \sqrt{2\pi K_n^* \widehat{\sigma}_i^2} \leq c(1+\varepsilon)\right]\right) \to 1 \tag{7}$$

Let us now introduce the term Y_j by

$$Y_j = 2 \exp\left(-K_n^* \widehat{\sigma}_i^2/2\right) / \sqrt{2\pi K_n^* \widehat{\sigma}_i^2}$$

Then we can easily see that (7) is fulfilled if

$$\sum_{j=\ell+1}^{n} Y_j \to c \tag{8}$$

in probability. By our definition of K_n^*, $EY_j = c/n$. Moreover, we know that $\widehat{\sigma}_i^2$ is distributed according to a scaled χ^2 distribution with ℓ degrees of freedom. Hence it is an elementary exercise to show that $EY_j^2 = O(1/n^2)$, and that Y_j and Y_k are independent if

$$|j-k| > \ell + 1.$$

As $\ell/n \to 0$, we can easily see that the variance of $\sum Y_j$ converges to zero.

This establishes (4). Now it is rather easy to establish the claim stated in our lemma: We have to show that

$$E\left(\prod_{i=\ell+1,..n} I\left\{\tau_i \leq K_n^*\right\}\right) \to \exp(-c)$$

Using (4), it is sufficient to show

$$E\left(\prod_{i\leq\nu} I\left\{\tau_i \leq K_n^*\right\}\right) \to \exp(-c)$$

Trivially,

$$E\left(\frac{I\left\{\tau_i \leq K_n^*\right\}}{E\left(I\left\{\tau_i \leq K_n^*\right\}|\mathcal{F}_{i-1}\right)}|\mathcal{F}_{i-1}\right) = 1.$$

A straightforward argument, perfectly analogous to the optional sampling theorem, yields

$$E\left(\frac{E\prod_{i\leq\nu} I\left\{\tau_i \leq K_n^*\right\}}{\prod_{i\leq\nu} E(I\left\{\tau_i \leq K_n^*\right\}|\mathcal{F}_{i-1})}\right) = 1. \tag{9}$$

According to the definition of ν,

$$-(1+\varepsilon)^2 \sum_{j=\ell+1}^{\nu} Y_j \leq \log \prod_{i\leq\nu} E(I\left\{\tau_i \leq K_n^*\right\}|\mathcal{F}_{i-1}) \leq -(1-\varepsilon)^2 \sum_{j=\ell+1}^{\nu} Y_j \tag{10}$$

and

$$\log \prod_{i\leq\nu} E(I\left\{\tau_i \leq K_n^*\right\}|\mathcal{F}_{i-1}) \geq -c(1+\varepsilon)^3 \tag{11}$$

Moreover, (4) implies that $P\left(\left[\sum_{j=\ell+1}^{\nu} Y_j = \sum_{j=\ell+1}^{n} Y_j\right]\right) \to 1$. Therefore $\sum_{j=\ell+1}^{\nu} Y_j \to c$ as well. Hence it can be seen that (11) and (10) allow us to deduce from (9) that

$$\begin{aligned}\exp(-(1+\varepsilon)^2 c) &\leq \liminf E \prod_{i\leq\nu} I\left\{\tau_i \leq K_n^*\right\} \\ &\leq \limsup E \prod_{i\leq\nu} I\left\{\tau_i \leq K_n^*\right\} \leq \exp(-(1-\varepsilon)^2 c).\end{aligned}$$

Now one can easily see that (4) allows us to replace ν with n in the preceding inequalities, which proves our lemma. ■

So far, we have computed the distribution of our test statistic for a very specific case, namely when $\mu_t = 0$ and $\sigma_t = 1$. We now have to show that the general case given by (1) can be reduced to the specific case discussed above. To achieve this goal, we have to impose some stronger assumptions on μ_t and σ_t.

Theorem 3 *Suppose μ_t and $\ln \sigma_t$ are diffusion processes with a.s. uniformly bounded diffusion coefficients. Then - provided that the conditions of Lemma 2 are satisfied and that $\ell_n / \ln n$ converges to a constant different from 0 - the difference between the test statistic applied to X_t and W_t converges to zero in probability.*

Proof. The proof is rather technical and is provided in the Appendix B. ■

Lemma 2 and Theorem 3 establish that our construct - rejecting when the max τ_i are larger than K_n^* - is indeed a test. Moreover, it is an easy, but tedious exercise to establish the order of magnitude of K_n^*. Because the distribution of $\widehat{\sigma}_i^2$ is a scaled χ^2 distribution, the left-hand side of equation (3) can be evaluated using the gamma function.[1] Then it is then an easy task to show

[1] Given an $\alpha-$level of significance, we can plug-in $c = -\log(1-\alpha)$. Then we can solve for the critical value K, which by definition satisfies

$$\left(\frac{K}{\ell}+1\right)^{-\frac{(\ell-1)}{2}} \left(\frac{K}{\ell}\right)^{-1/2} \frac{\Gamma(\ell/2-1/2)}{\Gamma(1/2)\Gamma(\ell/2)} + \frac{\log(1-\alpha)}{n} = 0 \tag{12}$$

If the sample size n is small and/or the average window size ℓ is small, then the approximation of Lemma 2 can be improved by employing a small sample corrected critical value K which satisfies

$$\left(\frac{K}{\ell}+1\right)^{-\frac{(\ell-1)}{2}} \left(\frac{K}{\ell}\right)^{-1/2} \frac{\Gamma(\ell/2-1/2)}{\Gamma(1/2)\Gamma(\ell/2)} \left(\frac{\ell-1}{\ell}\right) + \frac{\log(1-\alpha)}{n-\ell} = 0 \tag{13}$$

that

$$K_n^* / \left(2 \ln n\right) = 1.$$

Finally, it is elementary to establish our assertion that the test is consistent against jumps of the order

$$\left(1 + \varepsilon\right) \sigma_t \frac{\sqrt{2 \ln n}}{\sqrt{n}}.$$

4 Power of the Some Competing tests

As mentioned in the introduction, several tests for this testing problem already been published. Two of the best-known ones are the tests of BNS and of AJ. These tests are based on the following test statistics:

Definition 4 *BNS test statistic(Barndorff-Nielsen and Shephard (2006))*

$$\begin{aligned} \widehat{\tau}_{BNS}^{LIN} &= \frac{\sqrt{n}\left(RV - \frac{\pi}{2}BPV\right)}{\sqrt{\int_0^1 \sigma_u^4 du}}, \widehat{\tau}_{BNS}^{ADJ} = \frac{\sqrt{n}\left(1 - \frac{\pi BPV}{2RV}\right)}{\sqrt{\max\left[1, \int_0^1 \sigma_u^4 du / \left\{\int_0^1 \sigma_u^2 du\right\}^2\right]}} \textit{ where} \\ RV &= \sum_{j=1}^{1/\Delta} r_{t+j\Delta}^2 \textit{ and } BPV = \sum_{j=2}^{1/\Delta} \left|r_{t+j\Delta}\right| \left|r_{t+(j-1)\Delta}\right| \end{aligned}$$

Another alternative test has recently been proposed by Ait-Sahalia and Jacod. This test is based on the $p-$th power variation of the process X_t and it compares the estimates of this variation for different time scales.

