Modeling and Forecasting Financial Volatility Using Composite CARR Models

Isuru Ratnayake, V.A. Samaranayake
Missouri University of Science and Technology

Abstract
Recent studies have shown that when it comes to forecasting realized volatility, conditional autoregressive range (CARR) models, that utilize the daily range of a commodity price, outperforms the traditional GARCH approach that models the daily returns. The CARR models, however, assume that the unconditional mean range is constant over time, which holds only if the unconditional volatility remains fixed over the duration of the study period. As several authors have pointed out, there is strong empirical evidence suggesting the feasibility of modeling a slow-varying change in the unconditional volatility over the study period using long term volatility component. In this paper we propose a new composite range based component model to analyze both long term and short term volatility components in a daily price range data. The proposed CCARR models long term volatility changes as a stochastic component which itself exhibits conditional volatility and the application of the proposed model is illustrated by using S&P500 and FTSE 100 stock indices.

Key Words: CARR Models, Range Estimators, Financial Time Series, Market Volatility, Duration Models

1. Introduction

Financial volatility is defined as a measure of the dispersion of returns for a given asset. It is the conventional measure in assessing the risk of speculative assets. In general, riskiness of the market is directly proportional to the volatility. Volatility is closely linked with the stability of the financial market and plays a vital role in determining the level of economic activity. It is also used as a key input for asset pricing. Thus, financial volatility is an essential factor that policy makers and regulators should consider prior to any form of financial decision making. Modelling volatility is crucial in understanding the nature of the dynamics of the finical market.

Modeling financial volatility of asset prices has been discussed extensively in the financial and econometric literature over the years. One of the most successful volatility models used by researchers to model time series volatilities is the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model introduced by Bollerslev (1986). This paper, of course, in based on the ideas put forth in the seminal paper by Engle (1982), which proposed the Autoregressive Conditional Heteroscedasticity (ARCH) model in order to address the complexities of time varying volatility and volatility clustering in the financial time series. The ARCH approach models the error variance as a function of actual errors of the previous periods while GARCH method, which is an extension of ARCH method, models the variance of the error as a function of error terms and its conditional variance.

Owing to the significance of modeling and forecasting asset price volatilities, a wide range of empirical and theoretical investigations have been carried out within the context of econometric literature to select the ideal model. Akgiray (1989), mentioned that GARCH

(1, 1) models fit the daily return series data reasonably well. He stated this after considering the evidence from the time series behavior of stock prices. GARCH model used in this paper employed rate of return to study the volatility and found that daily return series demonstrated a significant level of second order dependence which cannot be modeled using merely a linear white noise process.

Due to the growing interest and developments in financial time series during the 1990's some researchers became heavily invested in modelling the time intervals between events. The first durational model was proposed by Engle and Russell (1998). In their publication, they introduced a new statistical model which is capable of analyzing irregularly spaced financial transaction data and they named the model the Autoregressive Conditional Durational Model (ACD). Since then, multiple authors have proposed related versions of ACD models such as logarithmic ACD (LACD) models by Bauwens and Giot (2000), Nonlinear ACD by Zhang et al. (2001), Box-Cox ACD by Hautsch (2002).

In many financial time series applications, standard deviation is the most common measure of stock return volatility since it not only calculates the dispersion of returns but also summarizes the probability of seeing extreme values in returns. Researchers have focused their attention to finding alternative measures of financial volatility such as range. It is a well-known fact in statistics that the range is a measure of variability of a random variable. Parkinson (1980) argued that volatility measures can be calculated by considering daily high, daily low, and opening price of a stock in addition to the traditional closing prices. He also compared traditional measures of volatility that are calculated simply by using closing prices, with extreme value methods by taking high and low prices of an asset. He concluded that range based method is far superior to the available standard methods. Beckers (1983) tested the validity of different volatility estimators. In his paper, he mentioned that the range of a stock price contains more important and fresh information. He also mentioned that using range of a stock price is better than using close to close changes. Hence range of an asset price for a given period can be used as a more informative proxy variable to measure the assets volatility during that period. Researchers studied this alternative approach to volatility modeling and developed new theoretical range based models with comprehensive empirical examples. For example, Brandt and Jones (2006) fitted effective exponential GARCH (EGARCH) models to range data from S&P500 index.

