Realized Volatility in Noisy Prices: a MSRV approach

Jalal Seifoddini
Department of Financial Management, Islamshahr Branch, Islamic Azad University, Islamshahr, Iran;
(corresponding author)
jalal.seifoddini@iau.ac.ir

Fraydoon Rahnamay Roodposhti
Department of Financial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran;
rahnama.roodposhti@gmail.com

ABSTRACT
Volatility is the primary measure of risk in modern finance and volatility estimation and inference has attracted substantial attention in the recent financial econometric literature, especially in high-frequency analyses. High-frequency prices carry a significant amount of noise. Therefore, there are two volatility components embedded in the returns constructed using high frequency prices: the true volatility of the unobservable efficient returns and the volatility from the existence of microstructure noise. Researchers proposed several methodologies for estimating these two components but each of these estimators has its own pros and cons. however, some of them have higher rate of convergence. Multi-Scale Realized Volatility (MSRV) is one of these estimators that reported to have a high efficiency in estimating true realized volatility. In this paper, after estimating these two components through the MSRV approach, we investigate the relation between them. Our results suggest that there is a positive meaningful relation between microstructure noise and true realized volatility.

Keywords:
Microstructure Noise, Volatility, High Frequency Data, Multi-Scale Realized Volatility
1. Introduction

Noise can exist anywhere in economy (Black, 1986), but in the financial literature, noise refers mostly to microstructure noise in prices. Noise in prices appears due to behavioral factors, or structural frictions, such as changes in supply and demand. In general, any form of temporary deviation in price from its fundamental value is called noise. The role of noise in financial markets is both positive and negative. Without noise, no financial market would exist because noise creates liquidity in financial markets (Black, 1986). Hu et al. research supports this hypothesis. They show that temporary price deviations involve important information about the level of liquidity in the overall market (Hu, Pan, & Wang, 2013).

Accurate real-time volatility forecasts are needed for many applications, such as the real-time pricing of options and real-time risk management of trading positions. In order to generate a forecast, however, we first need a good estimate of realized variance. On the one hand, without an efficient estimator of realized variance, it is hard to see how the performance of different forecasts can be reliably compared. On the other hand, it seems that a volatility forecasting model generates its best estimate of realized variance when supplied with its best measure. Microstructure noise can cause the series of price returns between trades to be auto-correlated; so the obvious estimator of realized variance is very biased (Gatheral & Oomen, 2010). The variance of continuously compounded returns depends on the variance of the underlying efficient returns and the variance of the microstructure noise components in returns. Both variance measures carry a fundamental economic significance. The variance of the efficient return process is a crucial ingredient in the practice and theory of asset valuation and risk management. The variance of the microstructure noise component reflects the market structure and the price setting behavior of market participants and thereby contains information about the market’s fine-grain dynamics (Bandi & Russell, 2006). Hence, in this paper, we try to determine the magnitude of the noise in the volatility estimates from high-frequency data and separate it from the price process.

The existence of traders who trade for reasons other than information provides the required diversity to the market. Thus, noise-based trades are essential for market liquidity (Morawski, 2008). However, noise trades make prices deviate from their fundamental values. Therefore, as the amount of noise-based trades increases, the profitability of information-based trades will increases; however this only happens because prices have more noise. Other researchers such as DeLong et al. also argue that noise trading can lead to a large divergence between market prices and fundamental values (De Long J. B., Shleifer, Summers, & Waldmann, 1990). Hence, more noise trading can be accompanied with more volatility in prices. Therefore, after separating fundamental volatility from microstructure noise we investigate the relationship between true volatility and noise to see if more noise leads to more volatility or vice versa.

