Long Memory and Structural Breaks in Extreme Value Estimators
- The Case of the U.S. Stock Indexes*-

극단값 변동성 추정치의 장기기억과 구조변화
- 미국의 주요 주가지수의 경우 -

Kwon Yong-Jae**
권 용 재

With numerous studies reporting long memory in financial volatilities, long memory became one of the stylized facts of volatility time series. Several researchers, however, including Granger and Hyung (2004) and Choi and Zivot (2007), argue that the long memory property of financial volatilities may be amplified by occasional structural breaks. This paper investigates the validity of the previous studies - whether long memory in extreme value estimators is overstated by structural breaks. I find an evidence that the degree of long memory in the extreme value estimators is inflated by structural breaks. I also find, however, that significant long memory is still discovered in the extreme value estimators even after the multiple breaks are controlled in the estimation.

Key words: Extreme Value Estimators, Long Memory, Structural Breaks, Volatility
학술진흥재단 분류 연구분야 코드: B030603, B050704

* I thank two anonymous referees for their thoughtful suggestions.
** 보험연구원 부연구위원(kwonyon3@kiri.or.kr)
논문투고일: 2010. 02. 17, 논문 최종 수정일: 2010. 03. 19, 논문 게재 확정일: 2010. 03. 26

I. Introduction

One stylized characteristic of financial volatility is the strong autocorrelation nature of it, which is referred to as long memory. The strong persistence, also known as long memory, in financial volatilities has been studied by many researchers in the fields of macroeconomics and empirical finance. This property has been commonly utilized for volatility modeling and forecasting.

Past studies have found that significant long memory is present in absolute returns (Ding et al., 1993; Granger and Ding, 1995), squared returns (Lobato and Savin, 1998), conditional volatility (Baillie et al., 1996), and realized volatility based on 5-minute intraday returns (Ebens, 1999; Andersen, Bollerslev, Diebold, and Ebens (ABDE hereafter), 2001; Andersen, Bollerslev, Diebold, and Labys (ABDL hereafter), 2001; Areal and Taylor, 2002; ABDL, 2003).

This paper deals with long memory in extreme value estimators. The extreme value estimators are historical volatility measures constructed with daily high, low, opening, and closing prices, which are reported by financial media everyday. Although studies on persistence of the extreme value estimators do not get much attention compared with those on the persistence of other types of estimators, there still are a number of studies reporting the long memory property in the extreme value estimators from various financial markets.

Taylor (1987) is the first to report the autocorrelation structure of daily price range. He finds that the logarithmic daily price range is highly autocorrelated compared to close-to-close absolute returns. Using Deutsche Mark-Dollar futures prices from 1977 to 1983, he shows that there is a high degree of autocorrelation in the daily range and that this information is helpful in forecasting currency volatility. Byers and Peel (2001) show that volatility series constructed from daily ratios of high and low prices for sterling exchange rates, S&P futures prices, the FT30 and gold prices exhibit the features of long memory processes. They used Lo (1991)'s modified R/S statistic and the ARFIMA model. By analyzing daily stock, currency and commodity futures prices, Chen et al. (2006) find that the Garman-Klass estimator along with the realized volatility estimator, based on 5-minute intraday returns, is significantly more autocorrelated than close-to-close squared and absolute returns. In addition, they show that the Garman-Klass volatility estimator and the realized volatility estimator are highly correlated with each other, and also that their degrees of long memory are not statistically different.

However, a number of recent research studies recently question the existence of long memory in financial volatilities. Granger and Hyung (2004) and Choi and Zivot (2007), for example, challenge the past studies, arguing that a part of long memory in financial time series may be due to the presence of occasional mean breaks. Long memory and structural breaks have been rarely studied on the same ground as Choi and Zivot (2007) mention, but it is necessary to distinguish structural breaks from long memory property because we have a higher chance to improve modeling and forecasting time series when we obtain a deeper understanding of the dynamics in financial variables. Although the same question can be asked to the extreme value estimators, there has been no research on these volatility estimators yet.

To fill the gap in the line of research, I estimate long memory and structural breaks in the extreme value estimators. The attempt made in this paper will be helpful in answering several questions regarding which types of econometric models should be employed to forecast future level of volatility, and whether long memory in the extreme value estimators is spurious.

The remainder of this paper is organized as follows. Section 2 briefly reviews studies on extreme value estimators and long memory of financial time series. Section 3 describes data and methodology used in this paper. Statistical results are explained in section 4. Finally, I conclude this paper in section 5.

II. Extreme Value Estimators and Long Memory of Financial Time Series

1. Extreme Value Estimators

Volatility is defined as a measure of price variability during a given period of time. According to this definition, volatility is a latent variable in the sense that we cannot observe it directly but only measure it. In practice, therefore, researchers have quantified it using proxy variables such as daily, weekly, or monthly returns. For that reason, both the size and the statistical properties of financial volatilities highly depend on how the measures are estimated. Volatilities of financial markets traditionally have been estimated by the method of moments, using daily, weekly or monthly close-to-close returns. These measures have been widely accepted as ex-post realized volatility in event studies or evaluation of forecasts. They are also employed for the ARCH-type models, stochastic volatility models, and the J.P. Morgan's RiskMetrics™ for modeling and forecasting volatility of financial assets.

A different approach employs the highest and lowest prices observed during a given time period. The volatility measures constructed with a price range are generally called extreme value estimators. As Wiggins (1991) points out, extreme value estimators are intuitively superior to the close-to-close estimator since they incorporate the range of prices observed over the entire day to the volatility estimators, while the traditional estimator is merely a "snapshot" price at the end of the day. As long as the assumption that trading is continuous and always monitored holds, these extreme value estimators are multiple times more efficient than the close-to-close estimator.

Parkinson (1980) is the first to suggest incorporating the price range to volatility estimators. Using a distribution originally derived by Feller (1951), he develops a volatility estimator based on the highest and lowest prices during a given time period,

$$\sigma^2_{P,i} = \frac{1}{4\ln(2)}\left(\ln\left(\frac{H_i}{L_i}\right)\right)^2 \tag{1}$$

where H_i and L_i stand for the highest and lowest prices respectively. According to Parkinson (1980), this estimator is at least 2.5 times more efficient than the traditional close-to-close estimator. After Parkinson (1980) first introduced this range-based volatility estimator, his idea has been modified by several researchers.

Garman and Klass (1980) argue that the Parkinson estimator does not utilize all the available market information by ignoring opening and closing prices in the estimator and suggest a new volatility estimator extending the Parkinson's approach. Essentially, this estimator is a weighted average of the Parkinson's estimator and the traditional close-to-close estimator.

$$\sigma^2_{GK,i} = 0.511\left(\ln\left(\frac{H_i}{L_i}\right)\right)^2 - 0.019\left(\ln\left(\frac{C_i}{O_i}\right)\left(\ln\left(\frac{H_i}{O_i}\right) + \ln\left(\frac{L_i}{O_i}\right)\right)\right.$$

$$-2\ln(\frac{H_i}{O_i})\ln(\frac{L_i}{O_i})\Big)-0.383\Big(\ln(\frac{C_i}{O_i})^2\Big) \quad (2)$$

where O_i and C_i is opening and closing prices, respectively. Using a simulation study, they show that the Parkinson and Garman-Klass estimators might be downward biased when price paths are not continuously monitored.

These two estimators are built under the assumption that the expected returns are equal to zero. In other words, Parkinson (1980) and Garman and Klass (1980) both assume that logarithmic price paths follow the Brownian motion with no drift. However, the price of a stock or a currency often has a trend or a trendy nature and this phenomenon is commonly seen in strong bull and bear markets[1). In this case, the Parkinson and Garman-Klass estimators tend to overestimate volatility. To cope with this potential problem, Rogers and Satchell (1991) suggest a new extreme value estimator independent of the drift term:

$$\sigma^2_{RS,i}=\ln(\frac{H_i}{C_i})\ln(\frac{H_i}{O_i})+\ln(\frac{L_i}{C_i})\ln(\frac{L_i}{O_i}) \quad (3)$$

In the Rogers-Satchell estimator, a volatility estimate is zero at a period when the opening price equals the lowest price and the closing price equals the highest price. Similarly, a volatility estimate also equals zero at a period when the opening price equals the highest price and the closing price equals the lowest price. This is the consequence from the notion that monotonic price movement during a market period is explained mostly by the nonzero

1) As Garman and Klass (1980) and Yang and Zhang (2000) mention, this trendy nature mostly matters in weekly and monthly time series.

drift in a price path.

Finally, Yang and Zhang (2000) propose a multi-period (n〉1) extreme value estimator that is independent of drift and consistent in the presence of opening jumps.

$$\sigma^2_{YZ,n} = \frac{1}{n-1}\sum_{i=1}^{n}\left[\left(\ln\left(\frac{O_i}{C_{i-1}}\right)\right)-\left(\frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{O_i}{C_{i-1}}\right)\right)\right)\right]^2$$
$$+\mathrm{k}\frac{1}{n-1}\sum_{i=1}^{n}\left[\left(\ln\left(\frac{C_i}{C_{i-1}}\right)\right)-\left(\frac{1}{n}\sum_{i=1}^{n}\left(\ln\left(\frac{C_i}{C_{i-1}}\right)\right)\right)\right]^2+(1-k)\sigma^2_{RS,n} \quad n>1 \qquad (4)$$

To minimize the variance, the constant k is set to be

$$k = \frac{0.34}{1.34+\frac{n+1}{n-1}}, \quad n>1$$

The Yang-Zhang estimator basically combines the overnight variance with the weighted average of the close-to-close variance and the Rogers-Satchell estimator.

As mentioned earlier, the efficiency of these extreme value estimators depends on the assumption that trading is continuously monitored. No drift assumption is also critical to the performance of the volatility estimators proposed by Parkinson (1980) and Garman and Klass (1980). If any of these assumptions is violated, the efficiencies of the estimators might be sub-optimal. Several studies evaluate their efficiency through simulation experiments, in-sample forecast evaluation, or the correlation with ex-post realized volatility. Early studies favor the extreme value estimators. However, several research studies find that the performances of the estimators vary when the assumptions are violated.

