
Chin Wen Cheong
Faculty of Management
Multimedia University
63100 Cyberjaya, Selangor, Malaysia

Zaidi Isa
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600 Bangi, Selangor, Malaysia

Tan Pei Pei
Faculty of Economics & Administration
University of Malaya
50603 Kuala Lumpur, Malaysia

Lee Min Cherng
Lee Kong Chian Faculty of Engineering and Science
Universiti Tunku Abdul Rahman
43300, Kajang, Selangor, Malaysia

Copyright © 2015 Chin Wen Cheong et al. This article is distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
This study evaluates the time-varying long range dependence behaviors of the S&P 500 volatility, index using the modified rescaled adjusted range (R/S) statistic which takes into account the possible covariances of the lags in the data.
Instead of a single estimation for the whole time span, the variations of the long range dependence are computed using the moving window rolling estimates approach. For a better computational result, a high frequency data set under the representation of bipower variation realized volatility is used to avoid possible abrupt jump. As for the empirical study, we have selected the period covered before and after the subprime mortgage crisis. The empirical results show that the long range dependence is influenced by the related economic events across the studied period. This time varying long range dependence analysis allow us to understand the informationally market efficiency before and after the subprime mortgage crisis.

**Keywords:** long range dependence, rescaled adjusted range, realized volatility

### 1. Introduction

Long range dependence (LRD) behavior of financial time series has been intensively studied [1, 2, 3] in some recent literature of finance. In the recent studies, this phenomenon has been observed [4, 5, 6] in the global stock, foreign exchange and commodity markets. The presence of LRD has important implication to redefine the efficient market hypothesis [7]. This includes the heterogeneous market hypothesis [8] which is based on the phenomenon of LRD. Besides this, the applications in finance such as risk management and portfolio investment [9] are also depending on the accuracy of this additional stylized fact in the model specification, estimation and forecasting.

This study aims to overcome the drawbacks of several issues in the LRD analysis in financial time series. These include the single determination of LRD for the whole time span, selection of high frequency volatility representations and the spurious LRD due to abrupt jump in the data. The *first issue* can be overcome by using a rolling estimation [3, 10] of LRD for a fix time-window. Based on the fitting window (says 512 or 1024 data points), the LRD is re-estimated and slide from the first till until the last data point. In this way, the LRD behavior can be observed for the studied time span. For the *second* issue, it is proven that high frequency volatility proxy is a better estimate [11] than daily closed data such as absolute and square returns for the unobservable latent volatility. The high frequency volatility is better known as the realized volatility which is commonly used in a 5-minute interval. *Lastly*, the spurious LRD volatility [12] can be triggered by the abrupt jumps in the time series. Thus, the selected volatility proxy must be robust to the abrupt jumps which may cause by the economic event such as the global credit crisis.

In this study, a rolling estimate of LRD volatility is based on the modified rescaled adjusted range method [13]. The bipower variation [14] volatility proxy is computed from the high frequency data which is designed to be robust to abrupt jump. With the combination of rolling estimation procedures and the usage of
high frequency data which is robust to jump, the LRD behavior can be better understand by the market practitioners, investors, policy-makers and academicians. The remaining of this study is organized as follows: Section 2 provides the high frequency data description; Section 3 describes the procedures of LRD estimation; Section 4 discusses the empirical results and finally, Section 5 concludes the findings of the study.

2. High frequency data

This study selects the mature market of United States S&P 500 index which consists of the top 500 active large-cap stocks that represent the U.S. equity market. The S&P 500 is a free-float capitalization-weighted index introduced in year 1957 which traded under the NYSE Euronext and NASDAQ OMX. With the vast development of digital storage database in financial market, it has encouraged the extensive utilization of high frequency data in financial analysis. The concept of high frequency volatility can also be referred to realized volatility. Let consider a continuous natural logarithmic price \( \ln P(t) \) follows a diffusion process as \( d\ln P(t) = \mu(t)dt + \sigma(t)dW(t) \), \( t \geq 0 \), where \( \mu(t) \), \( \sigma(t) \) and \( W(t) \) represent the instantaneous drift, volatility and a standard Brownian motion. In line with discrete-time return, \( R_t = \ln P(t) - \ln P(t - 1) \), the unit time interval normalized to a day. Thus, for a one-period daily return is defined as \( R_t = \ln P(t) - \ln P(t - 1) = \int_{t-1}^{t} \mu(s) ds + \int_{t-1}^{t} \sigma(s) dW(s) \). Under the condition of large sample size, the \( R_t \) is normally distributed as \( R_t \sim N\left( \int_{t-1}^{t} \mu(s) ds, IV_t \right) \), where \( IV_t = \int_{t-1}^{t} \sigma^2(s) ds \) represents the integrated variance or the underlying continuous-time stochastic volatility. In common practice, the \( IV \) is not observable, however it can be easily computed from high frequency data. According to Andersen and Bollerslev [11], the realized volatility (RV) on day \( t \) based on the \( N \) equally-spaced intraday return \((N \equiv 1/\Delta) \) is defined as \( RV_t(\Delta) \equiv \sum_{j=1}^{N} R_{t,j}^2 \), where the discrete intraday percentage continuously compounded return is defined as \( R_{t,j} = 100(\ln P_{t,j} - \ln P_{t,j-1}) \), where \( j = 1, \ldots, N \) and \( t = 1, \ldots, T \). In other words, each day \( t \) consists of \( N \) recorded trading activities. When \( \Delta \) approaches zero with the absence of jumps, \( RV_t(\Delta) \to \int_{t-1}^{t} \sigma^2(s) ds \). In other words, under these conditions the RV perfectly measures the IV. For this study, the durations for trading hours are from 9:30 to 16:00 (S&P 500) with \( N_{S&P 500} = 390 \). These time series are assumed to have \( \mathbb{E}[R_{t,q}] = 0 \) and finite \( \mathbb{E}_{\forall r,s,p,q}[R_{r,p}^2 R_{s,q}^2] \). For model-free proxy of volatility or often known as realized volatility, the daily squared compounded returns are