Chou (2005) first introduced the Conditional Auto Regressive Range (CARR) model, which is primarily an ACD model. While this ACD model is used to model the time intervals between events with positive observations , CARR is employed to model price volatility of an asset by considering range of the log prices for a given fixed time interval. CARR model is similar to the standard volatility models such as the GARCH model. However, one distinct difference between the two models is that, the GARCH model uses rate of return as its volatility measure while CARR model uses the range its volatility measure. The CARR model proposed by Chou is a simple but efficient tool for analyzing the volatility clustering property when compared to the GARCH models. This was proven empirically via out of sample forecasting of S&P500 data. Chou showed that the effectiveness of volatility estimates produced by CARR models is higher than the estimates of standard return based models such as GARCH models. Zou (2014) used CARR model and GARCH model to forecast volatility of the stock index in Shanghai stock market. He used Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), which were proposed by Chou (2005), in order to compare the Weibull-CARR model with GARCH-t models. In this study he concluded that Weibull CARR model outperforms GARCH-t model in the forecasting ability. Quiros (2011) discussed volatility forecasting with range models. He improved previous work done by Chou (2005) through extending the time period and analyzing the performance of CARR model in contrasting situations such as in periods with upward trends and in periods with downward trends. He proposed various range estimators to analyze the forecasting performance and further stated that Parkinson (1980) model is preferable to CARR model during periods with upward trends while CARR is recommended for periods with downward trend.

Chaing (2016) proposed the Lognormal Logarithmic Conditional Auto Regressive (Lognormal Log CARR) model with the aim of examining the volatility outliers and improving the accuracy of forecasting. This model was influenced by the Logarithmic Autoregressive Conditional Duration (Log ACD) model of Bauwens and Giot (2000). One major advantage of using either a Log ACD or a Log CARR model would be that these models relax positivity restrictions on the parameters of the conditional expectation function. Fernandes et al. (2005), in his works of multivariate extension of CARR model, derived the conditions for the existence of statistical properties such as first moment, stationarity of the model.

The broad scope of volatility models proposed by various academics provide us copious opportunities to model the volatility, but as a single component. Recent studies carried out on the subject leads us to examine the volatility of economic and financial variables as a function of long term and short term components.

Engle and Lee (1999) introduced an additive component to GARCH models with a long term and short term components. The Spline-GARCH model proposed by Engle and Rangel (2008) models equity market volatilities as a combination of macroeconomic activities and time series dynamics. In this same paper Engle and Rangel named the slow the moving trend in the volatility process as a low frequency volatility and presented the functional form of the low frequency volatility by adopting a non-parametric approach. In essence, they considered the low frequency component as deterministic. Instead of using an additive component, a multiplicative component was used in the Spline-GARCH model to separate low and high frequency volatilities. Therefore ‘high frequency return volatility’ is a product of a slow moving deterministic volatility component which can be represented by an exponential spline combined with a unit GARCH model. This model was able to capture short and long term behaviors of financial market volatilities. Slow moving volatility component can be used to model long run dynamic behavior of the market while unit GARCH model can be employed to capture short term dynamics. Based on the Spline GARCH model Engle et al. (2013) proposed a new component model with a direct link to the economic activities and this new class of models was named as GARCH MIDAS models. This paper explained long term volatility using an approach which can handle stock volatilities and economic activities recorded in different frequencies, namely, daily monthly or quarterly. The MIDAS technique was initially introduced by Ghysels (2006) and is used to build a link between the long run volatility component and macroeconomic variables. The unit GARCH process was used as in Spline GARCH approach to model the short run volatility component. The GARCH MIDAS model is a multiplicative model with differentiated short and long run components of volatility. The conditional volatility of returns in this model depends on macroeconomic variables and previous economic periods or lags. Engle et al. (2013) formulated long term movement with inflation and industrial production growth. They found that including macroeconomic variables to the model outperforms the traditional time series in terms of long and short horizon forecasting. With the motivation of GARCH MIDAS model, Swanson (2017) proposed CARR MIDAS model. In this study, volatility was decomposed in to short and long term components, and short run volatility component is explained by an exponential CARR (1, 1) model. Long run volatility component is computed by aggregating measures of scaled realized range over past *k* low frequency periods.

Several other authors also utilize empirical data to illustrate the modeling of short and long term volatility components using both Spline GARCH and GARCH MIDAS models. Nguyen and Walther (2017) conducted an empirical study using commodity futures which are traded in New York Stock Exchange (NYMEX). They fitted both Spline GARCH and GARCH MIDAS models. They found that disentangling high and low volatility components produced better results for in-sample fit in both models.