Wiggins (1991) conducts the first comparative analysis to examine the efficiency of extreme value estimators. Using the data from the U.S. individual stocks from 1970 to 1986, he shows that the Parkinson estimator is sensitive to recording errors because daily high and low prices tend to be more vulnerable to recording errors than closing prices. The efficiency of the estimator becomes significantly higher than that of the traditional estimator after the recording errors are removed. Wiggins(1991) also finds that stock prices and trading volume of individual stocks, which are proxies of the degree of continuous price monitoring, are positively related to the efficiency of the Parkinson estimator.

However, Wiggins (1992) shows that the Parkinson and Garman-Klass estimators are less biased and are more efficient than the traditional close-to-close estimator for the S&P 500 futures prices. He also shows that these extreme value estimators have richer information on future level of volatility than close-to-close estimators. Considering that the S&P 500 futures and highly-priced common stocks are very efficient in the sense that trading is continuous and always monitored, the two studies from Wiggins (1991, 1992) confirm the argument that the efficiency of extreme value estimators is dependent on the efficiency of the financial market, represented by a high level of trading volume and price.

Rogers et al. (1994) compare the performances of several extreme value estimators with those of the daily squared return measure using data from several British stocks. They find that error variances estimated with the extreme value estimators are significantly smaller than those estimated with the traditional variance estimator. In addition, they report that when there is a large drift in the price path, the Rogers-Satchell estimator is superior to the Parkinson and Garman-Klass estimators as well as the close-to-close estimator.

Using the S&P 500 index futures and several exchange rates data, Bali and Weinbaum (2005) find that extreme value estimators are less biased than the traditional close-to-close estimator for daily volatility and that the Parkinson estimator is the best among them across the dataset. To test efficiency of extreme value estimators, they use 5-minute realized volatility as benchmark volatility, recommended by Andersen and Bollerslev (1998). However, they argue that the extreme value estimators are not superior to the traditional close-to-close estimator in the cases of weekly and monthly volatility measurements.

Finally, Shu and Zhang (2006) investigate the relative performance of four extreme value estimators[2] with a Monte Carlo simulation experiment and a comparison with the realized volatility estimator. They find significant differences among the performances of the extreme value estimators when the asset price path involves a large drift or underlying volatility dynamics are different. For example, the Rogers-Satchell and Yang-Zhang estimators are not seriously affected by the size of the drift term while the Parkinson and Garman-Klass estimators tend to overestimate volatility. Another finding is that, under various underlying volatility settings[3], the Parkinson and Yang-Zhang estimators are less negatively biased than the Garman-Klass and Rogers-Satchell estimators although all four extreme value estimators have negative biases. It is interesting that the Parkinson estimator works well in terms of bias in spite of its simplicity. However, the Parkinson estimator performs poorly compared to the others regarding the efficiency criterion. In the end, Shu and Zhang (2006) show that the Yang-Zhang estimator is almost equivalent to that of 15-minute daily realized

2) Bali and Weinbaum (2005) and Shu and Zhang (2006) both study the Parkinson, Garman-Klass, Rogers-Satchell, and Yang-Zhang estimators.

3) Constant volatility, deterministic volatility, and jump volatility models are used for their experiment.

variance using the tick data of the S&P 500 cash index.

2. Definition of long memory

When economists handle financial variables such as prices and GDP, they commonly difference them. However, it appears that an integer difference is too much for some financial time series. This is the case when the autocorrelation function of the time series decays ultimately to zero but the decaying speed is too low, that values from long ago seem to significantly affect current and future values. Apparently, neither of ARMA nor ARIMA models is appropriate to describe this feature. At a stationary ARMA model, a large number of parameters are required in the rare persistence. ARIMA model is not proper either because differencing the data seems to go too far. To analyze this type of data in a parsimonious approach, alternative stationary processes, of which autocorrelations decay to zero at a slow rate, are needed. To bridge this gap, the long memory processes are introduced to macroeconomics and finance since the autocorrelation coefficients of the processes decline slowly at a hyperbolic rate.

There are several statistical definitions of the long memory property. In the time domain, a covariance stationary time series process $\{X_t\}$ is said to have long memory if

$$\sum_{k=-\infty}^{\infty} |\rho(k)| \qquad (5)$$

is non-finite, where $\rho(k)$ is the autocorrelation at lag k. Intuitively, the infinite sum of the autocorrelations implies that autocorrelations at long lags are non-negligible. Satisfying this condition, the long memory process has autocorrelation coefficients that behave like

$$\rho(k) \sim C_\rho k^{2d-1} \quad \text{as} \quad k \rightarrow \infty \tag{6}$$

where C_ρ is a constant and $d \in (0,0.5)$[4].

Long memory also can be defined in the frequency domain. A time series process has stationary long memory if the spectral density behaves like

$$f(\omega) \sim C_f \omega^{-2d} \quad \text{as} \quad \omega \rightarrow 0 \tag{7}$$

where C_f is a constant and $d \in (0,0.5)$. For a long memory process, the spectral density goes to infinity at the origin.

Granger and Joyeux (1980) and Hosking (1981) independently show that a long memory process can be modeled by employing a fractionally integrated process. In essence, it is a natural extension of the ARIMA models. A time series process $\{X_t\}$ follows an autoregressive fractionally integrated moving average process of order (p,d,q) where d is the fractional differencing parameter, denoted by ARFIMA (p,d,q), if

$$\phi(L)(1-L)^d(X_t-\mu) = \theta(L)\varepsilon_t \tag{8}$$

where $\phi(z)$ is autoregressive polynomial and $\theta(z)$ is moving average polynomial with roots outside the unit circle and ε_t is white noise. The fractional differencing operator, $(1-L)^d$, is defined by

4) To see how (7) is derived, see Granger and Joyeux (1980).

$$(1-L)^d = 1 + dL + \frac{1}{2}d(1+d)L^2 + \frac{1}{3}d(1+d)(2+d)L^3 + \cdots\cdots$$

$$= \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} L^k \tag{9}$$

where $\Gamma(\cdot)$ denotes the Gamma function. For $0<d<0.5$, the process is stationary and has long memory. If $0.5 \leqq d < 1$, X_t is not a stationary series but still mean-reverting. And when $-0.5<d<0$, the process is stationary and has short memory. When d equals zero, it simply becomes a short memory ARMA process. The parameter d captures long-run dynamics of time series while the autoregressive and moving average parameters describe the short-run behavior.

3. Long memory versus Structural breaks

Strong dependence in financial time series has been considered a helpful property for modeling and prediction. However, there have been studies arguing that persistent time series may not be strongly dependent, but the strong dependence is solely caused or overstated by occasional breaks in the mean or the slope of a stationary time series.

This question was first raised by researchers examining macroeconomic time series. Modifying the augmented Dickey-Fuller test, Perron (1989) finds that economic time series, of which the null hypothesis of nonstationarity is accepted in Nelson and Plosser (1982)'s study, are actually stationary with a trend after controlling for notable structural breaks such as the Great Crash and the Oil shock. Generalizing Perron (1989)'s approach, Zivot and Andrews (1992) endogenously find the most plausible break points in the series, and report weaker but robust results against nonstationarity null hypothesis. In the same spirit, Lastrapes (1989) and Lamoureux and Lastrapes (1990) show that strong dependence of the GARCH model might be an artifact of structural shifts by changes in economic policy, using the data from the U.S. interest rates and common stocks respectively.

Recently, several researchers argue that the long memory property can also be amplified by structural breaks in time series. For example, Granger and Hyung (2004) argue that long memory can be confused with occasional structural breaks or part of long memory may be caused by neglected breaks. They find that the degree of long memory tends to be higher when the structural breaks occur more often or when the size of the breaks are larger. Conversely, they show that spurious breaks are detected more often when the degree of long memory in time series gets higher. They find that the occasional break model fits the S&P 500 absolute returns marginally better than the I(d) model. However, long memory forecasts are better than forecasts from the occasional break model. They suggest that it will be a more complete study if the occasional break model and the long memory model are summarized into a single model.

Inspired by Granger and Hyung (2004), Choi and Zivot (2007) examine whether forward discount rates of the G-7 countries still exhibit significant long memory controlling for the mean breaks. The mean breaks are estimated by Bai and Perron (1998, 2003)'s structural break test. They find that although the degree of long memory is reduced to a large extent after the removal, there still remains significant evidence of long memory in the forward rates. Choi and Zivot (2007) conclude that long memory and structural breaks should be simultaneously considered in order to explain the forward rate biases.

Finally, Choi et al. (2006) apply Choi and Zivot (2007)'s approach to realized volatilities of foreign exchange markets. Choi et al. (2006) report that structural breaks can explain a part of the dependence of realized volatility although significant persistence is still found even after the structural breaks are removed from the series. In light of forecasting volatility, they show that the break-adjusted simple VAR model produces superior forecasts compared to any other prediction model. However, they conclude that the long memory VAR model, originally employed by ABDL (2003), still generate competitive forecasts even though the long memory VAR model has no consideration on the mean breaks.

III. Data and Methodology

1. Data

This paper utilizes data from three major U.S. stock indexes: the S&P 500 composite index (hereafter S&P 500), the Nasdaq-100 index (hereafter NDX100), and the Dow Jones Industrial Average (hereafter DJIA). S&P 500 is an index based on the stocks of the 500 American large-capitalization companies. NDX100 is a stock market index based on 100 domestic and international non-financial securities listed on the NASDAQ. Finally, the DJIA, the oldest U.S. market index, consists of 30 of the largest companies in the United States. These indexes are the most popular indicators for demonstrating the overall performance of the U.S. economy. This study utilizes intraday quotes of the three indexes.

I obtain daily historical data of S&P 500 and NDX100 from 'Yahoo Finance'[5]. For the DJIA, however, I purchase intraday price quotes from

5) http://finance.yahoo.com

Tickdata Inc., a private vendor[6]. The reason I purchase historical tick data for the DJIA is that publicly available DJIA historical indexes are constructed differently. In short, the highest and lowest values of the DJIA during a day, reported in the dataset, are not the highest and lowest values that we encounter in financial media. Traditionally, the highest and lowest values of the DJIA are constructed with the highest and lowest values of the component stocks of the DJIA. Since the component stocks neither hits the bottom nor hits the roof simultaneously during a day, DJIA's highest and lowest values obtained from the 'Yahoo Finance' website are plagued by significant biases. Therefore, we will have extreme value estimators with upward biases inevitably when these dataset are used to construct the estimators. To avoid this potential problem, I had to purchase intraday price quotes of the DJIA from the private vendor. Because the DJIA dataset starts from April 1993, volatility series for the DJIA is shorter than those for the other two indexes.

