Testing and Modelling Market Microstructure Effects with an Application to the Dow Jones Industrial Average*

Basel Awartani†
Queen Mary, University of London

Valentina Corradi‡
Queen Mary, University of London

Walter Distaso§
University of Exeter

January 2004
Preliminary and incomplete

Abstract

It is a well accepted fact that stock returns data are often contaminated by market microstructure effects, such as bid-ask spreads, liquidity ratios, turnover, and asymmetric information. This is particularly relevant when dealing with high frequency data, which are often used to compute model free measures of volatility, such as realized volatility. In this paper we suggest two test statistics. The first is used to test for the null hypothesis of no microstructure noise. If the null is rejected, we proceed to perform a test for the hypothesis that the microstructure noise variance is independent of the sampling frequency at which data are recorded. We provide empirical evidence based on the stocks included in the Dow Jones Industrial Average, for the period 1997-2002. Our findings suggest that, while the presence of microstructure induces a severe bias when estimating volatility using high frequency data, such a bias grows less than linearly in the number of intraday observations.

Keywords: Bipower variation, market microstructure, Realized Volatility.
JEL classification: C22, C12, G12.

*The authors gratefully acknowledge financial support from the ESRC, grant code R000230006.

†Queen Mary, University of London, Department of Economics, Mile End, London, E14NS, UK, email: b.awartani@qmul.ac.uk.

‡Queen Mary, University of London, Department of Economics, Mile End, London, E14NS, UK, email: v.corradi@qmul.ac.uk.

§University of Exeter, Department of Economics, Streatham Court, Exeter EX4 4PU, UK, email: w.distaso@ex.ac.uk.

1 Introduction

Dating back to Black's (1986) seminal paper, it is a well accepted fact that transaction data occurring in financial markets are often contaminated by market microstructure effects, such as bid-ask spreads, liquidity ratios, turnover, and asymmetric information (see also e.g. Hasbrouk (1993), Bai, Russell and Tiao (2000), Hasbrouck and Seppi (2001), O'Hara (2003) and references therein). It is argued in these papers that the observed transaction price can be decomposed into the efficient one plus a "noise" due to microstructure effects.

This fact is particularly relevant when dealing with high frequency data, which are often used to compute model free measures of volatility, such as realized volatility (see e.g. Barndorff-Nielsen and Shephard (2002, 2003, 2004c), Andersen, Bollerslev, Diebold and Labys (2001, 2003), Meddahi (2002) and Andersen, Bollerslev and Meddahi (2004)) and bipower variation (Barndorff-Nielsen and Shephard, 2004a,b). Although the relevant limit theory suggests that volatility estimates get more precise as the frequency of observations increases, this is not necessarily valid in the presence of microstructure noise which is not accounted for. The effect of microstructure noise on high frequency volatility estimators has been recently analyzed by Aït-Sahalia, Mykland and Zhang (2003), Zhang, Mykland and Aït-Sahalia (2003), and Bandi and Russell (2003a,b). These papers point out that changes in transaction prices over very small time intervals are mainly composed of noise and carry little information about the underlying return volatility. This is because, at least for the class of continuous semimartingale processes, volatility is of the same order of magnitude as the time interval, while the microstructure noise has a roughly constant variability. Therefore, as the time interval shrinks to zero, the signal to noise ratio related to the observed transaction prices also tends to zero, and using the estimators of volatility mentioned above one may run the risk of estimating the variance of the microstructure noise, rather than the underlying return volatility.

Hence, the need of measures of return volatility which are robust to the presence of market microstructure effects. An important contribution in this direction is that of Zhang, Mykland and Aït-Sahalia (2003), who suggest an asymptotically unbiased volatility estimator, based on subsampling techniques. The validity of their estimator hinges on the chosen model for the microstructure noise.

Our paper complements the papers cited above in two directions. First, we provide a test for the null hypothesis of no market microstructure effect, which is robust to the presence of possible large and rare jumps. Second, if the null is rejected, we test the null hypothesis of correct specification of the model of microstructure noise of Aït-Sahalia and Mykland and Zhang (2003), which assumes that the microstructure noise has a constant variance, regardless of the frequency at which data are sampled.

The first test statistic is based on the difference between two realized volatilities computed at different sampling intervals, say one minute and ten minutes. Under the null, both estimators will converge to the true integrated volatility process, though at a different speed. Given this, by properly scaling this difference we have a statistic with a normal limiting distribution under the null and unit asymptotic power. However, such a test statistic can diverge because of the presence of either microstructure noise or jumps. To overcome this problem, we also provide a jump robust version of the test, which is based on recent results by Barndorff-Nielsen and Shephard (2004a,b) and Corradi and Distaso (2004).

The second test statistic is based on the difference between two estimators of the microstructure noise computed again over different time intervals. Under the null model of a noise with constant variance, by properly scaling this difference we obtain a statistic with a normal limiting distribution. The test is consistent against the alternative of a noise with variance depending on the chosen sampling frequency. Indeed, an alternative model of economic interest would be one in which the microstructure noise variance is positively correlated with the time interval.

The proposed tests are then applied to transaction data recorded for the stocks included in the Dow Jones Industrial Average (DJIA) for the period 1997-2002, using a fixed time span equal to five days. The tests are computed over the different five days intervals. The empirical analysis suggests that while the presence of microstructure effects induces a severe bias when estimating volatility using high frequency data, such a bias grows less than linearly in the number of intraday observations.

This paper is organized as follows. Section 2 describes the employed methodology and derives the limiting behavior of the two test statistics. The empirical findings are reported in Section 3 and Section 4 contains some concluding remarks. All the proofs are contained in the Appendix.

2 Methodology

2.1 Set-up

Let $X_t = \log(S_t)$, where S_t denotes the price of a financial asset or derivative. Throughout the paper it is assumed that X_t follows a process of the type

$$X_t = \mu_t dt + c_t dq_t + \sigma_t dW_t, \tag{1}$$

where W_t is a standard Brownian motion. As for the jump component, $\Pr(dq_t = 1) = \lambda_t dt$, where λ_t is independent of σ_t^2, c_t is an *i.i.d.* process and is assumed to be independent of dq_t. This specification of the jump component covers the case of large and rare jumps, analyzed by Barndorff-Nielsen and Shephard (2004a). Although we cannot observe the trajectory of X_t, we still have data recorded at high frequency. Suppose that the number of daily observations is denoted by T and that, for each day, we have M intraday observations; therefore, over a fixed time span, say $\overline{T}$, we have a total of $\overline{T}M$ observations. Realized volatility is defined as

$$RV_{t,\overline{T},M} = \sum_{i=1}^{\overline{T}M} \left(X_{t+\frac{i}{M}} - X_{t+\frac{i-1}{M}} \right)^2, \;\; 0 \leq t \leq T - \overline{T}.$$

If X_t belongs to the class of continuous semimartingales (if $dq_t = 0$, a.s., $\forall\ t$, i.e. there are no jumps in the log price process), then (see e.g. Karatzas and Shreve (1991), Ch.1), as $M \to \infty$,

$$RV_{t,\overline{T},M} \overset{\Pr}{\to} \int_t^{t+\overline{T}} \sigma_s^2 ds = IV_{t,\overline{T}}. \tag{2}$$

Here σ_s^2 denotes the instantaneous volatility at time s, that is

$$\lim_{h\to 0} \frac{E\left((X_{s+h} - X_s)^2 \,|\Im_s \right)}{h} = \sigma_s^2,$$

where $\Im_s$ refers to the relevant conditioning set at time s.

