Dynamic Long Memory High Frequency Multipower Variation Volatility Evaluations for S&P500

Wen Cheong Chin1, Min Cherng Lee2, Tan Pei Pei3, Grace Lee Ching Yap4 & Christine Tan Nya Ling1
1 Faculty of Management, SIG Quantitative Economics and Finance, Multimedia University, Malaysia
2 Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Malaysia
3 Faculty of Economics and Administration, University of Malaya, Malaysia
4 Faculty of Engineering, University of Nottingham (Malaysia Campus), Malaysia
Correspondence: Wen Cheong Chin, Faculty of Management, SIG Quatitative Economics and Finance, Multimedia University, Malaysia, 63100 Cyberjaya, Selangor, Malaysia. E-mail: wcchin@mmu.edu.my

Received: September 4, 2015         Accepted: January 11, 2016       Online Published: February 4, 2016
doi:10.5539/mas.v10n5p1           URL: http://dx.doi.org/10.5539/mas.v10n5p1

The research is financed by FRGS, MOSTI, MALAYSIA.

Abstract
This study explores the multipower variation integrated volatility estimates using high frequency data in financial stock market. The different combinations of multipower variation estimators are robust to drastic financial jumps and market microstructure noise. In order to examine the informationally market efficiency, we proposed a rolling window estimate procedures of Hurst parameter using the modified rescale-range approach. In order to test the robustness of the method, we have selected the S&P500 as the empirical data. The empirical study found that the long memory cascading volatility is fluctuating across the studied period and drastically trim down after the subprime mortgage crisis. This time-varying long memory analysis allow us to understand the informationally market efficiency before and after the subprime mortgage crisis in U.S.

Keywords: integrated volatility, high frequency data, multiple power estimator, abrupt jump

1. Introduction
Volatility is one of the important elements in nowadays financial investment strategies such as asset management (Gormus, Soytas, & David Diltz, 2014), options pricing (Birkelund et al., 2015), forecasting future returns (Gallo & Otrando, 2015) and various risk management (Chkili, Hammoudeh, & Nguyen, 2014) applications. With the rapid growth of trading technology, high frequency trading has become an essential tool for recent financial market facilities. During bullish, hundreds or thousands of transactions are needed to be handled simultaneously within a single second. These heavy trading activities have led to enormous amounts of high frequency data in financial markets. The interest of high frequency volatility estimation has steadily increased after it has been proven (Andersen & Bollerslev, 1998; Blair, Poon, & Taylor, 2001) significantly improve the modelling and forecasting performances in foreign exchange and stock markets. Realized volatility, which is defined as the sum of the squared intraday returns over a specified time interval is a commonly used method in measuring volatility (Barndorff-Nielsen & Shephard, 2002). Recent studies on realized volatility can be found in Shirota, Hizu and Omori (2014) which includes the volatility leverage and long memory. One of the drawbacks of realized volatility estimator is its robustness to sudden jump due to unforeseen events such as natural disasters, financial crisis, political turmoil and etc that will affect the financial market. In order to overcome this shortcoming, the multipower variation estimator theory such as bi-power, tri-power and quad-power variation estimators, which was proposed by Barndorff-Nielsen and Shephard (2004) and has been used to estimate time-varying volatility and have shown robustness to jumps in asset price.

Understanding the predictability of volatility provides useful implication in efficiency market hypothesis (Fama, 1998) analysis. The availability of high frequency data and the discovery of nonlinearity in financial time series have led to new definitions such as heterogeneous market hypothesis (Dacorogna et al., 1998; Muller et al., 1997) and fractal market hypothesis (Peter, 1994). The market heterogeneity is against the assumption of homogeneous
investors in the traditional EMH where investors may differ from their endowments, interests, risk profiles, degree of information, contractual constraints and motivations of trading and reactions to news. Instead of responding to new market information the same way, heterogeneous investors react differently and operate on different investment time horizons (second, week, monthly, etc.). The time horizon investments are illustrated in Figure 1.