The Parkinson and Garman-Klass estimators are calculated based on the formula introduced by Parkinson (1980) and Garman and Klass (1980), respectively. A known potential problem using these two estimators is that they can overestimate volatility when there is a non-zero drift in a price path during a day. However, Yang and Zhang (2000) argue that no drift assumption is actually a good approximation as long as the time interval of each period is short. Since intraday time series satisfy this condition according to Garman and Klass (1980) and Yang and Zhang (2000), I employ the Parkinson and Garman-Klass estimators in this paper. Thus, the Rogers-Satchell estimator is excluded in this paper. The Yang-Zhang estimator is not included in this study either, not only because it estimates

5) http://www.tickdata.com

the volatility during multiple periods only, but also because it is basically an extension of the Rogers-Satchell estimator.

Since variance (s_t^2) and standard deviation (s_t) are the traditional volatility measures, Ebens (1999) and ABDL (2001) provide detailed information on distributions of the variance and standard deviation when they analyze the realized volatility measure. Following them, I investigate the autocorrelation structure in both measures of the extreme value estimators and report the basic statistics. However, for empirical analysis, logarithmic standard deviation is only employed. According to the realized variance studies employing stock indexes and individual stocks such as Ebens (1999), ABDE (2001) and Areal and Taylor (2002) and foreign currencies such as ABDL (2001, 2003), the logarithmic standard deviation of the realized volatility is approximately normally distributed. Due to this fact, most modeling and forecasting studies with realized volatility are implemented with the logarithmic standard deviation measure. Since the extreme value estimators are a type of model-free volatility measures just like the realized volatility estimator[7], I employ the logarithmic standard deviation for estimating long memory in the volatility estimators.

Descriptive statistics of the variances, standard deviations, and logarithmic standard deviations of the two extreme value estimators are reported in Table 1. In the table, it is clear that the empirical distribution of

7) As ABDE (2001) mention, existing studies on the distributional and dynamic properties of financial volatility have been based on the estimation of the parametric ARCHs or stochastic volatility models, or the analysis of implied volatilities from option prices. The validity of these volatility measures is strictly based on the distributional assumptions that each model contains. Unlike these measures, the realized variance or the extreme value estimators are the model-free volatility estimators because they are constructed with squared intraday returns or price range over the relevant horizon.

variance is not normal in all three indexes because they are strongly skewed to the right and highly kurtotic, which is caused by several outliers. Similarly, the distribution of standard deviation is also non-normal since their estimates of skewness and kurtosis are still far from the levels of the normal distribution. Unlike the variances, however, standard deviations are less skewed and less kurtotic. This is because outliers influence less on empirical distribution when the observations have standard deviation form.

The distribution of logarithmic standard deviations is quite different from those of variance and standard deviation. In all three indexes, logarithmic standard deviation seems to be distributed normally because the estimates of skewness and kurtosis are almost the same as those of normal distribution. This finding from the extreme value estimators is consistent with Ebens (1999), ABDE (2001), Areal and Taylor (2002), and ABDL (2001,2003) reporting that the logarithmic standard deviation (or logarithmic variance) of realized volatility is approximately normally distributed. Therefore, from the empirical distribution of the extreme value estimators in the three indexes, I conclude that these estimators share the distributional features with the realized volatility estimators based on 5-minute intraday returns.

〈Table 1〉 Descriptive statistics of extreme value estimators

A. S&P 500 (January 2 1990 – December 30 2005)

	Variance		Standard Deviation		Logarithmic Standard Deviation	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
Mean	0.781716	0.686384	0.749072	0.705826	-0.450347	-0.501625
Median	0.395289	0.362716	0.628720	0.602259	-0.464069	-0.507068
Max	25.93100	28.00735	5.092249	5.292197	1.627720	1.666233
Min	0.011355	0.004670	0.106559	0.068338	-2.239059	-2.683286
Std.*	1.323293	1.175409	0.469746	0.433866	0.563249	0.547272
Skewness	7.648019	9.071431	2.236573	2.343401	0.113111	0.134141
Kurtosis	100.1456	149.3090	12.53038	14.15982	2.854935	2.954516

B. DJIA (April 2 1993 – December 30 2005)

	Variance		Standard Deviation		Logarithmic Standard Deviation	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
Mean	0.826986	0.753623	0.783380	0.755374	-0.382799	-0.409317
Median	0.460329	0.427532	0.678476	0.653860	-0.387906	-0.424863
Max	26.12758	21.11009	5.111514	4.594572	1.631496	1.524876
Min	0.021702	0.013256	0.147317	0.115134	-1.915169	-2.161662
Std.*	1.351955	1.186108	0.461917	0.427889	0.517815	0.499446
Skewness	7.513769	7.619528	2.372273	2.364508	0.190938	0.174485
Kurtosis	94.27766	96.19648	13.55711	13.81248	3.048260	3.160777

C. NDX100 (January 2 1990 – December 30 2005)

	Variance		Standard Deviation		Logarithmic Standard Deviation	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
Mean	2.579642	2.298869	1.350418	1.276279	0.139206	0.088865
Median	1.222827	1.072215	1.105815	1.035478	0.100583	0.034863
Max	133.7186	132.6911	11.56368	11.51916	2.447869	2.444012
Min	0.070008	0.048117	0.264591	0.219356	-1.329570	-1.517059
Std.*	4.705576	4.228414	0.869599	0.818625	0.553811	0.538911
Skewness	10.34218	10.84996	2.440009	2.524234	0.304466	0.385057
Kurtosis	209.8791	252.3101	14.92386	15.35336	2.856099	3.036341

Note: * Std.: Standard Deviation

2. Methodology

Before estimating the long memory parameter of the extreme value estimators, a unit root test and a stationarity test will be implemented. One is the augmented Dickey-Fuller test, which has a nonstationarity null hypothesis, and the other is a stationarity test, which has a null hypothesis of stationarity, proposed by Kwiatkowski et al. (1992). (KPSS test hereafter) If the volatility series have long memory, both tests are expected to reject their null hypotheses.

In order to assess the presence of long memory in extreme value estimators, the GPH test and Whittle's estimator will be employed. The GPH test is used by many volatility studies including Ebens (1999) and Areal and Taylor (2002). The Whittle's method is employed by Kilic (2004) analyzing stock index return and volatility series, and Choi et al. (2006) estimating long memory in the realized volatilities of foreign currencies.

Then, I consider the effect of structural breaks to long memory property in the extreme value estimators[8]. Choi and Zivot (2007), and Choi et al. (2006) find significant long memory in forward rates of G-7 countries and realized foreign exchange volatility respectively, even after eliminating multiple mean breaks. Following their approach, I first examine if logarithmic standard deviations of the extreme value estimators have structural breaks using the structural break test, proposed by Bai and Perron (1998, 2003). By controlling for the estimated mean breaks from the original series, I construct the break-eliminated volatility series. Then, I re-estimate the long memory parameters in the break-eliminated series using the long memory tests which will be described in detail below.

Finally, I analyze the stability of the long memory parameter to determine whether long memory is present in the extreme value estimators. To achieve this goal, I estimate the parameter recursively over time with the Whittle's method.

(a) KPSS test

Let y_t be a time series for which we want to test stationarity. The KPSS stationarity test is based on the following time series regression model,

$$y_t = \alpha 1 + \beta t + \gamma \sum_{i=1}^{t} \varepsilon_t + u_t \tag{10}$$

where 1 and t are deterministic components, $\{\varepsilon_t\} \sim$ iid N $(0, \sigma_\varepsilon^2)$, and u_t is a

8) Logarithmic standard deviation is chosen for the Bai-Perron structural break test because it controls for outliers effectively. For the same reason, Choi et al. (2006) use it to realized volatility study.

stationary error. To test the null hypothesis that the series is trend stationary (TS), only γ is equal to zero in the above regression. In the case of testing level stationarity, $\beta = \gamma = 0$ in (10). Finally, $\alpha = \beta = \gamma = 0$ if we intend to test pure stationarity of the series. The KPSS test statistic is given by

$$KPSS = \left(n^{-2} \sum_{t=1}^{n} \hat{S}_t^2\right)\left(\hat{\omega}^{-2}\right) \tag{11}$$

where S_t is the partial sum of the residuals of a regression under the null hypothesis and ω^2 is a consistent estimate of the long-run variance of residuals from a regression. To have further information on limiting distributions and the corresponding critical values of the KPSS statistic, see Kwiatkowski et al. (1992).

(b) GPH test

The spectral density of the fractionally integrated process is given by:

$$f_y(\omega) = \{4\sin^2(\frac{\omega}{2})^{-d} f_u(\omega) \tag{12}$$

where ω is the Fourier frequency, u_t is a stationary short memory disturbance with zero mean and $f_u(\omega)$ is the spectral density of it. By taking the logarithm of the spectral density $f_y(\omega)$, we attain

$$\ln(f_y(\omega_\lambda)) = \ln(f_u(\omega_\lambda)) - d\ln\{4\sin^2(\frac{\omega_\lambda}{2})\} \tag{13}$$

where $\omega_\lambda = \dfrac{2\pi\lambda}{n}$, $\lambda = 1, 2, \cdots\cdots, m$.

This equation can be rewritten as follows:

$$\ln(f_y(\omega_\lambda)) = \ln(f_u(0)) - d\ln\{4\sin^2(\frac{\omega_\lambda}{2})\} + \ln\left[\frac{f_u(\omega_\lambda)}{f_u(0)}\right] \tag{14}$$

The parameter d will be estimated from a dataset $y_1, \cdots, y_n$. Geweke and Porter-Hudak (1983) propose the least square estimator of d, from the equation (14), based on the log periodogram of y. Therefore, we get

$$\ln(I(\omega_\lambda)) = \ln(f_u(0)) - d\ln\{4\sin^2(\frac{\omega_\lambda}{2})\} + \ln\left[\frac{I(\omega_\lambda)}{f_y(\omega_\lambda)}\right] + \ln\left[\frac{f_u(\omega_\lambda)}{f_u(0)}\right] \quad (15)$$

where the periodogram is $(I(\omega_\lambda) = \frac{1}{2\pi}|\sum_{i=1}^{n} y_t \exp(i\omega_\lambda t)|^2$ and $\omega_\lambda = \frac{2\pi\lambda}{n}$, $\lambda = 1,2,\cdots\cdots,m$. When frequencies (ω) are near zero, the last term on the right hand side of the equation (15) goes to zero, and the parameter d can be estimated with the simple least squares regression.