More recent provides a basic insight to different types of volatility models including range-based volatility models, and discusses the importance of analyzing the long term and short term volatility components in them. While CARR models assume a constant unconditional mean range over time, several other studies namely, Engle et al. (2013) and Conrad et al. (2018) suggest, with empirical evidence, that unconditional volatility in return series changed over the study periods.

In this Study we propose a new class of Composite Range Based Component Models for volatility to analyze long term and short term volatilities in daily price range data. We introduce a stochastic component to model the long term volatility in daily price range data, which in itself exhibits conditional volatility. Both long term and short term components are driven by the past realization of range price series. Further we introduce an estimator to estimate unobserved long term volatility component and discuss the parameter estimation procedure. Finally multiple indices such as S&P500 and FTSE 100 are used for the empirical study and compare the prediction and forecasting ability of proposed Composite CARR model against the single component CARR model.

2. The Model Specification and Discussion

2.1 The Conditional Autoregressive Range (CARR) Model

Chou (2005), proposed the CARR model which is primarily a range based model. CARR model is employed to fit the price volatility of an asset by considering range as a measure of price volatility. A CARR model of order (p, q) is presented as CARR (p, q) and defined as follows:

$$\begin{aligned} &R_t = \lambda_t \varepsilon_t, \\ &\mathrm{E}\left(R_t \mid \mathbb{F}_{t-1}\right) = \lambda_t = \omega + \sum_{i=1}^{p} \alpha_i R_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j}, \\ &\varepsilon_t \sim i.i.d.\ f(\phi, \varepsilon_t). \end{aligned} \tag{1}$$

Here λ_t is the conditional expectation of the range, based on all information up to time t. The non-negative disturbance term, also known as standardized range, is defined by ε_t $\left(=\dfrac{R_t}{\lambda_t}\right)$, which is independent and identically distributed with probability density function $f(.)$ with a unit mean. Since R_t and λ_t are positive, the coefficients of the conditional mean range equation have the following restrictions:

$\omega \geq 0, \alpha_i > 0, \beta_j > 0$ for all $i \in (1,2,3,...,p)$ and $j \in (1,2,3,...,q)$.

Let R_t be the price range defined over the time interval $\left[t_{open}, t_{close}\right]$ such that:

$R_t = \max(P_\tau) - \min(P_\tau)$, where $\tau \in \left[t_{open}, t_{close}\right]$. Here we let P_τ be the price of an asset at a given time τ.

2.2 The Composite Conditional Autoregressive Range (CCARR) Model

Let $P_{j,i,t}$ be the logarithmic price of a speculative asset defined at time j of a given short term period (i.e. day) t of any arbitrary long-term period t such as month, quarter and year. Here $j,i,t \in [i_{open,t}, i_{close,t}]$ and $i = 1,2,3,...,N_t$ where N_t is the number of days for the given long term period t. Here $t = 1,2,3,...,T$ where T be the number of long term periods in total time span. The observed price range over the short term time period i at a given long term period t is denoted as $R_{i,t}$ and it is defined as follows.

$$R_{i,t} = (\max(P_{j,i,t}) - \min(P_{j,i,t}))*100. \tag{2}$$

The Composite Conditional Autoregressive Range (CCARR) model for the range is defined as follows:

$$R_{i,t} = \tau_t g_{i,t} \varepsilon_{i,t}, \tag{3}$$

where $\varepsilon_{i,t} \sim i.i.d.\ f(\sigma, \varepsilon_{i,t})$ with a unit mean (i.e. $\mathrm{E}(\varepsilon_{i,t}) = 1$), $\forall i = 1,2,3,...,N_t$ and $t = 1,2,3,...,T$.

Observe that the daily price range $(= R_{i,t})$ is separated into short term and long term volatility components.

The long term volatility component τ_t is given by,

$$\begin{aligned} \tau_t &= \omega_t \eta_t, \\ \omega_t &= \gamma_0 + \gamma_1 \tau_{t-1} + \delta \omega_{t-1}, \\ \omega_t &= \mathrm{E}(\tau_t \mid \mathbb{F}_{(t-1)}). \end{aligned} \tag{4}$$

Here ω_t be the mean of the long term volatility component conditioned on all information up to time $t-1$ and $\mathbb{F}_{(t-1)}$ is the sigma field generated by the information set up to long term period $t-1$. The long term disturbance term is denoted by η_t, where $\eta_t \sim i.i.d.$ $f(\nu,\eta_t)$ and $\mathrm{E}(\eta_t)=1$. Long term volatility component τ_t is modeled as a stochastic component that itself exhibits conditional volatility according to the Conditional Autoregressive Range (CARR (1, 1)) process.