\[
R_t^2 = \left[ \sum_{a=1}^{N} R_{t,a} \right]^2 = \sum_{a=1}^{N} R_{t,a}^2 + 2 \sum_{a=1}^{N} \sum_{b=1}^{N} R_{t,a} R_{t,b} (a \neq b)
\]

The second term indicates the autocovariance which acts as the noise components in the realized volatility. If they are not correlated, then \( \mathbb{E}[\hat{R}_{t,a}^2] = \sum_{a=1}^{N} R_{t,a}^2 = \sigma_{t,actual}^2 \). This is defined as the unbiased estimator [11] of the daily
population variance or latent volatility. Unfortunately, the RV often encounters persistence issue when abrupt jumps occurred and according to Andersen et al. [11], this estimator is no longer consistent in the presence of jumps. In order to overcome this drawback, the bipower variation estimator is selected which introduced by Barndorff-Nielsen and Shephard [14] with the following definition:

$$\sigma_{BIP,t}^2(\Delta) = \frac{\pi}{2} \sum_{j=2}^{M} |R_{t,j}||R_{t,j-1}|$$

where $M$ is the total interval within a day. In order to account the possible jumps, the realized bi-power variation is constructed using the product of the two consecutive intraday returns. Under the assumption of Brownian semi-martingale process, if there is a jump in one of the returns, the product for the consecutive return has a small impact to the realized bi-power variation. For most of the studies, the 5-minute realized volatility are used in order to lessen the impact of market noise [11] whereas higher interval such as 15-minute data are recommended to reduce the estimation biasness issue [11].

3. Methodology

The long range dependence volatility can be measured either by Hurst [15] or using the fractional differencing parameter introduced by Granger and Joyeux [16], where they can be linked by the equation $d = H - 0.5$. In this study, we have selected the fractional differencing $d$ with the following interpretations:

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

The series generating process is stationary and invertible only when $|d| < 0.5$. For the estimation, the modified rescaled adjusted range (modified R/S) approach [13] had been used to overcome the sensitivity of short range dependence issue in the classical R/S. For a given set of volatility proxy $\{\sigma_{BIP,1}^2, \sigma_{BIP,2}^2, ..., \sigma_{BIP,t}^2\}$ and let $\tilde{\sigma}_{BIP}^2 = \frac{1}{\tau} \sum_{t} \sigma_{BIP,t}^2$, where $\tau$ is the studied time span, the modified R/S statistic is given by

$$\left(\frac{R_S}{\tilde{\sigma}}\right)_\tau^k = \frac{1}{\sigma_q} \left[ \max_{1 \leq t \leq \tau} \sum_{t=1}^{\tau} (\sigma_{BIP,t}^2 - \tilde{\sigma}_{BIP}^2) - \min_{1 \leq t \leq \tau} \sum_{t=1}^{\tau} (\sigma_{BIP,t}^2 - \tilde{\sigma}_{BIP}^2) \right]$$

with

$$\sigma_q^k = \left( \frac{1}{\tau} \sum_{t=1}^{\tau} (\sigma_{BIP,t}^2 - \tilde{\sigma}_{BIP}^2) \right)^2 + 2 \sum_{t=1}^{\tau} \omega_j(q) \left[ \sum_{t=1}^{\tau} (\sigma_{BIP,t}^2 - \tilde{\sigma}_{BIP}^2) (\tilde{\sigma}_{BIP,t-i} - \tilde{\sigma}_{BIP}^2) \right]^{1/2} = \sqrt{\sigma^2 + 2 \sum_{j=1}^{q} \omega_j(q) \tilde{\gamma}_j}$$

where the superscript $k$
Computation of high frequency S&P 500 LRD volatility

is the subdivided blocks, the weight \( \omega_j(q) \) \([17]\) is equal to \( 1 - \frac{j}{q+1} \). \( \sigma^2 \) is the sample variance and \( \hat{\gamma}_j \) 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 days or two weeks). 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 \( \left( \frac{R}{S} \right)^k \) versus their respective blocks to obtain the \( d \) from the relationship \( E \left( \frac{R}{S} \right) \sim cT^{d+0.5} \).