Geweke and Porter-Hudak (1983) show that if the number of frequencies used in the above regression, m, is a function of g(n) where $g(n)=n^\theta$ where $0<\theta<1$, the estimate of the regression is asymptotically normally distributed as:

$$\hat{d}=N\left(d, \frac{\pi^2}{6\sum_{\lambda=1}^{g(n)}(x_\lambda-\bar{x})^2}\right) \quad (16)$$

where $x_\lambda=\ln\{4\sin^2(\frac{\omega_\lambda}{2})\}$and $\bar{x}$ is the sample mean. Under the null hypothesis of no long memory, therefore, the t-statistic is calculated by the following formula:

$$t=\frac{\hat{d}}{\sqrt{\frac{\pi^2}{6\sum_{\lambda=1}^{g(n)}(x_\lambda-\bar{x})^2}}}. \quad (17)$$

The estimate of d depends on the selection of g(n), the number of periodogram ordinates, and this means that the size of the estimate of d relies on the size of θ. The common choice of θ had been between 0.5 and 0.6 since Geweke and Porter-Hudak (1983) use those numbers. However, Hurvich et al.(1998) establish that the optimal m is of order $o(n^{0.8})$. Since their study, many researchers have set θ to be around 0.8[9]. Following these studies, θ ranges between 0.5 and 0.8 in this paper.

(c) Whittle's Method

Whittle's method is based on a frequency domain maximum likelihood of fractionally integrated process. Suppose that we have a standard ARFIMA (p,d,q) model,

$$(1-L)^{d}(y_t-\mu)=u_t, \qquad u_t \sim ARMA(p,q). \tag{18}$$

The parameters of the model can be estimated by minimizing the following function numerically:

$$\sum_{i=1}^{m} \frac{I(\omega_i)}{f(\theta;\omega_i)} \tag{19}$$

where θ is the vector of parameters including the fractional differencing parameter, and $I(\omega_i)$ and $f(\theta;\omega_i)$ are the periodogram and the spectral density of a stationary process, respectively. It is assumed here that ut follows the standard normal distribution. Therefore, the fractional differencing parameter, d, will only be returned from the estimation based on the Whittle's method[10].

9) Ebens (1999), Luu and Martens (2003), and ABDL(2003) use 0.8. Li (2002) chooses 0.7, 0.75 and 0.8. Finally, Kilic (2004) and Choi and Zivot (2007) use 0.7 and 0.8.

10) This restriction essentially makes the Whittle's method equivalent to the ARFIMA (0,d,0) model, introduced by Granger and Joyeux (1980) and Hosking (1981).

(d) Bai and Perron (1998, 2003)'s Structural Break Test

To pinpoint the break dates, Bai and Perron (1998, 2003) starts to estimate unknown break points by the least square principle. Consider the following multiple regressions with m breaks:

$$y_t = \beta_j + u_t, \qquad (t = T_{j=1} + 1, \cdots, T_j) \tag{20}$$

for j=1,⋯,m+1. y_t is a time series, and β_j is the mean of volatility in (j+1)th regime. The u_t is the disturbance at time t. (T_1, T_2, ⋯, T_m) is a set of multiple break points. Since these break points are treated unknown, they should be estimated. For every possible m-partition, therefore, we get the least square estimates of $\beta_j(T_1, T_2, \cdots, T_m)$ by minimizing the following sum of squared residuals,

$$S_T(T_1, T_2, \ldots. T_m) = \sum_{j=1}^{m+1} \sum_{t=T_{j-1}+1}^{T_j} (y_t - \beta_j)^2. \tag{21}$$

The estimated parameters are mean estimates of regimes based on the partition. Putting the mean estimates back into the objective function and denoting the sum of squared residuals as $S_T(T_1, T_2, \cdots, T_m)$, we can obtain the estimated break points defined by

$$(\hat{T}_1,\ \hat{T}_2, \ldots., \hat{T}_m,) = \underset{T_1, T_2, \ldots, T_m}{\text{argmin}} S_T(T_1, T_2, \ldots, T_m), \tag{22}$$

considering all the possible m-partitions. After getting the break points, $(\hat{T}_1, \ldots., \hat{T}_m,)$, we finally identify the estimate of mean regimes, $\hat{\beta}_j(\{\hat{T}_1, \ldots, \hat{T}_m\})$.

To check whether there are structural breaks in the series, I will use the sup F test, double maximum (hereafter UDMAX) statistic, and weighted double maximum (hereafter WDMAX) statistic. Let $\sup F_T(k)$ be the F statistic for testing null of no structural breaks against alternative k breaks with the knowledge of the k break points. We can implement the sup F test using the $supF_T(k)$ statistic across different k to determine if there exists single or multiple mean breaks.

While the alternative hypothesis of the sup F test is that there are k breaks in the series, we can conduct generalized tests where the alternative hypothesis is an unknown number of breaks. The UDMAX statistic is to choose simply the maximum value of $supF_T(k)$ where k is 1 to M[11]. The WDMAX is similar to the UDMAX, but it gives weights to the individual statistics such that the marginal p-values are equal across values of k (number of break points). The critical values for the UDMAX and WDMAX statistics are provided by Bai and Perron (1998). These two statistics are used to determine whether there exists at least one mean break in volatility series.

If the no break hypothesis is rejected by the above tests, $supF_T(k+1|k)$ statistic, starting with k=1, will be used to get an idea of the exact number of breaks. To begin with, the k breaks model (k=1,⋯,M-1) with $supF_T(k)$ are chosen as a null. We reject the null of k breaks in favor of k+1 breaks when the sum of squared residuals from the alternative is sufficiently smaller than the sum of squared residuals from k breaks model. This test is conducted based on the sup F test described earlier.

In order to estimate multiple break points, I will use the sequential method proposed by Bai and Perron (1998, 2003). To run this procedure, I must first obtain the sum of squared residuals and a single break point of $supF_T(1)$. Then, $supF_T(k+1|k)$ test is conducted sequentially until the test fails to reject the null of no additional break. Although the $supF_T(k+1|k)$

11) Simply put, UDMAX = max $\{F_T(1),\cdots,F_T(M)\}$.

statistics, which are mentioned earlier, are disconnected, the $\sup F_T(k+1|k)$ statistics in this sequential method are connected. Therefore, under the same number of breaks (k), a single $\sup F_T(k+1|k)$ and $\sup F_T(k+1|k)$ of the sequential method usually provide us with different break points. Finally, the GAUSS™ code, written by Pierre Perron, will be used to implement all the tests and estimations described above[12).

Ⅳ. Results

As a preliminary procedure, I conduct nonstationarity and stationarity tests to the Parkinson and Garman-Klass estimators of the three indexes. Nonstationarity of the estimators is tested by the augmented Dickey-Fuller test of which the null hypothesis is that the series has a unit root. The KPSS test is employed to test the null hypothesis of stationarity against a unit root and long memory. Since researchers have investigated traditionally whether financial time series are characterized by trend-stationarity or nonstationarity, a constant and a linear time trend are controlled for when both tests are implemented to the extreme volatility estimators. When the augmented Dickey-Fuller test is conducted, the lag length is selected under the Schwartz Information Criterion.

The results of the two tests are reported in Table 2. As reported in this table, the augmented Dickey-Fuller test and the KPSS test reject their null hypotheses at 1% confidence level. These results demonstrate that the Parkinson and Garman-Klass estimators lie between stationarity and

12) The GAUSS™ code is obtained from Pierre Perron's website, http://people.bu.edu/perron.

nonstationarity in all three indexes. This is preliminary evidence that these volatility series have long memory.

〈Table 2〉 Nonstationarity and stationarity tests on extreme value estimators

Variance

	S&P500		DJIA		NDX100	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
ADF	-10.9012** (7)	-10.6310** (7)	-9.9104** (7)	-9.2196** (7)	-9.4042** (8)	-7.4968** (12)
KPSS	1.4585**	1.4959**	1.6186**	1.7552**	2.2719**	2.2717**

Standard deviation

	S&P500		DJIA		NDX100	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
ADF	-7.8477** (9)	-8.2796** (8)	-7.7164** (7)	-7.6623** (7)	-7.0149** (9)	-6.0435** (12)
KPSS	2.3154**	2.3608**	2.8558**	2.9658**	3.3718**	3.2439**

Logarithmic standard deviation

	S&P500		DJIA		NDX100	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
ADF	-7.1910** (9)	-7.2625** (9)	-7.2332** (7)	-7.4247** (7)	-6.6663** (9)	-6.6774** (9)
KPSS	KPSS	2.8833**	2.8856**	3.9004**	3.9452**	4.0402**

Note: 1) S&P500: Jan. 2nd 1990 - Dec. 30th 2005
2) DJIA: April 2nd 1993 - Dec. 30th 2005
3) NDX100: Jan. 2nd 1990 - Dec. 30th 2005
4) ADF: Augmented Dickey-Fuller test (Number in the parenthesis is the lag

length selected by the Schwartz Information Criterion.)
5) KPSS: Stationarity test suggested by Kwiatkowski et al. (1992)
6) *: significant at 5% level, **: significant at 1% level

Visual inspections of the correlograms provide insight on the long memory feature in the estimators. As Figure 1 show, the correlograms of the Parkinson and Garman-Klass estimators for the three stock indexes decline hyperbolically, which means that correlations between two distant observations are not negligible. These plots are similar to the correlograms reported in Ebens (1999), ABDL (2003), and Chen et al. (2006). This hyperbolic decay of autocorrelations is considered a common feature which indicates the presence of significant long memory.