The short term volatility component $g_{i,t}$ is given by,

$$g_{i,t} = (1-\alpha-\beta) + \alpha \frac{R_{(i-1,t)}}{\tau_t} + \beta g_{(i-1,t)}. \tag{5}$$

Here the short term volatility component $g_{i,t}$ is defined as obeying a unit CARR (1, 1) model similar to Engle and Rangel (2013). Following the derivation given by Engle et al. (2013) for the short term volatility component we can prove that unconditional expectation of short term volatility component is $\mathrm{E}(g_{i,t})=1$. Both short term and long term volatility components are driven by the past realization of the range series.

3. Estimations of CCARR Model

3.1 Estimating of the Long Term Volatility Component

Observe that the range observed on short term period i and long term period t is given by,

$$R_{i,t} = \tau_t . g_{i,t} . \varepsilon_{i,t},$$

$$\overline{R}_t = \frac{\sum_{i=1}^{N_t} R_{i,t}}{N_t} = \frac{\sum_{i=1}^{N_t} \tau_t . g_{i,t} . \varepsilon_{i,t}}{N_t} = \tau_t \frac{\sum_{i=1}^{N_t} g_{i,t} . \varepsilon_{i,t}}{N_t},$$

and since

$$\frac{\sum_{i=1}^{N_t} \left(E\left(E\left(g_{i,t} . \varepsilon_{i,t} \mid \mathbb{F}_{i-1,t} \right) \right) \right)}{N_t} = 1,$$

we can conclude that $\overline{R}_t \approx \tau_t$.

Therefore, long term unobserved volatility component can be estimated using mean range for the given fixed long term period.

3.2 Parameter Estimation of the CARR Model

In this section we will derive the log likelihood function for the proposed CCARR model. For this derivation, we will assume that model disturbance term $\varepsilon_{i,t}$ is independent and identically distributed as a lognormal distribution with mean $-\frac{\sigma^2}{2}$ and variance σ^2 (i.e. $\varepsilon_{i,t} \sim i.i.d\ LN\left(-\frac{\sigma^2}{2}, \sigma^2\right)$) where $\sigma^2 > 0$. The reason for this assumption is empirical evidence we gathered from the two data sets we will analyze in Section 4. However, other distributions may also be utilized. Under our assumption of the lognormal distribution,

$\mathrm{E}(\varepsilon_{i,t}) = 1$ and $\mathrm{var}(\varepsilon_{i,t}) = \exp(\sigma^2) - 1$. Long term disturbance term η_t is assumed to be independently and identically distributed as a lognormal distribution with mean $-\frac{\nu^2}{2}$ and variance ν^2 (i.e. $\eta_t \sim i.i.d\ LN\left(-\frac{\nu^2}{2}, \nu^2\right)$) where $\nu^2 > 0$. Hence the $\mathrm{E}(\eta_t) = 1$ and variance $\mathrm{var}(\eta_t) = \exp(\nu^2) - 1$. Further we assume that $\varepsilon_{i,t}$ and η_t are independent.

We consider equation $(2), (3)$ and (4) to obtain the following results:

$$\begin{aligned}
&R_{i,t} = \tau_t . g_{i,t} . \varepsilon_{i,t}, \\
&\ln(R_{i,t}) = \ln(\tau_t . g_{i,t} . \varepsilon_{i,t}), \\
&\ln(R_{i,t}) = \ln(\tau_t) + \ln(g_{i,t}) + \ln(\varepsilon_{i,t}), \\
&\ln(R_{i,t}) = \ln(\omega_t) + \ln(\eta_t) + \ln(g_{i,t}) + \ln(\varepsilon_{i,t}).
\end{aligned}$$

Since $\varepsilon_{i,t}$ and η_t are lognormal distributions, $\ln(\varepsilon_{i,t})$ and $\ln(\eta_t)$ are normal distributions, and $\ln(\varepsilon_{i,t}) + \ln(\eta_t)$ is normally distributed with mean $-\frac{\theta}{2}$ and variance θ where $\theta = \sigma^2 + \nu^2$.

Then the conditional distribution of $R_{i,t}$ given $\mathbb{F}_{i-1,t}$ is expressed as:

$$f(R_{i,t} \mid \mathbb{F}_{i-1,t}, \Phi) = \frac{1}{\sqrt{2\pi\theta} R_{i,t}} \exp\left(\frac{\left(\ln(R_{i,t}) - \ln(\omega_t) - \ln(g_{i,t}) + \frac{\theta}{2} \right)^2}{-2\theta} \right).$$

Here $\Phi = (\alpha, \beta, \gamma_0, \gamma_1, \delta, \sigma^2, \nu^2)$ is the parameter vector.