Since each group of time horizon investors consist of their own trading rules and strategies to inflowing information, these trading activities are creating a cascade of volatilities for all the time horizon investments. The combinations of these dissimilar volatilities (due to reaction times) are believed to produce hyperbolic autocorrelation decays or long memory property in financial markets. Figure 2 illustrates the structure of the heterogeneous market.

Since the core of the HMH is the long memory cascading volatility, this property has been intensively studied by Niu and Wang (2013), Su (2014) and Hull and McGroarty (2014) in some recent literature of finance. However, most of the long memory determination approaches are based on a single determination for the whole time span. In other words, the long memory is considered as constant over the whole period of studied period which not realistic in the complex dynamic financial market. This issue can be overcome by using a rolling estimation (Cajueiro & Tabak, 2005, 2008) over a fix time-window, says 512 or 1024 data points. The long memory is re-estimated and slid from the first till the last data point. In this way, the long memory can be observed over the studied time span.

This study is divided into two portions. First, we will look into the combination of multipower variation estimators which can be used to estimate the integrated volatility in the presence of jumps. These include realized volatility, bipower, tripower and quadpower variation volatility which are robust to abrupt jump and
microstructure issue in financial markets. Second, the heterogeneous market hypothesis will be examined based on the rolling window estimates of long memory Hurst parameter. The presence of long memory indicated that the volatilities are predictable to certain extents which are against the ideal efficient market. Unlike prior studies using realized volatility and static long memory estimation, our focus is to take into account the microstructure noise, possible abrupt jumps and also time-varying latent volatility measurements. Using a more accurate volatility estimator like multipower variation volatility should help explaining the market’s volatility and heterogeneous market hypothesis in a better way. The remaining of this study is organized as follows: Section 2 provides the description of multipower variation of volatility estimations and the rolling window Hurst parameter estimations; Section 3 discusses the empirical data and results and finally, Section 4 concludes the findings of the study.

2. Method

This study selects the mature market of United States S&P500 index which consists of the top 500 active large-cap stocks that represent the U.S. equity market. The S&P500 index is a free-float capitalization-weighted index introduced in year 1957 which traded under the NYSE Euronext and NASDAQ OMX. The empirical data is collected from Bloomberg from January 2008 to 2013 which consists of 1392 trading days with approximately 108576 5-minute data from 09.30 to 16.00. For one-day intervals, the continuously compounded intraday returns of day \( t \) with sampling frequency \( N \) can be written as

\[
r_{t,j} = 100(\ln p_{t_j}^{close} - \ln p_{t_{j-1}}^{close})
\]

(1)

\(j = 1, \ldots, N\) and \( t = 1, \ldots, T \). Thus, a full trading day for 5-minute interval consists of \( N \times \delta = 78 \) minutes with \( N \) equally-spaced subintervals of length \( \delta \). For the high frequency data background, we begin with the approximation of the realized volatility (RV) to latent volatility which is related to the theory of quadratic variation and integrated variance (Andersen and Bollerslev, 1998). Consider a stochastic-volatility process logarithmic price of financial asset,

\[d \sigma(t) = \mu(t)dt + \sigma(t)dW(t)\]

(2)

where \( \mu(t), \sigma(t) \) and \( W(t) \) are the drift, volatility and standard Brownian motion. The \( \mu(t), \sigma(t) \) may be time-varying but are assumed to be independent of \( dW(t) \). In an alternative representation

\[p_t = p_0 + \int_0^t \mu(t)dt + \int_0^t \sigma(t)dW(t).\]

(3)

The quadratic variation process for a sequence of partitions when \( m \to \infty \) is equivalent to the integrated variance

\[\lim_{m \to \infty} \sum_{t=0}^{m-1} (p_{t+1} - p_t)^2 = \int_0^t \sigma^2(t)dt.\]

(4)

In other words, the quadratic variation and hence the integrated variance can be consistently (Barndorff-Nielsen and Shephard, 2002) estimated by the sum of squares returns computed under the condition \( m \to \infty \) with very high frequency observations. However, high sampling frequencies reduce the extension of the estimator’s variance, but at the same time increase its bias component. It is noted that in the presence of abrupt jumps in the series, equation (3) is no longer consistent estimate for integrated variance. Thus, a more robust estimator which immunes to jump is needed to overcome this inconsistency issue.