Definition 5 *AJ test statistic(Ait-Sahalia and J. (2009))*[2]

$$\begin{aligned}
\widehat{\tau}_{AJ}^{p,k} &= \left(k^{p/2-1} - \widehat{S}(p,k,\Delta)\right) / \sqrt{\widehat{V}_{p,k}} \text{ where } p > 3, k \geq 2, \\
\widehat{S}(p,k,\Delta) &= \widehat{B}(p,k\Delta) / \widehat{B}(p,\Delta), \ \widehat{B}(p,k\Delta)_t = \sum_{i=1}^{n/k} |r_{t+ik\Delta}|^p \text{ and} \\
& \widehat{V}_{p,k} \text{ is the variance of } \widehat{S}(p,k,\Delta) \text{ under the null.}
\end{aligned}$$

The behavior of both sets of test statistics under our alternatives can easily be analyzed. We just add a specific jump to one of returns. One can easily see that if the jump is of $o(n^{-1/4})$, the difference between the BNS test statistic under the null and the alternative converges to 0. Hence their test will be much less powerful against the alternatives we consider. The same is true for AJ tests: After some calculations, we can see that the corresponding bound is $o\left(n^{-1/2+1/p}\right)$. So - in a bit of contrast to the views of the authors of the test - we think that larger orders p deserve more attention (Our simulation results, however, indicate that the limiting distribution does not well approximate the finite-sample distribution for higher order p). In any case, we think that this subject merits further research.

Despite the fact that both tests have "low" power against "our" alternatives, it should be noted that there are situations where the BNS and AJ tests have large advantages over our test. For instance, the relevant alternative could be the occurrence of many jumps in the sample. Assume there is not only one jump, but many. So let us assume that we have L jumps of size J (rather evenly distributed), and let us assume that the number of intervals between successive jumps is always greater than 1. Then it is easily seen from the definition of BNS statistics that the tests are consistent (i.e., its power converges to 1) if

$$\sqrt{n}LJ^2 \to \infty.$$

[2] For $\widehat{V}_{p,k}$, they suggest two estimators:
$\widehat{V}_{p,k}^c = \frac{\Delta_n M(p,k)\widehat{A}(2p,\Delta)_t}{\widehat{A}(p,\Delta)_t^2}, \widetilde{V}_{p,k}^c = \frac{\Delta_n M(p,k)\widetilde{A}\left(\frac{p}{p+1},2p+2,\Delta\right)_t}{\widetilde{A}\left(\frac{p}{p+1},p+1,\Delta\right)_t^2}$ where
$M(p,k) = \frac{1}{m_p^2}\left(k^{p-2}(1+k)m_{2p} + k^{p-2}(k-1)m_p^2 - 2k^{p/2-1}m_{k,p}\right)$,
$m_p = E(|Z_1|^p) = \pi^{-1/2}2^{p/2}\Gamma\left(\frac{p+1}{2}\right)$,
$m_{k,p} = E\left[|Z_1|^p \left|Z_1 + \sqrt{k-1}Z_2\right|^p\right]$,
$Z_i \sim^{iid} N(0,1)$,
$\widehat{A}(p,\Delta_n)_t = \frac{\Delta_n^{1-p/2}}{m_p}\sum |\Delta_i^n X|^p 1\{|\Delta_i^n X| \leq \alpha\Delta_n^{\varpi}\}, \varpi \in (0,1/2)$,
$\widetilde{A}(r,q,\Delta_n)_t = \frac{\Delta_n^{1-qr/2}}{m_r^q}\sum_{i=1}\Pi_{j=1}^q \left|\Delta_{i+j-1}^n X\right|^r$

An analogous result holds for AJ tests. These results are easily explainable by taking into account that both test statistics are constructed from sums: The contributions of several small jumps can cumulate, whereas our test does not allow for this. So one should employ their tests not as tests against the alternative of the occurrence of a single jump, but as tests against Levy-type alternatives. It might be worthwhile to investigate their power of the test against alternatives of this specific type.

5 Simulations

We first consider a very simple ideal process for returns:

$$r_{i/n} = \int_{(i-1)/n}^{i/n} dp_t = \int_{(i-1)/n}^{i/n} \sigma dW_t + \int_{(i-1)/n}^{i/n} J d\kappa_t \qquad \text{(Model 1)}$$

Second, we considered the model devised by Barndorff-Nielsen and Shephard (2006)

$$\begin{aligned} dp_t &= \sigma(s) W(ds) + J d\kappa_t & \text{(Model 2)} \\ \sigma^2(t) &= w_1 \sigma_1^2(t) + w_2 \sigma_2^2(t) \\ \sigma_k^2(t) &= -\int_0^t \lambda_k(s) \left\{ \sigma_k^2(s) - \xi(s) \right\} ds + \int_0^t \omega(s) \sigma_k(s) B_k(ds) \text{ where } k = 1, 2 \end{aligned}$$

We track the performance of four test statistics : Lee-Ploberger (LP), Barndorff-Nielsen & Shephard(BNS), Ait-Sahalia & Jacod(AJ), and Lee-Mykland (LM). We assume $J \sim N(0, \sigma_c^2)$ and consider three cases for jump sizes : no jump($\sigma_c^2 = 0$), 20% jump($\sigma_c^2 = 0.2\xi(s)$), and $\ln(n)/n$ jump($\sigma_c^2 = \ln(n)/n * \xi(s)$). The sample sizes considered are 72, 288, 1440, 2880, 8640; these sample sizes correspond to sampling interval lengths of 20 minutes, 5 minutes, 1 minute, 30 seconds, and 10 seconds, respectively, over a 24-hour trading day. We repeat this simulation 50,000 times. The parameters are calibrated by Barndorff-Nielsen and Shephard (2006) as follows : $\xi(s) = 0.509$, $\omega^2(s) = 0.461$, $w_1 = 0.218$, $w_2 = 1 - p_1$, $\lambda_1 = 0.0429$, $\lambda_2 = 3.74$.

Let us consider rejection probabilities under the null, no-jump case. Table 1 shows that our tests have better rejection probabilities than other tests in the continuous, i.e., pure diffusion model. In most cases, the other tests are equally precise with 10-second data but they are less precise for smaller samples. Our tests have precise size even with 20-minute data. The differences between our rejection probabilities and the nominal significance levels are less than 0.5 percentage point, which can be attributed to simulation variation. The next best test is the adjusted BNS test. It is better than other two tests but it tends to overreject the null in small samples.

