A Hausman Test for the Presence of Market Microstructure Noise in High Frequency Data*

Yacine Aït-Sahalia†
Department of Economics
and Bendheim Center for Finance
Princeton University and NBER

Dacheng Xiu‡
Booth School of Business
University of Chicago

This Version: February 17, 2017

Abstract

We develop tests that help assess whether a high frequency data sample can be treated as reasonably free of market microstructure noise at a given sampling frequency for the purpose of implementing high frequency volatility and other estimators. The tests are based on the Hausman principle of comparing two estimators, one that is efficient but not robust to the deviation being tested, and one that is robust but not as efficient. We investigate the asymptotic properties of the test statistic in a general nonparametric setting, and compare it with several alternatives that are also developed in the paper. Empirically, we find that improvements in stock market liquidity over the past decade have increased the frequency at which simple, uncorrected, volatility estimators can be safely employed.

Keywords: Hausman test, market microstructure noise, realized volatility, QMLE, TSRV, pre-averaging, super-efficiency, local power.
JEL Codes: C13, C14, C55, C58, G01.

1 Introduction

In theory, volatility estimation using high frequency data is a straightforward matter: summing the squares of log-returns should produce a consistent and efficient estimator, which is called as realized volatility (RV). In practice, however, matters are a bit more complicated. The idiosyncrasies of the

*We benefitted from the helpful comments of the Editor and referees, and from useful discussions with Jean Jacod, Yuta Koike and Jia Li. We thank Chaoxing Dai and Rui Da for excellent research assistance.

†Address: JRR Building, Princeton, NJ 08544, USA. E-mail address: yacine@princeton.edu.

‡Address: 5807 S Woodlawn Avenue, Chicago, IL 60637, USA. E-mail address: dacheng.xiu@chicagobooth.edu.
Xiu gratefully acknowledges financial support from the IBM Corporation Faculty Scholar Fund at the University of Chicago Booth School of Business.

trading process, including the facts that buy and sell orders execute at different prices, that prices live on a discrete tick grid, that some orders face sudden changes in the order book due to high frequency traders' actions, etc., all mean that the observed transaction prices as well as the price measures implied from quotes, are to be taken with some caution. As the frequency of observation increases, market microstructure noise typically becomes more significant, degrading the signal to noise ratio in the data to the point where the limit of the RV estimator changes to primarily reflect the presence of the noise rather than the volatility of the underlying price.

The first response in the literature has been to down-sample to a level considered "safe" from the worst effects of market microstructure noise, with the recommendation to sample every 5 or 15 minutes and not more frequently (see Andersen et al. (2001)). This approach has been criticized on the grounds that it is rarely optimal in econometrics to discard large amounts of data that are otherwise available, even if the data are known to be noisy (see Aït-Sahalia et al. (2005)). This has led to the development of many noise-robust estimators (see, e.g., Aït-Sahalia and Jacod (2014), Chapter 7) that make use of the full data sample, which as time passes is more and more likely to be available at a sub-second frequency, resulting in large quantities of data that would otherwise be discarded when sampling only every 5 or 15 minutes.

Noise-robust estimators are undoubtedly more complex to employ than the RV estimator, leading to the following natural question: is the effort worth it? In other words, can we determine whether a particular sample at a given frequency of observation is "safe" from market microstructure noise, so that RV is a suitable volatility estimator for that sample? This paper addresses this question by developing a test for this purpose. The test we propose compares RV to a noise-robust volatility estimator. The metric we employ for the comparison is based on an idea due to Hausman (1978).

Hausman (1978) proposed a series of specification tests for a null hypothesis against an alternative that rely on the following clever idea. Suppose that we have access to two estimators, A and B, of the same quantity or parameter, say β, with the following properties:

	Estimator A	Estimator B
$\mathbb{H}_0$: Null	Consistent and Efficient	Consistent
$\mathbb{H}_1$: Alternative	Inconsistent	Consistent

So both estimate the same β with ($\widehat{\beta}_A$) and without ($\widehat{\beta}_B$) imposing the restriction embedded in $\mathbb{H}_0$. The difference $\widehat{\beta}_B - \widehat{\beta}_A$ should be small under $\mathbb{H}_0$ but large under $\mathbb{H}_1$. But since $\widehat{\beta}_A$ and $\widehat{\beta}_B$ use the same data, they are likely to be correlated, leading a priori to a rather messy variance of $\widehat{\beta}_B - \widehat{\beta}_A$. One of key insights of Hausman (1978) is that the efficiency of $\widehat{\beta}_A$ under $\mathbb{H}_0$ implies that $\widehat{\beta}_B - \widehat{\beta}_A$ and $\widehat{\beta}_A$ must be asymptotically uncorrelated. Otherwise, a more efficient estimator could be constructed by linearly combining $\widehat{\beta}_A$ and $\widehat{\beta}_B$. So

$$\text{AVAR}(\widehat{\beta}_B - \widehat{\beta}_A) = \text{AVAR}(\widehat{\beta}_B) - \text{AVAR}(\widehat{\beta}_A). \tag{1}$$

With $V = \text{AVAR}(\widehat{\beta}_B) - \text{AVAR}(\widehat{\beta}_A)$ and $\widehat{V}$ a consistent estimator of it, a Hausman test statistic can be constructed in the form

$$H_n = n\left(\widehat{\beta}_B - \widehat{\beta}_A\right)^{\intercal} \widehat{V}^{-1} \left(\widehat{\beta}_B - \widehat{\beta}_A\right) \tag{2}$$

with $^{\intercal}$ denoting transposition and $^{-1}$ denoting the pseudo-inverse of the matrix. From (1), all that is needed to compute $\widehat{V}$ and hence H_n are the separate asymptotic variances of $\widehat{\beta}_A$ and $\widehat{\beta}_B$. Their covariance does not enter the calculation.

Many applications of this principle have been proposed, starting in Hausman (1978), where it was applied to the problem of detecting the potential endogeneity of a set of regressors, which is achieved by comparing (A), ordinary least squares (OLS) estimates, to (B), instrumental variables (IV) estimates; checking whether a set of extra instruments are valid, achieved by comparing (A), IV estimates with a large set of instruments, to (B), IV estimates with a subset of instruments; in panel data, comparing (A), random effects (RE) estimates, to (B), fixed effects estimates (FE) where in the RE case, the generalized least squares-type RE estimator is efficient for Gaussian errors but in the FE case, the RE estimator is inconsistent because of the omitted variable.

Hausman and Taylor (1981) extended the analysis to the case where V is possibly singular, hence the use of a pseudo-inverse above. White (1980) proposed to detect heteroscedasticity by comparing (A), OLS standard errors, to (B), heteroscedasticity-robust standard errors. Hausman and Pesaran (1983) proposed a $J-$test for testing two non-nested linear regression models, by comparing (A), OLS in the first model, to (B), OLS in an artificial model where the fitted values from the second model are added as regressors to the original regressors of the first model.

Aït-Sahalia (1996) constructed a test for the specification of continuous-time models using discrete data by comparing (A), the implied parametric density estimator at the frequency of observation from the assumed continuous-time model, to (B), a nonparametric density estimator constructed from the discrete data without reference to the assumed model. Hahn and Hausman (2002) tested whether first order asymptotics are satisfactory by comparing forward and reverse 2SLS, which have the same first order but different second order asymptotics due to second-order bias. Hausman et al. (2005) developed a test to determine whether instruments are strong or weak. A local power analysis of Hausman tests is due to Holly (1982) and equivalent formulations of the test to Holly and Monfort (1986).

The present paper applies the principles behind Hausman tests to the problem of testing for the presence of market microstructure noise in high frequency data. In a typical model for high frequency data, transaction log-prices observed at high frequency from 0 to T, at times $0, \Delta_n, 2\Delta_n, \ldots, n\Delta_n = T$, consist of an unobservable fundamental price $X_{i\Delta_n}$ plus some noise component U_i due to the imperfections of the trading process (see, e.g., for instance Black (1986))

$$\tilde{X}_{i\Delta_n} = X_{i\Delta_n} + U_i. \tag{3}$$

U summarizes a diverse array of market microstructure effects, either informational or not: bid-ask bounces, discreteness of price changes, differences in trade sizes or informational content of price changes, gradual response of prices to a block trade, the strategic component of the order flow, inventory control effects, transient liquidity issues, fleeting quotes by high frequency market makers, mini flash crashes, data feed errors, etc.

The "parameter" β of interest is now a random variable, the quadratic variation of the fundamental log-price process, denoted as σ^2_{QV}. Without noise, the realized volatility (RV) of the process, which is simply the sum of squares of log-returns

$$\widehat{\sigma}^2_{\text{RV}} = \frac{1}{T} \sum_{i=1}^{n} \left(\tilde{X}_{i\Delta_n} - \tilde{X}_{(i-1)\Delta_n} \right)^2, \tag{4}$$

estimates the quadratic variation.

In theory, sampling as often as possible ($\Delta_n \to 0$) will produce in the limit a perfect estimate $\widehat{\sigma}^2_{\text{RV}}$ of σ^2_{QV} in the absence of noise ($U \equiv 0$), as first shown theoretically in Jacod (1994). In the presence of market microstructure noise, however, $\widehat{\sigma}^2_{\text{RV}}$ diverges as $\Delta_n \to 0$ instead of converging to σ^2_{QV}. Indeed, since each transaction adds its own noise component, a log-return over a tiny time interval Δ_n is mostly composed of market microstructure noise, while the informational content of the log-return in variance terms is proportional to Δ_n. As Δ_n increases, the amount of noise in each log-return remains the same, since each price is measured with error, while the informational content of volatility increases and the estimator becomes less biased (see Aït-Sahalia et al. (2005)).

At what frequency does this effect start to matter, to the point that noise-robust estimators of σ^2_{QV} should be employed instead of $\widehat{\sigma}^2_{\text{RV}}$? The Hausman test we construct compares two estimators of σ^2_{QV}. The first, (A), is efficient if there is no noise, while the other, (B), is inefficient if there is no noise, but robust to the presence of noise. The test we propose can be considered a formalization and an improvement of the visual "signature plot" procedure of Andersen et al. (2000) (see also Patton (2011)), which depicts the divergence of $\widehat{\sigma}^2_{\text{RV}}$ as a function of the sampling frequency, just like Hausman (1978)'s test for the endogeneity of the regressors was a formalization and an improvement of Sargan (1958)'s recommendation to check whether OLS lies outside IV's confidence interval. This is illustrated in Figure 1: as the sampling frequency increases, $\widehat{\sigma}^2_{\text{RV}}$ diverges in the presence of noise whereas noise-robust estimators do not. The test compares the two and measures whether this divergence is significant or not.

The general recommendation to sample every 5 minutes when the data are noisy was thoroughly investigated in the recent paper Liu et al. (2015), using the ranking method of Patton (2011). Across a range of assets in different classes, they found that the subsampling approach of Zhang et al. (2005) employed to produce 5-minute daily returns volatilities is the preferred method for the purpose of estimating daily volatility. We find that the common practice of treating 5-minute returns as noise-free might be problematic in the earlier years for Dow Jones 30 index and S&P 100 index constituents, but is a reasonably safe choice for data sampled after 2009. For a large portion of S&P 500 index constituents, however, 5-minute returns cannot be treated as noise-free, even in the most recent part of the sample.

The paper is organized as follows. We start in Section 2 by motivating the problem of testing for the presence of market microstructure noise in a parametric context. We then construct and analyze Hausman tests using likelihood-based estimators in increasingly realistic yet complicated nonparametric settings. Section 3 discusses alternative tests that we propose for this problem, including an autocovariance-based test, a Student-t test, and a different Hausman test based on a pre-averaging estimator. Section 4 compares them in finite samples. Section 5 applies these tests to determine for the constituent stocks of the Dow Jones 30 index, as well as those of the S&P 100 and S&P 500, at which frequency $\widehat{\sigma}^2_{\mathrm{RV}}$ can safely be used, and relates the results of the tests to possible measures of market liquidity. Section 6 concludes. The appendix contains the proofs. A web appendix contains additional simulation results.

2 Noise-Robust Estimation with a Parametrically-Motivated Likelihood

2.1 The Parametric Case

We start by considering the simplest possible parametric model for the log price X_t:

$$X_t = \sigma_0 W_t, \tag{5}$$

where W is a Brownian motion and volatility σ_0 is constant. Observations $\tilde{X}$ are potentially contaminated by noise as follows

$$\tilde{X}_{i\Delta_n} = X_{i\Delta_n} + a_n U_i, \tag{6}$$

where U follows an i.i.d. Gaussian distribution with mean 0 and variance 1. Under $\mathbb{H}_0$, $a_n^2 = 0$ while under $\mathbb{H}_1$, $a_n^2 = a_0^2 > 0$ is a constant. This model is certainly too simplistic as a representation of the data, but it turns out that it is very useful to generate a surprisingly robust likelihood function not only to departures from the Gaussianity of the noise (see Aït-Sahalia et al. (2005)) but also from the constancy of the volatility parameter (see Xiu (2010)), and as we will see below provides a test that is applicable even in the presence of jumps.

In the absence of noise, $\widehat{\sigma}^2_{\mathrm{RV}}$ is consistent and achieves the parametric efficiency bound, that is, as $\Delta_n \to 0$,

$$\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\mathrm{RV}} - \sigma_0^2\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 2\sigma_0^4/T\right), \tag{7}$$

as shown in Gloter and Jacod (2001) and Zhang et al. (2005).

When noise is present, $\widehat{\sigma}^2_{\text{RV}}$ becomes inconsistent since

$$\widehat{\sigma}^2_{\text{RV}} = \sigma_0^2 + 2a_0^2\Delta_n^{-1} + o_p(1). \tag{8}$$

So $\widehat{\sigma}^2_{\text{RV}} \to \infty$ as $\Delta_n \to 0$. In this setting, a noise-robust parametric estimator is the maximum likelihood estimator (MLE) proposed by Aït-Sahalia et al. (2005). The observed log-returns

$$Y_i = \tilde{X}_{i\Delta_n} - \tilde{X}_{(i-1)\Delta_n}, \tag{9}$$

are such that

$$Y_i = \sigma_0 \left(W_{i\Delta_n} - W_{(i-1)\Delta_n} \right) + U_i - U_{(i-1)}, \tag{10}$$

$$Y_{i+1} = \sigma_0 \left(W_{(i+1)\Delta_n} - W_{i\Delta_n} \right) + U_{(i+1)} - U_i, \tag{11}$$

where the increments of the Brownian motion W are uncorrelated and the U_i's are independent, and so $\text{Cov}(Y_i, Y_{i+1}) = -a_0^2$, $\text{Cov}(Y_i, Y_{i+2}) = 0$, etc., because of the repetition of the same term U_i in both (10) and (11), but no further repetition of a common term occurs in log-returns more than one lag apart. This implies that the observed log-returns Y_i follow under $\mathbb{H}_1$ an MA(1) process with

$$\text{Var}(Y_i) = \sigma_0^2\Delta_n + 2a_0^2 \quad \text{and} \quad \text{Cov}(Y_i, Y_{i-1}) = -a_0^2. \tag{12}$$

So the proper log-likelihood function for the log-returns is

$$L(\sigma^2, a^2) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\det(\Sigma) - \frac{1}{2}Y^\intercal \Sigma^{-1} Y. \tag{13}$$

with $Y = (Y_1, Y_2, \ldots, Y_n)^\intercal$ and $\Sigma = \sigma^2\Delta_n\mathbb{I}_n + a^2\mathbb{J}_n$ where $\mathbb{I}_n$ to denote the $n \times n$ identity matrix and $(\mathbb{J}_n)_{ij} = -\mathbf{1}_{\{i=j\pm1\}} + 2 \times \mathbf{1}_{\{i=j\}}$ with $\mathbf{1}_{\{\bullet\}}$ denoting the indicator function, is the proper variance-covariance matrix to employ.

From Aït-Sahalia et al. (2005), the likelihood estimator $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{a}^2_{\text{MLE}})$ has the following asymptotic distribution under $\mathbb{H}_1$:

$$\begin{pmatrix} \Delta_n^{-1/4}\left(\widehat{\sigma}^2_{\text{MLE}} - \sigma_0^2\right) \\ \Delta_n^{-1/2}\left(\widehat{a}^2_{\text{MLE}} - a_0^2\right) \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 8a_0\sigma_0^4/T & 0 \\ 0 & 2a_0^4/T \end{pmatrix} \right). \tag{14}$$

Under the null hypothesis, i.e., when noise is absent, $\widehat{\sigma}^2_{\text{MLE}}$ remains asymptotically normal yet at a higher convergence rate, which matches that of $\widehat{\sigma}^2_{\text{RV}}$:

$$\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\text{MLE}} - \sigma_0^2\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 6\sigma_0^4/T\right). \tag{15}$$

The increase in the asymptotic variance in (15) compared to that of (7), by a factor 6 vs. 2, is due to $\widehat{\sigma}^2_{\text{MLE}}$ attempting to control for the presence of noise when in fact there is none.

The problem of testing the null hypothesis of $\mathbb{H}_0 : a_0^2 = 0$ against the alternative $\mathbb{H}_1 : a_0^2 > 0$ now falls into the classic Hausman test paradigm: $\widehat{\sigma}^2_{\text{MLE}}$ is consistent under both the null and alternative hypotheses, but $\widehat{\sigma}^2_{\text{MLE}}$ is not as efficient as $\widehat{\sigma}^2_{\text{RV}}$ under the null, which reaches the parametric efficiency bound under the null but is inconsistent under the alternative. We construct a Hausman test statistic accordingly as

$$H_{1n} = \Delta_n^{-1} \frac{\left(\widehat{\sigma}^2_{\text{MLE}} - \widehat{\sigma}^2_{\text{RV}}\right)^2}{\widehat{V}_{1n}}, \tag{16}$$

where $\widehat{V}_{1n}$ denotes a consistent estimator of the asymptotic variance

$$\text{AVAR}\left(\widehat{\sigma}^2_{\text{MLE}} - \widehat{\sigma}^2_{\text{RV}}\right) = \text{AVAR}\left(\widehat{\sigma}^2_{\text{MLE}}\right) - \text{AVAR}\left(\widehat{\sigma}^2_{\text{RV}}\right). \tag{17}$$

Such an estimator is for instance

$$\hat{V}_{1n} \equiv 4\left(\widehat{\sigma}^2_{\text{RV}}\right)^2 / T, \tag{18}$$

which is consistent under $\mathbb{H}_0$. We will show below (see Corollary 1) that

$$\begin{cases} H_{1n} \xrightarrow{\mathcal{L}} \chi_1^2, \text{ under } \mathbb{H}_0 \\ H_{1n} \xrightarrow{p} \infty, \text{ under } \mathbb{H}_1 \end{cases}, \tag{19}$$

which implies that the proposed Hausman test has asymptotic size control under the null, and is consistent under the alternative.

Next, we investigate the behavior of H_{1n} under the sequence of local alternatives $\mathbb{H}_n : a_n^2 = a_0^2 \Delta_n^{3/2}$. The fact that $\Delta_n^{3/2}$ is the right rate to consider will become apparent later. It follows from Corollary 1 below that the asymptotic distribution of H_{1n} under $\mathbb{H}_n$ is given by:

$$H_{1n} \xrightarrow{\mathcal{L}} \chi_1^2(a_0^4 \sigma_0^{-4} T) \text{ under } \mathbb{H}_n \tag{20}$$

where $\chi_1^2(a_0^4 \sigma_0^{-4} T)$ is a non-central Chi-squared distribution with one degree of freedom and non-centrality parameter $a_0^4 \sigma_0^{-4} T$. Intuitively, detecting the presence of noise becomes easier when the noise is larger ($a_0^2 \uparrow$), the signal volatility smaller ($\sigma_0 \downarrow$), or the time window longer ($T \uparrow$).

2.2 Robustness to Stochastic Volatility and Non-Gaussian Noise

The above analysis certainly relies on a very special model. We now investigate the asymptotic properties of the same Hausman test (16) in a more realistic setting where volatility is possibly stochastic and the microstructure noise is not necessarily Gaussian. Quite remarkably, the same estimator $\widehat{\sigma}^2_{\text{MLE}}$ from what is now potentially a misspecified likelihood can still be employed. In that scenario, the likelihood estimator $\widehat{\sigma}^2_{\text{MLE}}$ from above can be regarded as a quasi-maximum likelihood estimator (QMLE) in the sense of White (1982).

We generalize (5) by supposing that the log price X_t follows a continuous Itô semimartingale, namely

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \tag{21}$$

where b_s is locally bounded, and σ_s is another Itô semimartingale, potentially with jumps. The observed log prices continue to be given by (6)

$$\tilde{X}_{i\Delta_n} = X_{i\Delta_n} + a_n U_i, \quad \text{for} \quad 0 \le i \le n = T/\Delta_n,$$

where U is an i.i.d. noise, now not necessarily Gaussian, but with mean 0, variance 1, and a finite fourth moment.

We use the same estimators $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{a}^2_{\text{MLE}})$ obtained by maximizing the (now possibly misspecified) log-likelihood (13) and consider the same Hausman test statistic as above, i.e., (16)-(17), despite a different asymptotic variance estimator to be given below. In what follows, we investigate the asymptotic properties of Hausman test under the null hypothesis $\mathbb{H}_0 : a_n^2 = 0$, the alternative $\mathbb{H}_1 : a_n^2 = a_0^2 > 0$, and the sequence of local alternatives $\mathbb{H}_n : a_n^2 = a_0^2 \Delta_n^{3/2}$.

Under the alternative $\mathbb{H}_1$, $\widehat{\sigma}^2_{\text{RV}}$ remains inconsistent as before and explodes at the rate $O_p(\Delta_n^{-1})$, similarly to what happens in (8). As for the MLE estimator, Xiu (2010) derived the asymptotic distribution of $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{a}^2_{\text{MLE}})$:

$$\begin{pmatrix} \Delta_n^{-1/4}\left(\widehat{\sigma}^2_{\text{MLE}} - \frac{1}{T}\int_0^T \sigma_s^2 ds\right) \\ \Delta_n^{-1/2}\left(\widehat{a}^2_{\text{MLE}} - a_0^2\right) \end{pmatrix} \tag{22}$$

$$\xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{5a_0\left(\int_0^T \sigma_s^4 ds\right)}{T^{3/2}\left(\int_0^T \sigma_s^2 ds\right)^{1/2}} + \frac{3a_0\left(\int_0^T \sigma_s^2 ds\right)^{3/2}}{T^{5/2}} & 0 \\ 0 & \frac{2a_0^4 + \text{cum}_4[U]}{T} \end{pmatrix}\right),$$

where $\text{cum}_4[U]$ is the fourth cumulant of U, $\mathcal{L}$-s denotes stable convergence in law, and $\mathcal{MN}$ denotes a mixture of normals.