2. Literature Review

Volatility modeling has been a very active area of research in recent years. This interest is largely motivated by the importance of volatility in financial markets. Volatility estimates are widely studied as risk measures in many asset pricing models (e.g., see Adam et al. (2016) and Herskovic et al. (2016)). Also volatility enters option pricing models (e.g., see Song & Xiu (2016), Carr & Wu (2016) and Alós & León (2016)). This very active area of research resulted in the development of several types of models. These alternative models try to account for different stylized facts documented in the literature. Autoregressive Moving Average (ARMA) models, Autoregressive Conditional Heteroscedasticity (ARCH) models, Stochastic Volatility (SV) models, Regime Switching models and Threshold models are the most well-known ones in the literature. Stochastic volatility models treat volatility as an unobserved variable which is assumed to follow a certain stochastic process. These models are able to overcome some of the drawbacks of GARCH models (Satchell & Knight, 2011). These models are usually designed to estimate the daily, weekly, or monthly volatility using data sampled at the same frequency. However, thanks to the widespread availability of databases providing the intraday prices of financial assets econometricians have considered using data sampled at a very high frequency to compute ex-post measures of volatility at a lower frequency. This method is known as realized volatility approach (Bauwens, Hafner, & Laurent, 2012). In other words, when computing realized volatility, the aim is to use high-frequency price observations to construct an efficient ex post estimate
of the low-frequency return variance, i.e., the variance of returns measured over a horizon that is relatively long compared to the frequency of observation and that is unaffected by the microstructure effects potentially present in the high-frequency prices (Gatheral & Oomen, 2010). However, a major problem in estimating realized volatility using high-frequency data is how to estimate the volatility consistently and efficiently, when the observed asset returns contain error or noise, for example, in the form of microstructure noise (Zhang, 2006) because the realized volatility is sensitive to market microstructure noise and the distortions induced by market microstructure noise increase with the sampling frequency (Hansen & Huang, 2016). This issue has been addressed in the recent literature by separating microstructure noise from volatility.

There are two main approaches to separating noise from volatility. Studies performed in quote-driven markets usually use market makers quotes. They argue that it is common practice in the realized variance literature to use midpoints of bid-ask quotes or volume weighted mid-quote as measures of the true prices. While these measures are affected by residual noise, they are generally less noisy measures of the efficient prices than are transaction prices because they do not suffer from bid-ask bounce effects (see, e.g., Bandi & Russell (2008) and (2006), Hansen & Lunde (2006), Mancino & Sanfelici (2008), Griffin & Oomen (2011)). On the other hand, studies performed in order-driven markets use transaction prices; although using linear weighting of the best bid and ask prices is also applicable to order-driven markets. In studies performed to estimate noise by using transaction prices, usually two popular types of estimators have been used. The first one is a parametric estimator, and the second one is a non-parametric estimator (Ait-Sahalia & Xiu, 2012). The Maximum Likelihood Estimation (MLE) is the parametric estimator provided by Ait-Sahalia et al. (2005) which then improved by Xiu (2010) as a Quasi-Maximum Likelihood Estimation (QMLE) in order to emphasize model misspecification and keep the notation in line with the classic results of likelihood-based estimation under misspecified models. The non-parametric estimator is called the Two-Scales Realized Volatility (TSRV), which is provided by Zhang et al. (2005). However, TSRV is not efficient and converges to the true volatility only at the rate of $n^{-1/6}$. Therefore, Zhang modified the TSRV and proposed a new estimator called Multi-Scale Realized Volatility (MSRV) which converges to the true volatility at the rate of $n^{-1/4}$ (Zhang, 2006). Although, some other studies like Barndorff-Nielsen et al. (2008) and Barndorff-Nielsen et al. (2011) proposed a realized kernel estimator, however, it turns out that the realized kernel is closely linked with TSRV and MSRV (Wang, 2016). To determine which one is a better estimator of noise in stock prices, Gatheral and Oomen (2010) compared a comprehensive set of estimators including the above estimators. According to their study, the QMLE and MSRV are among the best in terms of efficiency and robustness. However, because MSRV does not depend on the probability distribution, in the present paper, we use MSRV to estimate noise and realized volatility. You can also see Seifoddini et al. (in press) for more information about the other approaches of estimating microstructure noise in high-frequency prices where we have estimated the noise through QMLE approach and compared the results with TSRV approach.