Table 2 also shows our tests are better in a continuous stochastic volatility model. With small sample sizes, our tests show moderate size distortion of around 1 percentage point, but this distortion is smaller than that of the other tests. Note that even though our test and the LM test statistics look similar, but their performance is different. The LM test needs larger window sizes than ours for normalization; their required order is the square root of the sample size. So its performance with small averaging windows is distorted a lot even in the pure diffusion case (Table 1). In contrast, our test is valid even with small windows. Both tests, of course, assume some continuity of volatility within the averaging window. However, because our test's averaging window is smaller, it is more robust in the case of rapidly changing volatility. This explains the performance of our test, when compared to the LM test, in the stochastic volatility model.

Next, we consider power of the test. Our tests have better power after controlling for size distortions. In some cases, the other tests have larger rejection probabilities under the alternative, but they also have larger rejection probabilities under the null hypothesis. If we adjust the critical values to remove size distortions, our test has better power in most cases. In particular, our tests have non-trivial power even with jump the order of which is $\ln(n)/n$, where n is the sample size, whereas the BNS and AJ tests have trivial power in that case. LM tests also have non-trivial power but their sizes are not as reliable as ours and their averaging window requirement is more restrictive than ours. As we see in figure 3 and table 3, the size-adjusted power curve of our tests envelop those of other tests, which is the sign of optimality of our test. In following section, we apply our tests to various important financial data.

6 Empirical Applications

We apply our test to stock index and foreign exchange rate in the FOREX database and individual stock data in the TAQ database. We consider USD index, USD/JPY, EUR/USD, GBP/USD, USD/CAN from 1999 to 2000. 1998-2000 DJDA, NASDAQ COMP, NASDAQ100, S&P500, S&P100, RUSSEL2000 are examined. Dow Jones 30 stocks in 2005 NYSE TAQ are also considered. The empirical results are reported in table 4-7. We find the followings.

First, we can find strong evidences of jumps in most markets. For the European and US foreign exchange market data, AJ test shows 15-45% trading days have jump. The ratios of jumping days increase to 20-35% for BNS test, 30-65% for LP test, and 70-85% for LM test. We observe similar patterns in other markets : The ratios of jumping days in the whole foreign exchange market are 20-60% for AJ test, 25-40% for BNS test, 45-85% for LP test, and 76-95% for LM test. Those in the stock indices are 2-8% for AJ test, 10-25% for BNS test, 15-40% for LP test, and 45-60% for LM test. Those in the Dow 30 stocks are 15-20% for AJ and BNS test, 20-30% for LP test, and 55-65% for LM test.

Second, we can order the tests by the ratio of jumping days. AJ test shows the least number of jumping days. The next is BNS tests and our LP test. LM test has largest ratio of jumping days. Our previous simulated power curve shows the similar pattern. Note that LM test has the largest power because of the size distortion under the null. So we can say our test has the best power in some sense.

Third, we can detect more jumps with more data. Five minutes data shows more jumping days than fifteen minutes data in most cases. In the foreign exchange market data, we consider 13 hours data and 24 hours data. We can detect more jumps with 24 hours data. We can see similar patterns in the previous simulated power curve example. Power of test increases as the sample size increases.

References

AIT-SAHALIA, Y. (2004): “Disentangling diffusion from jumps,” *Journal of Financial Economics*, 74(3), 487–528.

AIT-SAHALIA, Y., AND J. J. (2009): “Testing for jumps in a discretely observed process,” *Annals of Statistics*, 37(1), 184–222.

Andersen, T., T. Bollerslev, and F. Diebold (2007): "Roughing it Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89(4), 701–720.

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

Bandi, F., and J. Russell (2006): "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79(3), 655–692.

Barndorff-Nielsen, O., and N. Shephard (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64(2), 253–280.

——— (2003): "Realised power variation and stochastic volatility models," *Bernoulli*, 9(2), 243–265.

——— (2004): "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72(3), 885–925.

——— (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4(1), 1–30.

Boehmer, E., G. Saar, and L. Yu (2005): "Lifting the veil: An analysis of pre-trade transparency at the NYSE," *Journal of Finance*, 60(2), 783–815.

Bollerslev, T., T. Law, and G. Tauchen (2008): "Risk, jumps, and diversification," *Journal of Econometrics*, 144(1), 234–256.

Delattre, S., and J. Jacod (1997): "A Central Limit Theorem for Normalized Functions of the Increments of a Diffusion Process, in the Presence of Round-Off Errors," *Bernoulli*, 3(1), 1–28.

Drost, F. C., T. E. Nijman, and B. J. M. Werker (1998): "Estimation and Testing in Models Containing Both Jump and Conditional Heteroscedasticity," *Journal of Business & Economic Statistics*, 16(2), 237–43.

Eraker, B. (2004): "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices," *Journal of Finance*, 59(3), 1367–1404.

Eraker, B., M. Johannes, and N. Polson (2003): "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 58(3), 1269–1300.

Huang, X., and G. Tauchen (2005): "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3(4), 456–499.

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

Lee, S., and P. Mykland (2008): "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 21(6), 2535–2563.

Oomen, R. (2005): "Properties of bias corrected realized variance in calendar time and business time," *Journal of Financial Econometrics*, 3(4), 555–577.

Todorov, V. (2009): "Estimation of continuous-time stochastic volatility models with jumps using high-frequency data," *Journal of Econometrics*, 148(2), 131–148.

Zhang, L. (2006): "Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach," *Bernoulli*, 12(6), 1019–1043.

Zhang, L., P. Mykland, and Y. Ait-Sahalia (2005): "A tale of two time scales: Determining integrated volatility with noisy high-frequency data," *Journal of the American Statistical Association*, 100(472), 1394–1411.

A Normal and χ^2 distributions for small and large values

It is well known that for the standard distribution function $\Phi(x)$

$$\lim_{x\to\infty}(1-\Phi(x))\sqrt{2\pi}x\exp\left(x^2/2\right)=1$$

or equivalently

$$\lim_{x\to\infty}(\log(2\Phi(x)-1))\sqrt{2\pi}x\exp\left(x^2/2\right)/2=-1$$

we can define $M(\varepsilon)$ as the smallest value so that for all

$$x>M(\varepsilon) \tag{14}$$

$$(1-\varepsilon)\le\left|\sqrt{2\pi}x\exp\left(x^2/2\right)(1-\Phi(x))\right|\le(1+\varepsilon). \tag{15}$$

and

$$-2(1+\varepsilon)^2\le(\log(2\Phi(x)-1))\sqrt{2\pi}x\exp\left(x^2/2\right)\le-2(1-\varepsilon)^2 \tag{16}$$

Lemma 6 *So let us now choose an arbitrary $\varepsilon>0$, and let ℓ, n, w_i be the integers defined in the main section of the paper. If*

$$\ell\ge 2\ln n$$

and

$$K\to\infty,$$

then

$$p_n=P\left[w_i\le\ell M(\varepsilon)^2/K\right]=o\left(n^{-1}\right).$$

.