Since the setting is now nonparametric in the sense that the distribution of the log returns under (21) is not specified, the efficiency of $\widehat{\sigma}^2_{\text{RV}}$ is not as well defined as in the parametric setting, see, e.g., Renault et al. (2015). Therefore, we need to analyze the joint asymptotic behavior of $\widehat{\sigma}^2_{\text{MLE}}$ and $\widehat{\sigma}^2_{\text{RV}}$. We start by adopting a re-parametrization of the quasi-log-likelihood (13) from (σ^2, a^2) to (σ^2, η):

$$L(\sigma^2, \eta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\det(\Sigma) - \frac{1}{2}Y^\intercal \Sigma^{-1} Y, \tag{23}$$

where $\Sigma = \sigma^2 \Delta_n \mathbb{I}_n + \eta \Delta_n \mathbb{J}_n$. For $\widehat{\sigma}^2_{\text{MLE}}$ under $\mathbb{H}_0$, the true value of the nuisance parameter a^2 is on the boundary, i.e., $a^2 = 0$, which would lead to a non-Gaussian asymptotic distribution of $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{\eta}_{\text{MLE}})$. To avoid this complication, we extend the parameter space of η to allow negative values, to the extent that the covariance matrix Σ remains positive definite. Based on the new parametrization, it is easy to observe that $\eta > -\sigma^2/4$ is a sufficient condition. We thereby impose in our implementation that the parameter space is a compact set that satisfies this constraint, e.g., $\Theta = \{(\sigma^2, \eta) | 0 < \epsilon_1 \le \sigma^2 + 4\eta \le \epsilon_2, \epsilon_3 \le \sigma^2 \le \epsilon_4\}$.

Next we provide a general result, which we will use to derive the asymptotic distribution of the Hausman test under $\mathbb{H}_0$ as well as under $\mathbb{H}_n$:

Theorem 1. *Suppose that either $a_n^2 = a_0^2\Delta_n^{\gamma_0}$ with either $\gamma_0 \geq 3/2$ and $a_0^2 > 0$ or $\gamma_0 = 0$ and $a_0^2 = 0$ holds. Then the QMLE $(\widehat{\sigma}^2_{\mathrm{MLE}}, \widehat{\eta}_{\mathrm{MLE}})$ that maximizes (23) with respect to (σ^2, η) and the realized volatility $\widehat{\sigma}^2_{\mathrm{RV}}$ jointly satisfy:*

$$\Delta_n^{-1/2} \begin{pmatrix} \widehat{\sigma}^2_{\mathrm{MLE}} - \frac{1}{T}\int_0^T \sigma_s^2 ds \\ \widehat{\eta}_{\mathrm{MLE}} - a_0^2 \Delta_n^{\gamma_0 - 1} \\ \widehat{\sigma}^2_{\mathrm{RV}} - \frac{1}{T}\int_0^T \sigma_s^2 ds \end{pmatrix} \xrightarrow{\mathcal{L}-s} \mathcal{W}_T, \tag{24}$$

where $\mathcal{L}$-s denotes stable convergence in law towards a variable, $\mathcal{W}_T$ is defined on an extension of the original probability space, which conditionally on $\mathcal{F}$, is a three-dimensional Gaussian random variable with covariance matrix given by

$$\mathbb{E}(\mathcal{W}_T \mathcal{W}_T^\intercal | \mathcal{F}) = \frac{1}{T^2}\int_0^T \sigma_s^4 ds \times \begin{pmatrix} 6 & -2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Theorem 1 implies that

$$\mathrm{AVAR}\left(\widehat{\sigma}^2_{\mathrm{MLE}} - \widehat{\sigma}^2_{\mathrm{RV}}\right) = \mathrm{AVAR}\left(\widehat{\sigma}^2_{\mathrm{MLE}}\right) - \mathrm{AVAR}\left(\widehat{\sigma}^2_{\mathrm{RV}}\right) = \frac{4}{T^2}\int_0^T \sigma_s^4 ds, \tag{25}$$

hence the Hausman test statistic (16) in this nonparametric setting remains valid, except that we now need a new consistent estimator for $\mathrm{AVAR}\left(\widehat{\sigma}^2_{\mathrm{MLE}} - \widehat{\sigma}^2_{\mathrm{RV}}\right)$ that is valid under the model (21). For this purpose, we use the following quarticity estimator:

$$\widehat{V}_{2n} = \frac{4}{T^2}\widehat{Q}_{2n}, \text{ where } \widehat{Q}_{2n} = \frac{1}{3\Delta_n}\sum_{i=1}^n Y_i^4 \xrightarrow{p} \int_0^T \sigma_s^4 ds, \text{ under } \mathbb{H}_0. \tag{26}$$

With this $\widehat{V}_{2n}$, we define the test statistic as

$$H_{2n} = \Delta_n^{-1}\frac{(\widehat{\sigma}^2_{\mathrm{MLE}} - \widehat{\sigma}^2_{\mathrm{RV}})^2}{\widehat{V}_{2n}}. \tag{27}$$

We then show that

$$H_{2n} \xrightarrow{\mathcal{L}} \chi_1^2, \quad \text{under } \mathbb{H}_0. \tag{28}$$

Under $\mathbb{H}_1$, $\widehat{V}_{2n} = O_p(\Delta_n^{-2})$, hence

$$H_{2n} = O_p(\Delta_n^{-1}), \quad \text{under } \mathbb{H}_1. \tag{29}$$

We can also calculate the local power for H_{2n}. The size, power, and local power results for the test are summarized in the next corollary:

Corollary 1. *The test statistic H_{2n} has asymptotic size α under the null hypothesis $\mathbb{H}_0 : a_n^2 = 0$ and is consistent under the alternative hypothesis $\mathbb{H}_1 : a_n^2 = a_0^2 > 0$, that is,*

$$\mathbb{P}\left(H_{2n} > c_{1-\alpha} | \mathbb{H}_0\right) \to \alpha \quad \textit{and} \quad \mathbb{P}\left(H_{2n} > c_{1-\alpha} | \mathbb{H}_1\right) \to 1,$$

where $c_{1-\alpha}$ is the $(1-\alpha)$-quantile of the Chi-squared distribution with the degree of freedom being equal to 1. Moreover, H_{2n} follows a noncentral Chi-squared distribution with noncentrality parameter $a_0^4 T^2 \left(\int_0^T \sigma_s^4 ds\right)^{-1}$ and one degree of freedom, under the sequence of local alternative hypotheses $\mathbb{H}_n : a_n^2 = a_0^2 \Delta_n^{3/2}$.

2.3 Robustness to Jumps

Another important restriction we imposed in previous sections is the absence of jumps in log-prices. Yet, the presence of jumps is a salient feature of the data, see, e.g., Aït-Sahalia and Jacod (2012). In this section, we modify the test statistic so that it becomes robust to jumps.

Assume now that, as a further generalization to (16), the log-price X_t follows a possibly discontinuous Itô semimartingale, with the following standard representation:

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s dW_s + (\delta 1_{\|\delta\| \leq 1}) * (\mu - \nu)_t + (\delta 1_{\{\|\delta\| > 1\}}) * \mu_t, \tag{30}$$

where μ is a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}_d$ with the compensator $\nu(dt, dx) = dt \otimes \bar{\nu}(dx)$, and $\bar{\nu}$ is a σ-finite measure. Moreover, $\|\delta(\omega, t, z)\| \wedge 1 \leq \Gamma_n(z)$ for all (ω, t, z) with $t \leq \tau_n(\omega)$, where (τ_n) is a localizing sequence of stopping times and each function Γ_n satisfies $\int \Gamma_n(z)^\gamma \bar{\nu}(dz) < \infty$, for some $\gamma < 1$ (see Aït-Sahalia and Jacod (2014), Chapter 1, for further details, definitions and explanations about the model and the notation.)

Under the dynamics (30), the simple estimators $\widehat{\sigma}^2_{\text{RV}}$ and $\widehat{\sigma}^2_{\text{MLE}}$ we employed above no longer estimate the volatility of the process. Nevertheless, they remain useful candidates to test for the presence of noise, for the following reason. When jumps are present, $\widehat{\sigma}^2_{\text{MLE}}$ and $\widehat{\sigma}^2_{\text{RV}}$ under $\mathbb{H}_0$ now become consistent estimators of the total quadratic variation of the price process, including continuous and jump parts, with $\widehat{\sigma}^2_{\text{RV}}$ being a more efficient estimator. Under $\mathbb{H}_1$, the consistency remains true for $\widehat{\sigma}^2_{\text{MLE}}$ but not for $\widehat{\sigma}^2_{\text{RV}}$, and we can exploit this fact to construct a Hausman test statistic for the presence of noise. More specifically:

Theorem 2. *Suppose that either $a_n^2 = a_0^2 \Delta_n^{\gamma_0}$ with either $\gamma_0 \geq 3/2$ and $a_0^2 > 0$ or $\gamma_0 = 0$ and $a_0^2 = 0$ holds. Then the QMLE $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{\eta}_{\text{MLE}})$ that maximizes (23) with respect to (σ^2, η) and the realized volatility $\widehat{\sigma}^2_{\text{RV}}$ jointly satisfy:*

$$\Delta_n^{-1/2} \begin{pmatrix} \widehat{\sigma}^2_{\text{MLE}} - \frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s \leq T} (\Delta X_s)^2\right) \\ \widehat{\eta}_{\text{MLE}} - a_0^2 \Delta_n^{\gamma_0 - 1} \\ \widehat{\sigma}^2_{\text{RV}} - \frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s \leq T} (\Delta X_s)^2\right) \end{pmatrix} \stackrel{\mathcal{L}-s}{\longrightarrow} \mathcal{W}_T + \mathcal{Z}_T, \tag{31}$$

where $\mathcal{W}_T$ is the same as in Theorem 1, and $\mathcal{Z}_T$ is defined on the same extension of the original *probability space as $\mathcal{W}_T$ is, which is $\mathcal{F}$-conditionally centered, and its covariance matrix is given by:*

$$\mathbb{E}(\mathcal{Z}_T\mathcal{Z}_T^\intercal|\mathcal{F}) = \sum_{s\leq T}(\Delta X_s)^2\left(\sigma_{s-}^2+\sigma_s^2\right)\times\begin{pmatrix} 6 & -2 & 2\\ -2 & 1 & 0\\ 2 & 0 & 2\end{pmatrix}.$$

Moreover, $\mathcal{Z}_{1,T}-\mathcal{Z}_{3,T}$ and $\mathcal{Z}_{2,T}$ are $\mathcal{F}$-conditionally Gaussian random variables.

Therefore, we can construct a Hausman test statistic for $\mathbb{H}_0$ as follows:

$$H_{3n} = \Delta_n^{-1}\frac{(\widehat{\sigma}^2_{\text{MLE}}-\widehat{\sigma}^2_{\text{RV}})^2}{\widehat{V}_{3n}}, \tag{32}$$

where, writing $u_n = \tilde{\alpha}\Delta_n^{\varpi}$ with $1/(4-2\gamma)\leq\varpi<1/2$, and choosing k_n such that $k_n\Delta_n\to 0$,

$$\begin{aligned}
\widehat{V}_{3n} =& \frac{4}{T^2}\widehat{Q}_{3n},\\
\widehat{Q}_{3n} =& \frac{1}{3\Delta_n}\sum_{i=1}^{n}Y_i^4\cdot\mathbf{1}_{\{|Y_i|\leq u_n\}}+\sum_{i=k_n+1}^{n-k_n}Y_i^2\cdot\mathbf{1}_{\{|Y_i|>u_n\}}\cdot(\widehat{\sigma}^2_{i\Delta_n}+\widehat{\sigma}^2_{(i-k_n-1)\Delta_n}),\\
\widehat{\sigma}^2_{i\Delta_n} =& \frac{1}{k_n\Delta_n}\sum_{j=1}^{k_n}Y_{i+j}^2\cdot\mathbf{1}_{\{|Y_{i+j}|\leq u_n\}}.
\end{aligned} \tag{33}$$

By (10.24) and (10.27) of Aït-Sahalia and Jacod (2014), $\widehat{V}_{3n}$ is a consistent estimator of the asymptotic variance, i.e., AVAR $\left(\widehat{\sigma}^2_{\text{MLE}}-\widehat{\sigma}^2_{\text{RV}}\right)$.

As above, we have

$$H_{3n}\xrightarrow{\mathcal{L}}\chi_1^2 \quad \text{under } \mathbb{H}_0. \tag{34}$$

The behavior of H_{3n} is further characterized by:

Corollary 2. *The test statistic H_{3n} has asymptotic size α under the null hypothesis $\mathbb{H}_0: a_n^2=0$ and is consistent under the alternative hypothesis $\mathbb{H}_1: a_n^2=a_0^2>0$, that is,*

$$\mathbb{P}\left(H_{3n}>c_{1-\alpha}|\mathbb{H}_0\right)\to\alpha \quad \textit{and} \quad \mathbb{P}\left(H_{3n}>c_{1-\alpha}|\mathbb{H}_1\right)\to 1,$$

where $c_{1-\alpha}$ is the $(1-\alpha)$-quantile of the Chi-squared distribution with the degree of freedom being equal to 1. Moreover, H_{3n} follows a noncentral Chi-squared distribution with noncentrality parameter $a_0^4T^2\left(\int_0^T\sigma_s^4ds+\sum_{s\leq T}(\Delta X_s)^2(\sigma_{s-}^2+\sigma_s^2)\right)^{-1}$ and one degree of freedom, under the sequence of local alternative hypotheses $\mathbb{H}_n: a_n^2=a_0^2\Delta_n^{3/2}$.

We have shown that the Hausman tests have power to detect i.i.d. noise with sufficiently large variance. Our conjecture is that such tests also have power for stationary noise, an investigation which we leave for future work

3 Alternative Tests for the Presence of Noise

In this section, we develop alternative tests for the presence of noise in high frequency data. The first is based on comparing $\widehat{a}^2_{\text{MLE}}$ to 0; the second consists in testing whether the first-order autocorrelation of log-returns is 0, a distinctive implication in light of (12); the third is another Hausman test, but one that compares $\widehat{\sigma}^2_{\text{RV}}$ to a different noise-robust estimator that is based on pre-averaging the data, $\widehat{\sigma}^2_{\text{AVG}}$.

3.1 Testing Whether $a^2 = 0$

Given that the likelihood approach of Section 2 estimates jointly the volatility of the price and that of the noise, we can test for the presence of noise by testing directly whether the noise variance is zero. Recall that our estimator of noise variance $\widehat{a}^2_{\text{MLE}} = \widehat{\eta}_{\text{MLE}}\Delta_n$ can take negative values, as discussed above. By Theorem 2, the estimator satisfies the following central limit theorem:

$$\Delta_n^{-3/2}\left(\widehat{a}^2_{\text{MLE}} - a_0^2\Delta_n^{\gamma_0}\right) \xrightarrow{\mathcal{L}-s} \mathcal{W}_{2,T} + \mathcal{Z}_{2,T}, \tag{35}$$

under the null hypothesis $\mathbb{H}_0 : a_0^2 = 0$, as well as the local alternative hypothesis $\mathbb{H}_n : a_n^2 = a_0^2\Delta_n^{\gamma_0}$, where $a_0^2 > 0$ and $\gamma_0 \geq 3/2$. Therefore, we can construct a Student-t test using $\widehat{a}^2_{\text{MLE}}$, standardized by a quarticity estimator:

$$T_n = \Delta_n^{-3/2}\frac{|\widehat{a}^2_{\text{MLE}}|}{\widehat{Q}_{3n}^{1/2}}. \tag{36}$$

and we have

$$T_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \quad \text{under } \mathbb{H}_0. \tag{37}$$

By employing $\widehat{Q}_{3n}$ as the variance estimator in T_n, we achieve the same robustness as that of H_{3n} (stochastic volatility, non-Gaussian noise, and jumps). The following corollary summarizes the properties of T_n:

Corollary 3. *The Student-t test statistic T_n has asymptotic size α under the null hypothesis $\mathbb{H}_0$: $a_n^2 = 0$ and is consistent under the alternative hypothesis $\mathbb{H}_1 : a_n^2 > 0$, that is,*

$$\mathbb{P}\left(|T_n| > t_{1-\alpha/2}|\mathbb{H}_0\right) \to \alpha \quad \textit{and} \quad \mathbb{P}\left(|T_n| > t_{1-\alpha/2}|\mathbb{H}_1\right) \to 1,$$

where $t_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard Gaussian distribution. Moreover, T_n^2 follows a noncentral Chi-squared distribution with degree of freedom 1 and the noncentrality parameter $a_0^4T^2\left(\int_0^T \sigma_s^4 ds + \sum_{s\leq T}(\Delta X_s)^2(\sigma_{s-}^2 + \sigma_s^2)\right)^{-1}$, under the sequence of local alternative hypotheses $\mathbb{H}_n : a_n^2 = a_0^2\Delta_n^{3/2}$.

3.2 Testing for the Presence of First-Order Autocorrelation in Log-Returns

From (12), we see that the noise introduces a departure from the i.i.d. nature of log-returns in the form of a negative first order autocorrelation. The first order autocorrelation of log-returns can be estimated using:

$$\widehat{\rho}(1)_n = \frac{\sum_{j=1}^{n-1} Y_j Y_{j+1}}{\sum_{j=1}^{n} Y_j^2}, \tag{38}$$

which, in our high-frequency setting and under the null $\mathbb{H}_0 : a_0^2 = 0$, i.e., $\rho(1) = 0$, is an estimator with the following property:

$$\Delta_n^{-1/2} \widehat{\rho}(1)_n \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \frac{\int_0^T \sigma_s^4 ds + \sum_{s \le T} (\Delta X_s)^2 (\sigma_{s-}^2 + \sigma_s^2)}{\left(\int_0^T \sigma_s^2 ds + \sum_{s \le T} (\Delta X_s)^2\right)^2}\right). \tag{39}$$

This is shown in Lemma 1 in the appendix. Therefore, the asymptotic variance reduces to T^{-1} (corresponding to the the classical behavior of $\widehat{\rho}(1)_n$ in low frequency time series), if volatility is constant and jumps are absent.

To construct a feasible test for the null hypothesis that $\rho(1) = 0$, we propose the following statistic:

$$AC_n = \Delta_n^{-1/2} \frac{\sum_{j=1}^{n-1} Y_j Y_{j+1}}{\widehat{Q}_{3n}^{1/2}} \tag{40}$$

and we have

$$AC_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad \text{under } \mathbb{H}_0. \tag{41}$$

Again, by employing $\widehat{Q}_{3n}$ as the variance estimator in AC_n, we achieve the same robustness as that of H_{3n}. The following corollary summarizes the properties of AC_n:

Corollary 4. *The autocorrelation-based test statistic AC_n has asymptotic size α under the null hypothesis $\mathbb{H}_0 : a_n^2 = 0$ and is consistent under the alternative hypothesis $\mathbb{H}_1 : a_n^2 > 0$, that is,*

$$\mathbb{P}\left(|AC_n| > t_{1-\alpha/2} | \mathbb{H}_0\right) \to \alpha \quad \textit{and} \quad \mathbb{P}\left(|AC_n| > t_{1-\alpha/2} | \mathbb{H}_1\right) \to 1,$$

where $t_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard Gaussian distribution. Moreover, $(AC_n)^2$ follows a noncentral Chi-squared distribution with one degree of freedom and noncentrality parameter $a_0^4 T^2 \left(\int_0^T \sigma_s^4 ds + \sum_{s \le T} (\Delta X_s)^2 (\sigma_{s-}^2 + \sigma_s^2)\right)^{-1}$, under the sequence of local alternative hypotheses $\mathbb{H}_n : a_n^2 = a_0^2 \Delta_n^{3/2}$.

3.3 Hausman Test Based on a Pre-Averaging Volatility Estimator

In the same spirit as above, we can build an alternative Hausman test based now on comparing $\widehat{\sigma}^2_{\mathrm{RV}}$ with the pre-averaging volatility estimator $\widehat{\sigma}^2_{\mathrm{AVG}}$ of Jacod et al. (2009). Specifically, the test statistic we propose is

$$H_{4n} = \Delta_n^{-1/2} \frac{\left(\widehat{\sigma}^2_{\mathrm{RV}} - \widehat{\sigma}^2_{\mathrm{AVG}}\right)^2}{\widehat{V}(\widehat{\sigma}^2_{\mathrm{AVG}})}, \tag{42}$$

where, with $g = x \wedge (1-x)$, $x \in [0,1]$, $\phi(g) = \int_0^1 g(x)^2 dx$,

$$\widehat{\sigma}^2_{\mathrm{AVG}} = \frac{1}{k_n \phi(g) T} \sum_{i=1}^{n-k_n+1} \left((\overline{Y}_i^n)^2 - \frac{1}{2} \widehat{Y}_i^n \right),$$

$$\overline{Y}_i^n = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) Y_{i+j-1}, \quad \widehat{Y}_i^n = \sum_{j=1}^{k_n} \left(g\left(\frac{j}{k_n}\right) - g\left(\frac{j-1}{k_n}\right)\right)^2 (Y_{i+j-1})^2,$$

and k_n is a tuning parameter controlling the size of the window over which the averaging takes place, chosen such that $k_n \sqrt{\Delta_n} = \theta^{-1} \in (0, \infty)$.