2.1 Multipower Variation Volatility Estimators

Alternately, in order to overcome the inconsistency and asymptotic theory issue under the presence of abrupt jumps, one may use the higher multipower variation estimators. A general consistent realized multipower \( (p) \) variation estimator (Barndorff-Nielsen and Shephard, 2002) of the corresponding integrated power of the volatility is defined as

\[MPV_t(n,p) = p_{t/n}^{-n} \left( \frac{t}{t-n+1} \right) t^{p-1} \sum_{j=1}^{t-n+1} |r_j|^{p/n} ... |r_{j+n-1}|^{p/n}\]

(5)

where \( n \) and \( p \) are positive integers with the relationship \( n > p/2 \) with a finite sample correction of \( \left( \frac{t}{t-n+1} \right) \). \( n \) normally sets the window size of return blocks and \( p \) indicates the desired power variation. For i.i.d price changes, \( p_{t/n} = \left( \frac{p}{2n} \frac{\Gamma(p+1/2)}{\Gamma(1/2)} \right) \). If all the adjacent returns are i.i.d normally distributed, Barndorff-Nielsen et al. (2006) claimed that each term of \( MPV \) delivers an unbiased estimate of the power of volatility. For example, for realized volatility (Andersen et. al., 2001), the \( n=1 \) and \( p=2 \) is defined as
\[ RV_t = MPV_t(n = 1, p = 2) = \mu_2^{-1} \sum_{j=1}^{t} |r_j|^2 \] (6)

Other higher variations can be obtained by altering the values of \( n \) and \( p \). These include the bipower variation (Barndorff-Nielsen and Shephard, 2004), tripower variation (Andersen, Bollerslev, & Diebold, 2006) and Quadpower variation (Barndorff Nielsen & Shephard, 2004) volatility estimator as follows:

\[ BV_t = MPV_t(n = 2, p = 2) = \mu_2^{-2} \sum_{i=1}^{t} |r_i|^2 |r_{i+1}| \] (7)

\[ TV_t = MPV_t(n = 3, p = 2) = \mu_3^{-2/3} \sum_{i=1}^{t-2} |r_i|^2/3 |r_{i+1}|^{2/3} |r_{i+2}|^{2/3} \] (8)

\[ QV_t = MPV_t(n = 4, p = 2) = \mu_4^{-4} \sum_{i=1}^{t-3} |r_i|^4/3 |r_{i+1}|^{1/2} |r_{i+2}|^{1/2} |r_{i+3}|^{1/2} \] (9)

### 2.2 Multipower Variation Rolling Modified Rescaled-Range Estimation Procedures

The long memory volatility can be measured either by Hurst parameter (Hurst, 1951) or using fractional differencing parameter (Granger & Joyeux, 1980) where they can be related by the equation \( d = H - 0.5 \). In this study, we have selected the fractional differencing \( d \) with the following interpretation:

\[ \begin{align*}
    d &= \begin{cases} 
    (-0.5, 0), & \text{antipersistent} \\
    (0, 0.5), & \text{stationary long memory} \\
    (0.5, 1), & \text{nonstationary} \\
    1, & \text{nonstationary with infinite variance}
    \end{cases} 
\] (10)

The series generating process is stationary and invertible only when \( |d| < 0.5 \). For the estimation, the modified rescaled adjusted range (modified R/S) (Lo, 1991) approach had been used to overcome the sensitivity of short range dependence issue in the classical R/S. For a given set of volatility proxy \( \{\sigma_{MPV,1}, \sigma_{MPV,2}, ..., \sigma_{MPV,t}\} \) and let \( \sigma_{MPV}^2 = \frac{1}{t} \sum_{t} \sigma_{MPV,t}^2 \), where \( \tau \) is the studied time span, the modified R/S statistic is given by