Proof. Since w_i is distributed according to a χ^2distribution with ℓ degrees of freedom, we have

$$p_n=\frac{1}{\Gamma(\ell/2)}\int_0^{\ell M(\varepsilon)^2/(K)}x^{\ell/2-1}\exp(-x/2)\,dx.$$

Since $\exp(-x/2)\le 1$, we have with

$$C=\frac{M(\varepsilon)^2}{K}$$

$$p_n \le \frac{1}{\Gamma\left(\ell/2\right)} \frac{1}{\ell/2} C^{\ell/2} \ell^{\ell/2},$$

and therefore

$$\log p_n \le -\log \Gamma\left(\ell/2\right) - \log\left(\ell/2\right) + \frac{\ell}{2} \log C + \ell/2 \log \ell.$$

The well known formula of Stirling implies that for $\ell \to \infty$ (with $m = \ell/2 - 1$)

$$\log \Gamma\left(\ell/2\right) - \left(m\left(\log(m) - m + \log(\sqrt{2\pi m}\right) \to 0.$$

Therefore

$$\log p_n \le \left(\left(\ell/2\right)\log \ell - m\left(\log(m)\right) + \left(m + \frac{\ell}{2}\log C\right) + O(\log \ell)$$

One can easily see that $((\ell/2)\log \ell - m\left(\log(m)\right) = (\ell/2)\left(\log(\ell/m)\right) + O(\log m)$. So the terms linear in ℓ dominate the right hand side of the inequality, Moreover, as $M\left(\varepsilon\right)$ is fixed and $K \to \infty$, we may conclude that $C \to 0$. Therefore, $\frac{\ell}{2}\log C$ will become negative. Therefore, $\frac{\ell}{2}\log C$ will become negative and dominate other parts. Therefore it can immediately be seen that lim sup

$$\frac{-\log p_n}{\ell/2\left(-\log C\right)} \ge 1,$$

which implies $p_n \le \exp\left(-\ell/2\right) \frac{M(\varepsilon)^2}{K}$. ■

B The proof of Theorem 3

Our proof of Theorem 3 is based on the following lemma.

Lemma 7 *Suppose we have given a standard Wiener process W, an adapted process f and a constant α so that*

$$\int_z^b f^2 dt \le B.$$

Then, where $\int_a^b f dW$ is the usual Ito-integral,

$$P\left(\left[\left|\int_a^b f dW\right| \ge C\right]\right) \le 2\exp(-\frac{C^2}{2B}).$$

Proof. Novikov's theorem guarantees that for all u

$$E\left(\exp(u\int_a^b fdW - \frac{u^2}{2}\int_a^b f^2 dt\right) = 1$$

Hence

$$E(\exp(u\int_a^b fdW - \frac{u^2}{2}B) \leq 1$$

and therefore

$$\exp(uC - \frac{u^2}{2}B)P(\left[\int_a^b fdW > C\right]) \leq 1.$$

Setting

$$u = \frac{C}{B}$$

and repeating the same idea with $-\int_a^b fdW$ proves our proposition. ■

We are now prove Theorem 3 applying Lemma 7.

Proof of Theorem 3. We have

$$d\mu_t = A_t dt + B_t dV_t^{(1)},$$

$$d(\log\sigma_t) = C_t dt + D_t dV_t^{(2)},$$

where A_t, B_t, C_t, D_t are continuous processes and $V_t^{(1)}$,$V_t^{(2)}$ are (standard) Wiener processes. First of all let us demonstrate that without limitation of generality we can assume that A_t, B_t, C_t, D_t and $\mu_t, \log\sigma_t$ as well are uniformly bounded.

Since the processes A_t, B_t, C_t, D_t and $\mu_t, \ln\sigma_t$ are continuous, for every $\varepsilon > 0$ there exists a $M = M(\varepsilon)$ so that

$$P\left([\sup|A_t|, \sup|B_t|, \sup|C_t|, \sup|D_t|, \sup|\mu_t|, \sup|\ln\sigma_t| < M(\varepsilon)]\right) > 1 - \varepsilon.$$

Let us now define the stopping time $\tau^{(\varepsilon)}$ be defined as the first time one of the absolute values of A_t, B_t, C_t, D_t and $\mu_t, \ln\sigma_t$ becomes larger than $M(\varepsilon)$, or 1 if the absolute values of the processes remain below $M(\varepsilon)$ all the time. Then

$$P\left(\left[\tau^{(\varepsilon)} = 1\right]\right) > 1 - \varepsilon. \tag{17}$$

Let $r_{i,n} = \left(X_{i/n} - X_{(i-1)/n}\right)$ and $s_{i,n} = \left(W_{i/n} - W_{(i-1)/n}\right)$. Then let

$$\rho_n = \sup_i \frac{r_i^2}{(r_{i-1}^2 + r_{i-2}^2 + ..r_{i-\ell}^2)/\ell},$$

$$\rho_n^{(\varepsilon)} = \sup_{i\le\tau^{(\varepsilon)}} \frac{r_i^2}{(r_{i-1}^2 + r_{i-2}^2 + ..r_{i-\ell}^2)/\ell},$$

$$\xi_n = \sup_i \frac{s_i^2}{(s_{i-1}^2 + s_{i-2}^2 + ..s_{i-\ell}^2)/\ell},$$

and

$$\xi_n^{(\varepsilon)} = \sup_{i\le\tau^{(\varepsilon)}} \frac{s_i^2}{(s_{i-1}^2 + s_{i-2}^2 + ..s_{i-\ell}^2)/\ell}.$$

Then - by definition, ρ_n and ξ_n are our test statistics applied to $X_{i/n}$ and $W_{i/n}$, respectively. Moreover, (17) guarantees that

$$P\left(\left[\rho_n = \rho_n^{(\varepsilon)}\right]\right) > 1 - \varepsilon$$

and

$$P\left(\left[\xi_n = \xi_n^{(\varepsilon)}\right]\right) > 1 - \varepsilon,$$

too. Hence it is sufficient to show that for all $\varepsilon > 0$ the difference between converges to zero. For showing this, let us first observe that

$$\min(\frac{\sigma^2_{(i-k)/n}}{\sigma^2_{i/n}}) \le \frac{s_i^2}{(s_{i-1}^2 + s_{i-2}^2 + ..s_{i-\ell}^2)/\ell} / \frac{\sigma^2_{i/n}s_i^2}{(\sigma^2_{(i-1)/n}s_{i-1}^2 + \sigma^2_{(i-2)/n}s_{i-2}^2 + ..\sigma^2_{(i-\ell)}s_{i-\ell}^2)/\ell} \le \max(\frac{\sigma^2_{(i-k)/n}}{\sigma^2_{i/n}}).$$