Because the pre-averaging estimator $\widehat{\sigma}^2_{\mathrm{AVG}}$ converges at a slower rate $\Delta_n^{-1/4}$, we must adopt a different multiplier $\Delta_n^{-1/2}$. To obtain the asymptotic variance of $\widehat{\sigma}^2_{\mathrm{AVG}}$, we rely on Theorem 16.6.2 in Jacod and Protter (2012) which yields

$$\Delta_n^{-1/4} \left(\widehat{\sigma}^2_{\mathrm{AVG}} - \frac{1}{T} \left(\int_0^T \sigma_s^2 ds + \sum_{s \le T} (\Delta X_s)^2 \right)\right)$$
$$\stackrel{\mathcal{L}-s}{\longrightarrow} \mathcal{MN} \left(0, \frac{1}{\theta T^2} \left(\int_0^T R\left(\sigma_t^2, \theta^2 a_n^2\right) dt + \sum_{t \le T} S(\Delta X_t, \sigma_t^2, \theta^2 a_n^2) \right)\right), \tag{43}$$

where, writing g' as the derivative of g,

$$R(x,y) = \frac{4}{\phi(g)^2} \left(\Phi(g,g) x^2 + 2\Phi(g,g') xy + \Phi(g',g') y^2 \right),$$
$$S(x,y,z) = \frac{4}{\phi(g)^2} \left(\Psi_1(g,g) x^2 y + \Psi_2(g,g) x^2 y_- + \left(\Psi_1(g,g') + \Psi_2(g,g')\right) x^2 z \right),$$
$$\Phi(g,h) = \int_0^1 \int_t^1 g(u-t) g(u) du \int_t^1 h(v-t) h(v) dv dt,$$
$$\Psi_1(g,h) = \int_0^1 \left(\int_t^1 g(s) h(s+t) ds \right)^2 dt, \quad \Psi_2(g,h) = \int_0^1 \left(\int_0^{1-t} g(s) h(s-t) ds \right)^2 dt.$$

Moreover, (43) holds no matter whether $a_n^2 \ge 0$ or $a_n^2 \to 0$. By Theorems 16.4.2, 16.5.1, and 16.5.4 in Jacod and Protter (2012), we have a consistent estimator of the asymptotic variance:

$$\widehat{V}(\widehat{\sigma}^2_{\mathrm{AVG}}) = \widehat{V}_1(\widehat{\sigma}^2_{\mathrm{AVG}}) + \widehat{V}_2(\widehat{\sigma}^2_{\mathrm{AVG}}), \quad \text{where}$$
$$\widehat{V}_1(\widehat{\sigma}^2_{\mathrm{AVG}}) = \frac{\theta}{\phi(g)^2 T^2} \sum_{i=1}^{n-k_n+1} \left\{ \frac{4\Phi(g,g)}{3\phi(g)^2} (\overline{Y}_i^n)^4 + 4 \left(\frac{\Phi(g,g')}{\phi(g)\phi(g')} - \frac{\Phi(g,g)}{\phi(g)^2} \right) (\overline{Y}_i^n)^2 \widehat{Y}_i^n \right.$$
$$\left. + \left(\frac{\Phi(g,g)}{\phi(g)^2} - \frac{2\Phi(g,g')}{\phi(g)\phi(g')} + \frac{\Phi(g',g')}{\phi(g')^2} \right) (\widehat{Y}_i^n)^2 \right\} \cdot \mathbf{1}_{\{|\overline{Y}_i^n| \le v_n\}}$$

$$
\begin{aligned}
&\xrightarrow{p} \frac{1}{\theta T^2}\int_0^T R\left(\sigma_t^2, \theta^2 a_n^2\right) dt, \\
\widehat{V}_2(\widehat{\sigma}^2_{\text{AVG}}) =& \frac{4\theta}{k_n'\phi(g)^2T^2}\sum_{i=k_n'+1}^{n-k_n-k_n'+1} |\overline{Y}_i^n|^2 \cdot \mathbf{1}_{\{|\overline{Y}_i^n|>v_n\}} \Bigg\{ \frac{\Psi_1(g,g)}{\phi(g)^2}\sum_{j=1}^{k_n'}\left((\overline{Y}_{i+j}^n)^2 - \frac{1}{2}\widehat{Y}_{i+j}^n\right)\cdot \mathbf{1}_{\{|\overline{Y}_{i+j}^n|\le v_n\}} \\
&+\frac{\Psi_2(g,g)}{\phi(g)^2}\sum_{j=1}^{k_n'}\left((\overline{Y}_{i-j}^n)^2 - \frac{1}{2}\widehat{Y}_{i-j}^n\right)\cdot \mathbf{1}_{\{|\overline{Y}_{i-j}^n|\le v_n\}} + \frac{\Psi_1(g,g')}{2\phi(g)\phi(g')}\sum_{j=1}^{k_n'}\left(\widehat{Y}_{i+j}^n\right)\cdot \mathbf{1}_{\{|\overline{Y}_{i+j}^n|\le v_n\}} \\
&+\frac{\Psi_2(g,g')}{2\phi(g)\phi(g')}\sum_{j=1}^{k_n'}\left(\widehat{Y}_{i-j}^n\right)\cdot \mathbf{1}_{\{|\overline{Y}_{i-j}^n|\le v_n\}} \Bigg\} \\
&\xrightarrow{p} \frac{1}{\theta T^2}\sum_{t\le T} S(\Delta X_t, \sigma_t^2, \theta^2 a_n^2),
\end{aligned}
$$

$v_n = \bar{\alpha}(k_n\Delta_n)^{\bar{\varpi}}$, $k_n'/k_n \to \infty$, $k_n'\Delta \to 0$, and ϖ is some constant such that $1/(4-2\gamma) \le \varpi < 1/2$. By simple calculations based on our choice of the function $g(\cdot)$, we have $\Phi(g,g) = \Psi_1(g,g) = \Psi_2(g,g) = 151/80640$, $\Phi(g,g') = \Psi_1(g,g') = \Psi_2(g,g') = 1/96$, and $\Phi(g',g') = 1/6$.

Following the same pattern as before,

$$
H_{4n} \xrightarrow{\mathcal{L}} \chi_1^2 \quad \text{under } \mathbb{H}_0
$$

and furthermore:

Theorem 3. *The test statistic H_{4n} has asymptotic size α under the null hypothesis $\mathbb{H}_0 : a_n^2 = 0$ and is consistent under the alternative hypothesis $\mathbb{H}_1 : a_n^2 > 0$, that is,*

$$
\mathbb{P}\left(H_{4n} > c_{1-\alpha}|\mathbb{H}_0\right) \to \alpha \quad \textit{and} \quad \mathbb{P}\left(H_{4n} > c_{1-\alpha}|\mathbb{H}_1\right) \to 1,
$$

where $c_{1-\alpha}$ is the $(1-\alpha)$-quantile of the Chi-squared distribution with the degree of freedom being equal to 1. Moreover, H_{4n} follows a noncentral Chi-squared distribution with noncentrality parameter $4a_0^4\theta T^2\left(\int_0^T R(\sigma_t^2, 0) + \sum_{s\le T} S(\Delta X_t, \sigma_t^2, 0)\right)^{-1}$ and one degree of freedom, under the sequence of local alternative hypotheses $\mathbb{H}_n : a_n^2 = a_0^2\Delta_n^{5/4}$.

4 Small Sample Comparisons

We now employ Monte Carlo simulations to compare in small samples the different tests above. We simulate a Heston-type stochastic volatility model plus jumps in both price and volatility:

$$
dX_t = \mu dt + \sigma_t dW_t + dY_t, \tag{44}
$$

$$
d\sigma_t^2 = \kappa(\theta - \sigma_t^2)dt + \xi\sigma_t dB_t + J_t dN_t, \tag{45}
$$

where $\mathbb{E}(dW_t dB_t) = \rho dt$, $\rho = -0.75$, $\mu = 0.03$, $\kappa = 4$, $\theta = 0.1$, $\xi = 0.8$, $\log(J_t) \sim \mathcal{N}(-5, 1)$, N_t is a Poisson process with intensity $\lambda = 36$, and Y_t is a tempered-stable process (or CGMY process in Carr et al. (2002)) with Lévy jump measure given by:

$$\nu(x) = \frac{c}{|x|^{1+\beta}} e^{-\gamma_- |x|} \cdot \mathbf{1}_{\{x<0\}} + \frac{c}{x^{1+\beta}} e^{-\gamma_+ x} \cdot \mathbf{1}_{\{x>0\}}, \tag{46}$$

where $\gamma_+ = 3$, $\gamma_- = 5$, $c = 1.7$, and $\beta = 0.8$.

Additionally, we employ a log-volatility model specified as:

$$dX_t = \mu dt + \sigma_t dW_t + dY_t, \tag{47}$$

$$\sigma_t^2 = \exp(\alpha + \beta F_t), \quad dF_t = \kappa F_t dt + \sigma dB_t + J_t dN_t - \lambda \mu_J dt, \tag{48}$$

where $J_t \sim \mathcal{N}(\mu_J, \sigma_J^2)$, $\alpha = -2.8$, $\beta = 6$, $\kappa = -4$, $\rho = -0.8$, $\sigma = 0.8$, $\mu = 0.03$, $\lambda = 25$, $\mu_J = 0.02$, and $\sigma_J = 0.02$. The contribution of jumps to the total quadratic variation of the price process is about 40-60% using these parameters.

The data are sampled at frequencies ranging from one observation every 5 seconds to one every 5 minutes, spanning $T = 1$ week, and different values of the variance a^2 of the microstructure noise: 10^{-6}, 10^{-7} and 10^{-8} for power evaluation, and 0 for size of the test evaluation. We simulate 1,000 paths. The likelihood-based estimators are tuning parameter-free, but to implement their asymptotic variances, we need to choose a block size k_n, as well as α and ϖ in the truncation threshold u_n. We select $k_n = [n^{1/2}]$, $\varpi = 0.48$, $\tilde{\alpha} = \alpha_0(\frac{1}{T}\widehat{\int_0^T \sigma_t^2 dt})^{1/2}$, and $\alpha_0 = 3$ (see Aït-Sahalia and Jacod (2014) for a discussion). For the pre-averaging estimator, we use the same k_n, α_0, and ϖ. We choose $\bar{\alpha} = \alpha_0(\frac{1}{T}\widehat{\int_0^T \sigma_t^2 dt}\phi(g) + \theta^2 \widehat{a}_0^2 \phi(g'))^{1/2}$ and $k_n' = 2k_n$, as suggested in Aït-Sahalia and Xiu (2016). The results are not sensitive to the selection of these parameters, within a reasonable range.

We calculate H_{2n}, H_{3n}, H_{4n}, T_n, and AC_n respectively, for each sample simulated from above, i.e., (44), (45), (47), and (48). Tables B.1 and B.2 show that H_{3n}, T_n, and AC_n have desired size control at 5% level across all sampling frequencies. The power of these tests is also satisfactory when sampling frequency is high. By contrast, the performance of H_{4n} is not as desirable as that of H_{3n}, because the former relies on a volatility estimator that is not as efficient. H_{2n} does not perform well either, because it is not robust to the presence of jumps in the data. Using data re-generated from the same models above except with jumps excluded in prices, H_{2n} has good size control and is as powerful as other tests. Figures B.1 - B.4 in the web appendix compare the histograms of the test statistics H_{3n}, H_{4n}, T_n, and AC_n with their asymptotic distributions under $\mathbb{H}_0$ for both the Heston and log-volatility models. They match very well with the theoretical predictions. We thereby recommend in practice the use of H_{3n} for its robustness, as well as H_{2n} if one wishes to use a simpler test (without any tuning parameters) for certain dataset, for which price jumps can be ruled out by prior information. Additionally, T_n and AC_n can be employed to confirm the results of the Hausman tests.

5 Empirical Results: At What Frequency Does the Noise Start to Bite?

We implement the test on intraday observations of constituents of the Dow Jones 30 index (DJI), the S&P 100 index (OEX), and the S&P 500 index (INX) from January 2003 to December 2012, obtained from the TAQ database. Including deletions and additions to the index, there are in total 39 stocks for DJI, 153 for OEX, and 848 for INX. As common in the literature, we include transactions from three major exchanges: AMEX, NYSE, and NASDAQ. For each week and each stock of DJI and OEX, we subsample the intraday returns at frequencies ranging from every 5 seconds to every 5 minutes, whereas for the INX constituents, we subsample the returns within a monthly window at frequencies ranging from every 5 minutes to every 30 minutes. Overnight returns are excluded to avoid dividend issuance, stock splits and other issues. These do not matter for the question we address here.

One common practice to subsample the data uses the previous tick method, which relies on the last transaction price prior to or at the sampling times. To ensure the consistency of this approach, it is necessary to make sure that there is at least one transaction price between each sampling interval. Otherwise, a zero-return would be generated by the previous tick method, which then can introduce a bias to the autocorrelation of observations in the subsample. Typically, this is not a problem for liquid stocks when sampling frequency is beyond every few seconds. For illiquid ones, however, the "fabricated" zero returns are unavoidable using the previous tick sampling. Since Dow Jones constituents are rather liquid, sampling every few seconds does not create many zero returns. For the less liquid S&P 500 constituents, sampling beyond 5 minutes also reduces the amount of zero returns.

We compute the likelihood estimators based on subsampled returns with fabricated zero returns removed, while pretending that the sampling interval is constant. This amounts to estimating the tick-time volatility instead of the calendar-time volatility. Since the noise is associated with tick times, this strategy does not in principle affect the test results. In fact, if there exists a smooth time-change between tick times and calendar times, as discussed in Aït-Sahalia and Jacod (2014), realized volatility remains the same after the time-change.[1]

We first compute the volatility signature plots in Figure 2, which compares the average volatility estimates across different days and tickers for different sampling frequencies. We compare the estimates using RV, QMLE, TSRV (from Zhang et al. (2005)), and the pre-averaging estimator. Clearly, there is a large discrepancy between $\hat{\sigma}^2_{\mathrm{RV}}$ and the other, noise-robust estimators, and the difference

[1] The asymptotic distribution of $\widehat{\sigma}^2_{\mathrm{RV}}$ in a more general irregular sampling setting has been discussed by, e.g., Mykland and Zhang (2006) and Mykland and Zhang (2009), and further by Bibinger and Vetter (2015) for the case where jumps also exist. This distribution is in general non-Gaussian, due to the irregularity of the sampling intervals near where jumps occur.

becomes quite sizable when the sampling frequency increases. Empirically, the Hausman tests will quantify the extent to which the discrepancy visible on the curve between RV and the noise-robust estimators is statistically meaningful at a given frequency.

Additionally, we plot in Figure 3 the number of trades per week and the first-order autocorrelation of 1-second returns for 22 names that are part of the DJI throughout the sample. 1-second returns are not employed in the test statistics; this is simply to validate (12). For most stocks in the index, the numbers of trades peak during 2008, whereas the first-order autocorrelations shrink towards 0 over the sampling period. This shows that returns sampled every second become less noisy, based on that measure, despite the fact that the number of trades drops after the financial crisis. We also report in Figure 4 the monthly average national best bid and offer (NBBO) spread (in percentage) across the INX constituents, and the monthly average of total dollar volume over the sampling periods. These are potential measures of market liquidity that will be related below to the results of the test.

We then apply the five test statistics H_{2n}, H_{3n}, H_{4n}, AC_n and T_n to the sample at different frequencies of 5 and 30 seconds and 1 and 5 minutes, and summarize the percentage of rejections of $\mathbb{H}_0$ among all combinations of names and weeks for each year in Table 3 (resp. Table 4) for DJI (resp. OEX) constituents. The results show that at a given frequency, over time, the number of rejections decreases, consistent with an overall improvement in liquidity over the decade covered by the sample. The common practice of treating 5-minute returns as noise-free might be problematic in the earlier years of the sample for DJI and OEX constituents, but is a reasonably safe choice for data sampled after 2009. These numerical values are of course not meant to be treated as universal.

One major caveat is that they are obtained by examining a subset of the most liquid stocks in one of the most liquid market in the world. For comparison purposes, we report in Table 5 the corresponding results for the INX constituents then in Table 6 for the 30 least liquid INX constituents, in terms of the percentage bid-ask spreads, at frequencies of 5, 10, 15 and 30 minutes. The results show that, even among stocks that belong to the S&P 500 index, 5 minute returns cannot be treated as noise-free, even in the most recent part of the sample.

Finally, we compare in Figures 5, 6 and **??** the test statistic H_{3n} with standard measures of liquidity, i.e., the number of non-zero returns, bid-ask spreads (in percentage), and total dollar volume for 2003 and 2012, respectively. We find, reassuringly, that more liquidity in the traditional sense correlates with a lower value of the noise test statistic. In particular, the percentage bid-ask spread decreases dramatically from 2003 to 2012, leading to a shift of test statistics towards 0 (fewer rejections). However, the relationship between any one of these measures and the noise statistic is far from perfectly predictive, showing that the test for the presence of noise cannot be simply replaced by a computation of a liquidity measure.

6 Conclusions

Just like one is concerned in a regression about the potential endogeneity of the regressors, one is concerned here about the potential presence of microstructure noise. Just like one can rely on a statistical methodology to determine when endogeneity is sufficiently small to be tolerated, one would like a statistical methodology to determine at what sampling frequency the noise becomes sufficiently small to be tolerated or vice versa sufficiently large that noise-robust estimators should be employed. This paper proposes a variety of tests for this purpose.

We put the emphasis on assessing the deviations among different volatility estimators for the purpose of testing whether microstructure noise mattered in a given sample. The same testing idea can in principle be employed to compare other types of estimators, for instance estimators of the jump characteristics of the process. We find that the common practice of treating 5-minute returns as noise-free might be problematic in the earlier years of the sample for Dow Jones and S&P 100 constituents, but is a reasonably safe choice for data sampled after 2009 for these stocks, but not for all the S&P 500 stocks. The ease of implementing these tests means that repeating the exercise on any sample under consideration should become standard practice before considering any high frequency econometric procedure that is not noise-robust by construction.

References

Aït-Sahalia, Y., 1996. Testing continuous-time models of the spot interest rate. Review of Financial Studies 9, 385–426.

Aït-Sahalia, Y., Jacod, J., 2012. Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. Journal of Economic Literature 50, 1005–1048.

Aït-Sahalia, Y., Jacod, J., 2014. High Frequency Financial Econometrics. Princeton University Press.

Aït-Sahalia, Y., Mykland, P. A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial Studies 18, 351–416.

Aït-Sahalia, Y., Xiu, D., 2016. Increased correlation among asset classes: Are volatility or jumps to blame? Journal of Econometrics, forthcoming .

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2000. Great realizations. Risk , 105–108.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2001. The distribution of exchange rate realized volatility. Journal of the American Statistical Association 96, 42–55.

Bibinger, M., Vetter, M., 2015. Estimating the quadratic covariation of an asynchronously observed semimartingale with jumps. Annals of the Institute of Statistical Mathematics 67 (4), 707–743.

Black, F., 1986. Noise. The Journal of Finance 41, 529–543.

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

Gloter, A., Jacod, J., 2001. Diffusions with measurement errors: I - Local asymptotic normality. ESAIM P&S 5, 225–242.

Hahn, J., Hausman, J. A., 2002. A new specification test for the validity of instrumental variables. Econometrica 70, 163–189.

Hausman, J. A., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.

Hausman, J. A., Pesaran, M. H., 1983. The J-test as a Hausman specification test. Economics Letters 12, 277–281.

Hausman, J. A., Stock, J. H., Yogo, M., 2005. Asymptotic properties of the Hahn-Hausman test for weak-instruments. Economics Letters 89, 333–342.

Hausman, J. A., Taylor, W. E., 1981. A generalized specification test. Economics Letters 8, 239–245.

Holly, A., 1982. A remark on Hausman's specification test. Econometrica 50, 749–759.

Holly, A., Monfort, A., 1986. Some useful equivalence properties of Hausman's test. Economics Letters 20, 39–43.

Jacod, J., 1994. Limit of random measures associated with the increments of a Brownian semimartingale. Tech. rep., Université de Paris-6.

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

Jacod, J., Protter, P., 2012. Discretization of Processes. Springer-Verlag.

Liu, L. Y., Patton, A. J., Sheppard, K., 2015. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. Journal of Econometrics 187 (1), 293–311.

Mykland, P. A., Zhang, L., 2006. ANOVA for diffusions and Itô processes. Annals of Statistics 34, 1931–1963.

Mykland, P. A., Zhang, L., 2009. Inference for continuous semimartingales observed at high frequency. Econometrica 77, 1403–1445.

Patton, A. J., 2011. Data-based ranking of realised volatility estimators. Journal of Econometrics 161, 284–303.

Renault, E., Sarisoy, C., Werker, B. J., 2015. Efficient estimation of integrated volatility and related processes. Tech. rep., Tilburg University.

Rockafellar, R. T., 1997. Convex Analysis. Princeton University Press.

Sargan, J. D., 1958. The estimation of economic relationships using instrumental variables. Econometrica 26, 393–415.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

White, H., 1982. Maximum likelihood estimation of misspecified models. Econometrica 50, 1–25.

Xiu, D., 2010. Quasi-maximum likelihood estimation of volatility with high frequency data. Journal of Econometrics 159, 235–250.