Thus, the conditional log likelihood function can be derived as follows:

$$\begin{aligned}
&L(\Phi \mid R) = \prod_{\forall i,t}^{T} f(R_{i,t} \mid \mathbb{F}_{i-1,t}), \\
&l(\Phi \mid R) = \ln(L(\Phi \mid R)) = \sum_{\forall i,t}^{T} \ln\left(f(R_{i,t} \mid \mathbb{F}_{i-1,t}) \right), \qquad (6) \\
&l(\Phi \mid R) = -\frac{1}{2}\left(\sum_{\forall i,t}^{T} \left(\ln(2\pi) + \ln(\sigma^2) + 2\ln(R_{i,t}) + \left(\frac{\left(\ln(R_{i,t}) - \ln(\omega_t) - \ln(g_{i,t}) + \frac{\theta}{2} \right)^2}{\theta} \right) \right) \right).
\end{aligned}$$

Note that the Maximum Likelihood Estimation (MLE) is employed to obtain the model parameters for the proposed CCARR model. To utilize the MLE method, initial parameter values must be obtained. Determination of these initial values is discussed in the following sub-section.

3.3 Initial Value Estimation

Firstly, we need to find an estimator for the unobserved long term volatility component $(=\tau_t)$. In this study the unobserved long term volatility component is estimated by using monthly mean value of daily price ranges as derived in Section (2.2). The long term volatility component $(=\tau_t)$ is then modeled by the using CARR (1, 1) process as given in the Equation (4). After fitting a CARR (1, 1) we can find the initial values for parameters $\gamma_0, \gamma_1, \delta$ and σ^2. Next we need to find the initial values for the model parameters in the short term volatility component model $g_{i,t}$. Let $R^*_{i,t}$ be the daily adjusted price range which is defined as follows:

$$R^*_{i,t} = \frac{R_{i,t}}{\tau_t} = g_{i,t}\varepsilon_{i,t},$$

where $g_{i,t}$ is given by,

$$g_{i,t} = (1-\alpha-\beta) + \alpha\frac{R_{(i-1,t)}}{\tau_t} + \beta g_{(i-1,t)}. \tag{7}$$

We next fit a unit CARR (1, 1) model to the adjusted daily price range and find the initial parameter values for α, β and ν^2.

4. An Empirical Analysis

4.1 The Data Sets

In this study we use two stock indices, namely, the Standard and Poor's 500 (S&P500) index of United States and the Financial Times Stock Exchange 100 (FTSE 100) index on the London Stock Exchange. The sample periods for both S&P500 and FTSE 100 start on January 4, 1990 and end on December 31, 2018. Daily values for opening price, closing price, high price, low price and adjusted price are reported over the span of the study period. The data set is downloaded from the "Yahoo Finance" from the web site "https://finance.yahoo.com/" by using "quantmod" package in R software. The data set is divided in to two samples where one sample spans from January 4, 1990 to December 29, 2017 and is used for the model parameter estimation and in-sample predictions .The out of sample predictions are done by using the sample from January 1, 2018 to December 31, 2018. The same sample separation procedure is carried for the both stock indices.

Table 1 presents the summary statistics of the daily price range series for S&P500 and FTSE 100 indices. The daily price range $\left(=R_{i,t}\right)$ of a given day i on a month t is obtained as given in Equation (2).

Table 1: Summary Statistics for Daily S&P500 and FTSE 100, January 04, 1990 – December 29, 2017 (In-Sample)

Summary Statistics	S&P500	FTSE 100
Mean	1.2524	1.2926
Median	1.0165	1.0525
Maximum	10.9041	10.7532
Minimum	0.1456	0.0762
Standard Deviation	0.9185	0.9042
Skewness	3.2012	2.8668
Kurtosis	18.7175	14.9450
Jarque-Bera (Probability)	115093 (0.0000)	75987 (0.0000)
Ljung-Box Q-22 (Probability)	41938 (0.0000)	42047 (0.0000)

The high values for Kurtosis indicate a strong deviation from the normal distribution. Both price ranges have large positive skewness and it is suggested that a positively skewed density should be used to model disturbance term. Jarque-Bera test statistics fall far from zero and have extremely low p-values (<0.0000) leading to a rejection of the null hypotheses that the data is normally distributed. The Ljung-Box test null hypothesis is that the time series data are independently distrusted. In this study time lags of 22 trading days, which is the approximate number of trading dates for a month, was used for the test. After 22 lags of sample autocorrelations being examined, the large test statistic values and very small p-values (<0.0000) conclude that the data exhibit a strong persistence in daily price range data. Time series plots for the daily price range data of S&P500 and FTSE 100 over the in-sample period are given in Figure 1 and Figure 2.