Based on the evidence from statistical tests and correlograms discussed above, I directly estimate the degree of long memory in the volatility series. ARFIMA (0,d,0), Whittle's method, and the GPH test are employed for the estimations of long memory in the two extreme value estimators of the stock indexes. The results are presented in Table 4. In this table, all the estimates of the fractional integration parameter from ARFIMA (0,d,0) and Whittle's method ranges between 0.25 and 0.35. The sizes of these estimates are large enough to argue that the estimators have long memory and the estimates are all significantly different from zero at 1% confidence level.

Similar to these long memory tests, the GPH test estimates significant long memory in the extreme value estimators. However, the estimates of the GPH test are not stable across the size of the theta (θ) but tend to decrease as the size of the theta increases. The fact that long memory estimates are not constant may be puzzling. However, this finding is consistent with Areal and Taylor (2002), which reports a weak but inverse relationship between the number of ordinates and the long memory parameter estimates if we consider that the theta (θ) is positively related to the number of ordinates[13].

As for the next step, I estimate structural break dates from the two extreme value estimators. As mentioned earlier, the Bai-Perron structural break test is implemented to calculate structural break points. The test results are reported in Table 3 and the mean breaks are plotted with logarithmic standard deviations in Figure 2.

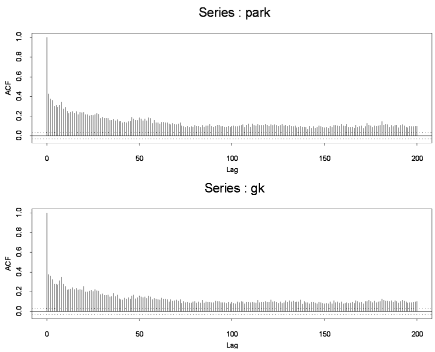
In the case of the S&P 500, the second regime (May 21 1992 - December 2 1996) has the lowest mean level of volatility while the fourth regime (January 2 2000 - August 6 2003) has the highest mean level of volatility. After the fourth regime, the average volatility decreases drastically.

13) Areal and Taylor (2002) displays the estimated degree of fractional integration d as a function of the number of ordinates. See Figure 8 in their paper.

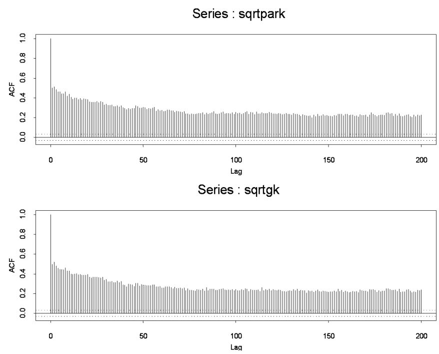
〈Figure 1〉 Correlograms of Extreme Value Estimators (200 Lags)

S&P 500

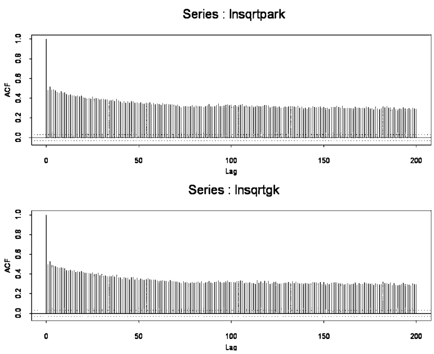
Variance (Parkinson / Garman-Klass)



Standard Deviation(Parkinson / Garman-Klass)

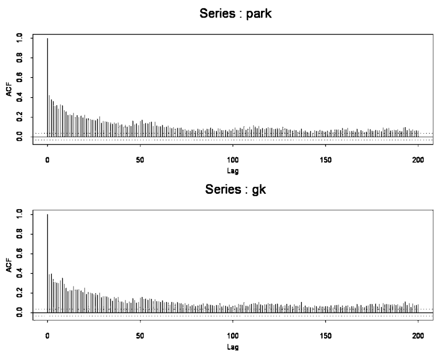


Logarithmic Standard Deviation
(Parkinson / Garman-Klass)

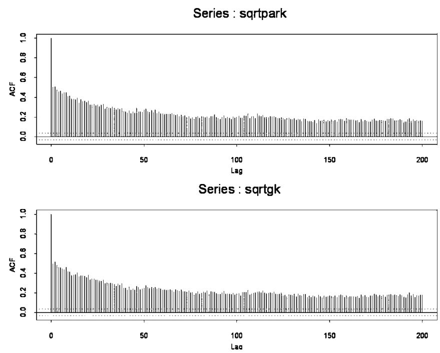


DJIA

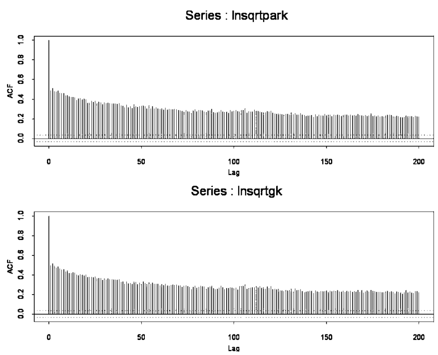
Variance (Parkinson / Garman-Klass)



Standard Deviation(Parkinson / Garman-Klass)

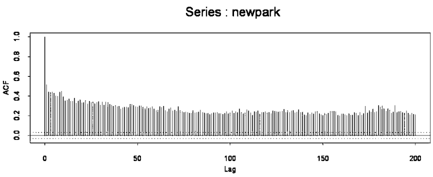


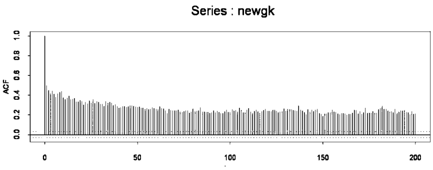
Logarithmic Standard Deviation
(Parkinson / Garman-Klass)



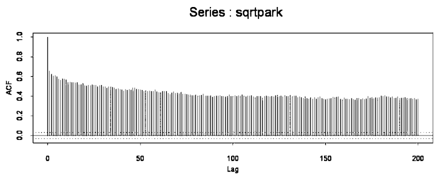
NDX100

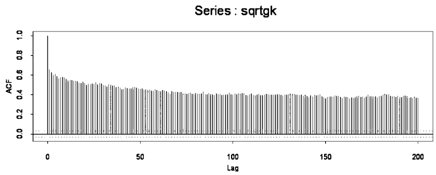
Variance (Parkinson / Garman-Klass)



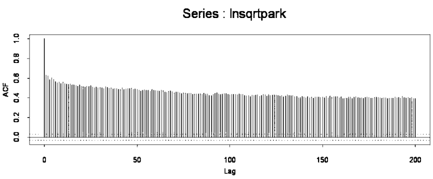


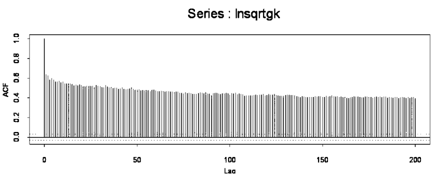
Standard Deviation(Parkinson / Garman-Klass)





Logarithmic Standard Deviation (Parkinson / Garman-Klass)





This statistical result from the S&P 500 index is consistent with that from Guo and Wohar (2006). Implementing the Bai-Perron structural break test to the VIX (Implied volatility index of the S&P 500), they find three distinct periods: pre-1992, 1992-1997, and post 1997. With longer volatility time series, I find additional evidence that the distinct mean break beginning in 1997 continues until early 2003.

〈Table 3〉 Bai and Perron (1998, 2003)'s multiple structural break test results

A. Statistics

	S&P500		DJIA		NDX100	
	Parkinson	Garman-Klass	Parkinson	Garman-Klass	Parkinson	Garman-Klass
$supF_T(1)$	386.45**	360.40**	350.26**	311.79**	210.63**	213.78**
$supF_T(2)$	444.44**	422.32**	426.82**	415.28**	557.01**	530.30**
$supF_T(3)$	384.68**	360.36**	320.05**	306.26**	415.67**	421.70**
$supF_T(4)$	351.19**	329.17**	250.09**	242.38**	328.77**	331.58**
$supF_T(5)$	293.43**	274.41**	201.99**	197.55**	276.31**	267.92**
UDMAX	444.44**	422.32**	426.82**	415.28**	557.01**	530.30**
WDMAX	734.47**	686.86**	560.43**	545.28**	731.38**	696.30**
$supF_T(2\|1)$	560.15**	548.42**	563.72**	573.71**	830.49**	753.31**
$supF_T(2\|1)$	161.07**	146.60**	36.80**	34.60**	113.68**	84.40**
$supF_T(2\|1)$	51.44**	49.62**	4.49	0	30.51**	5.24
$supF_T(2\|1)$	4.33	2.48	0	0	0	0

B. Estimated structural break dates (YY/MM/DD)[1)]

T_1	92/05/21	92/05/21	96/12/02	96/12/02	96/12/05	96/11/29
T_2	96/12/02	96/12/02	01/08/27	00/01/19	99/12/31	99/12/31
T_3	99/12/31	99/12/31	03/08/07	03/08/07	03/08/06	03/08/06
T_4	03/08/06	03/08/06				

C. Estimations of mean for each regime[2)]

Regime 1	-0.481454 (0.019)	-0.534610 (0.018)	-0.704851 (0.014)	-0.708564 (0.013)	-0.090372 (0.010)	-0.135318 (0.010)
Regime 2	-0.843552 (0.014)	-0.878425 (0.013)	-0.164230 (0.012)	-0.230790 (0.015)	0.320654 (0.015)	0.228832 (0.015)
Regime 3	-0.207924 (0.016)	-0.267147 (0.016)	-0.033315 (0.019)	-0.066567 (0.014)	0.697144 (0.014)	0.651329 (0.014)
Regime 4	0.004216 (0.015)	-0.056879 (0.015)	-0.653538 (0.017)	-0.687121 (0.017)	-0.258522 (0.017)	-0.280636 (0.017)
Regime 5	-0.664234 (0.019)	-0.719849 (0.018)				

Note: 1) Break dates and the mean for each regime are estimated by the sequential method suggested by Bai and Perron (1998, 2003). YY: Year, MM: Month, DD: Day

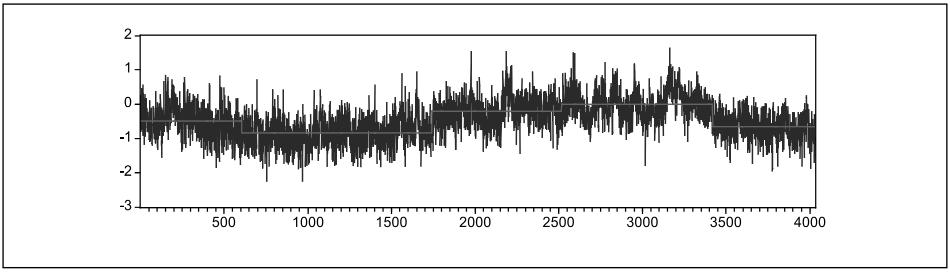
2) Numbers in the parentheses indicate standard errors.