Kupiec shows that in certain circumstances the risky asset's price exhibits excess volatility and agents engage in excess trading activity owing to the presence of destabilizing noise traders (Kupiec, 1996). Brown also shows that if noise traders affect prices, the risk they cause is volatility, then noise should be correlated with volatility (Brown, 1999). Blanchard et al. suggest that noise shocks explain a sizable fraction of short-run fluctuations (Blanchard, L'Huillier, & Lorenzoni, 2013). Orlitzky shows that greater noise in financial markets typically invites more noise trading, which in turn leads to excess market volatility (Orlitzky, 2013). However, we can also argue that when the volatility of the market increases the speculative behavior of market participants would increase and more investors would enter the market to exploit the opportunities and emotional behaviors in market would increase and the noise in market would subsequently increase. Researchers such as Bandi and Russel (2006) and Ait-Sahalia and Yu (2009) reported that microstructure noise is positively related with volatility.

### 3. Methodology

We estimate noise through a nonparametric approach, where volatility is left unspecified, stochastic, and we now explain the MSRV approach to separating the fundamental and noise volatilities in this
case. To do so, first we start with TSRV and then we extend it to MSRV.

To demonstrate the idea, let \( Y \) be the observed log price of stocks. To include noise in our model, suppose that the \( Y_{t_{n,i}} \) are noisy, and the corresponding true log prices are \( X_{t_{n,i}} \). Their relation can be modelled as:

\[
Y_{t_{n,i}} = X_{t_{n,i}} + \epsilon_{t_{n,i}}
\]  

(1)

Where the \( \epsilon_{t_{n,i}} \) are i.i.d noise with mean zero and variance of \( \alpha^2 \) and are independent from the \( X_{t_{n,i}} \) process. We assume that the process of log prices follows the Itô process (2006):

\[
dX_t = \mu dt + \sigma dW_t
\]

(2)

Where \( W_t \) represents a Brownian motion, \( \mu \) is a drift function, and \( \sigma \) is the diffusion coefficient.

Our goal is to estimate \( \int_0^T \sigma_t^2 dt \). For simplicity, we call \( \int_0^T \sigma_t^2 dt \) the integrated volatility, and denote it by \( \langle X, X \rangle_T = \int_0^T \sigma_t^2 dt \) over a fixed time period \([0,T]\).

The usual estimator of \( \langle X, X \rangle_T \) is the realized volatility (RV):

\[
\left\langle \frac{\hat{X}}{\hat{Y}} \rightangle^2 = \frac{\sum(Y_{t_{n,i}} - Y_{t_{n,i-1}})^2}{n}
\]

(3)

In the absence of noise, \( \left\langle \frac{\hat{Y}}{\hat{Y}} \right\rangle^{(n,1)} \) consistently estimates \( \langle X, X \rangle_T \). However, ignoring market microstructure noise leads to a dangerous situation when \( T \to \infty \). After suitable scaling, \( RV \) based on the observed log-returns is a consistent and asymptotically normal estimator—but of the quantity \( 2nE[e^2] \) rather than of the object of interest, \( \langle X, X \rangle_T \). With \( N = T/\Delta \), we have

\[
E[\hat{\sigma}^2] = \frac{2n\alpha^2}{T} + o(n) = \frac{2nE[e^2]}{T} + o(n)
\]

(4)

So \( (T/2n)[\hat{\sigma}^2] \) becomes an estimator of \( E[e^2] = \alpha^2 \).

Note, in particular, that \( \hat{\sigma}^2 \) estimates the variance of the noise, which is essentially unrelated to the object of interest \( \alpha^2 \) (Ait-Sahalia & Yu, 2009).

\[
\alpha^2 = \frac{T\hat{\sigma}^2}{2n}
\]

(5)

The recommendation in the literature has then been to sample sparsely at some lower frequency.