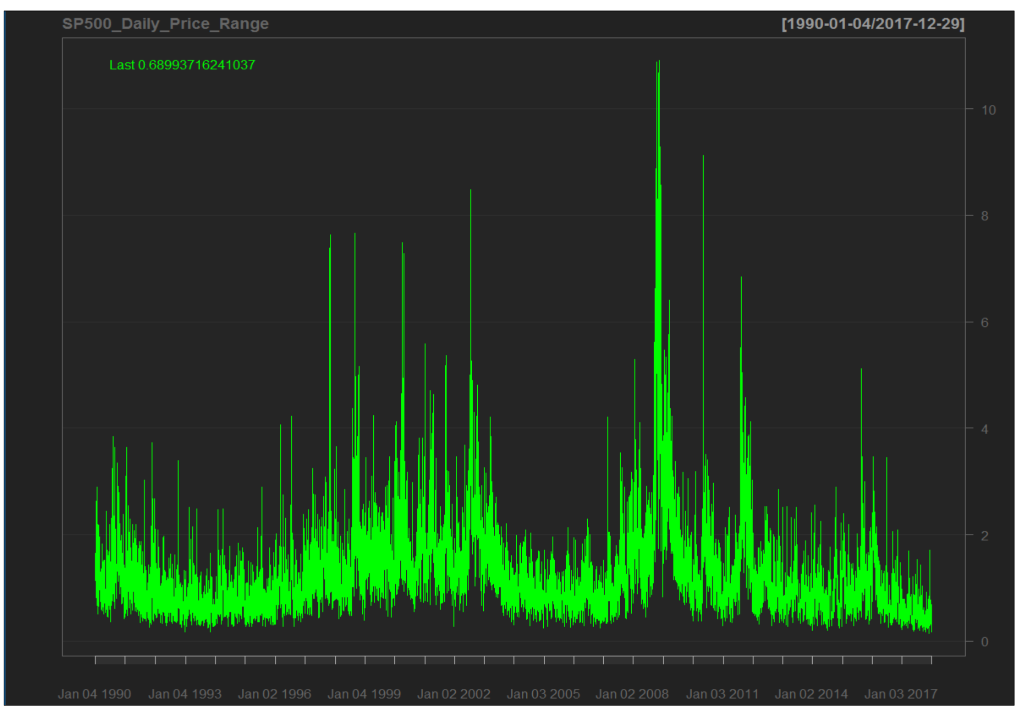

Figure 1: S&P500 daily price range from 01/04/1990 to 12/29/2017

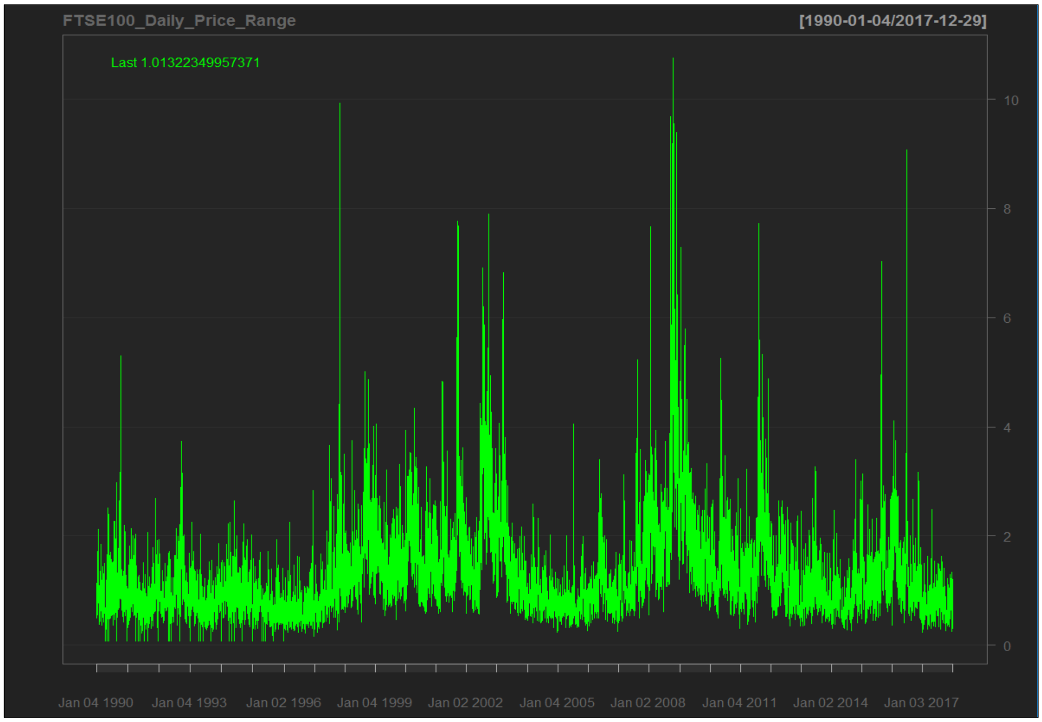

Figure 2: FTSE 100 daily price range from 01/04/1990 to 12/29/2017

Both graphs exhibit the same behavior over the period of study. Height of the spikes is an indication of price volatility and if the spikes are high during a certain period, then that period is considered to be highly volatile.

4.2 Estimation of CARR Model

Initially fitted a single component CARR model to daily price range data to explain price volatility over the study period. We assume the disturbance term ε_t in CARR $(p=1, q=1)$ model specified in the Equation (1) follows Exponential (ECARR), Weibull (WCARR) and Lognormal (LNCARR) distributions.

Table 2: Estimation of CARR (1, 1) Model Using Daily S&P500 Index Data

	ECARR (1,1)	WCARR(1,1)	LNCARR(1,1)
γ_0	0.0193 (0.0000)	0.0286 (0.0000)	0.0149 (0.0000)
γ_1	0.1679 (0.0000)	0.1780 (0.0000)	0.1653 (0.0000)
δ	0.8163 (0.0000)	0.7979 (0.0000)	0.8228 (0.0000)
AIC	15870.63	9868.95	8507.40

Table 3: Estimation of CARR (1, 1) Model Using Daily FTSE 100 Index Data

	ECARR (1,1)	WCARR(1,1)	LNCARR(1,1)
γ_0	0.0212 (0.0060)	0.0334 (0.0000)	0.0140 (0.0000)
γ_1	0.1715 (0.0000)	0.1953 (0.0000)	0.1677 (0.0000)
δ	0.8116 (0.0000)	0.7772 (0.0000)	0.8223 (0.0000)
AIC	16507.90	10050.41	9424.89

Since, daily price range data have large positive skewness, positively skewed distributions like Exponential, Weibull or lognormal should be used to model the data. According to the AIC values given in Table 2 and Table 3 LNCARR (1, 1) has lower AIC value, hence it fits better for the daily price range data for both stock indices.

4.3 Estimation of CCARR Model

The proposed CCARR process models daily price volatility by using short term and long term volatility components. In this study, day is considered as a short term time period while the month is taken as the long term period of interest. Initially we need to find an estimator for the unobserved long term volatility component, τ_t, and it is estimated by using monthly mean as previously derived in Section (3.1). Figures 3 and 4 presents the comparison of daily price range data and monthly mean as a long term volatility component for each of the indices. According to the Figures 3 and 4 monthly mean closely follows the long term changes in price volatility and it does a quite good job capturing the periods with high volatility.