3) *: significant at 5% level, **: significant at 1% level

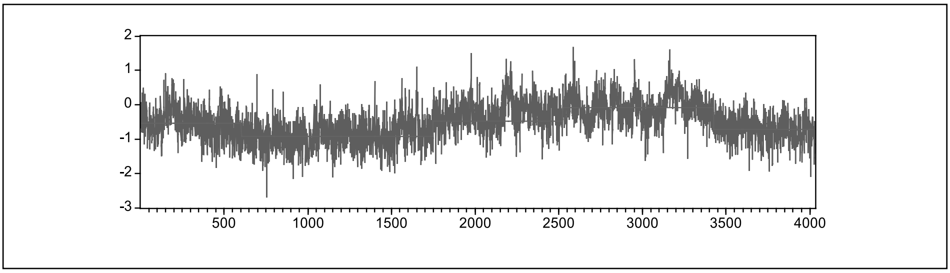
〈Figure 2〉 Logarithmic standard deviation of extreme value estimators and multiple mean breaks estimated by Bai-Perron (1998, 2003) test

S&P500
(January 2 1990 - December 30 2005)

Parkinson

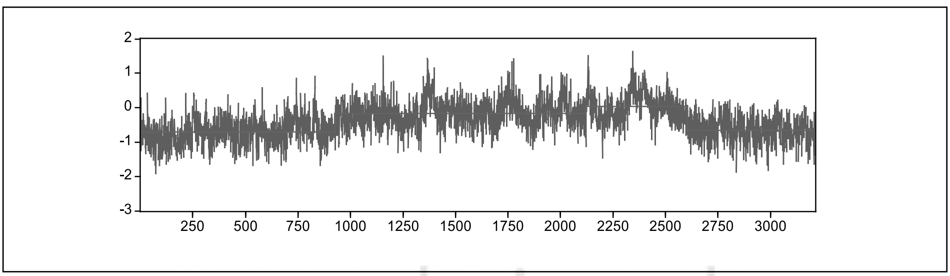


Garman-Klass

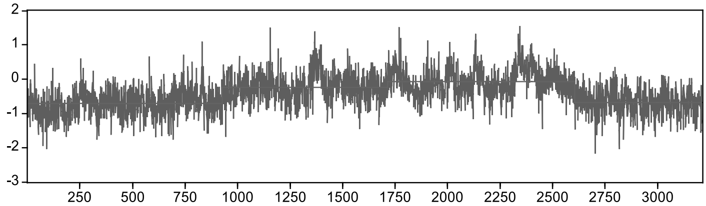


DJIA
(April 2 1993 - December 30 2005)

Parkinson

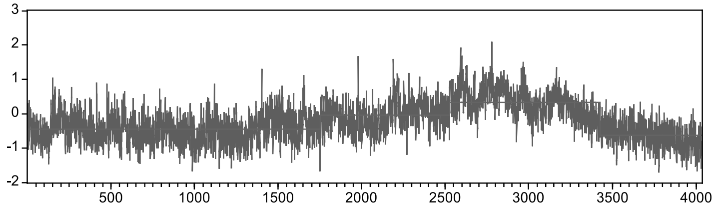


Garman-Klass

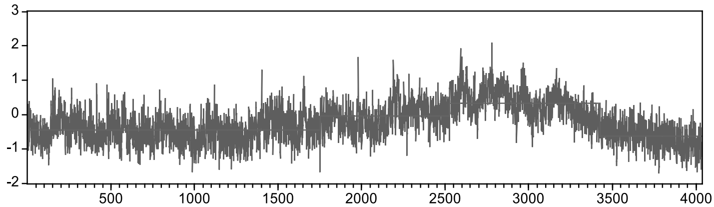


NDX100
(January 2 1990 - December 30 2005)

Parkinson



Garman-Klass



The break dates estimated from the extreme value estimators of the DJIA and the NDX100 are linked closely to the dates from the estimators of the S&P 500. For the two stock indexes, break dates are roughly located in

December 1996, around 2000, and August 2003.

Next, I examine whether long memory in the extreme value estimators exists after structural breaks are controlled for in the series. The mean breaks were estimated by the Bai-Perron test and I simply removed them from the original volatility series. Then, the degrees of long memory in the volatility series were re-estimated by the methods which have been used above. The estimation results are placed right below the estimates of long memory in the original series in Table 4.

According to this table, the degrees are reduced to a large extent when the test is performed by the ARFIMA (0,d,0) and Whittle's method. The size of the reductions mostly ranges between 0.06 and 0.08. Even after the removals, however, the degrees of long memory in the break-eliminated series are still sizable and statistically significant. These results are consistent with those in Choi et al. (2006), who analyze the long memory properties of several realized currency volatilities.

The GPH test results are similar to those from the ARFIMA (0,d,0) and Whittle's method although the sizes of the reductions in the GPH test differ across the size of the theta . As we can see in Figure 3, the autocorrelations of volatility series decrease to a large extent after the multiple mean breaks are removed. This decrease is consistent with the reductions in the degree of long memory appeared in Table 4.

Taking into account the statistical test results and the visual inspections of correlograms, I conclude that the degrees of long memory in the extreme value estimators are reduced after the mean breaks are eliminated from the volatility series but the break-eliminated series still remain highly persistent.

Finally, I examine how stable the estimated long memory parameters in the break-removed series are. Since persistence of a time series depends on the sample period that the series covers, it is natural that the degrees of long memory gradually vary when they are recursively estimated in the series. From the previous analysis, we find that structural breaks inflate the degree of long memory in volatility series. Based on that evidence, we can expect that the estimates of long memory in the break-removed series are more stable over time than those in the original series. If the long memory estimates from the break-eliminated series are large enough and stable over time, it might be evidence that a part of long memory in the original series is not spurious.

〈Table 4〉 Long memory estimates of logarithmic standard deviation and break-eliminated logarithmic standard deviation

A. S&P500

		ARFIMA (0,d,0)	Whittle	d_{GPH} (θ = 0.5, 0.6, 0.7, 0.8)			
Parkinson estimator	Log(SD)	0.2844 (23.05)	0.2680	0.5884 (6.5290)	0.5316 (9.3807)	0.4749 (13.0475)	0.3670 (15.3378)
	Break-adjusted log(SD)	0.2000 (16.21)	0.1999	0.2688 (2.9828)	0.3629 (6.4039)	0.3866 (10.6210)	0.3225 (13.4772)
Garman-Klass estimator	Log(SD)	0.2929 (23.74)	0.2769	0.5877 (6.5214)	0.5474 (9.6597)	0.4603 (12.6464)	0.3633 (15.1822)
	Break-adjusted log(SD)	0.2210 (17.91)	0.2191	0.2525 (2.8018)	0.3700 (6.5294)	0.3737 (10.2676)	0.3187 (13.3185)

B. DJIA

		ARFIMA (0,d,0)	Whittle	d_{GPH} (θ = 0.5, 0.6, 0.7, 0.8)			
Parkinson estimator	Log(SD)	0.2902 (20.99)	0.2765	0.5790 (6.0028)	0.4779 (7.8443)	0.4927 (12.4255)	0.3645 (13.8599)
	Break-adjusted log(SD)	0.2129 (15.40)	0.2135	0.2261 (2.3440)	0.2984 (4.8986)	0.3901 (9.8369)	0.3042 (11.5665)
Garman-Klass estimator	Log(SD)	0.2953 (21.36)	0.2817	0.5538 (5.7408)	0.4884 (8.0169)	0.4706 (11.8684)	0.3738 (14.2126)
	Break-adjusted log(SD)	0.2213 (16.01)	0.2219	0.2412 (2.5005)	0.3324 (5.4560)	0.3887 (9.8017)	0.3226 (12.2646)

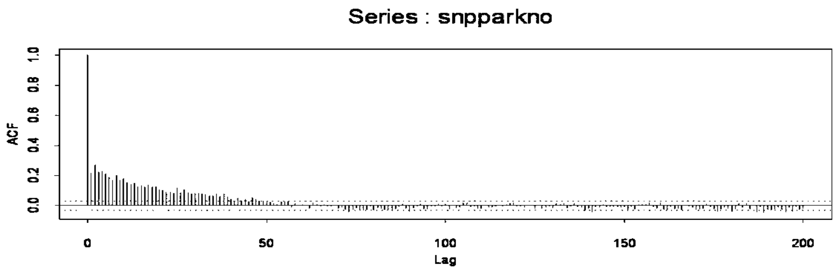
C. NDX100

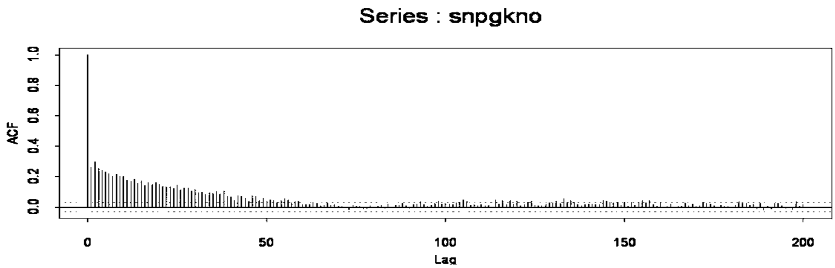
		ARFIMA (0,d,0)	Whittle	d_{GPH} (θ = 0.5, 0.6, 0.7, 0.8)			
Parkinson estimator	Log(SD)	0.3337 (27.05)	0.3126	0.5926 (6.5757)	0.4857 (8.5708)	0.4211 (11.5699)	0.3500 (14.6288)
	Break-adjusted log(SD)	0.2603 (21.10)	0.2589	0.4032 (4.4738)	0.4079 (7.1982)	0.3943 (10.8317)	0.3364 (14.0605)
Garman-Klass estimator	Log(SD)	0.3335 (27.03)	0.3133	0.6283 (6.9718)	0.4747 (8.3769)	0.4168 (11.4519)	0.3447 (14.4071)
	Break-adjusted log(SD)	0.2578 (20.89)	0.2571	0.4087 (4.5347)	0.3859 (6.8093)	0.3797 (10.4325)	0.3255 (13.6031)

Note: 1) Log(SD): Logarithmic standard deviation.
2) Break-adjusted log(SD): Logarithmic standard deviation where structural breaks, estimated by the Bai-Perron structural break test, are removed
3) The numbers in the parentheses indicate t-statistics.