However, if X_t is the sum of a continuous semimartingale component and a jump component, then the statement in (2) does no longer hold and, as pointed out by Barndorff-Nielsen and Shephard (2004a),

$$RV_{t,\overline{T},M} \overset{\Pr}{\to} IV_t + \sum_{i=N_t}^{N_{t+\overline{T}}} c_i^2, \tag{3}$$

where N_t is a counting process and c_i denotes the size of the jumps. Interestingly, Barndorff-Nielsen and Shephard (2004a) also suggest a measure of integrated volatility, namely bipower variation, which (when properly scaled) is a consistent estimator of integrated volatility and is robust to the presence of rare of large jumps.[1] Bipower variation is defined as

$$BV_{t,\overline{T},M} = \sum_{i=1}^{\overline{T}M-1} \left| X_{t+\frac{i+1}{M}} - X_{t+\frac{i}{M}} \right| \left| X_{t+\frac{i}{M}} - X_{t+\frac{i-1}{M}} \right|, \;\; 0 \leq t \leq T - \overline{T}.$$

Now, suppose that we can observe $S_{t+\frac{i}{M}}$ only up to an error, so that the observed price process is given by

$$\widetilde{S}_{t+\frac{i}{M}} = S_{t+\frac{i}{M}} \eta_{t+\frac{i}{M}}, \;\; t = 0, 1, \ldots, T-1. \tag{4}$$

Therefore

$$\begin{aligned} \log \widetilde{S}_{t+\frac{i}{M}} &= \log S_{t+\frac{i}{M}} + \log \eta_{t+\frac{i}{M}}, \;\; \text{or} \\ Y_{t+\frac{i}{M}} &= X_{t+\frac{i}{M}} + \epsilon_{t+\frac{i}{M}}. \end{aligned} \tag{5}$$

Here $\epsilon_{t+\frac{i}{M}}$ is interpreted as a noise capturing the market microstructure effect. Similarly to what is customarily assumed in the literature, $X_{t+i/M}$ and $\epsilon_{t+i/M}$ are independent and $E\left(\epsilon_{t+i/M}\right) = 0$. Now,

$$\begin{aligned} & E\left(Y_{t+\frac{i}{M}} - Y_{t+\frac{i-1}{M}}\right)^2 \\ = \; & E\left(X_{t+\frac{i}{M}} - X_{\frac{i-1}{M}}\right)^2 + E\left(\epsilon_{t+\frac{i}{M}} - \epsilon_{t+\frac{i-1}{M}}\right)^2 + E\left(\epsilon_{t+\frac{i}{M}} - \epsilon_{t+\frac{i-1}{M}}\right)\left(X_{t+\frac{i}{M}} - X_{t+\frac{i-1}{M}}\right) \\ = \; & E\left(X_{t+\frac{i}{M}} - X_{\frac{i-1}{M}}\right)^2 + E\left(\epsilon_{t+\frac{i}{M}} - \epsilon_{t+\frac{i-1}{M}}\right)^2. \end{aligned} \tag{6}$$

In the case of X_t being a continuous semimartingale, $E\left(X_{t+\frac{i}{M}} - X_{t+\frac{i-1}{M}}\right)^2 = O(M^{-1})$. Also, we will assume that

$$\lim_{M,N\to\infty, \;\; N/M\to 0} \frac{E\left(\sum_{i=1}^{\overline{T}M-1}\left(\epsilon_{t+\frac{i}{M}} - \epsilon_{t+\frac{i-1}{M}}\right)^2\right)}{N} = \infty.$$

As for bipower variation, note that, by Minkowski inequality,

$$E\left(\sum_{i=1}^{\overline{T}M-1} \left| Y_{t+\frac{i+1}{M}} - Y_{t+\frac{i}{M}} \right| \left| Y_{t+\frac{i}{M}} - Y_{t+\frac{i-1}{M}} \right| \right) \leq E\left(\sum_{i=1}^{\overline{T}M-1} \left| X_{t+\frac{i+1}{M}} - X_{t+\frac{i}{M}} \right| \left| X_{t+\frac{i}{M}} - X_{t+\frac{i-1}{M}} \right| \right)$$

[1] A viable alternative to this approach would be to follow Aït-Sahalia (2004), who proposes a method to disentangle the continuous Brownian and jump components, and then use the continuous component to compute an estimator of integrated volatility. Such an approach is valid for homoskedastic diffusion processes.

$$+E\left(\sum_{i=1}^{\overline{T}M-1}\left|\epsilon_{t+\frac{i+1}{M}}-\epsilon_{t+\frac{i}{M}}\right|\left|\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right|\right)+E\left(\sum_{i=1}^{\overline{T}M-1}\left|\epsilon_{t+\frac{i+1}{M}}-\epsilon_{t+\frac{i}{M}}\right|\left|X_{t+\frac{i}{M}}-X_{t+\frac{i-1}{M}}\right|\right).$$

Therefore, when using bipower variation it is not immediate how to decompose the total variability in integrated volatility and noise variance. Nevertheless, it is evident that, while bipower variation is robust to the presence of the jump component, it is not robust to microstructure effects.

As mentioned in the Introduction, our objective is perform a test for the null hypothesis of no microstructure effects over a sequence of finite time span (e.g. 5 working days) periods, and, for the cases in which the null is rejected, proceed to test the null hypothesis that the microstructure noise has constant variance, regardless of the sampling interval over which we compute realized volatility.

2.2 Testing the null hypothesis of no microstructure effects

In the sequel we shall assume that the true asset price process is generated as in (1). Define the following hypotheses:

$$H_0 : E\left(\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right)^2=0, \text{ for all } M,t \tag{7}$$

and

$$H_A : E\left(\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right)^2\neq 0, \tag{8}$$

where $E\left(\epsilon_{t+i/M}-\epsilon_{t+(i-1)/M}\right)^2$ and $\epsilon_{t+i/M}$ are defined respectively in (6) and (5).
Therefore the null hypothesis implies that there are no microstructure effects in the observed transaction prices, while the alternative is simply the negation of the null.

In order to test H_0 versus H_A, we propose the following statistic

$$Z_{M,N,\overline{T},t}=\frac{\sqrt{N\overline{T}}\left(\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2-\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2\right)}{\sqrt{\frac{2}{3}N\overline{T}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}, \tag{9}$$

where $N/M \to 0$ as $M,N\to\infty$ (i.e. M grows faster than N). The numerator in (9) can be expanded as

$$\sqrt{N\overline{T}} \tag{10}$$
$$\times\left(\left(\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2ds\right)-\left(\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2ds\right)\right).$$

Inspection of (10) reveals the logic behind the choice of the test statistic in (9); if there are no microstructure effects (under the maintained assumption of no jumps), then both $\sum_{i=1}^{\overline{T}M}\left(Y_{t+i/M}-Y_{t+(i-1)/M}\right)^2$ and $\sum_{i=1}^{\overline{T}N}\left(Y_{t+i/N}-Y_{t+(i-1)/N}\right)^2$ converge in probability to the true integrated volatility, $\int_t^{t+\overline{T}}\sigma_s^2 ds$. Also, by the central limit theorems in Barndorff-Nielsen and Shephard (2002 and 2004c),

$$\left(\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2 ds\right)=O_p(M^{-1/2})$$

and

$$\left(\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2 ds\right)=O_p(N^{-1/2}).$$

Therefore, under the null hypothesis, provided that as $N,M\to\infty$, $N/M\to 0$,

$$\sqrt{N\overline{T}}\left(\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2 ds\right)=o_p(1).$$

Under H_0, the limiting distribution of the test statistic is driven by

$$-\frac{\sqrt{N\overline{T}}\left(\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2 ds\right)}{\sqrt{\frac{2}{3}N\overline{T}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}},$$

which is asymptotically standard normal. Under the alternative, we expect $\sum_{i=1}^{\overline{T}M}\left(Y_{t+i/M}-Y_{t+(i-1)/M}\right)^2$ to be of a larger order of magnitude (in probability) than $\sum_{i=1}^{\overline{T}N}\left(Y_{t+i/N}-Y_{t+(i-1)/N}\right)^2$, and so we expect the statistic to diverge.

A possible problem with the statistic above is that standard normal critical values are no longer correct in the presence of jumps. In fact, in the presence of jumps, the numerator of (9) has a non zero mean and the square of the denominator is not a consistent estimator for the true integrated quarticity (i.e. $\int_t^{t+\overline{T}}\sigma_s^4 ds$). Therefore, we also suggest a statistic which is robust to jumps

$$ZB_{M,N,\overline{T},t}=\frac{\mu_1^{-2}\sqrt{N\overline{T}}}{\sqrt{2.6090\mu_1^{-4}}} \tag{11}$$

$$\times\ \frac{\left(\sum_{i=1}^{\overline{T}M-1}\left|Y_{t+\frac{i+1}{M}}-Y_{t+\frac{i}{M}}\right|\left|Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right|-\sum_{i=1}^{\overline{T}N}\left|Y_{t+\frac{i+1}{N}}-Y_{t+\frac{i}{N}}\right|\left|Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right|\right)}{\sqrt{N\overline{T}\sum_{i=1}^{\overline{T}M-3}\left|Y_{t+\frac{i+4}{M}}-Y_{t+\frac{i+3}{M}}\right|\left|Y_{t+\frac{i+3}{M}}-Y_{t+\frac{i+2}{M}}\right|\left|Y_{t+\frac{i+2}{N}}-Y_{t+\frac{i+1}{N}}\right|\left|Y_{t+\frac{i+1}{N}}-Y_{t+\frac{i}{N}}\right|}},$$

where $\mu_1=E\left(\left|N\left(0,1\right)\right|\right)$. Under the null hypothesis, the statistic above has a standard normal limiting distribution, regardless of the presence of possible jumps.