For analyzing the difference of the left and right side of the above inequality and one, it is sufficient to consider

$$\sup_{k\le\ell}\left|\ln(\frac{\sigma^2_{(i-k)/n}}{\sigma^2_{i/n}})\right|.$$

Now observe that $\ln(\sigma^2_{i/n}) - \ln(\sigma^2_{(i-k)/n}) = \int_{(i-k)/n}^{i/n} C_t dt + D_t dV_t^{(2)}$. For $i < \tau^{(\varepsilon)}$ $\left|\int_{(i-k)/n}^{i/n} C_t dt\right| \le kM/n$.Moreover, we have due to Lemma 7

$$P\left(\left[\left|\int_{(i-k)/n}^{i/n} D_t dV_t^{(2)}\right| > 2\sqrt{M\ell}\sqrt{\frac{\ln n}{n}}\right]\right) \le \frac{1}{n^2}$$

and hence

$$P\left(\left[\sup_{i\le\tau^{(\varepsilon)},k\le\ell}\left|\int_{(i-k)/n}^{i/n} D_t dV_t^{(2)}\right| > 2\sqrt{M\ell}\sqrt{\frac{\ln n}{n}}\right]\right) \le \frac{\ell}{n} \to 0.$$

Hence we can conclude that

$$P\left(\left[\sup_{k\le\ell}\left|\ln(\frac{\sigma^2_{(i-k)/n}}{\sigma^2_{i/n}})\right| > 4\sqrt{M\ell}\sqrt{\frac{\ln n}{n}}\right]\right) \to 0.$$

Since

$$\sup\frac{s_i^2}{(s_{i-1}^2+s_{i-2}^2+..s_{i-\ell}^2)/\ell} = O(\ln n),$$

we can conclude that the difference between

$$\sup\frac{s_i^2}{(s_{i-1}^2+s_{i-2}^2+..s_{i-\ell}^2)/\ell}$$

and

$$\sup\frac{\sigma^2_{i/n}s_i^2}{(\sigma^2_{(i-1)/n}s_{i-1}^2+\sigma^2_{(i-2)/n}s_{i-2}^2+..\sigma^2_{(i-\ell)}s_{i-\ell}^2)/\ell}$$

converges to zero.

It now remains to show that the differences

$$\left|r_{i,n}-\sigma_{(i-1)/n}s_{i,n}\right|$$

remain small. Now observe that

$$\begin{aligned}\left|r_{i,n}-\sigma_{(i-1)/n}s_{i,n}\right| &= \left|\int_{(i-1)/n}^{i/n}\left(\mu_t dt+\sigma_t dW_t-\sigma_{(i-1)/n}dW_t\right)\right| \\ &= \left|\int_{(i-1)/n}^{i/n}\mu_t dt\right| + \left|\int_{(i-1)/n}^{i/n}\left(\sigma_t-\sigma_{(i-1)/n}\right)dW_t\right| \\ &\le \max\left|\mu_t\right|\frac{1}{n} + \left|\int_{(i-1)/n}^{i/n}\left(\sigma_u-\sigma_{(i-1)/n}\right)dW_u\right|.\end{aligned}$$

For the analysis of

$$\left|\int_{(i-1)/n}^{i/n}\left(\sigma_u-\sigma_{(i-1)/n}\right)dW_u\right|$$

we will apply Lemma 7. Since σ_u is a diffusion process, where drift and diffusion coefficients were assumed to be bounded, we can conclude that for all $\alpha > 0$ there exists a M so that

$$P\left(\text{for all } i \text{ and } (i-1)/n \le u \le i/n \ \left|\sigma_u-\sigma_{(i-1)/n}\right| \le M\left|u-(i-1)/n\right|^{1/2-\alpha}\right) \to 1.$$

Hence

$$P\left[\left(\int_{(i-1)/n}^{i/n}\left(\sigma_u-\sigma_{(i-1)/n}\right)^2 du\right) \le 2Mn^{-2+\alpha}\right] \to 1.$$

To apply Lemma 7, however, we need to guarantee an uniform bound on the integral $\int_{(i-1)/n}^{i/n} \left(\sigma_u - \sigma_{(i-1)/n}\right)^2 du$. This can easily be achieved by using a stopping time.

We stop the process at time S, where

$$i/n \geq S \geq (i-1)/n,$$

if for the first time

$$\int_{(i-1)/n}^{S} \left(\sigma_u - \sigma_{(i-1)/n}\right)^2 du = 2Mn^{-2+\alpha},$$

otherwise we set

$$S = 1.$$

Obviously the definition of M guarantees that

$$P\left(S = 1\right) \geq 1 - \varepsilon$$

Hence if we define

$$\sigma_u^* = \begin{cases} \sigma_u \text{ for } u \leq S \\ \sigma_S \text{ otherwise,} \end{cases}$$

we have

$$P\left(\left[\sigma_u^* = \sigma_u \text{ for all } u\right]\right) \geq 1 - \varepsilon.$$

Hence it is sufficient to give estimates for $\int_{(i-1)/n}^{i/n} \left(\sigma_u^* - \sigma_{(i-1)/n}^*\right) dW_u$. For this task, however, we can apply Lemma 7 and conclude that

$$P\left(\left|\int_{(i-1)/n}^{i/n} \left(\sigma_u^* - \sigma_{(i-1)/n}^*\right) dW_u\right| > \sqrt{8Mn^{-2+\alpha}\ln n}\right) \leq \frac{2}{n^2}$$

Since $\alpha > 0$ was arbitrary, we can conclude that for arbitrary $\beta > 0$

$$P\left[\sup\left|\int_{(i-1)/n}^{i/n} \left(\sigma_u^* - \sigma_{(i-1)/n}^*\right) dW_u\right| > n^{-1+\beta}\right] \to 0,$$

which demonstrates that these terms are negligible. ■

C Data

We use the FOREX historical database which has intra-daily transactions data of stock indices and foreign exchange rates. We use the intra-daily data of US Dollar Index(DXA0), USD/JPY(Japanese yen, JPYA0), EUR/USD(EURA0), GBP/USD(British pound, GBPA0), and USD/CAN(Canadian dollar, CADA0). Since the European and US foreign exchange markets are larger than other markets, we first consider the data from 2:00 to 16:00 eastern time. We exclude the first and last 30 minutes data because we observe relatively infrequent trading. We also consider the whole 24 hour data including data from the Asian and pacific market.