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

**4. Empirical Results in U.S. Subprime Mortgage Crisis**

A quick review of long range dependence of the S&P 500 can be obtained from its sample autocorrelation (SACF). Figure 1 presents the overall data set SACF for the first 200 lags for the high frequency bi-power variation volatility \( (\sigma_{BIP,t}^2) \) and the daily return square \( (\sigma_{sqr,t}^2) \) proxy, which served as the comparison also is included in the plot. For the first 100 lags, the ACF for \( \sigma_{BIP,t}^2 \) is consistently higher than \( \sigma_{SQR,t}^2 \). This implies that the volatility proxy of \( \sigma_{BIP,t}^2 \) is more persistence than \( \sigma_{SQR,t}^2 \). The higher intense information (390 observations) within a day for the high frequency data has resulted in higher serial correlation as compared to the square return which based on one data per day. Higher volatility persistency implies higher predictability which against the concept of efficient market hypothesis. In other words, the absence of LRD volatility supports the presence of efficient market whereas the existence of LRD against it.
The procedures of rolling window modified R/S statistics are illustrated as follows:

**Step 1**: Fix the time windows for two years (512 observations).

**Step 2**: Subdivide \(k\) blocks aggregated series starts from \(2^2\) until \(2^8\).

**Step 3**: Regress \(\left(\frac{R}{S}\right)_\tau^k\) versus their respective blocks based on the equation:

\[
LOG\left[E\left(\frac{R}{S}\right)\right] = LOG(c) + (d + 0.5)LOG(k)
\]  

(5)

<table>
<thead>
<tr>
<th>Regression</th>
<th>Coefficients</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.498737</td>
<td>0.015959</td>
</tr>
<tr>
<td>Slope ((H))</td>
<td>0.968627</td>
<td>8.66E-06</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.9946</td>
<td></td>
</tr>
<tr>
<td>(d=H-0.5)</td>
<td><strong>0.468627</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Step 4**: Re-compute the next rolling window \(d\) by moving 10 steps ahead.
Table 1: Regression Analysis for d Parameter at Second Iteration.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.41153</td>
</tr>
<tr>
<td>Slope (H)</td>
<td>0.843509</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9897</td>
</tr>
<tr>
<td>$d=H-0.5$</td>
<td><strong>0.343509</strong></td>
</tr>
</tbody>
</table>

Step 5: By repeating Step 2 to Step 4, the time-varying $d$ can be obtained as follows:

Figure 3: The Rolling Estimates of $d$ Parameter.

For the volatility long range dependence discussion, the whole time span is divided into three sub-periods namely Period I, Period II and Period III as indicated in Figure 3. Overall, the estimated $d$ parameters are fluctuating around the downward, upward and downward drifting from Period I, II and III respectively. The tendency of upward or downward drifting has close relationship with the related economic events that influence the U.S. S&P 500 markets. For Period I, the most relevant one of the industries that contributed to the U.S. economy is the mortgage industry. 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. The values of $d$ have fluctuated from the range 0.200 to 0.600 with a downward drift before the mortgage bubble burst [18] in September 2008 when the Lehman Brothers filing for Federal government bailout. U.S. market can be considered as a mature market, therefore it is more efficient in updating the new information in the price as illustrated in Period I.
After the collapse of Lehman Brothers [19], the value $d$ has a tendency to increase from 0.300 until 0.500 and ended at the middle of year 2011. This tendency implies that the U.S. market become less informationally efficient with stronger long range dependence in their fluctuations of returns. The severe impact of the subprime mortgage crisis has further spread among the U.S. and also the global investors. 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 investors, the long range dependence becomes stronger during the Period II. After that the U.S. government has implemented various 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 Period III, we can observe a consistent cyclical movements of differencing parameter ($d$ values fluctuates around 0.41 to 0.43) after September of 2008 until the end of year 2011. This might be the adjustment period for the U.S. financial stock markets after the severe impact of subprime mortgage crisis and the collapse of Lehman Brothers.

5. Conclusion

Efficient market hypothesis postulates that there is no way to gain abnormal profit; and all the information should be reflected in the stock price and no speculations for return volatility are possible too. However this study found that there is long range dependence in the stock volatility by using S&P 500 index. This study uses the rolling window modified R/S approach to compute the time-varying long range dependence volatility of S&P500 index. Under the analysis procedures, we are able to observe the time-varying long range dependence volatility of S&P 500 across the study period. To observe the fluctuations of long range dependence volatility before and after the U.S. subprime mortgage crisis, we have selected the duration of study from year 2007 until year 2013. The results reveal a hidden pattern in the long range dependence behavior of the stock market, where small cycle was observed during the study period. The market becomes more efficient during the subprime crisis and the pattern change direction on the onset of Lehman Brothers case in 2008. With government intervention after 2011, the long range dependence level dropped to a lower level. Taken together, volatility is generally predictable however the effect is lessening during the crisis or critical period.

References


Received: August 24, 2015; Published: September 25, 2015