2.3 A simple specification test for the microstructure noise

For all the periods in which, according to either or both test statistics proposed in the previous subsection, we reject the null hypothesis of no microstructure noise, it may be interesting to perform a specification test for the microstructure error. In particular, the hypothesis of interest is that of the microstructure noise having a constant variance, independent of the frequency at which data are recorded. In fact, most of the recent literature on incorporating microstructure effects (see e.g. Aït-Sahalia, Mykland and Zhang (2003), Zhang, Mykland and Aït-Sahalia (2003), and Bandi and Russell (2003a,b)) postulates a model of noise with constant variance.

More precisely, let $E\left(\epsilon_{t+i/M} - \epsilon_{t+(i-1)/M}\right)^2 = 2\nu_{t,M}$ and $E\left(\epsilon_{t+i/M} - \epsilon_{t+(i-1)/M}\right)^2 = 2\nu_{t,N}$. The null and alternative hypotheses can be formulated as follows

$$H_0' : \nu_{t,M} = \nu_{t,N}, \quad \text{for all} \;\; M, N \tag{12}$$

and

$$H_A' : \nu_{t,M} < \nu_{t,N}. \tag{13}$$

Thus, the alternative of interest is that the variance of the microstructure noise is negatively correlated with the frequency at which data are recorded. The alternative hypothesis is compatible with the microstructure noise model outlined by Barndorff-Nielsen and Shephard (2004b), who specify a two component model; the first is a jump component and the second is an error process with the variance decaying to zero at a rate $1/\sqrt{M}$. As a consequence, the bias incurred in estimating volatility using realized volatility grows less than linearly in the number of intraday observations.[2] We propose the following test statistic

$$V_{M,N,\overline{T},t} = \sqrt{N\overline{T}} \left(\frac{ \frac{\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}} - Y_{t+\frac{i-1}{M}}\right)^2}{2\overline{T}M} - \frac{\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}} - Y_{t+\frac{i-1}{N}}\right)^2}{2\overline{T}N} }{ \sqrt{\frac{1}{N\overline{T}} \sum_{i=1}^{N\overline{T}} \left(Y_{t+\frac{i}{N}} - Y_{t+\frac{i-1}{N}}\right)^4} } \right). \tag{14}$$

The logic behind the statistic proposed above is the following. Under the null hypothesis, both $\sum_{i=1}^{\overline{T}M}\left(Y_{t+i/M} - Y_{t+(i-1)/M}\right)^2/2\overline{T}M$ and $\sum_{i=1}^{\overline{T}N}\left(Y_{t+i/N} - Y_{t+(i-1)/N}\right)^2/2\overline{T}N$ converge to the same

[2] The fact that the variance of measurement error is in general not independent of the sampling frequency has been kindly pointed out to us by Andrew Chesher.

probability limit, say ν, though the former converges faster than the latter. Thus,

$$\sqrt{N\overline{T}}\left(\frac{\frac{\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2}{2\overline{T}M}-\nu}{\sqrt{\frac{1}{N\overline{T}}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}\right)$$

is asymptotically negligible and the statistic in (14) is asymptotic equivalent to

$$-\sqrt{N\overline{T}}\left(\frac{\frac{\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2}{2\overline{T}N}-\nu}{\sqrt{\frac{1}{N\overline{T}}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}\right),$$

which has a standard normal limiting distribution (see Theorem A1 in Zhang, Mykland and Aït-Sahalia (2003)).

Under the alternative, $\sum_{i=1}^{\overline{T}M}\left(Y_{t+i/M}-Y_{t+(i-1)/M}\right)^2/2\overline{T}M$ and $\sum_{i=1}^{\overline{T}N}\left(Y_{t+i/N}-Y_{t+(i-1)/N}\right)^2/2\overline{T}N$ converge respectively to ν_M and ν_N, and thus the statistic diverges.

In the next subsection, the limiting distributions of the proposed test statistics are derived.

2.4 Main theoretical results

In the sequel we need the following assumptions

A1: X_t is generated as in (1).

A2: $\int_t^{t+\overline{T}}\sigma^4(s)ds<\infty$, almost surely, for any t and $\overline{T}$.

A3: $E\left(\epsilon_{\frac{i}{M}}-\epsilon_{\frac{i-1}{M}}\right)^4<\infty$.

A4: $\sum_{i=1}^{M\overline{T}}\left(\epsilon_{t+i/M}-\epsilon_{t+(i-1)/M}\right)^2/M\overline{T}=o_p(b_{t,M}^{-1})$, and $\sum_{i=1}^{N\overline{T}}\left(\epsilon_{t+i/N}-\epsilon_{t+(i-1)/N}\right)^2/N\overline{T}=o_p(b_{t,N}^{-1})$, where, as $M,N\to\infty$, $\frac{b_{t,M}}{b_{t,N}}M^{1-\alpha}\to\infty$, for $\alpha\in(0,1)$ and $N=M^\alpha$.

A4': $E\left(\sum_{i=1}^{M\overline{T}}\left(\epsilon_{t+i/M}-\epsilon_{t+(i-1)/M}\right)^2/M\overline{T}\right)=\nu_{t,M}$ and $E\left(\sum_{i=1}^{N\overline{T}}\left(\epsilon_{t+i/N}-\epsilon_{t+(i-1)/N}\right)^2/N\overline{T}\right)=\nu_{t,N}$, $\sqrt{N}\left(\nu_{t,M}/\nu_{t,N}\right)\to\infty$, as $M,N\to\infty$.

Notice that **A1** and **A2** are customary in the literature on realized volatility; **A3** requires a finite fourth moment of the market microstructure noise and seems to be trivially satisfied. Finally, **A4** allows the variability of the microstructure error to decrease with the sampling interval. In particular, the variance of the microstructure noise is allowed to approach zero as the sampling interval goes to zero, but at a slow enough rate.

Then, we can state the following Propositions.

Proposition 1

(i) Let A1-A2 hold and assume that $\lambda_t = 0$ *for all* t*. Under* H_0*, defined in (7), as* $M, N \to \infty$ *and* $N/M \to 0$,

$$Z_{M,N,\overline{T},t} \xrightarrow{d} N(0,1).$$

(ii) Let A1-A4 hold. Under H_A*, defined in (8), as* $M \to \infty$*, for* $\varepsilon > 0$,

$$\lim_{M\to\infty} \Pr\left(\frac{b_{t,N}}{b_{t,M}} M^{\alpha-1} Z_{M,N,\overline{T},t} > \varepsilon\right) = 1.$$

Therefore, we can just perform a one-sided test and reject the null hypothesis of no microstructure effects when we get a value for $Z_{M,N,\overline{T},t}$ larger than, say, the 95% percentile of a standard normal. In the proposition above, the statement under the null is robust to possible leverage effects, but not to possible jumps.

The null limiting distribution and power properties of the jump robust version of the test for no microstructure effects test are given in the next Proposition.

Proposition 2

(i) Let A1-A2 hold and assume that $\mu_t = 0$ *for all* t *and that* σ_t *in (1) is independent of* X_t*. Under* H_0*, defined in (7), as* $M, N \to \infty$ *and* $N/M \to 0$,

$$ZB_{M,N,\overline{T},t} \xrightarrow{d} N(0,1).$$

(ii) Let A1-A4 hold. Under H_A*, defined in (8), as* $M \to \infty$*, for* $\varepsilon > 0$,

$$\lim_{M\to\infty} \Pr\left(\frac{b_{t,N}}{b_{t,M}} M^{\alpha-1} ZB_{M,N,\overline{T},t} > \varepsilon\right) = 1.$$

As outlined in the previous subsection, every time we reject the null hypothesis in (7), according to either or both the suggested statistics, we may be interested in performing a specification test for the microstructure noise. Its properties are given in the next Proposition.

Proposition 3

(i) Let A1-A3 hold. Under H_0'*, defined in (12), as* $M, N \to \infty$, $N/M \to 0$,

$$V_{M,N,\overline{T},t} \xrightarrow{d} N(0,1).$$

(ii) Let A1-A3 hold. Also, assume that

$$E\left(\frac{1}{M\overline{T}}\sum_{i=1}^{M\overline{T}}\left(\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right)^2\right)=2\nu_{t,M} \quad and \quad E\left(\frac{1}{N\overline{T}}\sum_{i=1}^{N\overline{T}}\left(\epsilon_{t+\frac{i}{N}}-\epsilon_{t+\frac{i-1}{N}}\right)^2\right)=2\nu_{t,N}$$

are such that $\sqrt{M}\nu_{t,M}\to\infty$ and $\sqrt{N}\nu_{t,N}\to\infty$. Under H_A, defined in (13), as $M\to\infty$, for $\varepsilon>0$,

$$\lim_{M\to\infty}\Pr\left(\frac{1}{\sqrt{N\overline{T}}}V_{M,N,\overline{T},t}>\varepsilon\right)=1.$$

An application of the testing procedure outlined in this section to the stocks included in the Dow Jones Industrial Average is given in the next Section.