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

Panel A		Heston Model with Jumps									
		$\Delta_n = 5$ seconds					$\Delta_n = 30$ seconds				
5% level	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	0.0	4.7	4.8	4.7	7.4	0.3	5.9	5.6	5.8	8.0
	10^{-8}	14.9	96.0	95.9	95.9	38.6	0.0	12.0	11.6	11.9	7.3
Power	10^{-7}	60.4	100.0	100.0	100.0	97.9	25.8	90.1	90.0	90.1	37.3
	10^{-6}	93.0	100.0	100.0	100.0	98.8	70.5	99.6	99.6	99.6	98.0
		$\Delta_n = 1$ minute					$\Delta_n = 5$ minutes				
	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	0.2	5.5	5.6	5.4	6.2	0.9	5.7	5.8	6.0	8.6
	10^{-8}	0.2	6.7	6.5	6.8	8.8	1.0	5.3	5.3	5.2	7.8
Power	10^{-7}	10.5	54.9	53.8	54.8	16.9	1.4	7.9	7.6	8.1	9.1
	10^{-6}	57.2	99.0	99.0	99.0	91.4	22.0	49.6	48.3	49.2	27.6
Panel B		Heston Model without Jumps									
		$\Delta_n = 5$ seconds					$\Delta_n = 30$ seconds				
5% level	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	5.0	5.0	4.9	5.0	6.5	4.7	4.9	4.6	5.0	6.0
	10^{-8}	100.0	100.0	100.0	100.0	74.7	24.2	24.6	24.6	24.4	9.7
Power	10^{-7}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	79.4
	10^{-6}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		$\Delta_n = 1$ minute					$\Delta_n = 5$ minutes				
	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	5.8	6.3	6.2	6.5	6.2	4.4	4.4	5.3	4.6	9.0
	10^{-8}	7.7	7.7	7.3	7.9	7.0	4.7	4.8	5.1	4.8	8.8
Power	10^{-7}	92.8	92.8	92.5	92.9	33.9	6.7	7.0	7.9	7.1	10.6
	10^{-6}	100.0	100.0	100.0	100.0	100.0	79.2	79.5	78.1	79.5	51.7

Table 1: Simulation Results: Percentage of Rejections of $\mathbb{H}_0$ for the Heston Model

Panel A		Log-Volatility Model with Jumps,									
		$\Delta_n = 5$ seconds					$\Delta_n = 30$ seconds				
5% level	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	0.0	5.2	5.1	5.2	9.0	0.0	5.5	5.3	5.5	7.5
	10^{-8}	16.4	96.9	96.9	96.9	49.5	0.5	19.8	19.5	19.4	10.2
Power	10^{-7}	57.5	100.0	100.0	100.0	95.9	21.8	94.7	94.7	94.5	49.2
	10^{-6}	89.5	100.0	100.0	100.0	95.4	65.0	99.9	99.9	99.9	98.4
		$\Delta_n = 1$ minute					$\Delta_n = 5$ minutes				
	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	0.0	5.5	5.6	5.5	7.1	0.6	5.7	5.7	5.7	7.7
	10^{-8}	0.3	8.9	8.7	8.8	9.4	0.6	7.9	7.6	7.9	7.1
Power	10^{-7}	10.7	67.5	67.5	67.3	24.9	1.0	8.1	7.5	8.2	9.2
	10^{-6}	50.9	98.2	98.1	98.1	90.9	26.1	57.4	56.0	56.2	37.2
Panel B		Log-Volatility Model without Jumps									
		$\Delta_n = 5$ seconds					$\Delta_n = 30$ seconds				
5% level	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	5.4	5.6	5.5	5.6	5.4	4.9	5.3	5.4	5.3	5.4
	10^{-8}	99.9	99.9	99.9	99.9	90.3	46.7	47.7	47.0	47.5	13.2
Power	10^{-7}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	90.1
	10^{-6}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		$\Delta_n = 1$ minute					$\Delta_n = 5$ minutes				
	a_0^2	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	4.5	4.8	4.9	4.7	6.2	5.4	5.9	5.5	6.1	10.5
	10^{-8}	11.0	11.5	11.0	11.3	9.2	5.4	5.7	5.9	5.7	10.0
Power	10^{-7}	97.5	97.6	97.5	97.6	57.5	10.7	11.0	11.0	10.9	14.2
	10^{-6}	100.0	100.0	100.0	100.0	100.0	90.6	91.0	90.6	90.4	68.9

Table 2: Simulation Results: Percentage of Rejections of $\mathbb{H}_0$ for Log-Volatility Model

	$\Delta_n = 5$ seconds					$\Delta_n = 30$ seconds				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	85.10	87.06	86.67	88.10	73.33	51.83	52.09	52.29	52.03	53.99
2004	86.35	87.05	86.86	88.14	72.12	54.36	54.42	54.17	54.74	51.28
2005	82.68	83.33	83.14	85.75	75.42	57.84	57.84	57.45	58.43	56.99
2006	85.32	85.90	85.90	87.50	72.63	58.91	58.01	57.82	58.78	55.77
2007	72.76	73.85	73.97	74.74	70.58	40.13	39.36	39.10	40.45	53.78
2008	54.73	55.80	55.13	55.47	57.87	25.73	23.60	23.73	23.20	47.93
2009	56.99	59.93	59.28	60.46	47.06	31.37	30.98	30.65	30.00	38.50
2010	46.86	53.01	52.69	53.21	42.05	27.82	28.08	27.88	28.08	40.90
2011	42.88	48.72	48.46	49.23	35.58	23.85	23.46	23.27	23.53	33.53
2012	33.33	42.35	42.42	42.55	28.04	19.93	21.76	22.09	21.63	32.03
	$\Delta_n = 1$ minute					$\Delta_n = 5$ minutes				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	42.55	40.59	40.59	39.08	46.01	14.71	9.22	8.50	9.08	28.76
2004	43.53	41.35	41.28	41.79	42.31	13.46	9.62	8.65	8.85	23.72
2005	44.12	42.75	42.75	42.22	48.76	15.42	11.44	10.39	10.33	31.31
2006	38.33	35.96	35.71	35.32	46.60	12.05	8.33	8.08	7.88	32.50
2007	22.88	20.96	20.19	19.10	47.76	10.06	6.22	5.77	4.42	32.95
2008	17.60	15.20	14.53	14.00	43.13	8.80	5.27	5.00	4.07	28.13
2009	23.33	21.44	21.24	20.59	35.29	7.71	4.64	4.51	4.25	25.16
2010	23.59	22.31	22.37	21.22	39.23	9.49	6.28	5.90	4.81	31.60
2011	18.53	16.67	16.86	15.83	33.14	7.44	3.91	3.85	3.14	27.18
2012	14.58	13.86	13.59	13.59	33.14	7.45	4.71	4.05	3.92	30.20

Table 3: Percentage of Rejections Over Time of $\mathbb{H}_0$ at the 5% Level at Different Frequencies: DJA Stocks

	$\Delta_n = 5$ seconds					$\Delta_n = 30$ seconds				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	79.86	82.14	81.90	83.59	68.76	54.29	55.22	55.22	55.02	54.33
2004	76.33	78.59	78.41	80.57	64.69	53.10	54.06	53.92	54.92	51.04
2005	76.18	78.57	78.34	81.23	67.73	54.98	55.55	55.32	56.45	56.00
2006	80.20	82.04	82.00	83.91	66.87	56.51	56.27	56.11	57.38	53.53
2007	80.00	80.85	80.79	81.32	72.57	45.81	45.02	44.57	46.32	54.60
2008	53.74	54.79	54.19	54.95	61.56	29.79	27.88	27.72	27.95	50.26
2009	52.52	55.33	54.26	56.04	47.33	28.15	27.96	27.70	27.39	38.87
2010	45.27	50.96	51.16	51.98	39.80	27.61	27.96	27.69	27.57	38.33
2011	44.39	50.75	50.75	51.47	32.84	23.65	23.94	23.57	23.67	31.41
2012	33.84	40.80	40.94	40.92	27.94	18.70	19.48	19.54	19.06	32.22
	$\Delta_n = 1$ minute					$\Delta_n = 5$ minutes				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	44.78	43.84	43.78	42.90	47.33	17.33	12.80	12.71	12.71	31.39
2004	42.78	42.33	42.29	42.43	43.94	15.78	12.43	11.49	12.00	26.63
2005	43.30	42.45	42.57	41.98	49.32	19.02	15.18	14.59	14.18	34.82
2006	38.96	36.93	36.73	36.69	46.40	14.67	11.00	10.36	10.24	33.42
2007	25.66	23.83	23.09	22.23	47.77	10.43	6.89	6.43	5.96	33.68
2008	20.19	17.56	17.12	16.35	43.93	9.65	6.56	5.79	5.53	29.19
2009	22.09	20.70	20.30	19.87	36.13	7.52	5.07	4.89	4.80	25.93
2010	21.27	20.24	19.88	19.55	36.82	9.22	6.22	5.94	4.65	30.29
2011	17.14	15.35	15.08	14.92	31.25	7.25	3.88	3.86	3.02	26.53
2012	14.66	13.74	13.34	13.20	33.46	7.14	4.64	4.14	3.86	29.00

Table 4: Percentage of Rejections Over Time of $\mathbb{H}_0$ at the 5% Level at Different Frequencies: S&P 100 Stocks

	$\Delta_n = 5$ minutes					$\Delta_n = 10$ minutes				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	32.01	28.20	27.92	27.68	31.63	14.70	11.10	10.90	11.02	20.25
2004	29.57	26.22	25.96	26.02	27.16	13.22	9.89	9.46	10.23	17.94
2005	34.32	30.69	30.02	29.75	36.51	14.63	11.10	10.74	10.20	24.52
2006	26.98	23.02	22.65	22.12	32.43	11.68	8.44	8.37	7.86	22.67
2007	16.40	13.72	13.41	12.76	32.60	9.03	6.27	6.14	5.71	23.60
2008	17.68	13.45	13.09	12.66	24.90	7.79	4.35	4.47	4.33	16.62
2009	12.89	9.80	9.56	9.31	24.51	5.99	4.04	3.80	3.76	18.87
2010	12.63	9.74	9.36	8.59	27.91	8.18	5.29	5.19	5.16	22.21
2011	11.99	8.97	8.74	7.86	22.57	7.32	4.52	4.25	4.46	18.52
2012	13.63	11.03	10.57	9.90	25.57	7.70	5.41	5.21	5.16	19.83
	$\Delta_n = 15$ minutes					$\Delta_n = 30$ minutes				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	9.29	5.98	5.80	5.90	16.42	4.03	2.05	1.86	2.09	13.55
2004	7.97	5.11	5.03	5.17	15.71	3.40	1.54	1.38	1.61	11.45
2005	8.69	5.93	5.76	5.52	20.95	3.89	1.88	1.71	1.73	15.98
2006	9.13	6.02	5.87	5.60	20.25	4.78	2.76	2.61	2.60	16.39
2007	6.62	4.09	3.79	3.77	20.45	3.93	2.14	2.00	1.73	16.27
2008	4.89	2.50	2.39	2.47	13.71	2.94	1.34	1.17	1.42	10.96
2009	3.67	1.81	1.83	1.83	16.88	3.65	2.11	2.11	1.92	15.54
2010	4.86	2.92	2.71	2.73	19.73	2.94	1.36	1.26	1.13	16.12
2011	4.21	2.10	2.08	2.36	16.11	2.69	1.07	1.03	1.07	13.98
2012	4.39	2.70	2.61	2.46	16.73	2.77	1.36	1.15	1.52	14.16

Table 5: Percentage of Rejections Over Time of $\mathbb{H}_0$ at the 5% Level at Different Frequencies: S&P 500 Stocks

	$\Delta_n = 5$ minutes					$\Delta_n = 10$ minutes				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	56.11	53.89	53.61	52.78	46.94	31.94	28.61	28.61	29.17	28.61
2004	50.00	47.78	47.78	48.33	47.22	26.94	23.33	23.33	23.61	27.78
2005	57.50	56.11	54.72	55.56	57.78	31.94	29.72	28.33	30.28	39.44
2006	57.22	54.44	54.44	53.89	49.17	26.67	24.17	23.33	23.06	27.50
2007	40.00	38.61	37.22	37.50	36.94	15.83	13.89	14.44	14.44	23.33
2008	23.33	21.11	20.28	19.44	18.33	10.56	8.06	8.06	7.50	12.50
2009	29.17	27.22	26.67	26.39	30.28	12.50	11.39	10.56	10.56	21.39
2010	29.17	26.94	26.67	25.28	35.00	15.83	13.06	12.50	12.50	24.72
2011	21.67	18.33	18.06	16.67	25.83	10.83	8.89	8.89	8.89	20.83
2012	29.72	27.22	27.78	26.39	32.78	15.83	10.83	11.39	10.83	25.00
	$\Delta_n = 15$ minutes					$\Delta_n = 30$ minutes				
Year	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}	H_{2n}	H_{3n}	T_n	AC_n	H_{4n}
2003	17.22	13.89	13.89	14.17	21.11	5.28	3.89	3.06	4.17	13.06
2004	12.50	9.44	10.00	9.17	20.56	4.17	2.78	2.50	1.94	14.17
2005	21.11	18.61	18.61	17.22	30.56	6.67	6.11	6.11	6.11	21.67
2006	16.67	13.61	13.06	14.17	24.44	5.28	4.17	3.89	3.89	18.33
2007	12.78	10.00	9.17	9.72	20.00	5.00	3.89	3.61	3.06	15.28
2008	5.28	4.17	3.89	4.72	10.56	3.06	1.39	1.39	2.22	7.50
2009	7.78	5.28	5.28	6.11	14.72	3.33	1.94	1.94	1.94	11.39
2010	10.00	8.06	7.50	7.50	20.00	4.44	4.17	3.61	3.06	16.39
2011	4.44	3.33	3.89	3.89	15.00	4.17	1.39	1.67	1.11	15.28
2012	6.67	4.72	4.72	3.89	20.00	4.44	1.94	1.67	1.11	15.83

Table 6: Percentage of Rejections Over Time of $\mathbb{H}_0$ at the 5% Level at Different Frequencies: Least Liquid S&P 500 Stocks

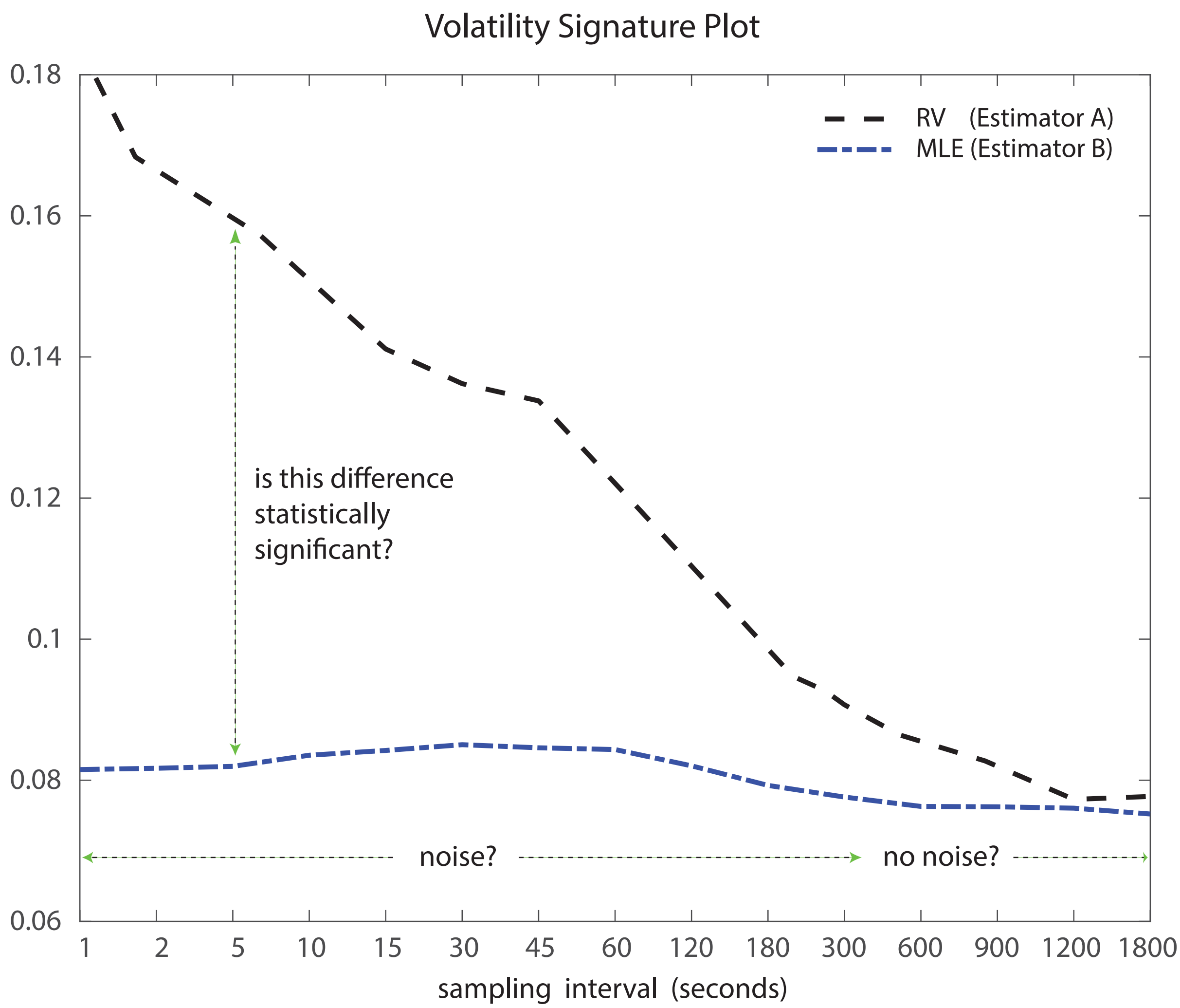
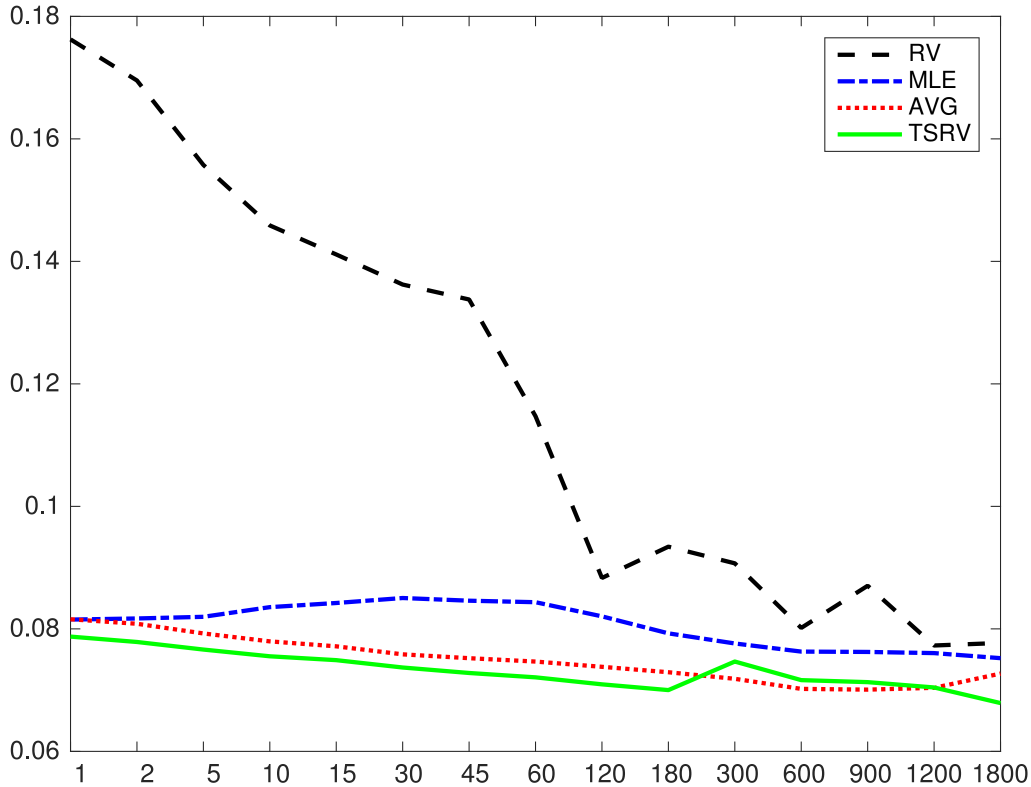


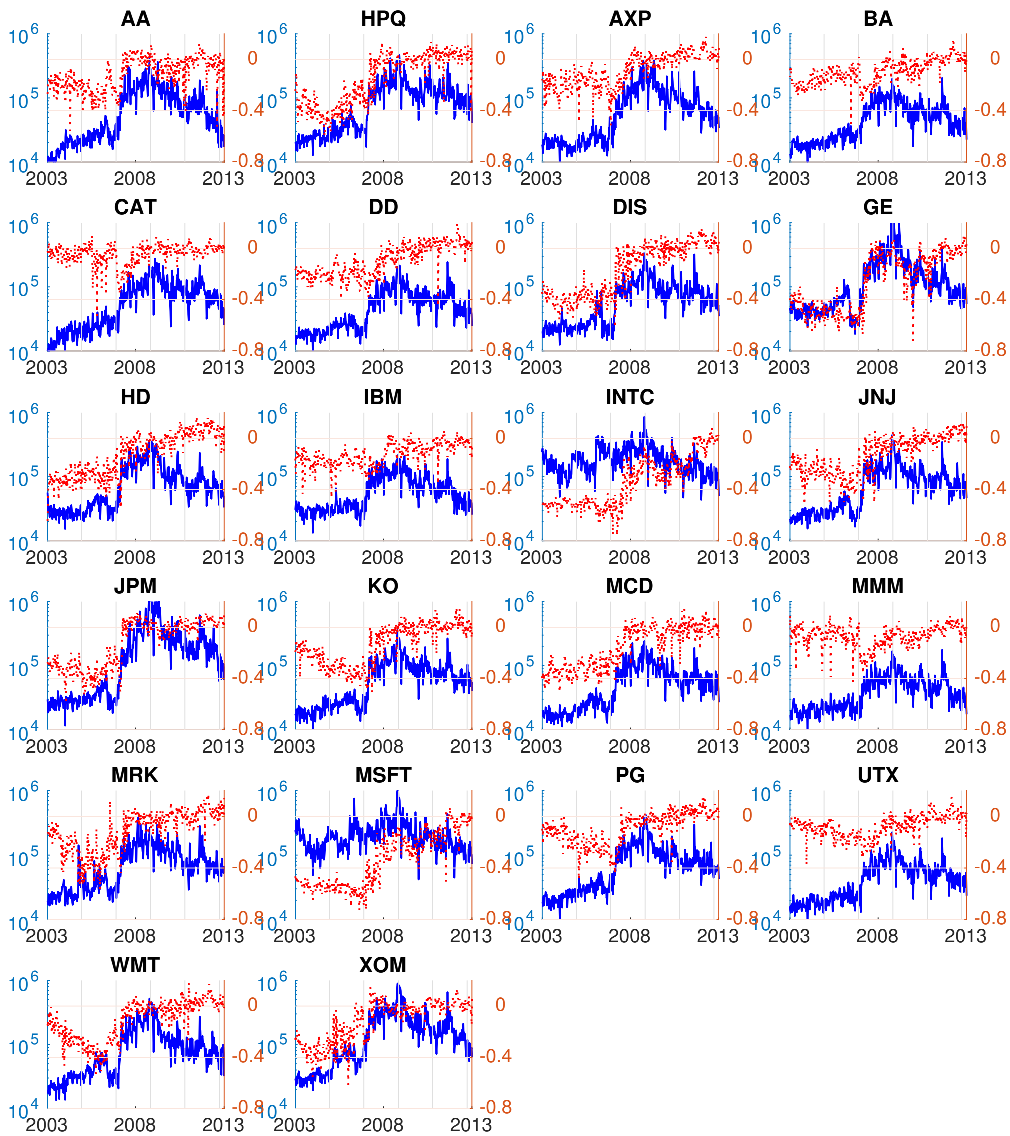
Figure 1: The test measures whether the divergence of $\widehat{\sigma}^2_{\mathrm{RV}}$ due to the noise, relative to a noise-robust estimator such as $\widehat{\sigma}^2_{\mathrm{MLE}}$, is significant