For stock indices, we analyze Dow Jones Industrial Average (DJIA), NASDAQ Composite Index (COMPQ), NASDAQ-100 Index(NDX)[3], S&P 500 Index(SPX), S&P 100 Index(OEX), and Russell 2000 Index (RUT). Contrary to other market value weighted indices, the Dow Jones index is a price weighted index and represents well-established blue-chip stocks. The NASDAQ Composite is the index of all of the common stocks and similar securities listed on the NASDAQ stock market, so it measures the performance of technology stocks. The NASDAQ-100 is the index of 100 of the largest non-financial companies listed on the NASDAQ. The S&P 500 is a large-cap stock market index of 500 of largest common stocks actively traded in the US stock market. The S&P 100 chooses 100 largest companies in the S&P 500 considering sector balance. The Russell 2000 Index is a small-cap stock market index of the bottom 2,000 stocks in the Russell 3000 Index which measures the performance of the small-cap segment of the US stock market. Since US stock market opens at 9:30 and closes at 16:00 eastern time, we consider that time span. But we exclude the first and last 30 minutes because of infrequent transactions.

We also use the New York Stock Exchange (NYSE) Trade and Quote (TAQ) database which covers intra-daily transactions data for securities listed on the major stock exchanges. Because of limited accessibility, we mainly consider data of 2005 year. Dow Jones 30 stocks are chosen as main subjects because they are generally leading blue-chip stocks represent their industry and constitute a popular stock market indicator, Dow Jones Industrial Average (DJIA). We use the transaction data from the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and NASDAQ National Market System (NMS). (We choose data whose Ex field is

[3]Note that NASDAQ100(NDX) starts from 2.24. 1998 in FOREX database. So number of sample is smaller.

"N","A", or "T") Furthermore, only regular way sales are selected. We exclude special sales like Bunched sales (B), Automatic Executed sales(E), and Burst Basket Executed sales(F). TAQ database deals with intra-daily data which may have trading error or canceled transactions. By choosing trades whose CORR field is equal to either zero or one, we exclude erroneous data like cancelled trades and obvious error records. About 99.71 percentage data have the proper CORR field. When we see multiple trades with different prices at the same time, we choose a volume-weighted average of the trade price. Since we consider jumps, we did not filter data based on the size of price change.[4] During the opening and closing hours, we observe volatile movement of prices. We consider rather a clean time horizon excluding near opening and closing hours : from 10:30 to 15:30. We also exclude holidays and some trading days which had few transactions, say Labor day, Thanksgiving day, Black Friday, Christmas, and so on. On December 1th, AT&T substitutes SBC. We used the return of SBC until Nov. 30th and used that of AT&T after December 1st.

[4]Standard filtering rules exlcude trades which are less than 50% or greater 150% of the previous prices. (Boehmer, Saar, and Yu (2005))

D Tables and Figures

Following tables summarize rejection probabilities under 5% size.

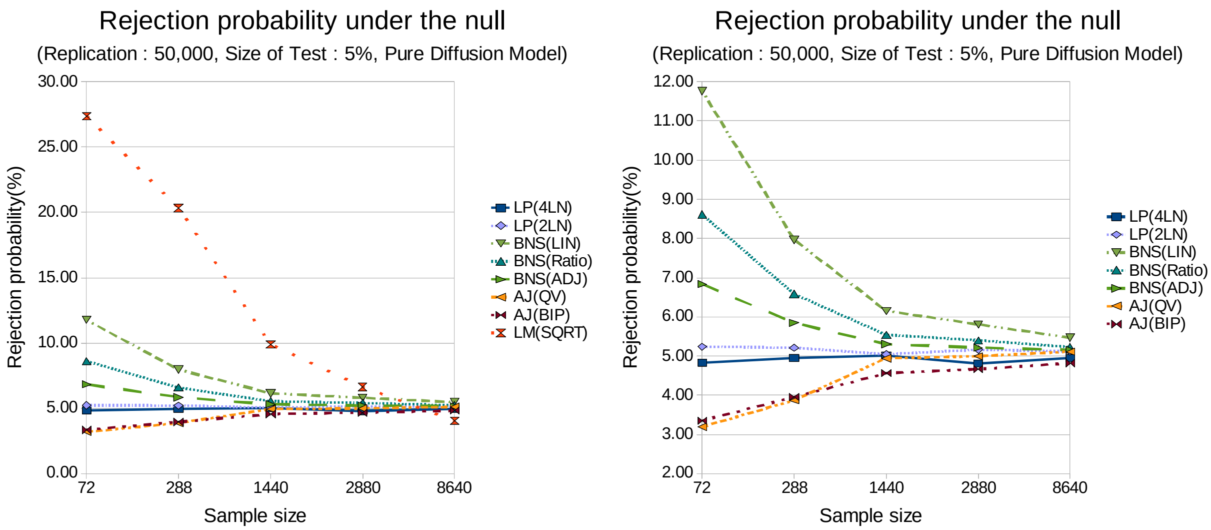

Figure 1 : Simulated rejection probability under the null

(model 1, 5% level of significance)

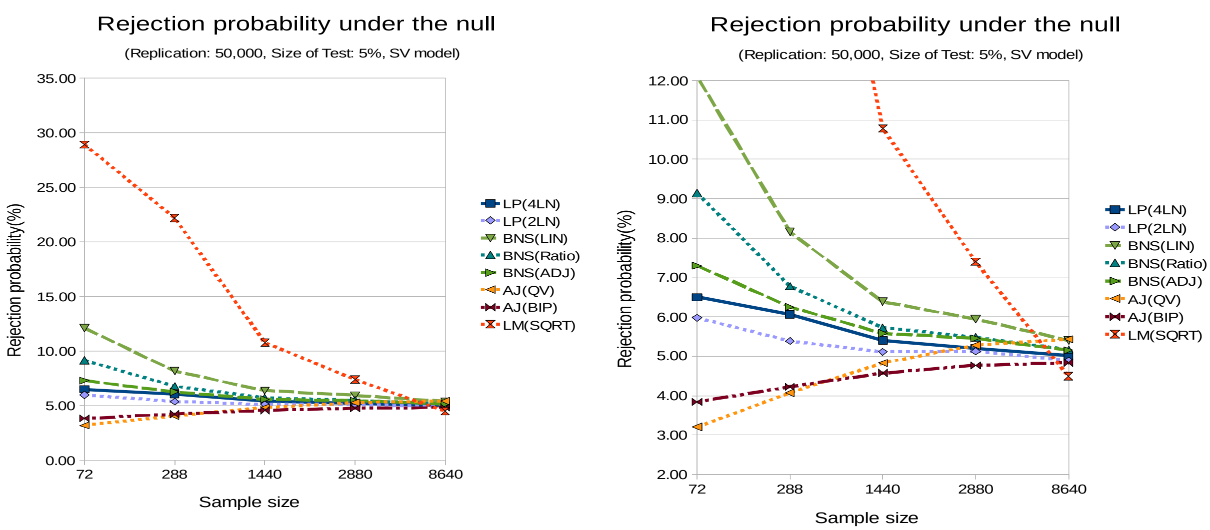

Figure 2 : Simulated rejection probability under the null

(model 2, 5% level of significance)