3 Empirical evidence from the Dow Jones Industrial Average

3.1 Data description

The empirical analysis of market microstructure effects is based on data retrieved from the Trade and Quotation (TAQ) database at the New York Stock Exchange. The TAQ database contains intraday trades and quotes for all securities listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX) and the Nasdaq National Market System (NMS). The data is published monthly on CD-ROM since 1993 and on DVD since June 2002. Our sample contains the DJIA stocks (30 stocks in total[3]) and extends from January 1, 1997 until December 24, 2002, for a total of 1505 trading days.[4] Also, in our empirical example $\overline{T}=5$ and therefore we have a total of 301 five-days periods.

Table 2 shows the average number of quotations per minute for all individual stocks. The table presents a spectrum of liquidity, ranging from as low as 3 quotations per minute for United Technologies Corp. (UTX) to as high as 94 for Intel Corp.. The two most liquid stocks in the sample

[3]It is worth mentioning that the 30 companies included in the DJIA are not the same throughout the sample period. Woolworth, Bethlehem Steel, Texaco and CBS (Westinghouse Electric) have been replaced by Wal-Mart, Johnson & Johnson, Hewlett-Packard and Citigroup (Travelers Group) in 1997. In addition SBC Communications, Microsoft, Intel and Home Depot have replaced Union Carbide, Chevron, Goodyear, and Sears in 1999. Our sample contains the DJIA of individual firms as it was in 2000. The names and the symbols of the stocks included in the sample are reported in Table 1.

[4]Trading days are divided over the different years as follows: 253, 252, 252, 252, 248, 252 from 1997 to 2002. Note that there are 5 days missing in 2001 due to September 11th.

are Intel and Microsoft and it takes only a fraction of a second to have a fresh quote. Liquidity has increased substantially during the sample period and, as we approach 2002, the number of quotations per minute almost doubles for some stocks (e.g. 3M Company (MMM), Citigroup Inc. (C), Home Depot Corp. (HD), Microsoft (MSFT)).

From the original data set, which includes prices recorded for every trade, we extracted 1 minute and 10 minutes interval data, similarly to Andersen and Bollerslev (1997). Provided that there is sufficient liquidity in the market, the 5 minutes frequency is generally accepted as the highest frequency at which the effect of microstructure biases are not too distorting (see Andersen, Bollerslev, Diebold and Labys (2001), Andersen, Bollerslev and Lang (1999) and Ebens (1999)); hence the choice of the two mentioned frequencies to calculate the test statistics, in order to highlight the full extent of the microstructure noise effects.

The price figures for each 1 and 10 minutes intervals are determined as the interpolated average between the preceding and the immediately following quotes, weighted linearly by their inverse relative distance to the required point in time. For example, suppose that the price at 15:29:56 was 11.75 and the next quote at 15:30:02 was 11.80, then the interpolated price at 15:30:00 would be $\exp(1/3 \times \log(11.80) + 2/3 \times \log(11.75)) = 11.766$. From the 1 and 10 minutes price series we calculated 1 and 10 minutes intradaily returns as the difference between successive log prices expressed in percentages; then

$$R_{t+\frac{i}{N}} = 100 \times \left(\log(Y_{t+\frac{i}{N}}) - \log(Y_{t+\frac{i-1}{N}}) \right),$$

where $R_{t+i/N}$ denotes the return for intraday period i/N on trading day t, with $t = 0, \ldots, T-1$. The New York Stock Exchange opens at 9:30 a.m. and closes at 4.00 p.m.. Therefore a full trading day consists of 391 (resp. 40) intraday returns calculated over an interval of one minute (resp. ten minutes). For some stocks, and in some days, quotations arrive some time after 9:30; in these cases we always set the first available trading price after 9:30 a.m to be the price at 9:30 a.m.. Not all the days in our sample consists of 391 (resp. 40) price observations; this is attributable to the fact that the NYSE closes early on certain days, such as on Christmas Eve[5]; for all these intervals without price quotes we insert zero return values. Highly liquid stocks may have more than one price at certain points in time (for example 5 or 10 quotations at the same time stamp is normal

[5]In addition to Christmas Eve, there are 12 other short days in the sample, making a total of 17 days.

for INTC and MSFT); when there exists more than one price at the required interval, we select the last provided quotation. For interpolating a price from a multiple price neighborhood, we select the closest provided price for the computation.

3.2 Testing for the null of no microstructure effects

Table 3, columns 2 and 3, reports the findings for the test based on the statistic defined in (9). More precisely, we report the number and the percentages of rejections, based on a one-sided test.

We first notice that only for six stocks the percentage of rejection is below 20% of the cases. For twelve stocks the percentage of rejection is between 20% and 40% of the cases, and for the remaining twelve stocks it is higher than 40% with a maximum of 66%. This indicates that, though microstructure effect plays an important role, its contribution is quite variable over time. Overall, we do not find evidence of a clear relationship between liquidity and microstructure effects. For example, taking two rather liquid stocks like IBM and Intel, for the former we reject the null about 20% of the times, while for the latter we reject about 39% of the times. In Figure 1, we report the plot of the test statistic over the different 5 days intervals considered, with the dotted line representing the 5% upper tail critical value of a standard normal. We notice that there are a few instances in which the statistic takes a large and negative values. This happens mainly for stocks characterized by low liquidity, such as EK and HON. The reason for this finding is that, since these stocks are not very liquid, and therefore are not traded often enough, a lot of returns over 1 minute interval are zero, while are not zero over 10 minutes interval. Hence a negative value for the test statistic.

Table 3, columns 4 and 5, reports the results for the test based on the statistic defined in (11). We notice that the rejection rates are comparable with those obtained using the previous test, although slightly lower (the only exception being Intel). As the latter statistic is robust to the presence of large and rare jumps, the obtained results seem to provide evidence in favor of the fact that most of the 5 days periods are not characterized by jumps. The fact that only a small number of days is characterized by jumps is also confirmed by the empirical findings of Andersen, Bollerslev and Diebold (2003), and Huang and Tauchen (2003). The plot of the test statistic, for each 5 days interval, is inserted in Figure 2 and displays a very similar pattern to the one observed in Figure 1.

As explained in the previous Section, each time we reject the hypothesis of no microstructure noise effect, we perform a specification test for the hypothesis of a microstructure noise with constant variance (independent of the sampling interval). The relevant empirical results are contained in the next subsection.

3.3 A specification test for the variability of the microstructure noise

For all the periods in which we reject the null hypothesis of no market microstructure effects, we then test H_0' versus H_A', defined respectively in (12) and (13), using the statistic suggested in (14). We perform two sequences of test, the first one conditional on rejecting the null of no microstructure using the test statistic in (9) and the second conditional on the same outcome using the statistic in (11). The results are reported in Table 4, columns 2 to 5, and the plots are given in Figures 3 and 4. It is immediate to see that the null hypothesis is rejected in almost all the cases. This provides strong evidence that while the presence of microstructure induces a severe bias when estimating volatility using high frequency data, such a bias grows less than linearly in the number of intraday observations. In fact, given (6), acceptance of the null hypothesis would imply that

$$\begin{aligned} E\left(\sum_{i=1}^{M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2\right) &= E\left(\sum_{i=1}^{M}\left(X_{t+\frac{i}{M}}-X_{t+\frac{i-1}{M}}\right)^2\right)+E\left(\sum_{i=1}^{M}\left(\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right)^2\right) \\ &\simeq \int_t^{t+1}\sigma_s^2 ds+2M\nu_M=\int_t^{t+1}\sigma_s^2 ds+2M\nu. \end{aligned}$$

However, our findings strongly suggest that $\nu_M < \nu_N$, thus indicating that the microstructure bias grows at a slower than linear rate as the number of intraday observations increases.

4 Concluding remarks

To be done.

References

Aït-Sahalia, Y. (2004), Disentangling Diffusion from Jumps, *Journal of Financial Economics*, forthcoming.

Aït-Sahalia, Y., P.A. Mykland and L. Zhang (2003), How Often to Sample a Continuous Time Process in the Presence of Market Microstructure Noise, Working Paper, Princeton University.

Andersen, T.G., and T. Bollerslev, (1997) Intraday Periodicity and Volatility Persistence in Financial Markets, *Journal of Empirical Finance*, 4, 115-158.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, (2001) The Distribution of Realized Exchange Rate Volatility, *Journal of the American Statistical Association*, 96, 42-55.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, (2003) Modelling and Forecasting Realized Volatility, *Econometrica*, 71, 579-625.