\[ \left( \frac{R}{S} \right)_\tau = \frac{1}{\sigma_q} \left[ \max_{1 \leq s \leq t} \sum_{i=1}^{t} (\sigma_{MPV,t}^2 - \sigma_{MPV}^2) - \min_{1 \leq s \leq t} \sum_{i=1}^{t} (\sigma_{MPV,t}^2 - \sigma_{MPV}^2) \right] \] (11)

with

\[ \sigma_k^2 = \left( \frac{1}{t} \sum_{t=1}^{t} (\sigma_{MPV,t}^2 - \sigma_{MPV}^2)^2 + \frac{2}{t} \sum_{j=1}^{t} \omega_j(q) \left[ \sum_{i=t+1}^{t} (\sigma_{MPV,t}^2 - \sigma_{MPV}^2)(\sigma_{MPV,t-i}^2 - \sigma_{MPV}^2) \right] \right)^{1/2} \]

where the superscript \( k \) is the subdivided blocks, the weights (Newey and West, 1987) \( \omega_j(q) \) is equal to \( 1 - \frac{j}{q+1} \).

\( \sigma^2 \) is the sample variance and \( \tilde{\gamma}_j \) is the autocovariance of bi-power variation volatility respectively. It is obvious that if \( q = 0 \), the R/S become the normal classical definition. The estimation of fractional differencing parameter can be summarized by the following procedures:

**Step 1:** Fix the time windows for two years (512 observations) each and re-estimate after \( k \)-step (for this study the \( k \) is selected as ten day or two week). The \( d \) is computed for the first time window and the sample is rolled forward \( k \)-step by eliminating the first \( k \) observations and adding the next \( k \) observations.

**Step 2:** Determine the modified R/S statistics for subdivided \( k \) blocks aggregated series.

**Step 3:** Regress the \( \left( \frac{R}{S} \right)_\tau^k \) versus their respective blocks to obtain the \( d \) from the relationship \( E \left( \frac{R}{S} \right) \sim \text{CT} d + 0.5 \).

**Step 4:** Re-compute the next time window \( d \) by moving \( k \) steps ahead until the data exhausted.
3. Empirical Study

Figure 3 presents the plots for four types of multiple power variation of volatility estimators. It is found that both \( RV \) and \( BV \) indicated similar pattern however the value of \( BV \) is almost half of the magnitude of \( RV \). The averaging processes in \( BV \) have smoothed the fluctuation of volatility as compared to \( RV \). Similar to \( BV \), \( TV \) and \( QV \) on the other hand, have shown significantly reduction in the magnitude as compared to \( RV \). In short, the average smoothing process to a higher power has greater impact to the magnitude of volatility fluctuations.

![Figure 3. Volatility estimators for \( RV \), \( BV \), \( TV \) and \( QV \)](image)

Another interesting property which had been received great attentions is the long memory behavior of volatility. The preliminary long memory property can be examined using sample autocorrelation function (SACF) for each volatility estimators. Figure 4 indicates the SACF for all the volatility estimators. The hyperbolic decaying for the first 200 lags of SACF for all the estimators indicates the presence of long memory. However, the \( TV \) and \( QV \) show higher intensity of long memory at the first 100 lags and follows \( RV \) and \( BV \) afterwards. In order to
observe the variation of long memory over the years, we have conducted a dynamic modified R/S estimation on each of the multiple variation estimators. Figure 5 shows that the long memory behavior is not constant over the years and it is observed that the majority long memory estimations fluctuated between 0.1 and 0.5. The presence of long memory property in financial market has important implication to redefine the efficient market hypothesis (Fama, 1998). This includes heterogeneous market hypothesis (Dacorogna et al., 2001) which based on the phenomenon of long memory. In short, the stronger the long memory (close to 0.5), the less efficient the market will be. If the volatility series follows an i.i.d. random walk, then the estimated parameter will indicate d as 0.500 which implies that the market is ideally efficient with no component of predictability.