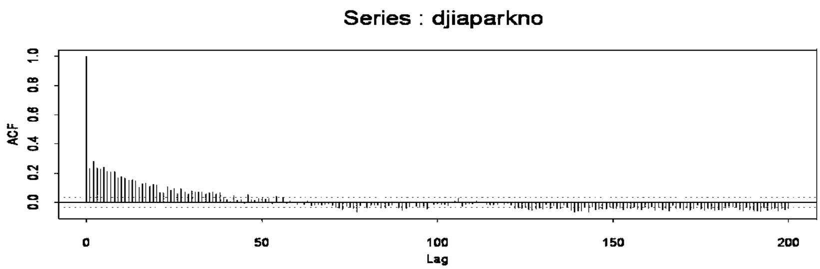
〈Figure 3〉 Sample Autocorrelations of break-adjusted logarithmic standard deviation of extreme value estimators (200 lags)

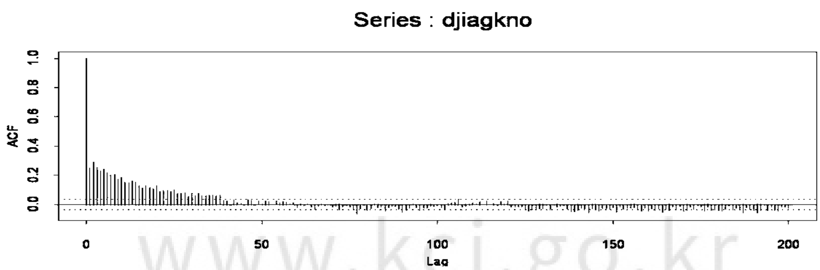
S&P 500



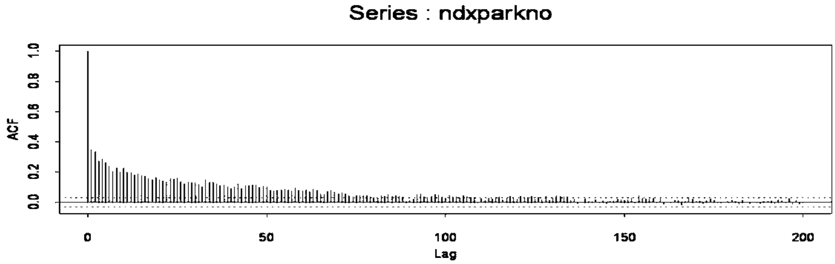


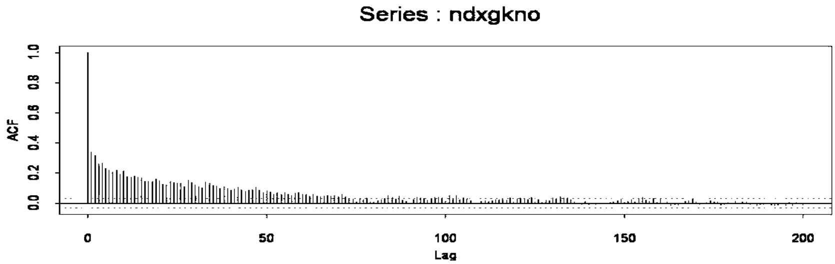
DJIA





NDX100





Note: The first correlogram of each index is that of the Parkinson estimator and the second is that of the Garman-Klass estimator.

To investigate this further, I recursively estimated the long memory parameters using Whittle's method. The S&P500 and NDX100 in my dataset have more observations than the DJIA. This mismatch of samples may prevent us from determining whether the reductions of long memory estimates occur in the same way of the three indexes. To see if the reductions in long memory occur consistently across the extreme value estimators of all the indexes, it will be better that the recursive estimations start and finish at the same time period. To achieve this goal, the long memory parameters were estimated recursively after the first 1,500 samples for the S&P500 and the NDX100 and after the first 677 samples for the DJIA. By doing this, the estimations start December 6 1995 and finish

December 30 2005.

The recursive estimations of the long memory parameters are displayed in Figure 4. In the three indexes, the degree of long memory become flatter after the removal of the multiple mean breaks in the two estimators. However, significant long memory still appears in the break-eliminated estimators. As Choi and Zivot (2007) argue, this is evidence that a part of long memory in the estimators is not spurious.

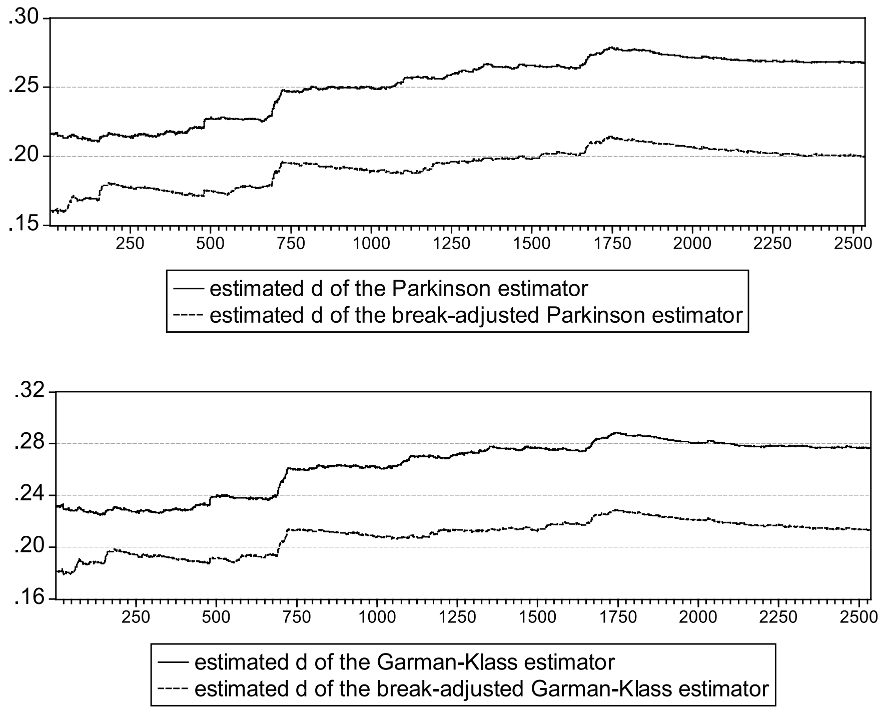
Considering all the results reported, I conclude that the extreme value estimators of the U.S. stock indexes possess significant long memory even after multiple structural breaks are controlled for.

V. Conclusion

Based on several long memory tests, I estimated long memory in the extreme value estimators of the three U.S. stock indexes. From the analyses, I found ample evidence that long memory is present in the extreme value estimators just as previous studies found long memory in other types of volatility time series. Along with the several conventional long memory tests, I reflected recent research investigating the relation between long memory and occasional structural breaks. Conducting the so-called 'break-adjusted' long memory test proposed by Choi and Zivot (2007), I still found significant long memory in the break-eliminated estimators although the sizes of long memory are reduced compared to those in the original series.

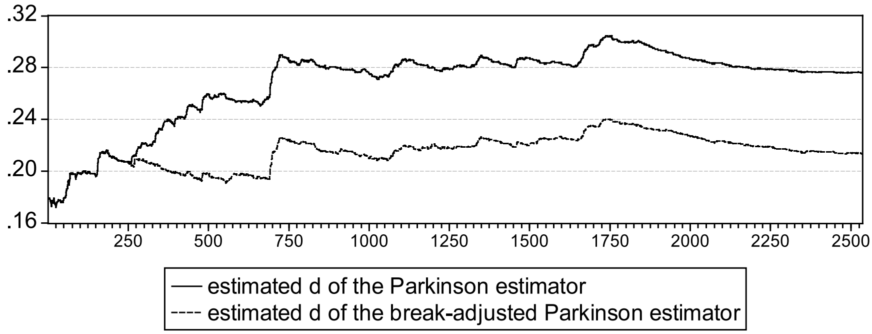
Figure 4. Recursive estimation of the long memory parameter with the Whittle's method in the extreme value estimators