Andersen, T.G., T. Bollerslev and F.X. Diebold (2003), Some Like it Smooth and Some Like it Rough: Untangling Continuous and Jump Components in Measuring, Modelling and Forecasting Asset Return Volatility, Working Paper, Duke University.

Andersen, T.G., T. Bollerslev and S. Lang (1999), Forecasting Financial Market Volatility. Sample Frequency vs Forecast Horizon, *Journal of Empirical Finance*, 6, 457-477.

Andersen, T.G., T. Bollerslev and N. Meddahi (2004), Analytic Evaluation of Volatility Forecasts, *International Economic Review,* forthcoming.

Bai, X., J.R. Russell and G. Tiao (2000), Effects on Nonnormality and Dependence on the Precision of Variance Estimates Using High Frequency Data, Working Paper, University of Chicago, Graduate School of Business.

Bandi, F.M., and J.R. Russell (2003a), Microstructure Noise, Realized Volatility, and Optimal Sampling, Working Paper, University of Chicago, Graduate School of Business.

Bandi, F.M., and J.R. Russell (2003b), Volatility or Noise? Working Paper, University of Chicago, Graduate School of Business.

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

Barndorff-Nielsen, O.E., and N. Shephard (2003), Realized Power Variation and Stochastic Volatility, *Bernoulli*, 9, 243-265.

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

Barndorff-Nielsen, O.E., and N. Shephard (2004b), Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation, Working Paper, University of Oxford.

Barndorff-Nielsen, O.E., and N. Shephard (2004c), A Feasible Central Limit Theory for Realized Volatility under Leverage, Working Paper, University of Oxford.

Black, F. (1986), Noise, *Journal of Finance*, XLI, 529-543.

Corradi, V., and W. Distaso (2004), Specification Tests for Daily Integrated Volatility, in the Presence of Possible Jumps, Working Paper, Queen Mary, University of London.

Ebens, H. (1999), Realized Stock Volatility, Working Paper, John Hopkins University.

Hasbrouck, J. (1993), Assessing the Quality of a Security Market. A New Approach to Transaction Cost Measurement, *Review of Financial Studies*, 6, 191-212.

Hasbrouck, J., and D. Seppi (2001), Common Factors in Prices, Order Flows, and Liquidity, *Journal of Financial Economics,* 59, 383-411.

Huang, X., and G. Tauchen (2003), The Relative Contribution of Jumps to Total Price Variation, Working Paper, Duke University.

Karatzsas, I., and S.E. Shreve (1991), *Brownian Motion and Stochastic Calculus,* Springer and Verlag, New York.

Meddahi, N. (2002), A Theoretical Comparison between Integrated and Realized Volatility, *Journal of Applied Econometrics*, 17, 475-508.

O'Hara, M. (2003), Presidential Address: Liquidity and Price Discovery, *Journal of Finance*, LVIII, 1335-1354.

Zhang L., P.A. Mykland and Y. Aït-Sahalia (2003), A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High Frequency Data, Working Paper, Carnegie Mellon University.

Appendix

Proof of Proposition 1.

(i) The statistic defined in (9) can be expanded into

$$\begin{aligned} Z_{M,N,\overline{T},t} &= \sqrt{\frac{N}{M}} \frac{\sqrt{M\overline{T}}\left(\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2 ds\right)}{\sqrt{\frac{2}{3}N\overline{T}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}} \\ &- \frac{\sqrt{N\overline{T}}\left(\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2-\int_t^{t+\overline{T}}\sigma_s^2 ds\right)}{\sqrt{\frac{2}{3}N\overline{T}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}. \end{aligned} \quad (15)$$

Theorems 1 in Barndorff-Nielsen and Shephard (2002, 2004c) establish that, as $M \to \infty$,

$$\sqrt{M\overline{T}}\left(RV_{t,\overline{T},M,T}-\int_t^{t+\overline{T}}\sigma_s^2 ds\right) \xrightarrow{d} N\left(0, 2\int_t^{t+\overline{T}}\sigma_s^4 ds\right).$$

Therefore the first term in (15) is $o_p(1)$, given that, as $M, N \to \infty$, $N/M \to 0$. As a consequence, the limiting distribution of $Z_{M,N,\overline{T},t}$ under H_0 will be determined by the second component of (15). By a similar argument as before, since, as $N \to \infty$,

$$\frac{1}{3}N\overline{T}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4 \xrightarrow{\Pr} \int_t^{t+\overline{T}}\sigma_s^4 ds,$$

the statement follows immediately.

(ii) Under the alternative hypothesis, it is possible to expand the numerators of the components of (15) respectively as

$$\sqrt{N\overline{T}}\left[\sum_{i=1}^{\overline{T}M}\left(\left(X_{t+\frac{i}{M}}-X_{t+\frac{i-1}{M}}\right)^2+\left(\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right)^2+2\left(X_{t+\frac{i}{M}}-X_{t+\frac{i-1}{M}}\right)\left(\epsilon_{t+\frac{i}{M}}-\epsilon_{t+\frac{i-1}{M}}\right)\right)\right.$$
$$\left.-\int_t^{t+\overline{T}}\sigma_s^2 ds\right]$$

and

$$-\sqrt{N\overline{T}}\left[\sum_{i=1}^{\overline{T}N}\left(\left(X_{t+\frac{i}{N}}-X_{t+\frac{i-1}{N}}\right)^2+\left(\epsilon_{t+\frac{i}{N}}-\epsilon_{t+\frac{i-1}{N}}\right)^2+2\left(X_{t+\frac{i}{N}}-X_{t+\frac{i-1}{N}}\right)\left(\epsilon_{t+\frac{i}{N}}-\epsilon_{t+\frac{i-1}{N}}\right)\right)\right.$$

$$- \int_t^{t+\overline{T}} \sigma_s^2 ds \Bigg] .$$

Similarly

$$\sum_{i=1}^{N\overline{T}} \left(Y_{t+\frac{i}{N}} - Y_{t+\frac{i-1}{N}} \right)^4$$

$$= \sum_{i=1}^{N\overline{T}} \left[\left(X_{t+\frac{i}{N}} - X_{t+\frac{i-1}{N}} \right)^4 + \left(\epsilon_{t+\frac{i}{N}} - \epsilon_{t+\frac{i-1}{N}} \right)^4 + 6 \left(X_{t+\frac{i}{N}} - X_{t+\frac{i-1}{N}} \right)^2 \left(\epsilon_{t+\frac{i}{N}} - \epsilon_{t+\frac{i-1}{N}} \right)^2 \right.$$

$$\left. +4 \left(X_{t+\frac{i}{N}} - X_{t+\frac{i-1}{N}} \right)^3 \left(\epsilon_{t+\frac{i}{N}} - \epsilon_{t+\frac{i-1}{N}} \right) + 4 \left(X_{t+\frac{i}{N}} - X_{t+\frac{i-1}{N}} \right) \left(\epsilon_{t+\frac{i}{N}} - \epsilon_{t+\frac{i-1}{N}} \right)^3 \right] .$$

Therefore the statistic will diverge to plus infinity and the test will be consistent if, as $N, M \to \infty$,

$$\frac{\sum_{i=1}^{\overline{T}M} \left(\epsilon_{t+\frac{i}{M}} - \epsilon_{t+\frac{i-1}{M}} \right)^2 - \sum_{i=1}^{\overline{T}N} \left(\epsilon_{t+\frac{i}{N}} - \epsilon_{t+\frac{i-1}{N}} \right)^2}{\sqrt{\sum_{i=1}^{N\overline{T}} \left(\epsilon_{t+\frac{i}{N}} - \epsilon_{t+\frac{i-1}{N}} \right)^4}} \to \infty,$$

which holds under Assumption A4. ■

Proof of Proposition 2. From Barndorff-Nielsen and Shephard (2004a), under the null hypothesis, as $M \to \infty$,

$$\sum_{i=1}^{\overline{T}M-1} \mu_1^{-2} \left| Y_{t+\frac{i+1}{M}} - Y_{t+\frac{i}{M}} \right| \left| Y_{t+\frac{i}{M}} - Y_{t+\frac{i-1}{M}} \right| \xrightarrow{\Pr} \int_t^{t+\overline{T}} \sigma_s^2 ds,$$

$$\mu_1^{-4} \sum_{i=1}^{\overline{T}M-3} \left| Y_{t+\frac{i+4}{M}} - Y_{t+\frac{i+3}{M}} \right| \left| Y_{t+\frac{i+3}{M}} - Y_{t+\frac{i+2}{M}} \right| \left| Y_{t+\frac{i+2}{N}} - Y_{t+\frac{i+1}{N}} \right| \left| Y_{t+\frac{i+1}{N}} - Y_{t+\frac{i}{N}} \right| \xrightarrow{\Pr} \int_t^{t+\overline{T}} \sigma_s^4 ds$$

and from Theorem 1 in Barndorff-Nielsen and Shephard (2004b)

$$\sqrt{M\overline{T}} \left(\mu_1^{-2} BV_{t,\overline{T},M,T} - \int_t^{t+\overline{T}} \sigma_s^2 ds \right) \xrightarrow{d} N \left(0, 2.6090 \int_t^{t+\overline{T}} \sigma_s^4 ds \right).$$

The statements then come by the same argument as above. ■

Proof of Proposition 3.