Figure 5. Dynamic long memory using modified rescaled-range approach

Table 1. Total Annual return of S&P500 based on compound annual growth rate

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Annual Return Including Dividends</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>−37.00%</td>
</tr>
<tr>
<td>2009</td>
<td>26.46%</td>
</tr>
<tr>
<td>2010</td>
<td>15.06%</td>
</tr>
<tr>
<td>2011</td>
<td>2.11%</td>
</tr>
<tr>
<td>2012</td>
<td>16.00%</td>
</tr>
<tr>
<td>2013</td>
<td>32.39%</td>
</tr>
</tbody>
</table>

First of all, Table 1 indicates the total annual return of S&P500 based on CAGR. The analysis will only conducted on the post subprime mortgage crisis during the period of 2008 when the Lehman Brothers (Brueckner, Calem, & Nakamura, 2012) filing for Federal government bailout. Especially from year 2002 until 2007, the U.S. market has shown remarkable mortgage boom with the estimated subprime mortgages worth more than USD 1.3 trillion. It is observed that there is a huge negative (-37%) total return in year 2008, however bounded back to positive total return and dropped to 2.11% in year 2011. From year 2011 to 2013, there is an upward trend for the total return.

In this study, the whole time span is divided into two sub-periods namely the region I (February 2010 - October 2011) and region II (October 2011-May 2013) as indicated in Figure 5. Overall, the estimated d parameters are fluctuating around 0.40 and an upward drifting for region I and region II respectively. Region I can be considered as the recovery period after the subprime mortgage crisis for U.S. After the severe impact of the subprime mortgage crisis in year 2008, it has further spread among the U.S. and also to the global financial markets. Under this condition, most of the market participants become followers to the speculated negative news of subprime mortgage bubble. Due to these similar reactions of particular group of investors, the long memory has been created among them. After the crisis, the U.S government has implemented various immediate policies and regulations to ease the severe damages by the mortgage crisis. These include lower down the Federal funds rate, improve the bank liquidities, economic stimulus packages and bailouts or merged of banks. Thus in region I, we can observe a consistent cyclical movements of differencing parameter (d values fluctuates around 0.40) from 2010 until the October of year 2011. This might be the adjustment period for the U.S. financial stock markets after the severe impact of subprime mortgage crisis.

Before entering region II which is before October year 2011, the S&P500 recorded not very impressive total return over the first 3 quarters of year 2011 with negative total returns. However in the 2011Q4, the S&P500 total return indicated gains of 11.82% (Taylor Woods, 2012) bringing the total year’s return into positive territory at 2.11%. Thus, in Figure 5, the long memory indicated a huge drop during this transition period from region I to region II. This tendency implies that the U.S. market become more informationally efficient with weaker long memory in region II at the beginning. This is in line with the nature of U.S. market which had been categorized as a mature market. Therefore it is more efficient in updating the in flowing information among the investors. After that we saw a upward trend for the d parameter which indicated the market become less efficient with stronger long memory until year 2013 where the U.S. market become less influence by the impact of credit crisis. This is also shown in Table 1 where the S&P500 recorded gains in total return until year 2013. In other words, the investors have restructured to the common group based on their investment opportunities time window under the normal market condition. Overall, the time varying long memory estimated parameter allows us to observe the movement of long memory or in term of informationally market efficiency after the subprime mortgage crisis.

As a conclusion, the various types of volatility representations provide alternative comparison regarding the estimated time-varying long memory parameters. The rescaled-range method with the refinement using multipower variation volatility representations can benefit both academic and industrial sectors with better information for decision making and future research enhancement in risk management analysis.

4. Discussion

The usage of high frequency data in financial market analysis become crucial due to its accuracy in estimation which can be used in finance application such as forecasting, risk analysis and portfolio management. This study includes these data with the representation of multipower variation volatility estimators. The empirical findings show that the intensity of estimated volatility reduced with the higher power variations. In other words, the estimators become less noisy when the power of variations increase from realized volatility to quadpower variation volatility. We also found that the long memory measurement is not constant over the selected studied period. This information allow us to understand better the dynamic long memory behavior and it is useful to provide better model specification, forecasting and finance application analysis.

References


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).