However, one of the most basic lessons of statistics is that discarding data is, in general, not advisable. Zhang, Mykland and Ait-Sahalia (2005) proposed the TSRV which makes use of the full data sample, however delivers consistent estimators of both \( \langle X, X \rangle_T \). TSRV, is based on subsampling, averaging, and bias-correction. By evaluating the quadratic variation at two different frequencies, averaging the results over the entire sampling, and taking a suitable linear combination of the result at the two frequencies, one obtains a consistent and asymptotically unbiased estimator of \( \langle X, X \rangle_T \). TSRV has the form

\[
\langle X, X \rangle^{(TSRV)} = \left[ Y, Y \right]^{(n,K)} - \frac{2}{nK} \left[ Y, Y \right]^{(n,1)}
\]

(6)

Where

\[
\left[ Y, Y \right]^{(n,K)} = \frac{1}{K} \sum(Y_{t_{n,i}} - Y_{t_{n,i-1}})^2
\]

(7)

TSRV’s construction is quite simple: first calculate \( \left[ Y, Y \right]^{(n,1)} \) using all data. Then, partition the original grid of observation times, \( G = \{t_0, ..., t_n\} \) into subsamples, \( G^{(k)}, k = 1, ..., K, \) where \( n/K \to \infty \) as \( n \to \infty \). For example, for \( G^{(1)} \) start at the first observation and take an observation every 5 minutes; for \( G^{(2)} \) start at the second observation and take an observation every 5 minutes, etc. Then we average the estimators obtained on the subsamples to calculate \( \left[ Y, Y \right]^{(n,K)} \).

The TSRV estimator has many desirable features, but its rate of convergence is not satisfactory (Ait-Sahalia & Yu, 2009), (Zhang, 2006). Therefore, Zhang (2006) modified TSRV and proposed the Multi-Scale Realized Volatility (MSRV). MSRV has the form

\[
\langle X, X \rangle^{(MSRV)} = \sum_{i=1}^M c_i \left[ Y, Y \right]^{(n,K_i)}
\]

(8)

where \( M \) is a positive integer greater than 2. Comparing to \( \langle X, X \rangle^{(TSRV)} \), which uses two time scales (1 and K), \( \langle X, X \rangle^{(MSRV)} \) combines M different time scales. The weights \( c_i \) are selected so that \( \langle X, X \rangle^{(MSRV)} \) is asymptotically unbiased and has optimal convergence rate. The rationale is that by combining more than two time scales, we can improve the efficiency of the estimator.

After estimating the noise part and the realized volatility of the price series, we investigate to see
whether volatility can explain the noise in the market. At the present, our hypothesis is that their relation has the form

$$a_{j,t} = c_0 + a_j s_j + e_{j,t} \quad j=1,...,N$$  (9)

where $N$ is the number of companies in our sample.

4. Results

We performed our study on the shares of companies listed on the Tehran Stock Exchange, which is an order-driven market with no specialist or market maker. In studies conducted on financial markets microstructure, using high-frequency data, considering a 1-year time period can provide enough data for the results’ robustness (see for example, Doman (2010), Tissaoui (2012)) and Ahn & Cheung (1999). Accordingly, the time period considered in this research was from the beginning 2015 to the beginning of 2016. We set our maximum frequency to 5 minutes.

Because we need high-frequency data, the main criterion in the selection of stocks was that they should have the lowest number of closing days, the highest amount of trade volume and the highest number of trading days because we require stocks whose trade volume is high enough to obtain the required observations at the determined frequency. Therefore, in order to select the sample stocks, we use stocks included in the list of the 50 most active companies provided by the Tehran Stock Exchange. The list of these companies is seasonally announced by the Tehran Stock Exchange, which ranks companies based on trade volume and the number of trading days. In our research, we select companies represented on this list for four subsequent seasons. The same measures were considered by studies that use high-frequency data, for example, Ait-Sahalia and Yu (2009) removed stocks with fewer than 200 daily trades from their sample. Table (1) provides the standard deviation, minimum and maximum number of trading days, number of trades per day, daily trading volume and daily volume of orders of the sample in the time period under study.

We estimated the true realized volatility and noise level via the MSRV method. We set $M=4$ and we consider equal weights $c_i=0.25$. Table (2) describes the specifications of the estimated noise and true realized volatility.