(i) Under the null hypothesis, the test statistic can be rewritten as

$$V_{M,N,\overline{T},t} = \sqrt{N\overline{T}} \left(\frac{\left(\frac{\sum_{i=1}^{\overline{T}M} \left(Y_{t+\frac{i}{M}} - Y_{t+\frac{i-1}{M}} \right)^2}{2\overline{T}M} - \nu \right) - \left(\frac{\sum_{i=1}^{\overline{T}N} \left(Y_{t+\frac{i}{N}} - Y_{t+\frac{i-1}{N}} \right)^2}{2\overline{T}N} - \nu \right)}{\sqrt{\frac{1}{N\overline{T}} \sum_{i=1}^{N\overline{T}} \left(Y_{t+\frac{i}{N}} - Y_{t+\frac{i-1}{N}} \right)^4}} \right).$$

By Theorem A1 in Zhang, Mykland and Aït-Sahalia (2003),

$$\sqrt{M\overline{T}}\left(\frac{\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2}{2\overline{T}M}-\nu\right)=O_p(1),$$

since it converges in distribution. Thus, asymptotically,

$$V_{M,N,\overline{T},t}=-\sqrt{N\overline{T}}\left(\frac{\left(\frac{\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2}{2\overline{T}N}-\nu\right)}{\sqrt{\frac{1}{N\overline{T}}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}\right)\xrightarrow{d}N(0,1).$$

(ii) Under the alternative hypothesis, the statistic can be rearranged as

$$\begin{aligned}&V_{M,N,\overline{T},t}\\=\;&\sqrt{N\overline{T}}\frac{\left(\frac{\sum_{i=1}^{\overline{T}M}\left(Y_{t+\frac{i}{M}}-Y_{t+\frac{i-1}{M}}\right)^2}{2\overline{T}M}-\nu_{t,M}\right)-\left(\frac{\sum_{i=1}^{\overline{T}N}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^2}{2\overline{T}N}-\nu_{t,N}\right)}{\sqrt{\frac{1}{N\overline{T}}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}\\&+\frac{\sqrt{N\overline{T}}\left(\nu_{t,M}-\nu_{t,N}\right)}{\sqrt{\frac{1}{N\overline{T}}\sum_{i=1}^{N\overline{T}}\left(Y_{t+\frac{i}{N}}-Y_{t+\frac{i-1}{N}}\right)^4}}.\end{aligned}\tag{16}$$

Note that the first term of the right hand side of (16) is asymptotically normal. As the denominator in the second term of the right hand side of (16) is $O_p(\nu_{t,N})$, the statement then follows. ■

Table 1: Names and Symbols of the companies included in the DJIA

Company	Ticker Symbol
3M Company	MMM
Alcoa Inc.	AA
Altria Group, Inc.	MO
American Express Co.	AXP
AT&T Corp.	T
Boeing Co.	BA
Caterpillar, Inc.	CAT
Citigroup Inc.	C
Coca-Cola Co.	KO
DuPont (E.I.) de Nemours	DD
Eastman Kodak Co.	EK
Exxon Mobile Corp.	XOM
General Electric Co.	GE
General Motors	GM
Hewlett-Packard Co.	HPQ
Home Depot Inc.	HD
Honeywell Int'l. Inc.	HON
Intel Corp.	INTC
International Bus. Mach.	IBM
International Paper Co.	IP
J.P. Morgan Chase & Co.	JPM
Johnson & Johnson	JNJ
McDonald's Corp.	MCD
Merck & Co.Inc.	MRK
Microsoft Corp.	MSFT
Procter & Gamble Co.	PG
SBC Communications, Inc.	SBC
United Technologies Corp.	UTX
Wal-Mart Stores, Inc.	WMT
Walt Disney Co.	DIS

Table 2: Average number of trade quotations per minute of DJIA stocks

Stock	Average	1997	1998	1999	2000	2001	2002
MMM	3.65	1.86	2.22	2.60	3.08	4.73	7.35
AA	3.37	0.99	1.39	2.30	3.75	5.10	6.69
MO	9.39	7.28	7.46	9.51	9.51	8.70	13.74
AXP	5.68	2.12	3.03	4.01	5.80	9.32	9.74
T	12.26	8.96	6.80	16.39	20.99	10.09	10.07
BA	7.50	6.59	10.20	6.23	5.88	7.03	8.95
CAT	3.60	2.00	2.94	3.33	3.58	4.13	5.56
C	11.31	4.07	3.29	11.70	12.15	14.28	22.23
KO	7.25	5.94	5.97	8.41	7.45	6.55	9.04
DD	5.47	3.54	4.96	4.86	5.79	6.21	7.36
EK	3.48	3.40	2.79	2.73	3.16	3.85	4.91
XOM	8.02	4.44	4.77	5.72	7.41	9.95	15.70
GE	19.45	8.73	10.19	11.72	19.66	26.19	39.99
GM	4.90	3.49	3.76	3.76	4.32	4.81	9.18
HPQ	8.30	5.93	6.75	7.12	8.82	10.26	10.84
HD	11.32	2.98	7.45	8.19	14.80	13.25	21.07
HON	5.89	0.64	0.84	15.81	5.66	5.51	6.77
INTC	94.99	41.51	45.48	64.29	128.33	128.48	160.65
IBM	12.53	6.82	5.92	14.06	13.09	14.88	20.24
IP	3.75	2.09	2.42	3.01	4.20	4.62	6.12
JPM	6.16	1.56	2.66	2.74	4.25	10.26	15.46
JNJ	6.96	4.97	4.59	4.33	6.70	8.49	12.57
MCD	5.53	4.10	3.37	4.06	6.20	6.72	8.69
MRK	8.48	5.96	6.11	8.91	9.13	8.87	11.79
MSFT	84.60	22.18	38.47	72.08	100.45	114.07	159.27
PG	6.35	3.63	4.38	4.40	9.50	6.69	9.38
SBC	6.25	1.99	2.88	4.29	8.05	8.40	11.81
UTX	3.07	1.11	1.32	2.07	2.71	4.64	6.53
WMT	10.15	3.88	5.17	11.22	14.15	11.28	15.02
DIS	9.68	2.96	11.36	13.43	8.30	9.57	12.30

Table 3: Results of the tests for no microstructure effects

Stock	Test based on $Z_{M,N,\overline{T},t}$		Test based on $ZB_{M,N,\overline{T},t}$	
	# of Rej.	% of Rej.	# of Rej.	% of Rej.
MMM	53	17.6	41	13.6
AA	45	14.9	33	10.9
MO	200	66.4	197	65.4
AXP	45	14.9	22	7.3
T	142	47.1	132	43.8
BA	140	46.5	122	40.5
CAT	70	23.2	43	14.2
C	165	54.8	160	53.1
KO	130	43.1	101	33.5
DD	92	30.5	72	23.9
EK	63	20.9	29	9.6
XOM	163	54.1	124	41.1
GE	110	36.5	93	30.8
GM	60	19.9	49	16.2
HPQ	79	26.2	62	20.5
HD	128	42.5	123	40.8
HON	32	10.6	21	6.9
INTC	118	39.2	129	42.8
IBM	65	21.5	60	19.9
IP	112	37.2	84	27.9
JPM	63	20.9	37	12.2
JNJ	97	32.2	70	23.2
MCD	179	59.4	152	50.4
MRK	82	27.2	66	21.9
MSFT	145	48.1	145	48.1
PG	105	34.8	71	23.5
SBC	120	39.8	85	28.2
UTX	16	5.3	5	1.6
WMT	123	40.8	103	34.2
DIS	164	54.4	152	50.4