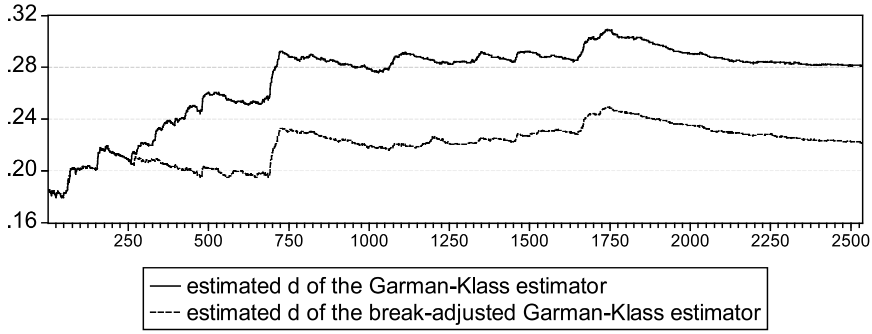
S&P500



Estimation period: December 6th 1995 – December 30th 2005

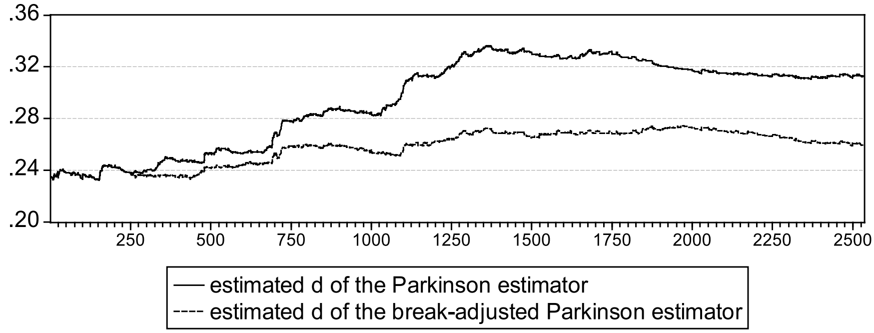
DJIA

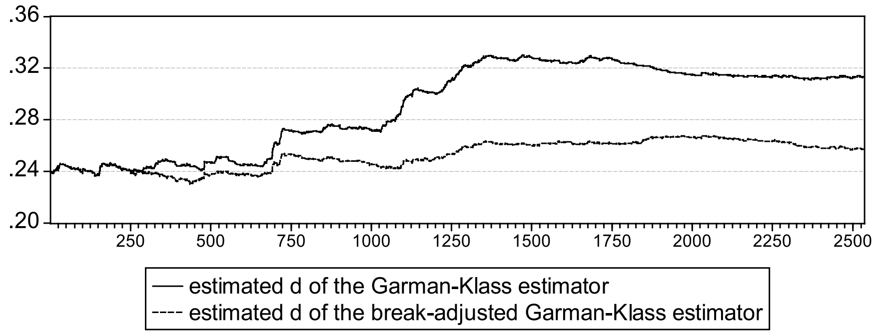




Estimation period: December 6th 1995 – December 30th 2005

NDX100





Estimation period: December 6th 1995 – December 30th 2005

A possible way of extending this paper is to look into the origin of long memory in the extreme value estimators. Past studies were merely reporting strong persistence of various volatility estimators. However, new approach should be launched if the degree of long memory is inflated by occasional structural breaks. For instance, econometric models, which reflect long memory and structural breaks simultaneously, should be considered for modeling and predicting future level of volatility.

One plausible alternative is a hybrid model that allows switching from one model to another according to market status. For example, Dufrenot et al. (2005) propose a two-regime self exciting threshold autoregressive model, also known as a two-regime SETAR, to reflect this phenomenon. In their model, a volatility process switches from the long memory process to weakly dependent process according to the level of volatility at the very last period. By applying this kind of time series models to forecasting, we may expect additional improvement in prediction performance. Furthermore, future research should generate additional efforts in developing a new long memory model because it will deepen our understanding about dynamics of financial volatilities.

References

Andersen, T.G., Bollerslev, T., Diebold, F.X., and Ebens, H., "The distribution of realized stock return volatility", *Journal of Financial Economics*, 61, 2001, pp.43~76.

Andersen, T.G., Bollerslev, T., Diebold, F.X., and Labys, P., "The distribution of realized exchange rate volatility", *Journal of American Statistical Association*, 96, 2001, pp.42~55.

______, "Modeling and forecasting realized volatility", *Econometrica*, 71, 2003, pp.579~625.

Areal, N.M.P.C., and Taylor, S.J., "The realized volatility of FTSE-100 futures prices", *Journal of Futures Markets*, 22, 2002, pp.627~648.

Bai, J., and Perron, P., "Estimating and testing linear models with multiple structural changes", *Econometrica*, 66, 1998, pp.47~78.

______, "Computation and analysis of multiple structural change models", *Journal of Applied Econometrics*, 18, 2003, pp.1~22.

Baillie, R.T., Bollerslev, T., and Mikkelsen, H.O., "Fractionally integrated generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, 74, 1996, pp.3~30.

Bali, T., and Weinbaum, D., "The empirical performance of alternative extreme-value volatility estimators", *Journal of Futures Markets*, 25, 2005, pp.873~892.

Byers, J.D., and Peel, D.A., "Volatility persistence in asset markets: long memory in high/low prices", *Applied Financial Economics*, 11, 2001, pp.253~260.

Chen, Z., Daigler, R.T., and Parhizgari, A.M., "Persistence of volatility in futures markets", *Journal of Futures Markets*, 26, 2006, pp.571~594.

Choi, K., and Zivot, E., "Long memory and structural changes in the forward discount: An empirical investigation", *Journal of International Money and Finance*, 26, 2007, pp.342~363.

Choi, K., Yu, W-C., and Zivot, E., "Long memory versus structural breaks in modeling and forecasting realized volatility", Working paper, Winona State University, 2006.

Ding, Z., Granger, C.W.J., and Engle, R.F., "A long memory property of stock market returns and a new model", *Journal of Empirical Finance,* 1, 1993, pp.83~106.

Dufrenot, G., Guegan, D., and Peguin-Feissolle, A., "Modelling squared returns using a SETAR model with long-memory dynamics", *Economics Letters*, 86, 2005, pp.237~243.

Ebens, H., "Realized stock volatility", Johns Hopkins University, 1999.

Feller, W., "The asymptotic distribution of the range of sums of independent random variables", *Annals of Mathematical Statistics,* 22, 1951, pp.427~432.

Garman, M., and Klass, M., "On the estimation of security price volatilities from historical data", *Journal of Business,* 53, 1980, pp.67~78.

Geweke, J., and Porter-Hudak, S., "The estimation and application of long memory time series models", *Journal of Time Series Analysis*, 4, 1983, pp.221~237.

Granger, C.W.J., and Ding, Z., "Some properties of absolute returns: An alternative measure of risk", *Annales d'Economie et de Statistique*, 40, 1995, pp.67~91.

Granger, C.W.J., and Hyung, N., "Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns", *Journal of Empirical Finance*, 11, 2004, pp.399~421.

Granger, C.W.J., and Joyeux, R., "An introduction to long memory time series and fractional differencing", *Journal of Time Series Analysis*, 1, 1980, pp.15~29.

Guo, W., and Wohar, M.E., "Identifying regime changes in market volatility", *Journal of Financial Research*, 29, 2006, pp.79~93.

Hosking, J., "Fractional differencing", *Biometrika*, 68, 1981, pp.165~176.

Hurvich, C.M., Deo, R., and Brodsky, J., "The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series", *Journal of Time Series Analysis*, 19, 1998, pp.19~46.

Kilic, R., "On the long memory properties of emerging capital markets: evidence from Istanbul stock exchange", *Applied Financial Economics*, 14, 2004, pp.915~922.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., "Testing the null hypothesis of stationarity against the alternative of a unit root", *Journal of Econometrics*, 54, 1992, pp.159~178.

Lamoureux, C.G., and Lastrapes, W.D., "Persistence in variance, structural change, and the GARCH model", *Journal of Business and Economic Statistics*, 8, 1990, pp.225~234.

Lastrapes, W.D., "Exchange rate volatility and U.S. monetary policy: An ARCH application", *Journal of Money, Credit and Banking*, 21, 1989, pp.66~77.

Li, K., "Long-memory vs. Option-implied volatility predictions", *Journal of Derivatives*, 9, 2002, pp.9~25.

Lo, A.W., "Long term memory in stock market prices", *Econometrica*, 59, 1991, pp.1279~1313.

Lobato, I.N., and Savin, N.E., "Real and spurious long memory properties of stock market data", *Journal of Business and Economic Statistics*, 16, 1998, pp.261~283.

Luu, J., and Martens, M., "Testing the mixture of distributions hypothesis using "realized" volatility", *Journal of Futures Markets*, 23, 2003, pp.661~679.

Nelson, C.R., and Plosser, C.I., "Trends and random walks in macroeconomic time series: some evidence and implications", *Journal of Monetary Economics*, 10, 1982, pp.139~162.

Parkinson, M., "The extreme value method for estimating the variance of the rate of return", *Journal of Business*, 53, 1980, pp.61~68.

Perron, P., "The Great Crash, the Oil Price shock, and the unit-root hypothesis", *Econometrica*, 57, 1989, pp.1361~1401.

Rogers, L.C.G., and Satchell, S.E., "Estimating variance from high, low and closing prices", *Annals of Applied Probability*, 1, 1991, pp.504~512.

Rogers, L.C.G., Satchell, S.E. and Yoon, Y., "Estimating the volatility of stock prices: a comparison of methods that use high and low prices", *Applied Financial Economics*, 4, 1994, pp.241~247.

Shu, J., and Zhang, J.E., "Testing range estimators of historical volatility", *Journal of Futures Markets*, 26, 2006, pp.297~313.

Taylor, S.J., "Forecasting the volatility of currency exchange rates", International *Journal of Forecasting*, 3, 1987, pp.159~170.

Wiggins, J.B., "Empirical tests of the bias and efficiency of the extreme-value variance estimator for common stocks", Journal of Business, 64, 1991, pp.417~432.

______, "Estimating the volatility of S&P 500 futures prices using the extreme-value method", *Journal of Futures Markets*, 12, 1992, pp.265~273.

Yang, D., and Zhang, Q., "Drift independent volatility estimation based on high, low, open and close prices", *Journal of Business*, 73, 2000, pp.477~491.

Zivot, E., and Andrews, D.W.K., "Further evidence on the Great Crash, the Oil-Price Shock, and the unit-root hypothesis", *Journal of Business and Economic Statistics*, 10, 1992, pp.251~270.

요 약

본 연구는 극단값 변동성 측정치(extreme value estimators)의 장기기억과 구조변화를 측정하였다. 장기기억은 변동성 측정치의 정형화된 사실 중 하나로서 다수의 연구들이 변동성 시계열이 장기기억의 특성을 가지고 있음을 보고한 바 있다. 그러나 Granger and Hyung (2004)과 Choi et al. (2006)의 연구는 기존의 연구들에서 측정된 변동성 시계열의 장기기억이 실은 구조변화에 의해 그 크기가 부풀려져 있을 수 있다고 주장한 바 있다. 본 연구는 이들의 연구를 극단값 변동성 측정치의 장기기억 연구에 적용하여 극단값 변동성 측정치의 장기기억의 수준이 부풀려져 있는지를 분석해 보았다. 분석해 본 결과 극단값 변동성 측정치의 비정기적인 구조변화가 장기기억의 정도를 증폭시키고 있음을 발견하였으나 이 효과를 감안하더라도 일정한 수준의 장기기억이 여전히 존재하고 있음을 또한 발견하였다.

※ 국문 색인어: 구조변화, 극단값 변동성 측정치, 변동성, 장기기억