Figure 2: Average Signature Plot for the Dow Jones 30 Index Constituents



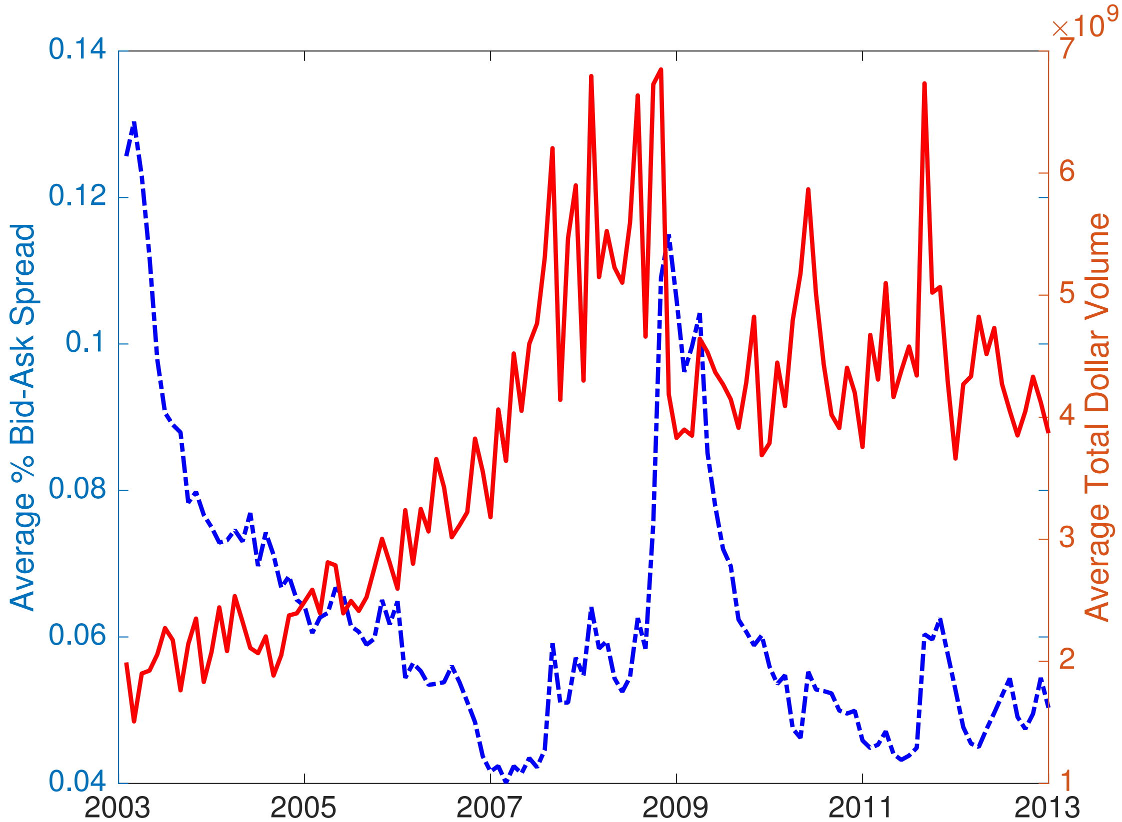
Note: This figure reports the average estimates of annualized daily volatility across all Dow Jones 30 index constituents from 2003 to 2012. Four volatility estimators are implemented: RV, MLE, TSRV and AVG.

Figure 3: Time Series of the Number of Trades and First-Order Autocorrelation, DJI



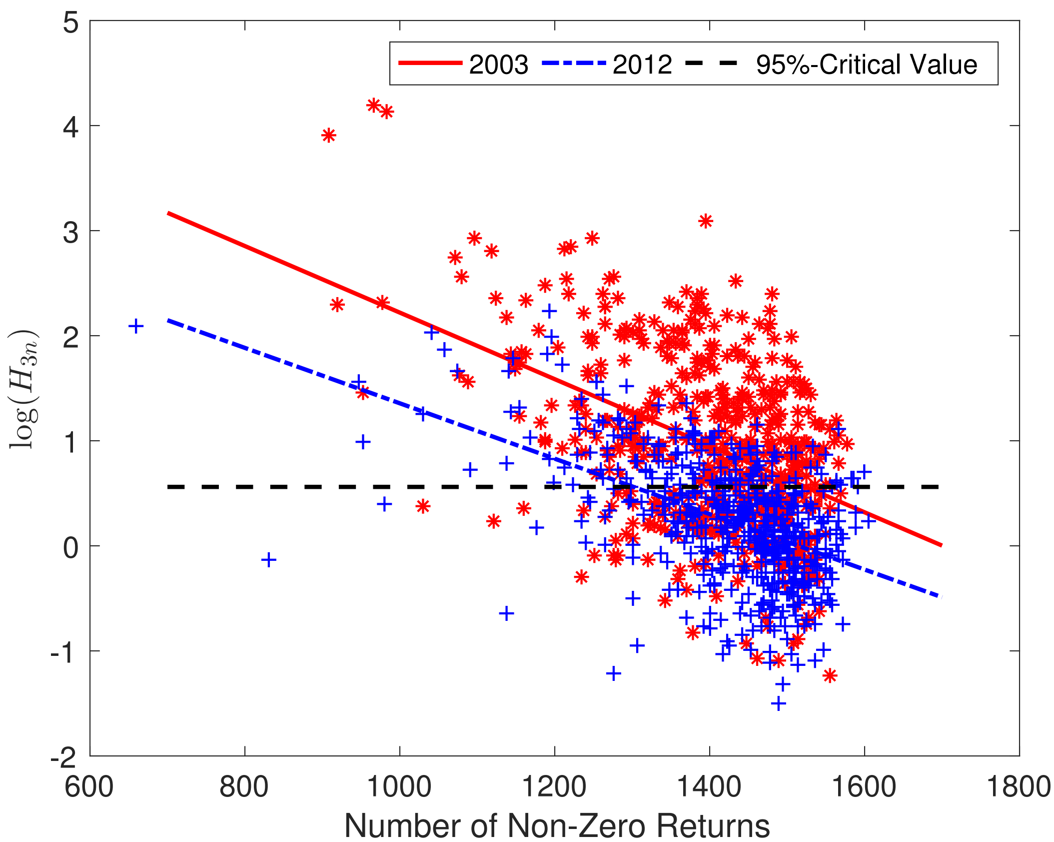
Note: This figure plots the time series of number of trades (blue, solid, left y-axis) and the first-order autocorrelation of 1-second returns (red, dotted, right y-axis) for the 22 stocks that were included in the Dow-Jones Industrials 30 index throughout the sample period 2003-2012.

Figure 4: Bid-Ask Spread and Dollar Volume



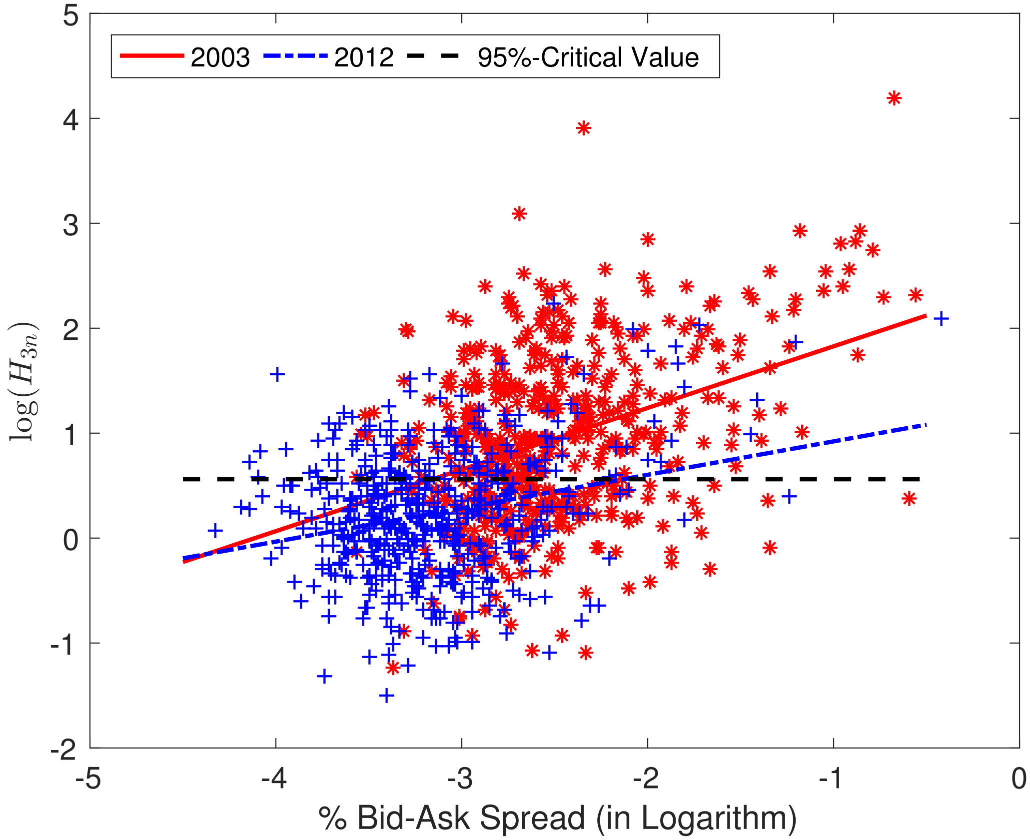
Note: This figure plots the time series of the monthly cross-sectional average NBBO bid-ask spread (in percentage of NBBO price), (blue, dot-dashed) and the monthly average dollar volume (red, solid) for S&P 500 index constituents from 2003 to 2012. The y-axes on the left is the percentage, whereas the y-axes on the right is the dollar volume.

Figure 5: Test Statistics and Number of Non-Zero Returns



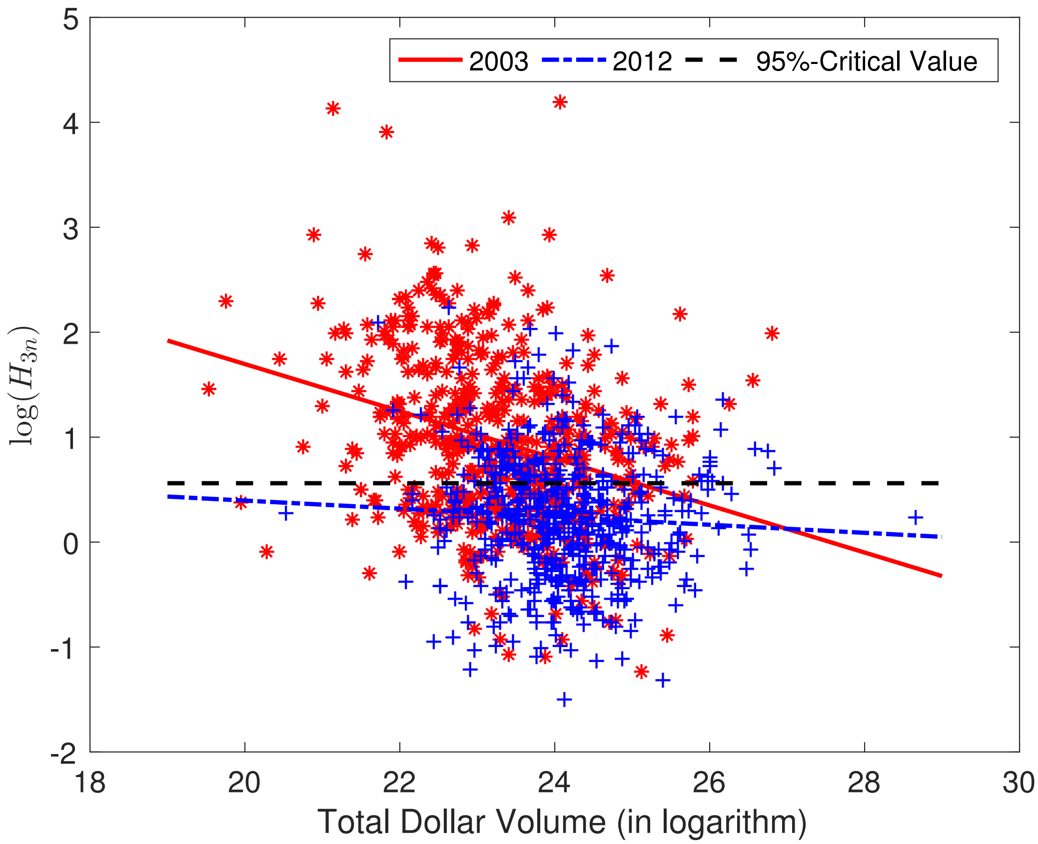
Note: This figure is a scatterplot of the test statistic H_{3n} averaged over 12 months (in logarithm) against the average number of non-zero returns per month at 5-minute frequency for the S&P 500 index constituents of 2003 and 2012, respectively, along with the regression line for each year. The black dashed line marks the 95%-critical value. As expected, a higher number of non-zero returns is associated with a lower value of the noise test statistic.

Figure 6: Test Statistics and Bid-Ask Spreads



Note: This figure plots the test statistic H_{3n} averaged over 12 months (in logarithm) against the NBBO bid-ask spread (in logarithm of the percentage of NBBO price) for S&P 500 index constituents of 2003 and 2012, respectively, along with the regression line for each year. The black dashed line marks the 95%-critical value. As expected, a lower bid-ask spread is associated with a lower value of the noise test statistic.

Figure 7: Test Statistics and Dollar Volume



Note: This figure plots the test statistic H_{3n} averaged over 12 months (in logarithm) against the total dollar volume (in logarithm) for the S&P 500 index constituents of 2003 and 2012, respectively, along with the regression line for each year. The black dashed line marks the 95%-critical value. As expected, a higher trading volume is associated with a lower value of the noise test statistic.

Appendix A Proofs

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space. K denotes a generic constant which may change from line to line.

Appendix A.1 Proof of Theorem 1

Proof. To prove this, we follow the strategy suggested in Xiu (2010), i.e., we investigate the asymptotic behavior of the score vector. We only need to prove for the case $\gamma_0 > 0$ and $a_0^2 > 0$. The case with $\gamma_0 = a_0^2 = 0$ is similar yet easier.

For any $(\sigma^2, \eta) \in \Theta$, we define

$$\gamma^2 = \eta + \frac{1}{2}\sqrt{\sigma^2(4\eta + \sigma^2)} + \frac{1}{2}\sigma^2, \quad \phi = 1 - \frac{1}{2\eta}\sqrt{\sigma^2(4\eta+\sigma^2)} + \frac{1}{2\eta}\sigma^2, \tag{A.1}$$

and $\phi = 0$ when $\eta = 0$.

It is easy to verify that $-1 < \phi < 1$, and that $\Sigma = \Delta_n\left(\gamma^2(1-\phi)^2\mathbb{I}_n + \gamma^2\phi\mathbb{J}_n\right)$, since

$$\sigma^2 = \gamma^2(1-\phi)^2, \quad \text{and} \quad \eta = \gamma^2\phi. \tag{A.2}$$

By the calculations in Aït-Sahalia et al. (2005), we have

$$\det(\Sigma) = \gamma^{2n}\frac{1-\phi^{2n+2}}{1-\phi^2}\Delta_n^n, \quad \Sigma^{-1} = \Delta_n^{-1}\gamma^{-2}\frac{\phi^{|i-j|} - \phi^{i+j} - \phi^{2n+2-(i+j)} - \phi^{2n+2-|i-j|}}{(1-\phi^2)(1-\phi^{2n+2})},$$

where we use the notation that $\phi^k = \mathbf{1}_{\{k=0\}}$ if $\phi = 0$ (Note that when $\phi = 0$, Σ is diagonal). Therefore, the likelihood can be rewritten as

$$\begin{aligned} L(\gamma^2, \phi) = &- \frac{n}{2}\log 2\pi - \frac{n}{2}\log\gamma^2\Delta_n + \frac{1}{2}\log(1-\phi^2) - \frac{1}{2}\log(1-\phi^{2n+2}) \\ &- \frac{1}{2\gamma^2\Delta_n}\sum_{i=1}^n\sum_{j=1}^n \frac{\phi^{|i-j|} - \phi^{i+j} - \phi^{2n+2-(i+j)} - \phi^{2n+2-|i-j|}}{(1-\phi^2)(1-\phi^{2n+2})} Y_iY_j. \end{aligned}$$

We now focus on the asymptotic result of $(\widehat{\gamma}^2, \widehat{\phi})$, which will lead to the result for $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{\eta}_{\text{MLE}})$, since the change of variable in (A.1) and its inverse (A.2) are smooth.

Because $-1 < \phi < 1$, ϕ^{2n+2} is exponentially small. Therefore, we obtain the following score functions, up to some exponentially small terms:

$$\begin{aligned} \Psi_1 = &\frac{1}{n}\frac{\partial L(\gamma^2,\phi)}{\partial \gamma^2} = -\frac{1}{2\gamma^2} + \frac{1}{2\gamma^4(1-\phi^2)T}\sum_{i=1}^n\sum_{j=1}^n\left(\phi^{|i-j|} - \phi^{i+j} - \phi^{2n+2-(i+j)} - \phi^{2n+2-|i-j|}\right)Y_iY_j, \\ \Psi_2 = &\frac{1}{n}\frac{\partial L(\gamma^2,\phi)}{\partial \phi} = -\frac{\phi}{n(1-\phi^2)} - \frac{\phi}{\gamma^2(1-\phi^2)^2T}\sum_{i=1}^n\sum_{j=1}^n\left(\phi^{|i-j|} - \phi^{i+j} - \phi^{2n+2-(i+j)} - \phi^{2n+2-|i-j|}\right)Y_iY_j \\ &- \frac{1}{2\gamma^2(1-\phi^2)T}\sum_{i=1}^n\sum_{j=1}^n\left(|i-j|\phi^{|i-j|-1} - |i+j|\phi^{i+j-1}\right)Y_iY_j \end{aligned}$$

$$+\frac{1}{2\gamma^2(1-\phi^2)T}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big((2n+2-(i+j))\phi^{2n+1-(i+j)}+(2n+1-|i-j|)\phi^{2n+1-|i-j|}\Big)Y_iY_j.$$

Suppose A_{ij} can take values from

$$\Big\{\phi^{|i-j|},\quad \phi^{i+j},\quad \phi^{2n+2-(i+j)},\quad \phi^{2n+2-|i-j|},\quad |i-j|\phi^{|i-j|-1},\quad |i+j|\phi^{i+j-1},$$
$$(2n+2-(i+j))\phi^{2n+1-(i+j)},\quad (2n+1-|i-j|)\phi^{2n+1-|i-j|}\Big\},$$

and denote that

$$\zeta(1)_i^n=\frac{(\Delta_i^nX)^2}{\Delta_n}-\sigma^2_{(i-1)\Delta_n},\quad \zeta(2)_i^n=\frac{\Delta_i^nX}{\sqrt{\Delta_n}},\quad \zeta'(r)_i^n=\mathbb{E}\left(\zeta(r)_i^n|\mathcal{F}_{(i-1)\Delta_n}\right),\quad \zeta''(r)_i^n=\zeta(r)_i^n-\zeta'(r)_i^n.$$

By the standard estimates for continuous X, see e.g., Aït-Sahalia and Jacod (2014), we have

$$|\zeta'(1)_i^n|\le K\sqrt{\Delta_n},\quad \mathbb{E}\left(|\zeta(1)_i^n|^m|\mathcal{F}_{(i-1)\Delta_n}\right)\le K.$$

Moreover, the following estimates hold even when X contains jumps:

$$|\zeta'(2)_i^n|\le K\sqrt{\Delta_n},\quad \mathbb{E}\left(|\zeta(2)_i^n|^m|\mathcal{F}_{(i-1)\Delta_n}\right)\le K\Delta_m^{0\wedge(1-\frac{m}{2})}. \tag{A.3}$$

Therefore, it then follows from Doob's inequality that

$$\begin{aligned}\mathbb{E}\left(\sup_{l\le n}\left|\sum_{i=1}^{l}A_{ii}\zeta(1)_i^n\right|\right)\le&\mathbb{E}\left(\sum_{i=1}^{n}A_{ii}\left|\zeta'(1)_i^n\right|\right)+\left(\sum_{i=1}^{n}A_{ii}^2\mathbb{E}\left(\zeta(1)_i^n\right)^2\right)^{1/2}\\ \le&K\Delta_n^{1/2}\sum_{i=1}^{n}A_{ii}+K\left(\sum_{i=1}^{n}A_{ii}^2\right)^{1/2}\\ \le&\begin{cases}K\Delta_n^{-1/2}, & \text{if}\quad A_{ij}=\phi^{|i-j|},\\ K, & \text{otherwise.}\end{cases}\end{aligned}\tag{A.4}$$