Table 4: Results of the specification tests for the microstructure noise

Stock	Test based on $V_{M,N,\overline{T},t}$, conditional on rejecting H_0 using $Z_{M,N,\overline{T},t}$		Test based on $V_{M,N,\overline{T},t}$, conditional on rejecting H_0 using $ZB_{M,N,\overline{T},t}$	
	# of Rej.	% of Rej.	# of Rej.	% of Rej.
MMM	50	94.3	35	85.3
AA	44	97.7	27	81.8
MO	165	82.5	152	77.1
AXP	44	97.7	20	90.9
T	118	83.0	103	78.0
BA	130	92.8	103	84.4
CAT	64	91.4	33	76.7
C	113	68.4	100	62.5
KO	122	93.8	90	89.1
DD	85	92.3	62	86.1
EK	58	92.0	22	75.8
XOM	157	96.3	113	91.1
GE	108	98.1	86	92.4
GM	47	78.3	36	73.4
HPQ	71	89.8	49	79.0
HD	114	89.0	97	78.8
HON	30	93.7	18	85.7
INTC	103	87.2	104	80.6
IBM	64	98.4	47	78.3
IP	107	95.5	81	96.4
JPM	59	93.6	29	78.3
JNJ	88	90.7	56	80.0
MCD	157	87.7	121	79.6
MRK	82	100	57	86.3
MSFT	133	91.7	121	83.4
PG	103	98.0	66	92.9
SBC	113	94.1	73	85.8
UTX	16	100	5	100
WMT	106	86.1	83	80.5
DIS	145	88.4	121	79.6

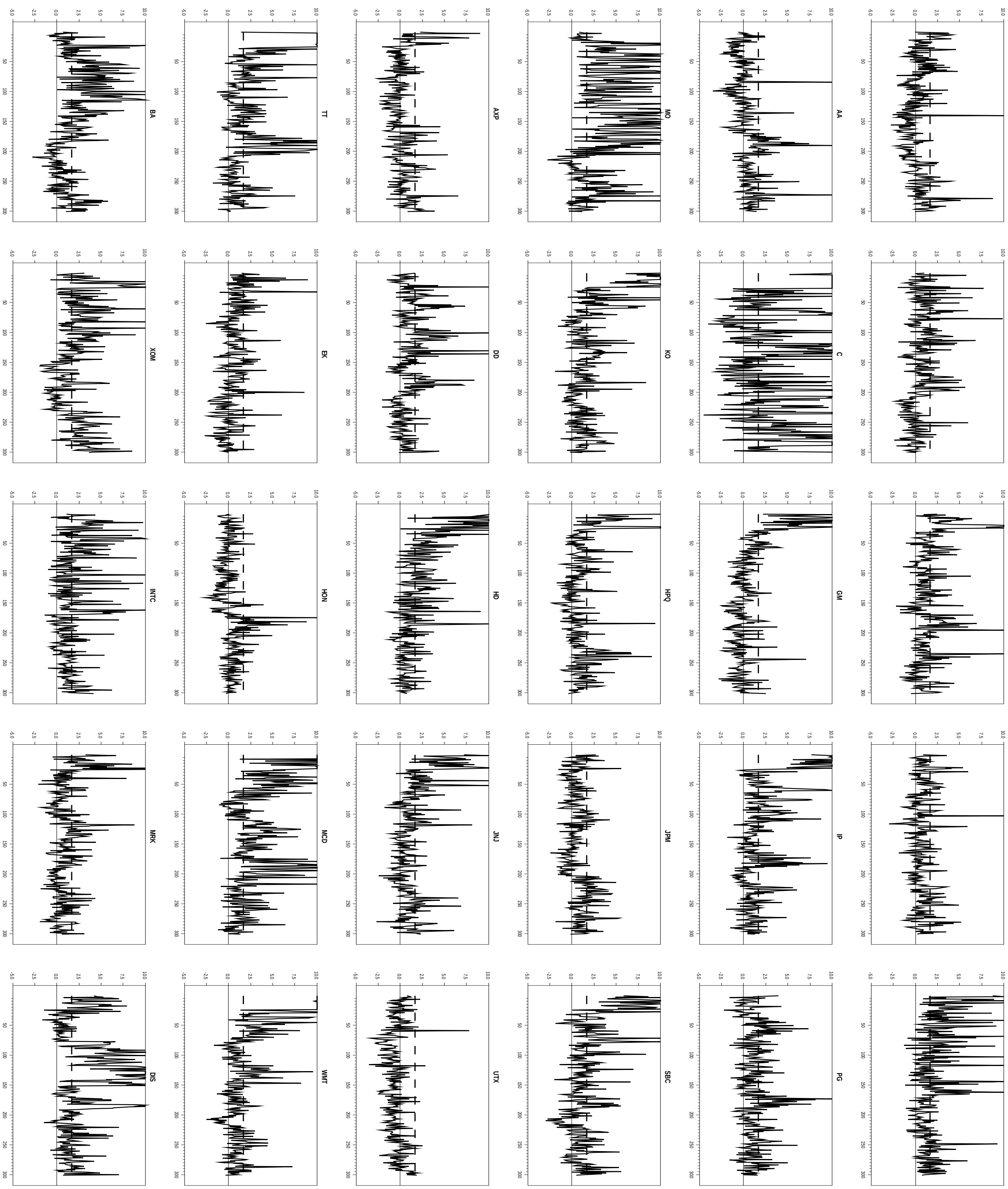

Figure 1: Plot of the test statistic defined in (9), with the upper 95% percentile of the standard normal (dotted line)

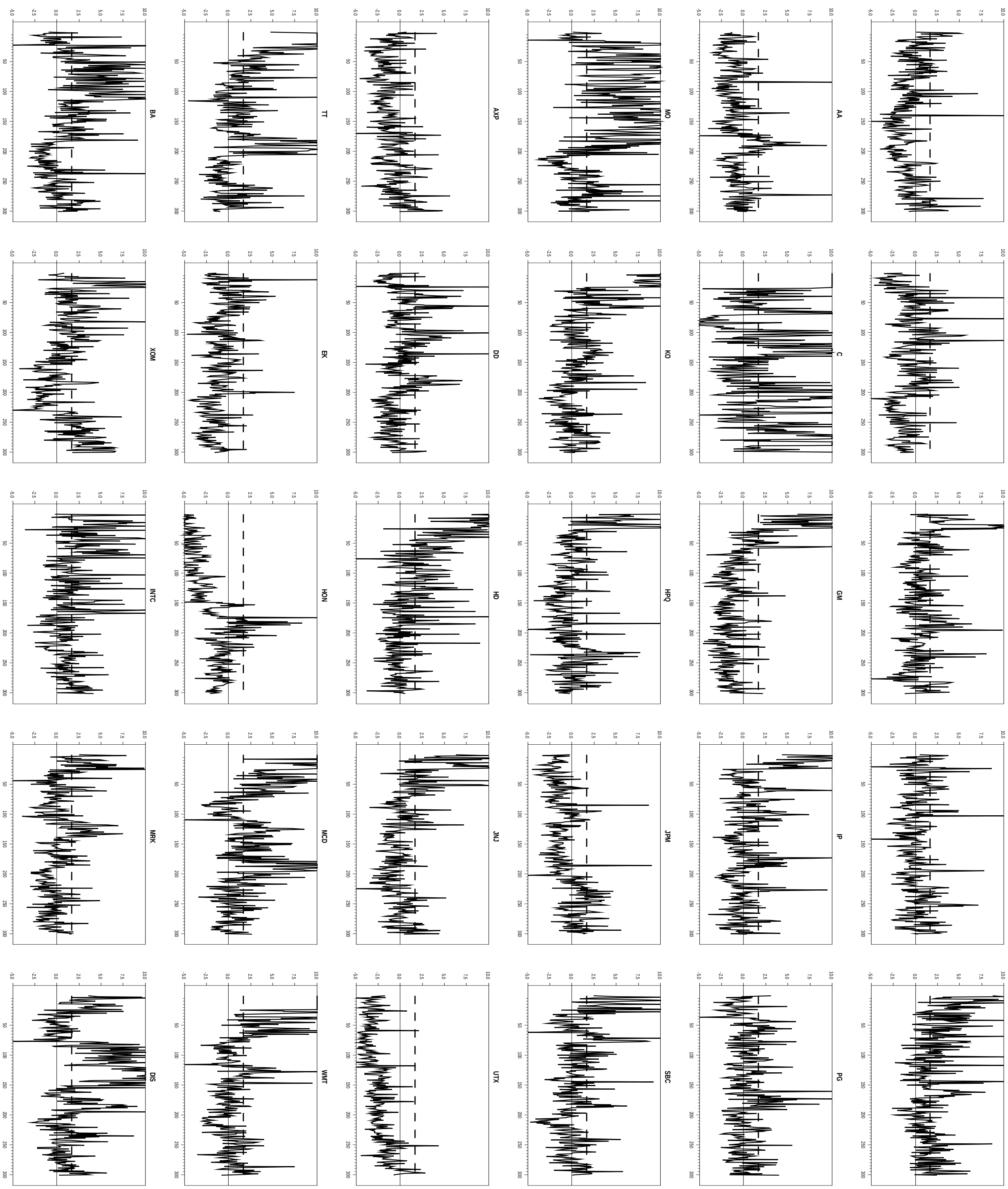

Figure 2: Plot of the test statistic defined in (11), with the 95% percentile of the standard normal (dotted line)