Model1-1 : Pure Diffusion W/O JUMP										
	LP		BNS			AJ		LM		
n	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
72	4.83	5.24	11.76	8.61	6.83	3.19	3.35	27.37	6.03	27.37
288	4.95	5.21	7.97	6.58	5.85	3.87	3.95	20.34	11.03	40.13
1440	5.02	5.05	6.15	5.54	5.31	4.94	4.56	9.90	15.73	61.03
2880	4.81	5.16	5.81	5.40	5.21	4.99	4.67	6.63	19.23	71.74
8640	4.95	5.12	5.47	5.23	5.15	5.11	4.82	4.05	23.10	80.10
Model1-2 : Pure Diffusion W 20% JUMP										
	LP		BNS			AJ		LM		
n	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
72	35.55	27.89	32.84	28.88	26.36	6.00	9.65	50.71	35.33	50.71
288	58.83	52.54	47.96	46.13	45.32	19.13	35.31	66.41	62.32	74.81
1440	78.67	74.84	64.92	64.31	64.06	34.56	67.75	80.79	82.15	91.81
2880	84.23	81.38	70.71	70.33	70.20	38.68	76.21	85.34	87.44	95.65
8640	90.36	88.75	78.27	78.09	78.05	42.75	84.93	90.84	92.63	98.14
Model1-3 : Pure Diffusion W LN(N)/N JUMP										
	LP		BNS			AJ		LM		
n	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
72	14.18	10.86	17.42	13.74	11.46	3.83	5.11	35.01	14.89	35.01
288	14.27	11.18	12.16	10.33	9.45	5.23	6.94	29.52	20.79	47.41
1440	14.69	11.12	8.36	7.64	7.31	6.30	7.73	20.64	25.87	66.13
2880	14.45	11.00	7.49	7.02	6.81	6.22	7.51	17.83	29.24	75.63
8640	14.86	11.43	6.39	6.13	6.05	6.17	7.10	15.90	33.18	82.89

Table 1 : Simulated rejection probability of Model 1

Model2-1 : CIR SV-Diffusion W/O JUMP										
	LP		BNS			AJ		LM		
n	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
72	6.50	5.98	12.10	9.14	7.30	3.20	3.84	28.91	7.85	28.91
288	6.07	5.39	8.15	6.78	6.25	4.07	4.23	22.18	12.80	41.27
1440	5.40	5.11	6.38	5.73	5.58	4.83	4.57	10.78	16.77	61.29
2880	5.20	5.12	5.94	5.48	5.46	5.28	4.77	7.40	19.82	71.99
8640	5.02	4.89	5.40	5.17	5.15	5.43	4.84	4.49	23.64	80.36
Model2-2 : CIR SV-Diffusion W 20% JUMP										
	LP		BNS			AJ		LM		
n	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
72	37.56	29.86	33.74	29.85	27.69	5.74	9.72	52.50	37.66	52.50
288	59.69	53.28	48.15	46.28	45.70	18.33	33.97	67.73	63.67	75.70
1440	79.03	75.34	64.94	64.28	64.18	34.22	66.94	81.09	82.50	91.73
2880	84.40	81.67	70.78	70.39	70.35	38.18	75.41	85.59	87.62	95.64
8640	90.46	88.86	78.10	77.93	77.92	42.32	84.41	91.02	92.84	98.19
Model2-3 : CIR SV-Diffusion W LN(N)/N JUMP										
	LP		BNS			AJ		LM		
n	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
72	16.53	12.50	18.18	14.57	12.38	3.71	5.45	37.27	17.33	37.27
288	16.03	12.10	12.59	10.88	10.18	5.26	6.94	31.85	23.20	48.94
1440	16.38	12.26	8.65	7.88	7.73	6.19	7.54	22.65	28.01	67.12
2880	16.27	12.18	7.79	7.23	7.19	6.57	7.60	19.88	30.98	76.15
8640	16.27	12.42	6.38	6.11	6.10	6.42	7.07	17.59	34.57	83.38