Similarly, by successively applying Doob's inequality and Cauchy-Schwarz inequality, we have

$$\begin{aligned}&\mathbb{E}\left(\left|\sum_{i=1}^{n}\sum_{j=1}^{i-1}A_{ij}\zeta(2)_j^n\zeta(2)_i^n\right|\right)\\ \le&\mathbb{E}\left(\sum_{i=1}^{n}\left|\sum_{j=1}^{i-1}A_{ij}\zeta(2)_j^n\right|\left|\zeta'(2)_i^n\right|\right)+\mathbb{E}\left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{i-1}A_{ij}\zeta(2)_i^n\right)^2\mathbb{E}\left((\zeta(2)_i^n)^2|\mathcal{F}_{(i-1)\Delta_n}\right)\right\}^{1/2}\\ \le&K\Delta_n^{1/2}\sum_{i=1}^{n}\left\{\sum_{j=1}^{i-1}A_{ij}\Delta_n^{1/2}+\left(\sum_{j=1}^{i-1}A_{ij}^2\right)^{1/2}\right\}+\left\{\sum_{i=1}^{n}\left(\left(\sum_{j=1}^{i-1}A_{ij}\right)^2\Delta_n+\sum_{j=1}^{i-1}A_{ij}^2\right)\right\}^{1/2}\\ \le&\begin{cases}K\Delta_n^{-1/2}, & \text{if}\quad A_{ij}=\phi^{|i-j|}\text{ or }|i-j|\phi^{|i-j|-1},\\ K, & \text{otherwise.}\end{cases}\end{aligned}\tag{A.5}$$

Further, by Doob's inequality and Cauchy-Schwarz inequality again, we have, for all A_{ij},

$$
\begin{aligned}
&\mathbb{E}\left(\left|\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}(U_j - U_{j-1})\zeta(2)_i^n\right|\right)\\
\leq &\sum_{i=1}^{n}\mathbb{E}\left|\sum_{j=1}^{n} A_{ij}(U_j - U_{j-1})\right| |\zeta'(2)_i^n| + \mathbb{E}\left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}(U_j - U_{j-1})\right)^2 \mathbb{E}\left((\zeta(2)_i^n)^2|\mathcal{F}_{(i-1)\Delta_n}\right)\right\}^{1/2}\\
\leq &K\Delta_n^{(\gamma_0-1)/2}.
\end{aligned}
\tag{A.6}
$$

Finally, using Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\Big\{(U_j - U_{j-1})(U_i - U_{i-1}) - \mathbb{E}\left((U_j - U_{j-1})(U_i - U_{i-1})\right)\Big\}\right)^2\\
= &\sum_{i,j,i',j'=1}^{n} A_{ij}A_{i'j'}\big(\mathrm{cum}_4\left[U_i - U_{i-1}, U_j - U_{j-1}, U_{i'} - U_{i'-1}, U_{j'} - U_{j'-1}\right]\\
&+\left(2\cdot\mathbf{1}_{\{k=k'\}} - \mathbf{1}_{\{k=k'-1\}} - \mathbf{1}_{\{k=k'+1\}}\right)\left(2\cdot\mathbf{1}_{\{l=l'\}} - \mathbf{1}_{\{l=l'-1\}} - \mathbf{1}_{\{l=l'+1\}}\right)\big)\\
\leq &K\Delta_n^{2\gamma_0-1}.
\end{aligned}
\tag{A.7}
$$

The above calculations (A.4) - (A.7) lead to

$$
\Psi_1 = -\frac{1}{2\gamma^2} + \frac{1}{2\gamma^4(1-\phi^2)T}\sum_{i=1}^{n}\sum_{j=1}^{n}\phi^{|i-j|}\left(\Delta_i^n X\Delta_j^n X + \mathbb{E}(U_i - U_{i-1})(U_j - U_{j-1})\right) + O_p(\Delta_n),
\tag{A.8}
$$

$$
\begin{aligned}
\Psi_2 = &-\frac{\phi}{\gamma^2(1-\phi^2)^2T}\sum_{i=1}^{n}\sum_{j=1}^{n}\phi^{|i-j|}\left(\Delta_i^n X\Delta_j^n X + \mathbb{E}(U_i - U_{i-1})(U_j - U_{j-1})\right)\\
&-\frac{1}{2\gamma^2(1-\phi^2)T}\sum_{i=1}^{n}\sum_{j=1}^{n}|i-j|\phi^{|i-j|-1}\left(\Delta_i^n X\Delta_j^n X + \mathbb{E}(U_i - U_{i-1})(U_j - U_{j-1})\right) + O_p(\Delta_n),
\end{aligned}
\tag{A.9}
$$

because the other terms do not contribute to the asymptotic variances on the leading order, given that $\gamma_0 \geq 3/2$.

We introduce the following two functions:

$$
\overline{\Psi}_1 = -\frac{1}{2\gamma^2} + \frac{1}{2\gamma^4(1-\phi^2)T}\left(\int_0^T \sigma_s^2 ds + 2(1-\phi)Ta_0^2\Delta_n^{\gamma_0-1}\right),
\tag{A.10}
$$

$$
\overline{\Psi}_2 = -\frac{\phi}{\gamma^2(1-\phi^2)^2T}\left(\int_0^T \sigma_s^2 ds + 2(1-\phi)Ta_0^2\Delta_n^{\gamma_0-1}\right) + \frac{a_0^2\Delta_n^{\gamma_0-1}}{\gamma^2(1-\phi^2)}.
\tag{A.11}
$$

Since $\mathbb{E}(U_i - U_{i-1})(U_j - U_{j-1}) = a_0^2\Delta_n^{\gamma_0}\left(2\mathbf{1}_{\{i=j\}} - \mathbf{1}_{\{i=j-1\}} - \mathbf{1}_{\{i=j+1\}}\right)$, and

$$
\mathbb{E}\left(\sum_{i=1}^{n}\sigma_{(i-1)\Delta_n}^2\Delta_n - \int_0^T \sigma_s^2 ds\right)^2 \leq K\sum_{i=1}^{n}\int_{(i-1)\Delta_n}^{i\Delta_n}\mathbb{E}\left(\left\|\sigma_s^2 - \sigma_{(i-1)\Delta_n}^2\right\|^2 |\mathcal{F}_{(i-1)\Delta_n}\right) ds \leq K\Delta_n,
$$

it follows from (A.4) - (A.7) that

$$\left\|\Psi_i - \overline{\Psi}_i\right\| = O_p(\Delta_n^{1/2}), \quad i = 1, 2.$$

This also holds uniformly over the compact parameter space Θ in terms of (ϕ, η), i.e., $-1 + \epsilon_1 \leq \phi \leq 1 - \epsilon_1$, $\epsilon_2 \leq \gamma^2 \leq \epsilon_3$, because Ψ_i and $\overline{\Psi}_i$ can be written as the differences of convex functions and a slight modification of Theorem 10.8 in Rockafellar (1997). Suppose (γ^{2*}, ϕ^*) is the solution of $\overline{\Psi} = 0$. By Theorem 2 of Xiu (2010), we establish: $\widehat{\gamma}^2 - \gamma^{2*} = o_p(1)$ and $\widehat{\phi} - \phi^* = o_p(1)$. Moreover, using (A.1), (A.10), and (A.11), we derive that

$$\sigma^{2*} = \gamma^{2*}(1 - \phi^*)^2 = \frac{1}{T}\int_0^T \sigma_s^2 ds, \quad \eta^* = \gamma^{2*}\phi^* = a_0^2 \Delta_n^{\gamma_0 - 1},$$

hence the desired consistency of $(\widehat{\sigma}^2_{\mathrm{MLE}}, \widehat{\eta}_{\mathrm{MLE}})$ follows from the consistency of $(\widehat{\gamma}^2, \widehat{\phi})$.

We now derive the joint CLT of $\Psi_i - \overline{\Psi}_i$, $i = 1, 2$, evaluated at (γ^{2*}, ϕ^*), along with $\widehat{\sigma}^2_{\mathrm{RV}}$, so the CLT of $(\widehat{\sigma}^2_{\mathrm{MLE}}, \widehat{\eta}_{\mathrm{MLE}}, \widehat{\sigma}^2_{\mathrm{RV}})$ will follow directly from the "sandwich" formula and the "Delta" method, using Theorem 3 of Xiu (2010).

Using (A.1), we have $\gamma^{2*} = \sigma^{2*} + o(1)$, and $\phi^* = a_0^2 \Delta_n^{\gamma_0 - 1} + O(\Delta_n^{1/2}) = O(\Delta_n^{1/2})$. Therefore, evaluated at (γ^{2*}, ϕ^*), the scaled difference between Ψ_i and $\overline{\Psi}_i$ satisfies

$$\begin{aligned}
\Delta_n^{-1/2}\Psi_1(\gamma^{2*}, \phi^*) =& \frac{1}{2\gamma^{*4}T}\left(\sum_{i=1}^n (\Delta_i^n X)^2 - \int_0^T \sigma_s^2 ds\right) + O_p(\Delta_n^{1/2}), \\
\Delta_n^{-1/2}\Psi_2(\gamma^{2*}, \phi^*) =& -\frac{1}{\gamma^{*2}T}\sum_{i=1}^n \Delta_i^n X \Delta_{i-1}^n X + O_p(\Delta_n^{1/2}),
\end{aligned}$$

so that we can apply Theorem 11.2.1 in Jacod and Protter (2012), and obtain

$$\Delta_n^{-1/2}\left(\Psi_1(\gamma^{2*}, \phi^*), \Psi_2(\gamma^{2*}, \phi^*), \widehat{\sigma}^2_{\mathrm{RV}} - \frac{1}{T}\int_0^T \sigma_s^2 ds\right)^{\intercal} \xrightarrow{\mathcal{L}-s} \bar{\mathcal{W}}_T,$$

where $\bar{\mathcal{W}}_T$ is defined on the extension of the original probability space, which is a $\mathcal{F}$-conditional Gaussian variable with the covariance matrix given by

$$\mathbb{E}\left(\bar{\mathcal{W}}_T \bar{\mathcal{W}}_T^{\intercal} | \mathcal{F}\right) = \begin{pmatrix} (2\gamma^{8*})^{-1} & 0 & (\gamma^{4*})^{-1} \\ 0 & (\gamma^{4*})^{-1} & 0 \\ (\gamma^{4*})^{-1} & 0 & 2 \end{pmatrix} \times \frac{1}{T^2}\int_0^T \sigma_s^4 ds.$$

Moreover, by (A.10) and (A.11) we have

$$\begin{pmatrix} -\nabla\overline{\Psi}(\gamma^{2*}, \phi^*) & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} (2\gamma^{4*})^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + o_p(1). \tag{A.12}$$

By entry-wise Taylor expansion of the vector-valued function (Ψ_1, Ψ_2) and the "sandwich" formula, we have the asymptotic distribution of $(\widehat{\gamma}^2, \widehat{\phi}, \widehat{\sigma}^2_{\text{RV}})$, which is an $\mathcal{F}$-conditional centered Gaussian variable with covariance matrix given by

$$\begin{pmatrix} -\nabla\overline{\Psi}(\gamma^{2*},\phi^*) & 0 \\ 0 & 1 \end{pmatrix}^{-1} \mathbb{E}\left(\bar{\mathcal{W}}_T\bar{\mathcal{W}}_T^\intercal|\mathcal{F}\right) \begin{pmatrix} -\nabla\overline{\Psi}(\gamma^{2*},\phi^*) & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & \gamma^{-4*} & 0 \\ 2 & 0 & 2 \end{pmatrix} \times \frac{1}{T^2}\int_0^T \sigma_s^4 ds.$$

Finally, using (A.2) and the "Delta" method, we have the asymptotic distribution of $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{\eta}_{\text{MLE}}, \widehat{\sigma}^2_{\text{RV}})$, $\mathcal{W}_T$, defined on the extension of the original probability space. Conditional on $\mathcal{F}$, it is Gaussian variable with covariance matrix given by

$$\begin{aligned}\mathbb{E}\left(\mathcal{W}_T\mathcal{W}_T^\intercal|\mathcal{F}\right) &= \begin{pmatrix} 1 & -2\gamma^{2*} & 0 \\ 0 & \gamma^{2*} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 2 \\ 0 & \gamma^{-4*} & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2\gamma^{2*} & \gamma^{2*} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \frac{1}{T^2}\int_0^T \sigma_s^4 ds \\ &= \begin{pmatrix} 6 & -2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix} \times \frac{1}{T^2}\int_0^T \sigma_s^4 ds.\end{aligned}$$

∎

Appendix A.2 Proof of Corollary 1

Proof. By Theorem 1, we have

$$\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\text{MLE}} - \widehat{\sigma}^2_{\text{RV}}\right) \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \frac{4}{T^2}\int_0^T \sigma_s^4 ds\right).$$

It then follows from the consistency of $\widehat{V}_{2n}$ that

$$H_{2n} \xrightarrow{\mathcal{L}} \chi_1^2, \quad \text{under } \mathbb{H}_0.$$

Moreover, under $\mathbb{H}_1$, by (22) we have $\widehat{\sigma}^2_{\text{MLE}} = O_p(1)$, whereas

$$\widehat{\sigma}^2_{\text{RV}} = \frac{1}{T}\int_0^T \sigma_s^2 ds + \frac{2a_0^2}{\Delta_n} + o_p(1).$$

Also, $\widehat{V}_{2n} = O_p(\Delta_n^{-2})$. Therefore, $H_{2n} = O_p(\Delta_n^{-1})$. Finally, under $\mathbb{H}_n$, we have

$$\widehat{\sigma}^2_{\text{RV}} = \frac{1}{T}\left\{\sum_{i=1}^n (\Delta_i^n X)^2 + 2a_n\sum_{i=1}^n \Delta_i^n X(U_i - U_{i-1}) + a_n^2\sum_{i=1}^n (U_i - U_{i-1})^2\right\}.$$

It is easy to derive that

$$\sum_{i=1}^n (U_i - U_{i-1})^2 = 2n + O_p(n^{1/2}), \quad \sum_{i=1}^n (\Delta_i^n X)(U_i - U_{i-1}) = O_p(1),$$

hence we have

$$\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\mathrm{RV}} - \frac{1}{T}\int_0^T \sigma_s^2 ds\right) = 2a_0^2 + \Delta_n^{-1/2}\left(\frac{1}{T}\sum_{i=1}^{n}(\Delta_i^n X)^2 - \frac{1}{T}\int_0^T \sigma_s^2 ds\right) + o_p(1).$$

Combining this with Theorem 1, we have:

$$\begin{aligned}\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\mathrm{MLE}} - \widehat{\sigma}^2_{\mathrm{RV}}\right) =& \Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\mathrm{MLE}} - \frac{1}{T}\int_0^T \sigma_s^2 ds\right) + \Delta_n^{-1/2}\left(\frac{1}{T}\int_0^T \sigma_s^2 ds - \widehat{\sigma}^2_{\mathrm{RV}}\right)\\ &\xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(2a_0^2, \frac{4}{T^2}\int_0^T \sigma_s^4 ds\right).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\widehat{V}_{2n} =& \frac{4}{3\Delta_n T^2}\sum_{i=1}^{n}\Big\{(\Delta_i^n X)^4 + a_n^4(U_i - U_{i-1})^4 + 4a_n\left(\Delta_i^n X\right)^3\left(U_i - U_{i-1}\right)\\ &+6a_n^2\left(\Delta_i^n X\right)^2\left(U_i - U_{i-1}\right)^2 + 4a_n^3\left(\Delta_i^n X\right)\left(U_i - U_{i-1}\right)^3\Big\}.\end{aligned}$$

It is easy to show that under $\mathbb{H}_n$,

$$\begin{aligned}&\sum_{i=1}^{n}(U_i - U_{i-1})^4 = 12n + O_p(n^{1/2}), \quad \sum_{i=1}^{n}\left(\Delta_i^n X\right)^3\left(U_i - U_{i-1}\right) = O_p(n^{-1}),\\ &\sum_{i=1}^{n}\left(\Delta_i^n X\right)\left(U_i - U_{i-1}\right)^3 = O_p(1), \quad \sum_{i=1}^{n}\left(\Delta_i^n X\right)^2\left(U_i - U_{i-1}\right)^2 = 2\int_0^T \sigma_s^2 ds + O_p(n^{-1/2}),\end{aligned}$$

hence it follows that

$$\widehat{V}_{2n} = \frac{4}{T^2}\int_0^T \sigma_s^4 ds + O_p(\Delta_n) + O_p(\Delta_n^{3/4}) + O_p(\Delta_n^{5/4}) + O_p(\Delta_n^{1/2}),$$

which establishes the desired result. ∎

Appendix A.3 Proof of Theorem 2

Proof. The proof is similar to that of Theorem 1, but we need the following lemma.

Lemma 1. *Suppose that X is an Itô semimartingale given by (30), and that $\bar{\gamma}_t^n$ is the $(q+1)$-dimensional vector-valued centered autocovariance function:*

$$\begin{aligned}\bar{\gamma}_t^n &= \left(\gamma(0)_t^n - \int_0^t \sigma_s^2 ds - \sum_{0\le s\le t}(\Delta X_s)^2, \gamma(1)_t^n, \ldots, \gamma(q)_t^n\right)^{\intercal}, \quad \textit{where}\\ \gamma(l)_t^n &= \sum_{i=1}^{[t/\Delta_n]-l}(\Delta_i^n X)(\Delta_{i+l}^n X), \quad 0 \le l \le q.\end{aligned}$$

Then we have as $\Delta_n \to 0$,

$$\Delta_n^{-1/2}\bar{\gamma}_t^n \xrightarrow{\mathcal{L}-s} \mathcal{W}_t + \mathcal{Z}_t,$$

where $\mathcal{W}$ is a continuous process defined on an extension of the original probability space, which conditionally on $\mathcal{F}$, is a centered Gaussian martingale with covariance given by

$$\mathbb{E}\left(\mathcal{W}_{i,t}\mathcal{W}_{j,t}|\mathcal{F}\right) = (\delta_{i,j} + \delta_{i,0}\delta_{j,0})\int_0^t \sigma_s^4 ds, \quad 0 \leq i, j \leq q;$$

and $\mathcal{Z}$ is a purely discontinuous process defined on the same extension of the original probability space, which conditionally on $\mathcal{F}$, is a centered martingale with covariance given by

$$\mathbb{E}\left(\mathcal{Z}_{i,t}\mathcal{Z}_{j,t}|\mathcal{F}\right) = (\delta_{i,j} + \delta_{i,0}\delta_{j,0})\sum_{0\leq s\leq t}(\Delta X_s)^2(\sigma_{s-}^2 + \sigma_s^2), \quad 0 \leq i, j \leq q,$$

and $\mathcal{Z}$ is independent of $\mathcal{W}$. Moreover, $\mathcal{Z}_i$ is a Gaussian process for $i \geq 1$. $\mathcal{Z}_0$ is Gaussian if X and σ^2 do not co-jump.

Proof of Lemma 1 . We set $\mathcal{Q}_- = \{-q, -q+1, \ldots, -1\}$, $\mathcal{Q}_+ = \{1, 2, \ldots, q\}$, and $\mathcal{Q} = \mathcal{Q}_- \cup \mathcal{Q}_+$. We introduce a family of variables $((\Psi_{n,j})_{j\in\mathcal{Q}}, \Psi_{n-}, \Psi_{n+}, \kappa_n)_{n\geq 1}$, defined on an auxiliary space $(\Omega', \mathcal{F}', \mathbb{P}')$, all independent, and with the following laws: $\Psi_{n,j}$, Ψ_{n-}, Ψ_{n+} are i.i.d. $\mathcal{N}(0,1)$. κ_n is uniform on $[0,1]$. $(\tau_n)_{n\geq 1}$ is an arbitrary weekly exhausting sequence for the jumps of X. We define a very good filtration $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$, such that

$$\widetilde{\Omega} = \Omega \times \Omega', \quad \widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \quad \widetilde{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}',$$

where $(\widetilde{\mathcal{F}}_t)_{t\geq 0}$ is the smallest filtration containing $(\mathcal{F}_t)_{t\geq 0}$ and such that $((\Psi_{n,j})_{j\in\mathcal{Q}}, \Psi_{n-}, \Psi_{n+}, \kappa_n)$ is $\widetilde{\mathcal{F}}_{\tau_n}$ measurable for all n. We also define

$$R_{n,j} = \begin{cases} \sigma_{\tau_n-}\Psi_{n,j} & \text{if} \quad j \in \mathcal{Q}_-; \\ \sqrt{\kappa_n}\sigma_{\tau_n-}\Psi_{n-} + \sqrt{1-\kappa_n}\sigma_{\tau_n-}\Psi_{n+} & \text{if} \quad j = 0; \\ \sigma_{\tau_n}\Psi_{n,j} & \text{if} \quad j \in \mathcal{Q}_+. \end{cases}$$

W'_t is a $(q+1)$-dimensional standard Brownian motion in $(\Omega', \mathcal{F}', \mathbb{P}')$, adapted to $(\widetilde{\mathcal{F}}_t)_{t\geq 0}$, and independent of $((\Psi_{n,j})_{j\in\mathcal{Q}}, \Psi_{n-}, \Psi_{n+}, \kappa_n)_{n\geq 1}$. Finally, we define two $(q+1)$-dimensional processes on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$:

$$\mathcal{W}_t = \int_0^t \sigma_s^2 dW'_s \times \begin{pmatrix} \sqrt{2} & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix},$$

$$\mathcal{Z}_t = \sum_{n=1,\tau_n\leq t}^{\infty} \Delta X_{\tau_n}\left(2R(n,0), R(n,1) + R(n,-1), \ldots, R(n,q) + R(n,-q)\right)^{\intercal},$$

which are a.s. càdlàg, adapted, and conditionally on $\mathcal{F}$ have centered and independent increments.