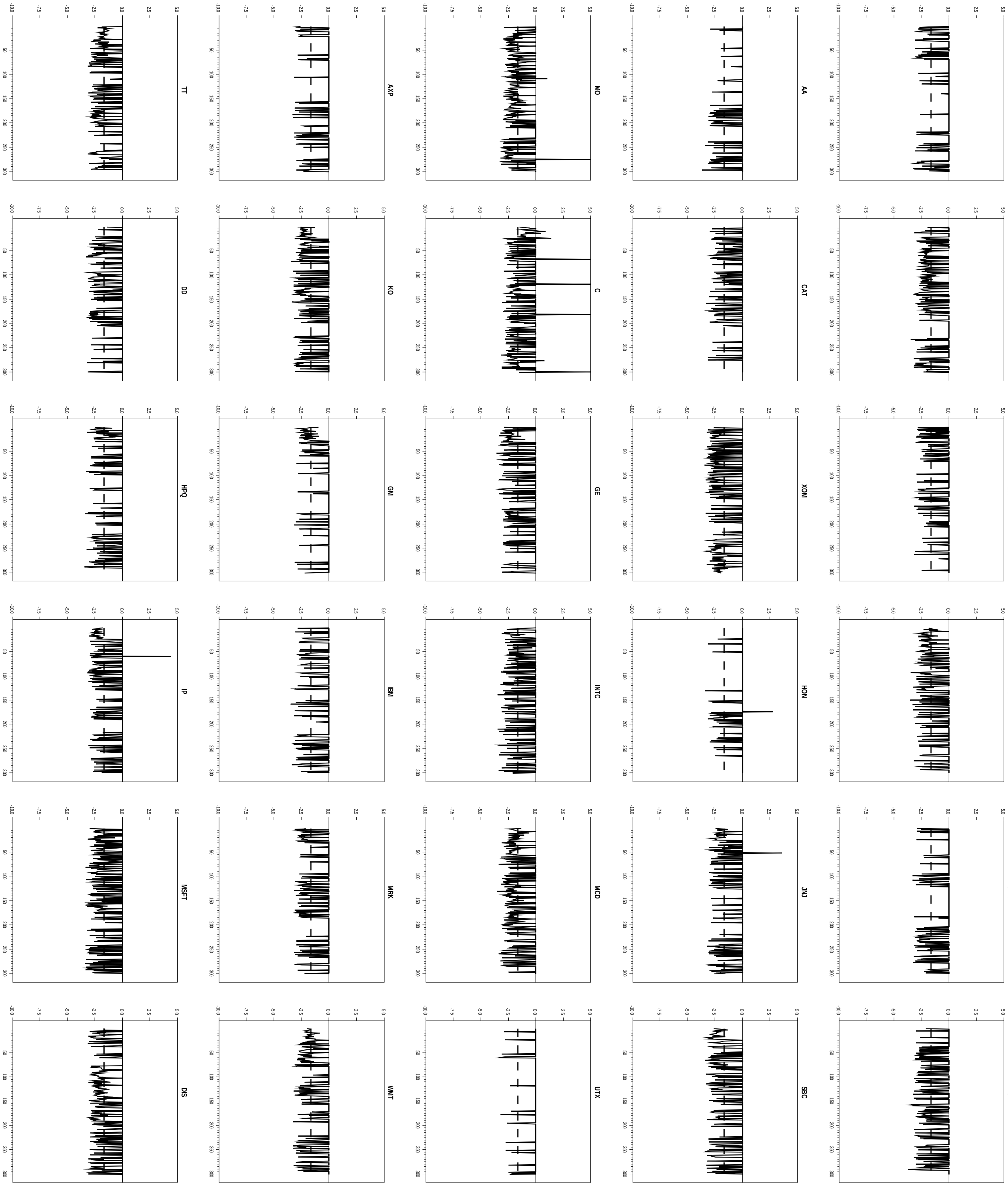

Figure 3: Plot of the test statistic defined in (14), conditionally on rejections of the null hypothesis in (7) using $RV_{t,M,\overline{T}}$, with the 5% percentile of the standard normal (dotted line)

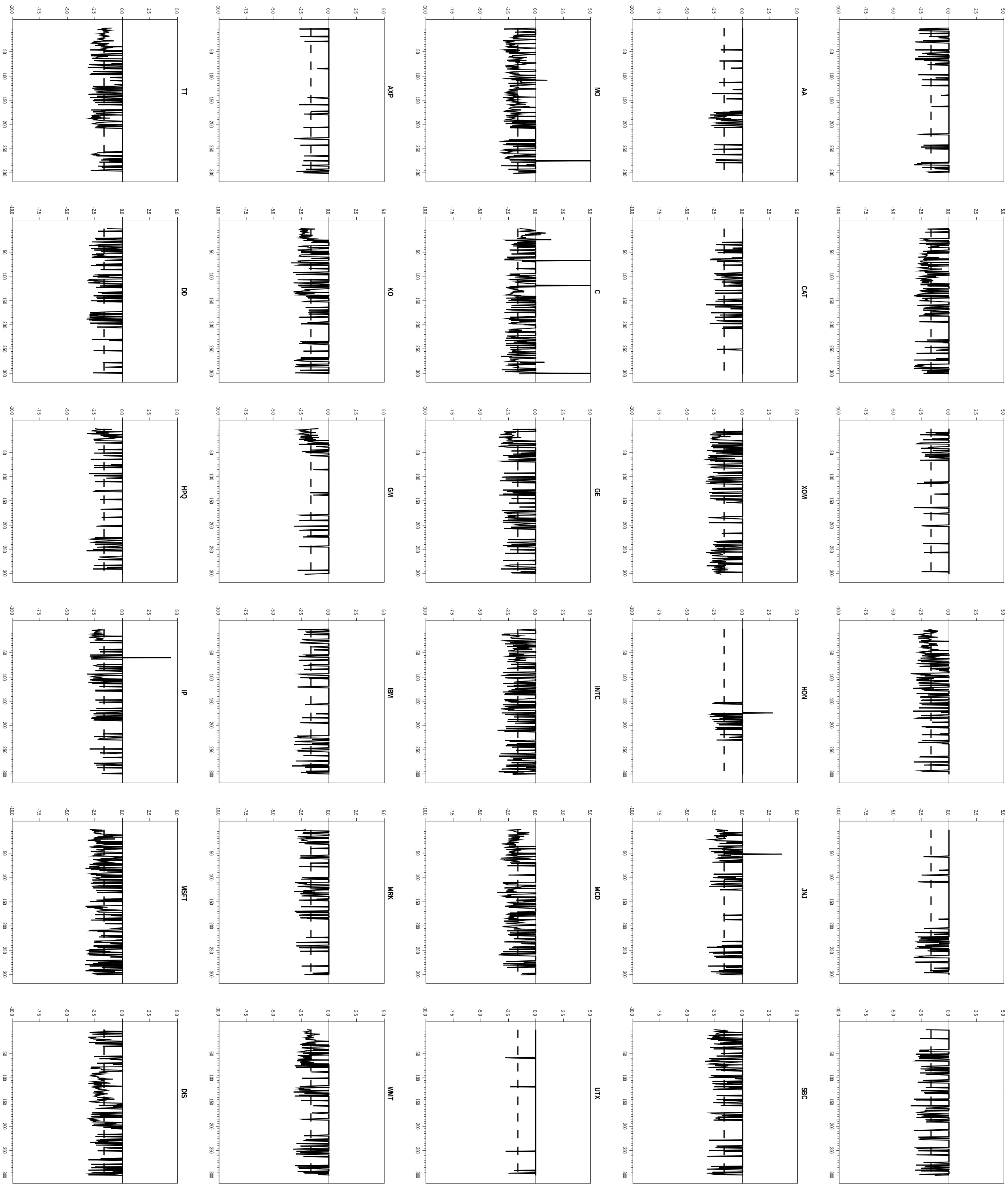

Figure 4: Plot of the test statistic defined in (14), conditionally on rejections of the null hypothesis in (7) using $BV_{t,M,\overline{T}}$, with the 5% percentile of the standard normal (dotted line)