Table 2 : Simulated rejection probability of Model2

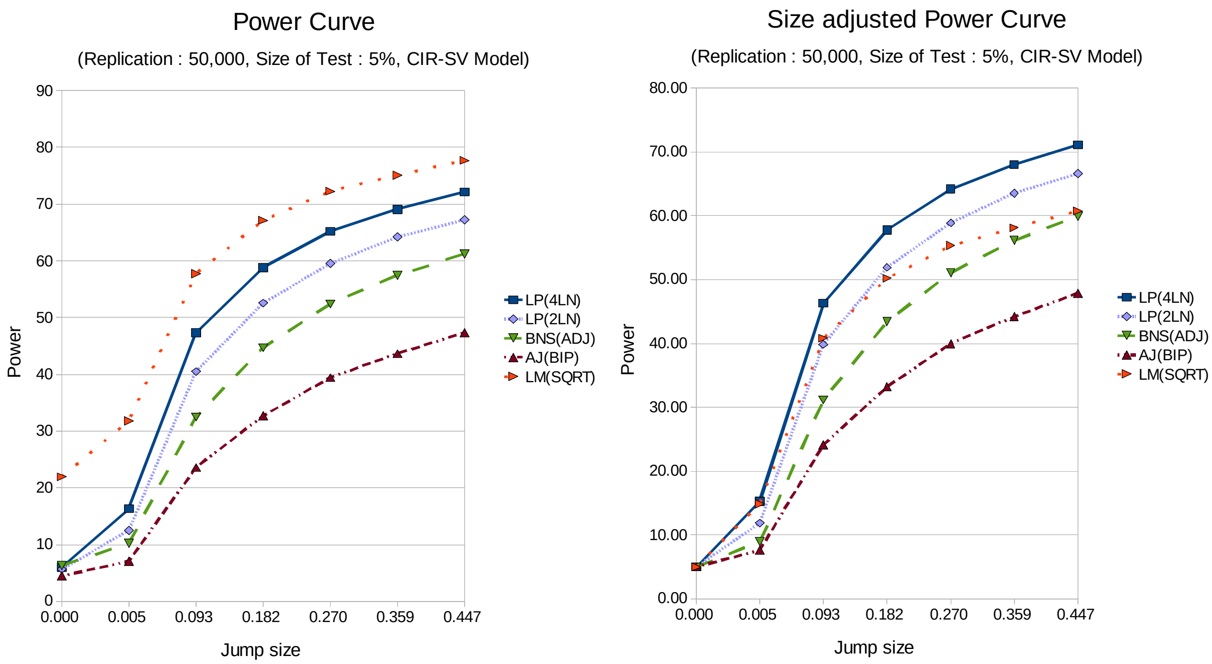

Figure 3 : Simulated power curve of Model 2 with 5 min frequency

		LP		BNS			AJ		LM		
Jump/MeanVol	Meaning	4LN	2LN	LIN	Ratio	ADJ	QV	BIP	SQRT	4LN	2LN
0.000	No Jump	6.04	5.65	8.38	6.88	6.31	4.48	4.48	21.93	12.46	41.40
0.005	LN(N)/N	16.31	12.50	12.77	10.85	10.26	5.44	7.09	31.82	23.22	49.53
0.093		47.31	40.50	35.31	33.19	32.48	13.30	23.61	57.75	52.29	68.49
0.182		58.81	52.55	47.25	45.35	44.72	18.00	32.69	67.13	62.87	75.52
0.270		65.21	59.52	54.66	52.92	52.37	20.88	39.43	72.26	68.53	79.29
0.359		69.08	64.19	59.46	57.96	57.44	23.33	43.67	75.05	72.12	81.43
0.447	20%	72.11	67.22	63.18	61.70	61.24	25.17	47.34	77.70	74.90	83.27
Size Distortion		1.04	0.65	3.38	1.88	1.31	-0.52	-0.52	16.93	7.46	36.40
Size Adjusted Power											
Jump/MeanVol											
0.000	No Jump	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
0.005	LN(N)/N	15.27	11.85	9.38	8.97	8.95	5.96	7.61	14.88	15.76	13.13
0.093		46.27	39.85	31.92	31.31	31.17	13.82	24.13	40.81	44.83	32.09
0.182		57.77	51.90	43.87	43.47	43.41	18.53	33.21	50.20	55.41	39.11
0.270		64.17	58.87	51.28	51.04	51.06	21.41	39.96	55.33	61.07	42.88
0.359		68.04	63.54	56.08	56.09	56.13	23.86	44.20	58.11	64.66	45.03
0.447	20%	71.07	66.58	59.79	59.82	59.93	25.70	47.87	60.76	67.44	46.87

Table 3 : Simulated rejection prob of Model 2 with 5 min frequency

15min	CODE	LP(4LN)	LP(2LN)	BNS	AJ	LM	n/day	days
USD	DXA0	46.22(4H)	36.61(2H)	28.83	16.56	70.76	51.3	489
JPY	JPYA0	44.02(4H)	34.06(2H)	22.11	20.32	69.52	51.7	502
EUR	EURA0	48.21(4H)	39.64(2H)	30.08	15.94	73.51	51.7	502
GBP	GBPA0	39.84(4H)	31.47(2H)	23.11	16.33	66.73	51.7	502
CAN	CADA0	48.70(4H)	36.73(2H)	27.15	19.16	74.45	51.6	501
5min	CODE	LP(4LN)	LP(2LN)	BNS	AJ	LM	n/day	days
USD	DXA0	65.03(54M)	50.31(44M)	33.33	22.09	83.03	151.6	489
JPY	JPYA0	64.34(55M)	54.38(45M)	32.67	37.65	85.66	154.1	502
EUR	EURA0	62.95(55M)	52.19(45M)	35.06	27.29	85.06	154.4	502
GBP	GBPA0	55.18(55M)	46.61(45M)	30.68	31.08	80.28	153.7	502
CAN	CADA0	60.68(54M)	55.89(44M)	26.75	45.51	85.23	150.9	501

Table 4 : Empirical Rejection ratio of FX rates (13HR)

15min	CODE	LP(4LN)	LP(2LN)	BNS	AJ	LM	n/day	days
USD	DXA0	63.80(4H31M)	47.65(2H20M)	31.29	19.22	80.98	82.9	489
JPY	JPYA0	58.17(4H41M)	47.61(2H26M)	27.49	24.50	78.29	89.8	502
EUR	EURA0	66.33(4H41M)	49.00(2H26M)	31.27	20.32	84.46	89.9	502
GBP	GBPA0	63.55(4H40M)	50.80(2H26M)	28.88	21.31	85.06	89.1	502
CAN	CADA0	65.07(4H37M)	52.69(2H23M)	27.74	28.34	89.62	87.2	501
5min	CODE	LP(4LN)	LP(2LN)	BNS	AJ	LM	n/day	days
USD	DXA0	72.8(1H53M)	61.35(58M)	41.31	25.77	90.59	238.9	489
JPY	JPYA0	83.67(1H54M)	71.71(59M)	39.64	50.00	95.02	265.6	502
EUR	EURA0	79.88(1H54M)	66.53(59M)	41.43	35.26	92.43	265.5	502
GBP	GBPA0	75.3(1H54M)	66.93(59M)	36.25	36.45	91.24	259.5	502
CAN	CADA0	76.45(1H52M)	66.47(58M)	27.35	57.49	93.41	240.7	501

Table 5 : Empirical Rejection ratio of FX rates (24HR)

15min	CODE	LP(4LN)	LP(2LN)	BNS	AJ	LM	n/day	days
DJIA	INDU	17.38(3H14M)	12.03(1H45M)	13.37	7.49	42.51	21.8	748
NASDAQ COMP	COMPQ	23.93(3H15M)	13.1(1H45M)	12.17	4.01	44.39	21.9	748
NASDAQ100	NDX	21.09(3H15M)	12.85(1H45M)	10.89	5.45	38.13	21.9	716
S&P500	SPX	17.87(3H15M)	10.93(3H15M)	14.53	5.87	39.73	21.9	750
S&P100	OEX	18.00(3H15M)	10.93(1H45M)	12.80	8.00	41.73	21.9	750
RUSSEL2000	RUT	29.12(3H15M)	20.61(3H15M)	10.24	3.32	52.26	21.9	752
5min	CODE	LP(SQRT)	LP(2LN)	BNS	AJ	LM	n/day	days
DJIA	INDU	24.87(1H25M)	17.25(45M)	21.12	8.16	47.46	65.3	748
NASDAQ COMP	COMPQ	31.02(1H25M)	18.98(45M)	19.52	3.61	53.34	64.6	748
NASDAQ100	NDX	24.16(1H25M)	14.11(45M)	15.92	4.89	46.09	65.5	716
S&P500	SPX	27.73(1H25M)	20.4(45M)	26.27	6.53	50.67	65.5	750
S&P100	OEX	26.4(1H25M)	17.47(45M)	20.53	8.00	47.07	65.5	750
RUSSEL2000	RUT	41.89(1H25M)	30.32(45M)	9.97	1.33	61.04	65.4	752

Table 6 : Empirical Rejection ratio of stock indices

freq	LP(4LN)	LP(2LN)	BNS	AJ	LM
15min	0.2223(3H15M)	0.1823 (1H45M)	0.1471	0.1404	0.5559
5min	0.3197(1H25M)	0.2418(45M)	0.1701	0.1694	0.6315

Table 7 : Empirical Rejection ratio of stocks of Dow30 (Average)