By localization, there exists a constant A and a non-negative function Γ such that

$$\|b_t(\omega)\| \leq A, \quad \|\sigma_t(\omega)\| \leq A, \quad \|X_t(\omega)\| \leq A, \quad \|\delta(\omega,t,z)\| \leq \Gamma(z) \leq A, \quad \int \Gamma(z)^2 \bar{\nu}(dz) \leq A.$$

Let $A_m = \{z : \Gamma(z) > 1/m\}$, $(S(m,j) : j \geq 1)$ are the successive jump times of the Poisson process $1_{A_m \backslash A_{m-1}} * \mu$, $(S_p)_{p\geq 1}$ is a reordering of the double sequence $(S(m,j))$, and let $\mathcal{P}_m$ denote the set of all indices p such that $S_p = S(m',j)$ for some $j \geq 1$ and some $m' \leq m$. We also introduce

$$\begin{aligned}
&S_-(n,p) = (i-1)\Delta_n, \quad S_+(n,p) = i\Delta_n, \quad \text{if } (i-1)\Delta_n < S_p \leq i\Delta_n. \\
&R_-(n,p) = \frac{1}{\sqrt{\Delta_n}}(X_{S_p} - X_{(i-1)\Delta_n}), \quad R_+(n,p) = \frac{1}{\sqrt{\Delta_n}}(X_{i\Delta_n} - X_{S_p}), \\
&R(n,p,0) = \frac{1}{\sqrt{\Delta_n}}(\Delta_i^n X - \Delta X_{S_p}), \quad R(n,p,j) = \frac{1}{\sqrt{\Delta_n}}\Delta_{i+j}^n X, \quad -q \leq j \leq q, \\
&b(m)_t = b_t - \int_{A_m \cap \{z : \|\delta(t,z)\| \leq 1\}} \delta(t,z)\bar{\nu}(dz), \\
&X(m)_t = X_0 + \int_0^t b(m)_s ds + \int_0^t \sigma_s dW_s + (\delta 1_{A_m}) * \mu, \\
&X'(m) = (\delta 1_{A_m}) * \mu, \quad X''(m) = X(m) - X'(m), \quad X^{\#}(m) = X - X(m) = (\delta 1_{A_m^c}) * (\mu - \nu)_t.
\end{aligned}$$

We set $\Omega_n(T,m)$ to be the set of all ω such that the jumps of $X'(m)$ in $[0,T]$ are spaced by more than $q\Delta_n$, and no such jump occurs in $[0, q\Delta_n]$ or $[T - q\Delta_n, T]$. Note that $\mathbb{P}(\Omega_n(T,m)) \to 1$, as $n \to \infty$. We also use the notation:

$$\begin{aligned}
&\beta_{i,j}^n = \Delta_n^{-1/2}\sigma_{(i-1)\Delta_n}\Delta_{i+j}^n W, \quad \bar{\beta}_i^n = (\beta_{i,0}^n, \beta_{i,1}^n, \ldots, \beta_{i,q}^n)^\intercal \times 1_{\{i+q \leq [t/\Delta_n]\}}, \\
&X(m)_{i,j}^{\prime\prime,n} = \Delta_n^{-1/2}\Delta_{i,j}^n X''(m), \quad \bar{X}(m)_i^{\prime\prime,n} = \left(X(m)_{i,0}^{\prime\prime,n}, X(m)_{i,1}^{\prime\prime,n}, \ldots, X(m)_{i,q}^{\prime\prime,n}\right)^\intercal \times 1_{\{i+q \leq [t/\Delta_n]\}},
\end{aligned}$$

We use $\bar{\gamma}_t^n(X,Y)$ to denote the vector-valued cross-autocovariance function between X and Y. Therefore, $\bar{\gamma}_t^n \equiv \bar{\gamma}_t^n(X,X)$. We set

$$\begin{aligned}
\mathcal{Z}_t^n(m) &= \sum_{p \in \mathcal{P}_m : S_p \leq \Delta_n[t/\Delta_n]} \zeta_p^n, \quad \text{where} \quad \zeta_p^n = (\zeta_{p,0}^n, \zeta_{p,1}^n, \ldots, \zeta_{p,q}^n)^\intercal, \quad \text{and} \\
\zeta_{p,j}^n &= (R(n,p,j) + R(n,p,-j))\,\Delta X_{S_p},
\end{aligned}$$

We have on the set $\Omega_n(T,m)$,

$$\frac{1}{\sqrt{\Delta_n}}\bar{\gamma}_T^n(X(m),X(m)) = \frac{1}{\sqrt{\Delta_n}}\bar{\gamma}_T^n(X''(m),X''(m)) + \mathcal{Z}_T^n(m).$$

By Proposition 4.4.10 and Lemma 11.1.3 of Jacod and Protter (2012), we have

$$\left\{(R(n,p,j))_{-q \leq j \leq q}\right\}_{p \geq 1} \xrightarrow{\mathcal{L}-s} \left\{(R_{p,j})_{-q \leq j \leq q}\right\}_{p \geq 1},$$

therefore, writing $\zeta_p = (\zeta_{p,0}, \zeta_{p,1}, \ldots, \zeta_{p,q})^\intercal$ and $\zeta_{p,j} = (R(p,j) + R(p,-j))\,\Delta X_{S_p}$, we have

$$(\zeta_p^n)_{p \geq 1} \xrightarrow{\mathcal{L}-s} (\zeta_p)_{p \geq 1}.$$

Since the set $\{S_p : p \in \mathcal{P}_m\} \cap [0,t]$ is finite, we have, as $n \to \infty$,

$$\mathcal{Z}_t^n(m) \xrightarrow{\mathcal{L}-s} \mathcal{Z}_t(m),$$

where

$$\mathcal{Z}_t(m) = \sum_{n=1,\tau_n \le t}^{\infty} \Delta X'(m)_{\tau_n} \left(2R(n,0), R(n,1)+R(n,-1), \dots, R(n,q)+R(n,-q)\right)^{\intercal}.$$

Note that by Doob's inequality, we have

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left(\sup_{s\le t} \|\mathcal{Z}_s - \mathcal{Z}_s(m)\|^2\right) \\
=&\mathbb{E}\left(\widetilde{\mathbb{E}}\left(\sup_{s\le t}\left\| \sum_{n=1,\tau_n\le t}^{\infty} \Delta X^{\#}(m)_{\tau_n} \left(2R(n,0), R(n,1)+R(n,-1), \dots, R(n,q)+R(n,-q)\right)^{\intercal}\right\|^2 \mid \mathcal{F}\right)\right) \\
\le& K\mathbb{E}\left(\sum_{s\le t}\left\|\Delta X^{\#}(m)_s\right\|^2 \left(\|\sigma_s^2\| + \|\sigma_{s-}^2\|\right) \mid \mathcal{F}\right) \\
\le& K\mathbb{E}\left(\sum_{s\le t}\left\|\Delta X^{\#}(m)_s\right\|^2\right) \le Kt\alpha(m) \longrightarrow 0,
\end{aligned}$$

as $m \to \infty$, where $\alpha(m) = \int_{z:\Gamma(z)\le m^{-1}} \Gamma(z)^2 \bar{\nu}(dz)$. This implies that $\mathcal{Z}(m) \overset{u.c.p.}{\Longrightarrow} \mathcal{Z}$.

Since we have

$$\frac{\Delta_{i+j}^n X''(m)}{\sqrt{\Delta_n}} - \beta_{i,j}^n = \frac{1}{\sqrt{\Delta_n}} \int_{(i+j-1)\Delta_n}^{(i+j)\Delta_n} b_s ds + \frac{1}{\sqrt{\Delta_n}} \int_{(i+j-1)\Delta_n}^{(i+j)\Delta_n} (\sigma_s - \sigma_{(i-1)\Delta_n}) dW_s,$$

it follows from (11.2.22) of Jacod and Protter (2012) that, for all $l \ge 2$,

$$\mathbb{E}\left(\|\bar{\beta}_i^n\|^l\right) + \mathbb{E}\left(\|\bar{X}(m)_i^{\prime\prime,n}\|^l\right) \le K_l, \quad \mathbb{E}\left(\|\bar{X}(m)_i^{\prime\prime,n} - \bar{\beta}_i^n\|^l\right) \le K_l \Delta_n.$$

We can also write

$$\frac{1}{\sqrt{\Delta_n}} \bar{\gamma}_T^n(X''(m), X''(m)) = \mathcal{W}_t^n + A_t^n(m,0) + A_t^n(m,1) - B_t^n,$$

where

$$\begin{aligned}
\mathcal{W}_t^n =& \sqrt{\Delta_n} \sum_{i=1}^{[t/\Delta_n]} \left(\beta_{i,0}^n \bar{\beta}_i^n - \left(\sigma_{(i-1)\Delta_n}^2, 0, \dots, 0\right)^{\intercal}\right), \\
A_t^n(m,0) =& \sqrt{\Delta_n} \sum_{i=1}^{[t/\Delta_n]} \chi_i^n(m,0), \quad A_t^n(m,1) = \sqrt{\Delta_n} \sum_{i=1}^{[t/\Delta_n]} \chi_i^n(m,1), \\
B_t^n =& \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{[t/\Delta_n]} \int_{(i-1)\Delta_n}^{i\Delta_n} \left(\sigma_s^2 - \sigma_{(i-1)\Delta_n}^2\right) ds,
\end{aligned}$$

and

$$\chi_i^n(m,0) = \mathbb{E}\left(X(m)_{i,0}^{\prime\prime,n}\bar{X}(m)_i^{\prime\prime,n} - \beta_{i,0}^n\bar{\beta}_i^n|\mathcal{F}_{(i-1)\Delta_n}\right),$$
$$\chi_i^n(m,1) = X(m)_{i,0}^{\prime\prime,n}\bar{X}(m)_i^{\prime\prime,n} - \beta_{i,0}^n\bar{\beta}_i^n - \mathbb{E}\left(X(m)_{i,0}^{\prime\prime,n}\bar{X}(m)_i^{\prime\prime,n} - \beta_{i,0}^n\bar{\beta}_i^n|\mathcal{F}_{(i-1)\Delta_n}\right).$$

By Lemmas 11.2.5 and 11.2.7 of Jacod and Protter (2012), we have

$$\mathcal{W}_t^n = \Delta_n^{-1/2}\sum_{i=1}^{[t/\Delta_n]}\sigma_{(i-1)\Delta_n}^2\Delta_i^n W_0'\Delta_i^n W' \xrightarrow{\mathcal{L}-s} \mathcal{W}_t,$$
$$A_t^n(m,j) \overset{u.c.p.}{\Longrightarrow} 0, \quad j=0,1.$$

By e.g., (5.3.24) of Jacod and Protter (2012), we have $B_t^n \overset{u.c.p.}{\Longrightarrow} 0$. Combining the above results yield

$$\frac{1}{\sqrt{\Delta_n}}\bar{\gamma}_t^n(X(m),X(m)) \xrightarrow{\mathcal{L}-s} \mathcal{W}_t + \mathcal{Z}_t,$$

for each m.

By Proposition 2.2.4 of Jacod and Protter (2012), it remains to prove that for any $\eta > 0$ and $t > 0$, as $m \to \infty$,

$$\sup_n \mathbb{P}\left(\sup_{s\le t}\|\bar{\gamma}_s^n(X(m),X(m)) - \bar{\gamma}_s^n(X,X)\| > \eta\sqrt{\Delta_n}\right) \longrightarrow 0. \tag{A.13}$$

We only need to prove this result for $\gamma_s^n(j)(X(m),X(m)) - \gamma_s^n(j)(X,X)$, for some $1 \le j \le q$, because the case with $j=0$ has been shown in (5.4.26) of Jacod and Protter (2012). Note that

$$\sum_{i=1}^{[t/\Delta_n]-j}\left(\Delta_i^n X(m)\Delta_{i+j}^n X(m) - \Delta_i^n X\Delta_{i+j}^n X\right) = \sum_{i=1}^{[t/\Delta_n]-j}\left(\Delta_i^n X\Delta_{i+j}^n X^{\#}(m) + \Delta_i^n X^{\#}(m)\Delta_{i+j}^n X(m)\right).$$

By Doob's inequality, we have

$$\mathbb{E}\left(\sup_{s\le t}\left|\sum_{i=1}^{[s/\Delta_n]-j}\Delta_i^n X\Delta_{i+j}^n X^{\#}(m)\right|\right) \le \left(\mathbb{E}\sum_{i=1}^{[t/\Delta_n]-j}\left(\Delta_i^n X\Delta_{i+j}^n X^{\#}(m)\right)^2\right)^{1/2} \le K(\alpha(m)\Delta_n)^{1/2}, \tag{A.14}$$

where we use the fact that for any $s \in [(i-1)\Delta_n, i\Delta_n]$,

$$\mathbb{E}\left(X^{\#}(m)_s - X^{\#}(m)_{(i-1)\Delta_n}\right)^2 \le K\alpha(m)\Delta_n, \quad \mathbb{E}\left(X_s - X_{(i-1)\Delta_n}\right)^2 \le K\Delta_n.$$

Similarly, by Doob's inequality and the Cauchy-Schwarz inequality, we have:

$$\mathbb{E}\left(\sup_{s\le t}\left|\sum_{i=1}^{[s/\Delta_n]-j}\Delta_i^n X^{\#}(m)\Delta_{i+j}^n X(m)\right|\right) \tag{A.15}$$
$$\le \left(\mathbb{E}\sum_{i=1}^{[t/\Delta_n]-j}\left(\Delta_i^n X^{\#}(m)\Delta_{i+j}^n X(m)\right)^2\right)^{1/2} + \sum_{i=1}^{[t/\Delta_n]-j}\left|\mathbb{E}\left(\Delta_i^n X^{\#}(m)\Delta_{i+j}^n X(m)|\mathcal{F}_{(i-1)\Delta_n}\right)\right|$$

$$\leq K(\alpha(m)\Delta_n)^{1/2}.$$

Combining (A.14) and (A.15) yields (A.13), which concludes the proof. ∎

To prove Theorem 2, we now adopt the same change of variable as in (A.1), and the same definition of Ψ_i, $i = 1, 2$. Note that (A.5), (A.6), and (A.7) still hold, when X contains jumps, due to (A.3). Therefore, (A.8) and (A.9) also hold. We need a change in the definition of $\overline{\Psi}_i$, $i = 1, 2$.

Define

$$\overline{\Psi}_1 = -\frac{1}{2\gamma^2} + \frac{1}{2\gamma^4(1-\phi^2)T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\leq T}(\Delta X_s)^2 + 2(1-\phi)Ta_0^2\Delta_n^{\gamma_0-1}\right),$$

$$\overline{\Psi}_2 = -\frac{\phi}{\gamma^2(1-\phi^2)^2T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\leq T}(\Delta X_s)^2 + 2(1-\phi)Ta_0^2\Delta_n^{\gamma_0-1}\right) + \frac{a_0^2\Delta_n^{\gamma_0-1}}{\gamma^2(1-\phi^2)}.$$

Then, we have

$$\begin{aligned}\Psi_1 - \overline{\Psi}_1 =& \frac{1}{2\gamma^4(1-\phi^2)T}\left\{\left(\sum_{i=1}^n(\Delta_i^n X)^2 - \int_0^T \sigma_s^2 ds - \sum_{s\leq T}(\Delta X_s)^2\right) + 2\sum_{i=1}^n\sum_{j<i}\phi^{|i-j|}\Delta_i^n X\Delta_j^n X\right\} \\ &+ O_p(\Delta_n) \\ \Psi_2 - \overline{\Psi}_2 =& -\frac{\phi}{\gamma^2(1-\phi^2)^2T}\left(\sum_{i=1}^n(\Delta_i^n X)^2 - \int_0^T \sigma_s^2 ds - \sum_{s\leq T}(\Delta X_s)^2\right) \\ &- \frac{1}{\gamma^2(1-\phi^2)^2T}\sum_{i=1}^n\sum_{j<i}\left(2\phi^{|i-j|+1} + (1-\phi^2)|i-j|\phi^{|i-j|-1}\right)\Delta_i^n X\Delta_j^n X + O_p(\Delta_n)\end{aligned}$$

By Theorem 5.4.2 of Jacod and Protter (2012),

$$\sum_{i=1}^n(\Delta_i^n X)^2 - \int_0^T \sigma_s^2 ds - \sum_{s\leq T}(\Delta X_s)^2 = O_p(\Delta_n^{1/2}).$$

Combining this with (A.5), (A.6), and (A.7), it follows that

$$\left\|\Psi_i - \overline{\Psi}_i\right\| = O_p(\Delta_n^{1/2}), \quad i = 1, 2.$$

As before, we can then establish the desired consistency of $(\widehat{\sigma}^2_{\text{MLE}}, \widehat{\eta}_{\text{MLE}})$ with respect to $(\sigma^{2*}, \eta^*) = \left(\frac{1}{T}\left\{\int_0^T \sigma_s^2 ds + \sum_{s\leq T}(\Delta X_s)^2\right\}, a_0^2\Delta_n^{\gamma_0-1}\right)$.

To find the CLT, we consider $\Psi_i - \overline{\Psi}_i$ evaluated at (γ^{2*}, ϕ^*). Note that $\overline{\Psi}_i(\gamma^{2*}, \phi^*) = 0$, so

$$\begin{aligned}\Delta_n^{-1/2}\Psi_1(\gamma^{2*}, \phi^*) =& \frac{1}{2\gamma^{*4}T}\left\{\sum_{i=1}^n(\Delta_i^n X)^2 - \int_0^T \sigma_s^2 ds - \sum_{s\leq T}(\Delta X_s)^2\right\} + O_p(\Delta_n^{1/2}), \\ \Delta_n^{-1/2}\Psi_2(\gamma^{2*}, \phi^*) =& -\frac{1}{\gamma^{*2}T}\sum_{i=2}^n \Delta_i^n X\Delta_{i-1}^n X + O_p(\Delta_n^{1/2}).\end{aligned}$$

By Lemma 1 below, we have

$$\Delta_n^{-1/2}\begin{pmatrix}\Psi_1(\gamma^{2*},\phi^*)\\ \Psi_2(\gamma^{2*},\phi^*)\\ \widehat{\sigma}^2_{\mathrm{RV}}-\frac{1}{T}\left(\int_0^T\sigma_s^2 ds+\sum_{s\le T}(\Delta X_s)^2\right)\end{pmatrix}\xrightarrow{\mathcal{L}-s}\bar{\mathcal{W}}_T+\bar{\mathcal{Z}}_T, \tag{A.16}$$

where $\bar{\mathcal{W}}_T$ is the same as in the proof of Theorem 1, and $\bar{\mathcal{Z}}_T$ is defined on the same extension as $\bar{\mathcal{W}}_T$. The conditional covariance of $\bar{\mathcal{Z}}_T$ is given by:

$$\mathbb{E}\left(\bar{\mathcal{Z}}_T\bar{\mathcal{Z}}_T^{\intercal}|\mathcal{F}\right)=\begin{pmatrix}(2\gamma^{8*})^{-1} & 0 & (\gamma^{4*})^{-1}\\ 0 & (\gamma^{4*})^{-1} & 0\\ (\gamma^{4*})^{-1} & 0 & 2\end{pmatrix}\times\frac{1}{T^2}\sum_{s\le T}(\Delta X_s)^2(\sigma_{s-}^2+\sigma_s^2).$$

Moreover, $\bar{\mathcal{Z}}_{2,T}$ is $\mathcal{F}$-conditional Gaussian. By entry-wise Taylor expansion of Ψ_i, and the convergence of $\nabla\Psi$ to $\nabla\overline{\Psi}$, we have

$$\begin{aligned}\Delta_n^{-1/2}\begin{pmatrix}\widehat{\gamma}^2-\gamma^{2*}\\ \widehat{\phi}-\phi^*\\ \widehat{\sigma}^2_{\mathrm{RV}}-\sigma^{2*}\end{pmatrix}=&-\Delta_n^{-1/2}\begin{pmatrix}\partial\overline{\Psi}_1/\partial\gamma^2 & \partial\overline{\Psi}_1/\partial\phi & 0\\ \partial\overline{\Psi}_2/\partial\gamma^2 & \partial\overline{\Psi}_2/\partial\phi & 0\\ 0 & 0 & 1\end{pmatrix}^{-1}\begin{pmatrix}\Psi_1(\gamma^{2*},\phi^*)\\ \Psi_2(\gamma^{2*},\phi^*)\\ \widehat{\sigma}^2_{\mathrm{RV}}-\sigma^{2*}\end{pmatrix}+o_p(1)\\ =&-\Delta_n^{-1/2}\begin{pmatrix}(2\gamma^{4*}) & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\Psi_1(\gamma^{2*},\phi^*)\\ \Psi_2(\gamma^{2*},\phi^*)\\ \widehat{\sigma}^2_{\mathrm{RV}}-\sigma^{2*}\end{pmatrix}+o_p(1).\end{aligned}$$

Applying the change of variable from (γ^2,ϕ) to (σ^2,η), we have

$$\begin{aligned}\Delta_n^{-1/2}\begin{pmatrix}\widehat{\sigma}^2_{\mathrm{MLE}}-\sigma^{2*}\\ \widehat{\eta}-\eta^*\\ \widehat{\sigma}^2_{\mathrm{RV}}-\sigma^{2*}\end{pmatrix}=&\Delta_n^{-1/2}\begin{pmatrix}1 & -2\gamma^{2*} & 0\\ 0 & \gamma^{2*} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\widehat{\gamma}^2-\gamma^{2*}\\ \widehat{\phi}-\phi^*\\ \widehat{\sigma}^2_{\mathrm{RV}}-\sigma^{2*}\end{pmatrix}+o_p(1)\\ =&\Delta_n^{-1/2}\begin{pmatrix}2\gamma^{4*} & -2\gamma^{2*} & 0\\ 0 & \gamma^{2*} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\Psi_1(\gamma^{2*},\phi^*)\\ \Psi_2(\gamma^{2*},\phi^*)\\ \widehat{\sigma}^2_{\mathrm{RV}}-\sigma^{2*}\end{pmatrix}+o_p(1).\end{aligned}$$

A simple matrix multiplication gives the desired asymptotic covariance matrix. Note that Ψ_2 is $\mathcal{F}$-conditional Gaussian, it implies that the asymptotic distribution of $\widehat{\eta}-\eta^*$ is also $\mathcal{F}$-conditionally Gaussian. Moreover,

$$\begin{aligned}\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\mathrm{MLE}}-\widehat{\sigma}^2_{\mathrm{RV}}\right)=&\Delta_n^{-1/2}\left(2\gamma^{4*}\Psi_1(\gamma^{2*},\phi^*)-2\gamma^{2*}\Psi_2(\gamma^{2*},\phi^*)-\widehat{\sigma}^2_{\mathrm{RV}}\right)+o_p(1)\\ =&-2\Delta_n^{-1/2}\gamma^{2*}\Psi_2(\gamma^{2*},\phi^*)+o_p(1).\end{aligned}$$

Therefore, $\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\mathrm{MLE}}-\widehat{\sigma}^2_{\mathrm{RV}}\right)$ is $\mathcal{F}$-conditionally Gaussian, which concludes the proof. ∎

Appendix A.4 Proof of Corollary 2

The proof is similar to that of Corollary 1. By Theorem 2, we have

$$\Delta_n^{-1/2}\left(\widehat{\sigma}^2_{\text{MLE}} - \widehat{\sigma}^2_{\text{RV}}\right) \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \frac{4}{T^2}\left(\int_0^T \sigma_s^4 ds + \sum_{t\leq T}(\Delta X_s)^2(\sigma^2_{s-} + \sigma^2_s)\right)\right).$$

Moreover, by (10.24) and (10.27) of Aït-Sahalia and Jacod (2014), $\widehat{V}_{3n}$ is a consistent estimator of the asymptotic variance. Therefore, H_{3n} has the desired distribution under $\mathbb{H}_0$. Under $\mathbb{H}_1$, $\widehat{\sigma}^2_{\text{RV}} = 2a_0^2\Delta_n^{-1} + O_p(1)$, and $\widehat{\sigma}^2_{\text{MLE}} = O_p(1)$. Moreover, as $u_n \to 0$, $\mathbf{1}_{\{|Y_i|\leq u_n\}} \to 0$, it follows that $\widehat{V}_{3n} \xrightarrow{p} 0$, hence $H_{3n} \xrightarrow{p} \infty$. Finally, under $\mathbb{H}_n$, because $a_nU_i = O_p(\Delta_n^{3/4})$, the noise term is not affected by jump truncation, hence similar calculations as in Corollary 1 yields that $\widehat{V}_{3n}$ remains consistent, which concludes the proof.