<table>
<thead>
<tr>
<th>Stocks characteristics</th>
<th>Max</th>
<th>Min</th>
<th>Average</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trading days</td>
<td>221</td>
<td>184</td>
<td>201</td>
<td>8</td>
</tr>
<tr>
<td>Number of trades per day</td>
<td>13,128</td>
<td>40</td>
<td>355</td>
<td>4</td>
</tr>
<tr>
<td>Daily trading volume</td>
<td>226 m</td>
<td>1,940</td>
<td>2 m</td>
<td>7</td>
</tr>
<tr>
<td>Daily volume of orders</td>
<td>98,019</td>
<td>215</td>
<td>12,300</td>
<td>9</td>
</tr>
<tr>
<td>The average time interval between orders</td>
<td>240</td>
<td>0</td>
<td>196</td>
<td>10</td>
</tr>
</tbody>
</table>

After estimating the microstructure noise in prices and their true realized volatility, first we investigate the existence of correlation between them.

As the correlation analysis presented in Table (3) demonstrates, there is a meaningful correlation between realized volatility and estimated noise. Now, we conduct a regression model to study the explanatory power of volatility on noise. We run the Redundant Fixed Effects Tests to account for firm-specific heterogeneity.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Variable symbol</th>
<th>Average</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise</td>
<td>$\alpha_j$</td>
<td>0.0022</td>
<td>0.0019</td>
</tr>
<tr>
<td>Realized volatility</td>
<td>$\sigma_j$</td>
<td>0.3169</td>
<td>0.2289</td>
</tr>
</tbody>
</table>

Redundant Fixed Effects Test results presented in Table (4) suggest that fixed effects model should be...
used instead of pooled model. Now we run the stock fixed-effect regression.

Table 5. The results of panel regression on noise and realized volatility

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coef.</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.0014</td>
<td>0.00</td>
</tr>
<tr>
<td>SIGMA</td>
<td>0.0013</td>
<td>0.00</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.056</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>780</td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>8.033</td>
<td></td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.805</td>
<td></td>
</tr>
</tbody>
</table>

According to the regression results presented in Table (5) there is a meaningful positive relation between noise and realized volatility. Although, the results also show a meaningful relation between noise and the intercept of the regression and that means volatility is not the only effective factor on microstructure noise. However, that actually makes sense because noise, by its nature, can arise due to a lot of known and unknown factors, and volatility is just one of them.

5. Discussion and Conclusions

In this paper we estimated the microstructure noise in prices and the true realized volatility using high-frequency data. Then, we investigated the hypothesis that volatility is positively related to microstructure noise in prices. We conducted our research through Multi-scale realized Volatility approach and panel regression. According to our findings, we can conclude that more volatility in the prices of stocks is positively associated with an increase in the level of noise in prices. Our findings are in line with the findings of Ait-Sahalia and Yu (2009), and Bandi and Russell (2006) who reported a positive and significant relationship between estimated noise variance and efficient price variance. We infer that this could be due to the intensification of emotional decisions and the presence of noise traders in volatile market situations since there is a long history of evidence arguing that irrationality and noise trading increases along with market volatility. In an early study De long et al. (1989) asserted that empirical research has identified a significant amount of volatility in stock prices that cannot easily be explained by changes in fundamentals; one interpretation is that asset prices respond not only to news but also to irrational “noise trading.” since then, a lot of studies have been conducted on the effect of noise trading on market volatility, most of them in support of this hypothesis.

In a recent study Aabo et al. (2017) decomposed volatility and investigated its relationship with noise trading and their findings shows that high market volatility on its own is associated with more mispricing and more specifically they find that larger values of absolute idiosyncratic volatility reflect an increasing role of noise traders.

Our findings suggest that stock market investors should be more cautious of noisy prices in the volatile market situations and be aware of the increase in noise trading and irrational behaviors in these situations. We suggest that it would be better for investors to avoid making impulsive decisions in these situations because they may take unnecessary excess risks due to noise in prices that would decrease in less volatile market conditions.

References