Appendix A.5 Proof of Corollary 3

Proof. The desired result follows from Theorem 2 and the proof of Corollary 2. ∎

Appendix A.6 Proof of Corollary 4

Proof. The desired result follows from Lemma 1 and the proof of Corollary 2. ∎

Appendix A.7 Proof of Theorem 3

Proof. Recall that we have

$$\Delta_n^{-1/4}\left(\widehat{\sigma}^2_{\text{AVG}} - \frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\leq T}(\Delta X_s)^2\right)\right)$$
$$\xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \frac{1}{\theta T^2}\left(\int_0^T R\left(\sigma_t^2, \theta^2 a_n^2\right)dt + \sum_{t\leq T} S(\Delta X_t, \sigma_t^2, \theta^2 a_n^2)\right)\right),$$
$$\widehat{V}(\widehat{\sigma}^2_{\text{AVG}}) \xrightarrow{p} \frac{1}{\theta T^2}\left(\int_0^T R\left(\sigma_t^2, \theta^2 a_n^2\right)dt + \sum_{t\leq T} S(\Delta X_t, \sigma_t^2, \theta^2 a_n^2)\right),$$

which hold whether $a_n^2 \geq 0$ or $a_n^2 \to 0$, as long as $k_n\Delta_n^{1/2} = \theta^{-1} \in (0, \infty)$. Since under $\mathbb{H}_0$, we have $\widehat{\sigma}^2_{\text{RV}} = \frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\leq T}(\Delta X_s)^2\right) + O_p(\Delta_n^{-1/2})$, it follows that

$$H_{4n} \xrightarrow{\mathcal{L}} \chi_1^2, \quad \text{under } \mathbb{H}_0.$$

Under $\mathbb{H}_1$, we have

$$\widehat{\sigma}^2_{\text{RV}} = \frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\leq T}(\Delta X_s)^2\right) + \frac{2a_0^2}{\Delta_n} + o_p(1),$$

it implies that $H_{4n} = O_p(\Delta_n^{-5/2})$.

Finally, under the local alternative $\mathbb{H}_n : a_n^2 = a_0^2 \Delta_n^{5/4}$, we have

$$\begin{aligned}
&\Delta_n^{-1/4}\left(\widehat{\sigma}^2_{\text{AVG}} - \widehat{\sigma}^2_{\text{RV}}\right) \\
=&\Delta_n^{-1/4}\left(\widehat{\sigma}^2_{\text{AVG}} - \frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\le T}(\Delta X_s)^2\right)\right) + \Delta_n^{-1/4}\left(\frac{1}{T}\left(\int_0^T \sigma_s^2 ds + \sum_{s\le T}(\Delta X_s)^2\right) - \widehat{\sigma}^2_{\text{RV}}\right) \\
\xrightarrow{\mathcal{L}-s}&\mathcal{MN}\left(2a_0^2, \frac{1}{\theta T^2}\left(\int_0^T R\left(\sigma_t^2, 0\right)dt + \sum_{t\le T} S(\Delta X_t, \sigma_t^2, 0)\right)\right),
\end{aligned}$$

it follows that H_{4n} converges to a non-central Chi-squared distribution with one degree of freedom and the non-centrality parameter $4a_0^4\theta T^2\left(\int_0^T R(\sigma_t^2, 0) + \sum_{s\le T} S(\Delta X_t, \sigma_t^2, 0)\right)^{-1}$. ∎

Appendix B Additional Simulation Results

	a_0^2	level	$\Delta_n = 5$ seconds H_{3n}	T_n	AC_n	H_{4n}	$\Delta_n = 30$ seconds H_{3n}	T_n	AC_n	H_{4n}
Size	0	1%	1.2	1.2	1.2	2.9	1.2	1.1	1.2	1.9
	10^{-8}	1%	93.2	93.0	93.2	25.1	3.6	3.6	3.5	2.9
Power	10^{-7}	1%	100.0	100.0	100.0	97.4	83.7	84.0	83.6	23.2
	10^{-6}	1%	100.0	100.0	100.0	98.8	99.2	99.1	99.1	97.3
Size	0	10%	11.0	10.9	11.0	11.2	10.6	10.4	10.5	13.1
	10^{-8}	10%	96.9	96.9	96.9	45.6	19.2	19.4	19.1	11.6
Power	10^{-7}	10%	100.0	100.0	100.0	97.9	92.7	92.6	92.7	47.2
	10^{-6}	10%	100.0	100.0	100.0	98.8	99.8	99.8	99.8	98.4

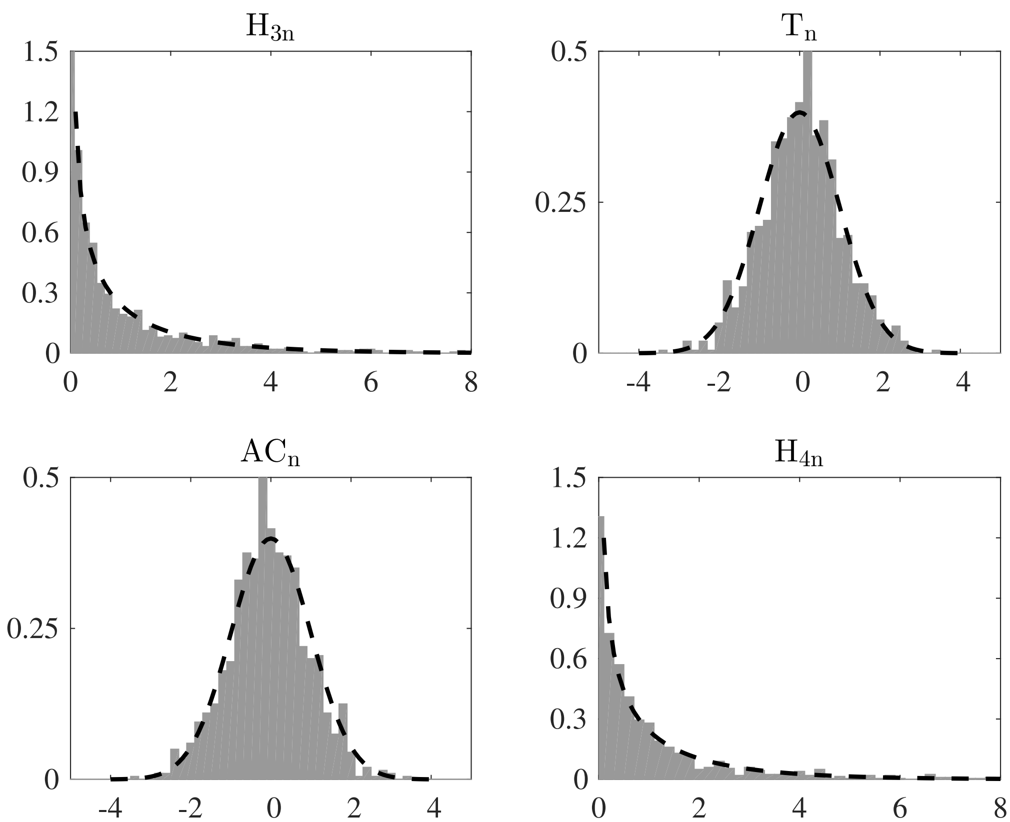
	a_0^2	level	$\Delta_n = 1$ minute H_{3n}	T_n	AC_n	H_{4n}	$\Delta_n = 5$ minutes H_{3n}	T_n	AC_n	H_{4n}
Size	0	1%	1.3	1.4	1.3	2.4	1.6	1.8	1.6	4.5
	10^{-8}	1%	1.5	1.5	1.7	3.1	1.7	1.8	1.7	3.5
Power	10^{-7}	1%	37.3	36.9	36.9	9.0	2.1	2.9	2.2	4.8
	10^{-6}	1%	98.5	98.3	98.3	87.3	29.8	30.4	28.7	17.8
Size	0	10%	9.8	9.5	9.7	9.6	11.3	10.1	11.6	12.5
	10^{-8}	10%	12.1	12.0	12.2	13.6	10.2	10.1	10.3	12.6
Power	10^{-7}	10%	63.5	62.8	63.4	24.1	13.2	13.6	13.2	13.9
	10^{-6}	10%	99.6	99.5	99.5	92.9	59.0	57.7	58.0	34.2

Table B.1: Simulation Results: Percentage of Rejections of $\mathbb{H}_0$ for the Heston Model at the 1% and 10% Levels

	a_0^2	level	$\Delta_n = 5$ seconds				$\Delta_n = 30$ seconds			
			H_{3n}	T_n	AC_n	H_{4n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	1%	1.1	1.1	1.1	4.6	1.1	1.0	1.1	2.7
	10^{-8}	1%	94.9	94.9	94.9	34.8	9.3	8.6	9.1	3.4
Power	10^{-7}	1%	100.0	100.0	100.0	95.6	89.7	89.6	89.5	36.3
	10^{-6}	1%	100.0	100.0	100.0	95.4	99.9	99.9	99.9	97.3
Size	0	10%	9.7	9.6	9.5	14.7	10.5	10.9	10.8	11.3
	10^{-8}	10%	97.2	97.2	97.2	56.4	28.4	28.3	28.1	14.7
Power	10^{-7}	10%	100.0	100.0	100.0	95.9	95.9	95.7	95.9	55.8
	10^{-6}	10%	100.0	100.0	100.0	95.4	99.9	99.9	99.9	98.5
			$\Delta_n = 1$ minute				$\Delta_n = 5$ minutes			
	a_0^2	level	H_{3n}	T_n	AC_n	H_{4n}	H_{3n}	T_n	AC_n	H_{4n}
Size	0	1%	1.0	1.0	1.0	3.0	2.0	2.0	1.9	4.4
	10^{-8}	1%	2.3	2.2	2.2	5.7	3.3	3.2	3.4	3.6
Power	10^{-7}	1%	54.8	54.8	54.7	13.8	2.1	2.4	2.0	5.3
	10^{-6}	1%	97.3	97.3	97.3	87.0	43.5	41.7	43.1	25.4
Size	0	10%	11.2	10.8	11.1	10.5	10.8	10.6	11.0	11.7
	10^{-8}	10%	14.1	13.7	13.5	13.2	13.5	13.1	13.4	10.7
Power	10^{-7}	10%	74.4	73.7	74.0	31.6	13.6	13.6	12.9	13.5
	10^{-6}	10%	98.7	98.7	98.7	92.6	67.3	66.1	66.2	44.2

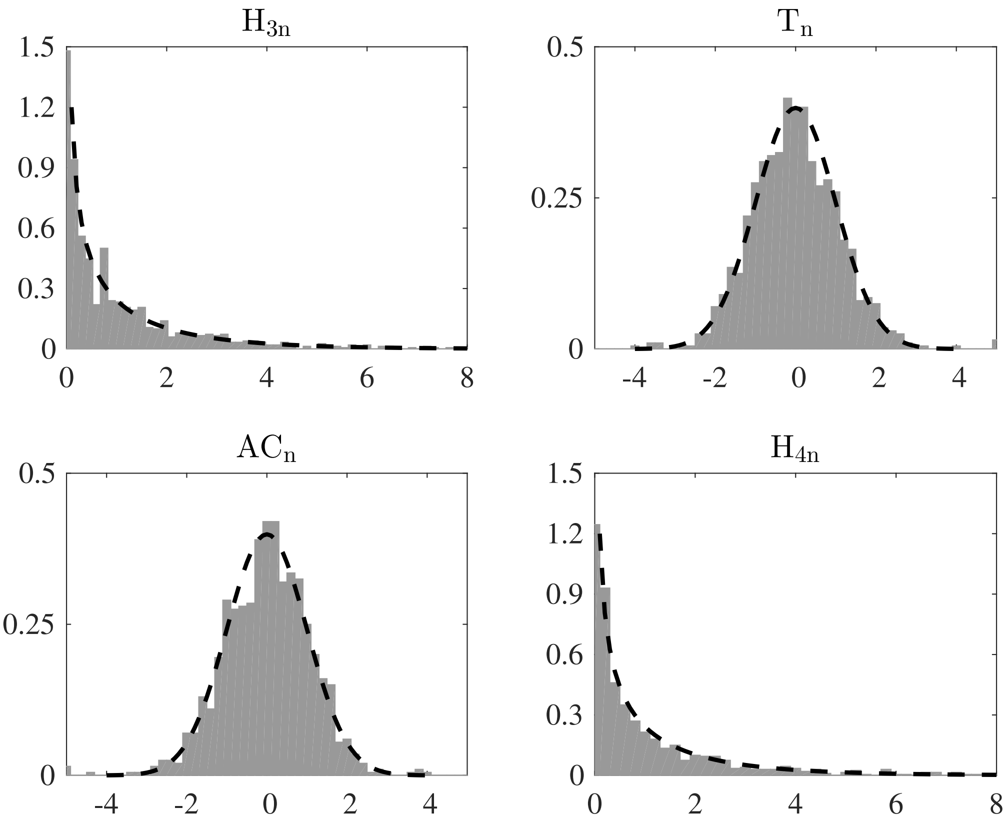
Table B.2: Simulation Results: Percentage of Rejections of $\mathbb{H}_0$ for the Log-Volatility Model at the 1% and 10% Levels

Figure B.1: Small Sample and Asymptotic Distributions of the Test Statistics under $\mathbb{H}_0$: Heston Model with Jumps, 5-second Sampling



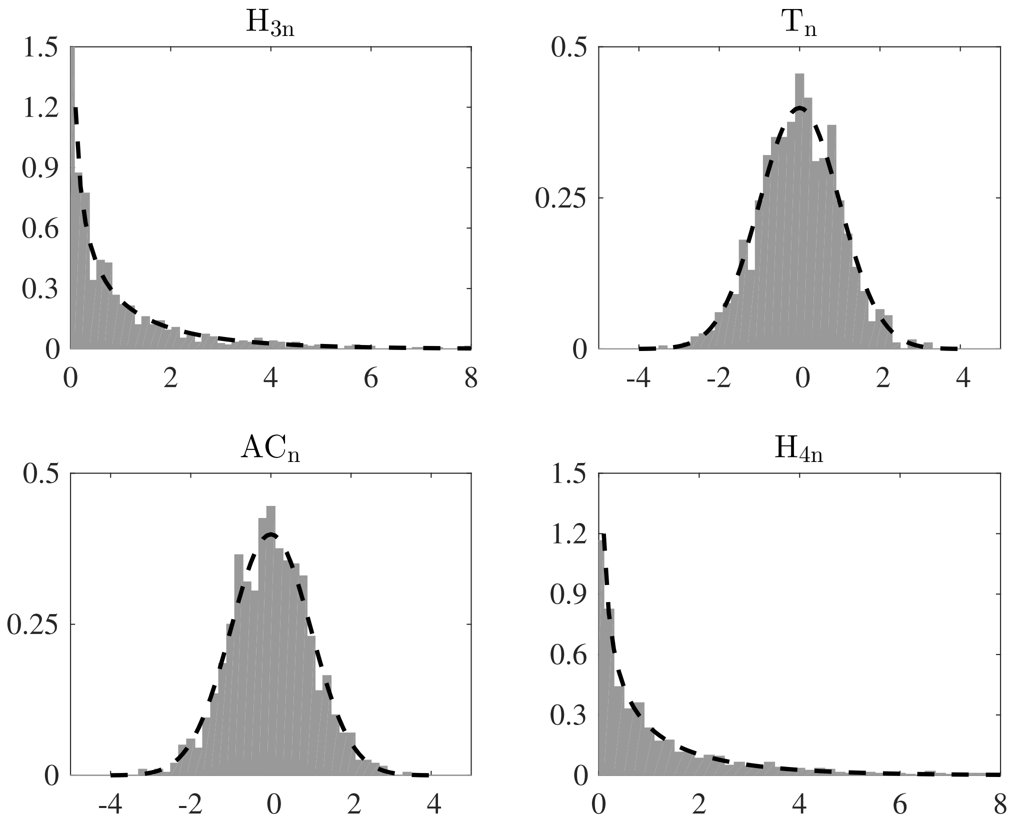
Note: The dashed curves provide the densities of the asymptotic distributions, χ_1^2 for H_{3n}, H_{4n} and $N(0,1)$ for T_n, AC_n respectively. A Heston-style model with jumps (47) and (48) is simulated for H_{3n}, T_n, AC_n, and H_{4n}. The average length of the sampling intervals is 5 seconds. The sampling window is 1 week.

Figure B.2: Small Sample and Asymptotic Distributions of the Test Statistics under $\mathbb{H}_0$: Heston Model with Jumps, 5-minute Sampling



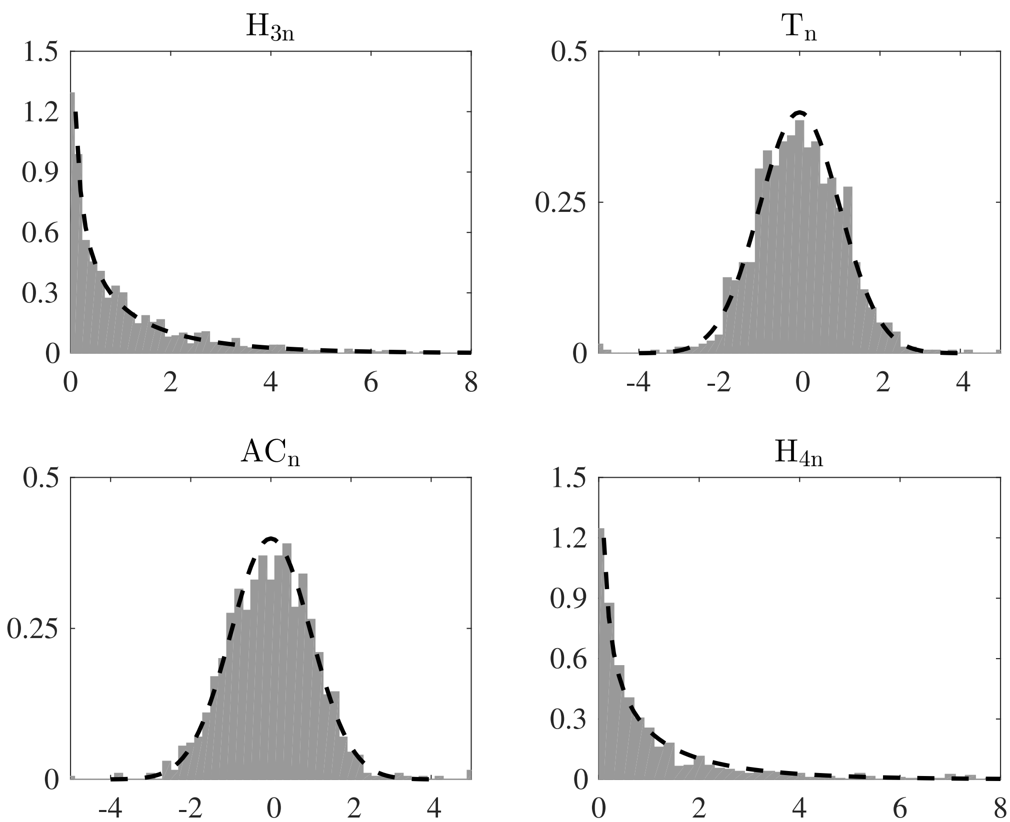
Note: The dashed curves provide the densities of the asymptotic distributions, χ_1^2 for H_{3n}, H_{4n} and $N(0,1)$ for T_n, AC_n respectively. A Heston-style model with jumps (47) and (48) is simulated for H_{3n}, T_n, AC_n, and H_{4n}. The average length of the sampling intervals is 5 minutes. The sampling window is 1 week.

Figure B.3: Small Sample and Asymptotic Distributions of the Test Statistics under $\mathbb{H}_0$: Log-volatility Model, 5-second Sampling



Note: The dashed curves provide the densities of the asymptotic distributions, χ_1^2 for H_{3n}, H_{4n} and $N(0,1)$ for T_n, AC_n respectively. A Log-volatility model with jumps (47) and (48) is simulated for H_{3n}, T_n, AC_n, and H_{4n}. The average length of the sampling intervals is 5 seconds. The sampling window is 1 week.

Figure B.4: Small Sample and Asymptotic Distributions of the Test Statistics under $\mathbb{H}_0$: Log-volatility Model, 5-minute Sampling



Note: The dashed curves provide the densities of the asymptotic distributions, χ_1^2 for H_{3n}, H_{4n} and $N(0,1)$ for T_n, AC_n respectively. A Log-volatility model with jumps (47) and (48) is simulated for H_{3n}, T_n, AC_n, and H_{4n}. The average length of the sampling intervals is 5 minutes. The sampling window is 1 week.