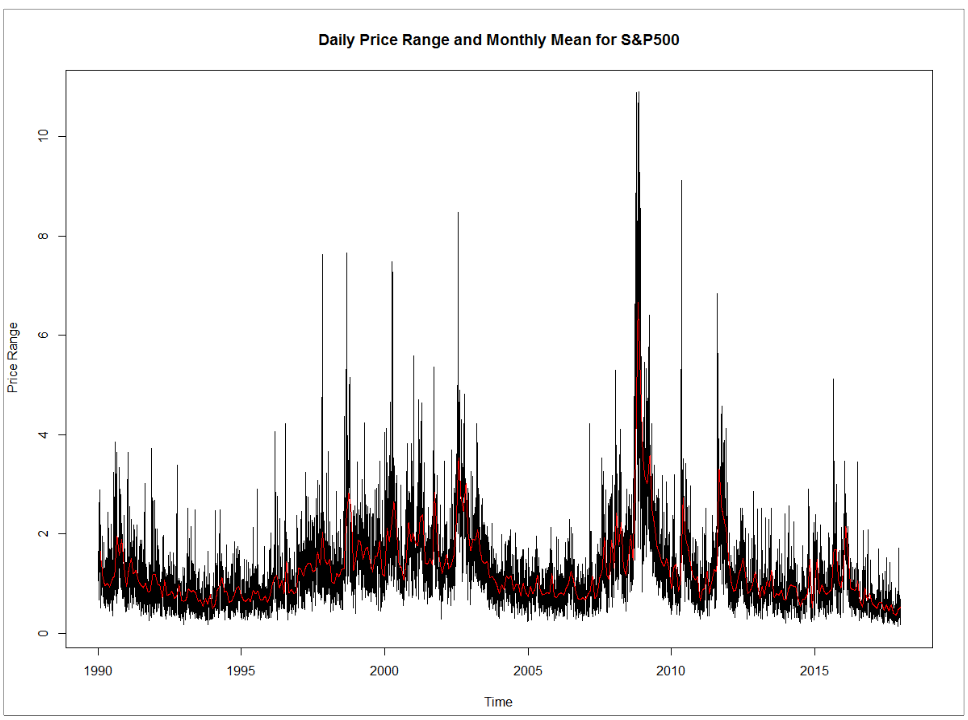

Figure 3: Daily Price Ranges (black) and Monthly Observed Mean for S&P500 (red) from 01/04/1990 to 12/29/2017

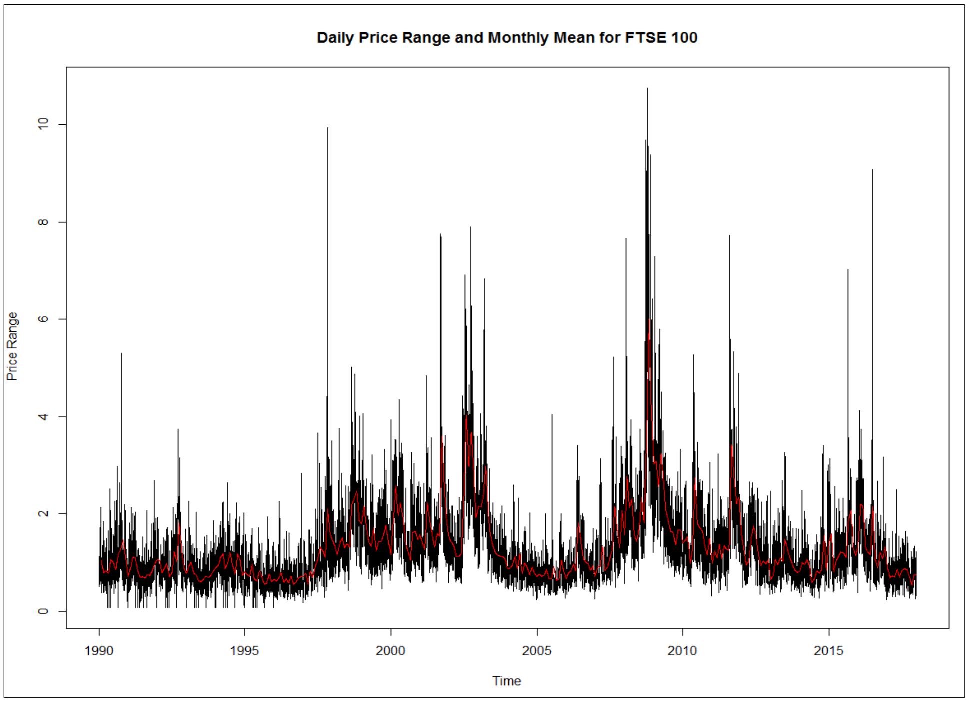

Figure 4: Daily Price Ranges (black) and Monthly Observed Mean for FTSE 100 (red) from 01/04/1990 to 12/29/2017

Based on the method describe the Section 3.3 we estimate the initial values for the both indices. Initial values for the S&P500 parameters are (0.20, 0.63, 0.27, 0.20, 0.70, 0.25) and that of the FTSE 100 are (0.13, 0.64, 0.25, 0.20, 0.68, 0.26). After determining the initial values for the model parameters, the partial log likelihood function (6) is maximized by using ‘nloptr’ package which is a nonlinear optimization algorithm in R. Table 4 presents the MLE results for the CCARR model.

Table 4: Estimation of the CCARR Model Using Daily S&P500 Index Data and FTSE 100 Index Data

	S&P500 Estimated Coefficients	FTSE 100 Estimated Coefficients
γ_0	0.0353	0.0359
γ_1	0.1595	0.2214
δ	0.8053	0.7536
α	0.1753	0.1855
β	0.7758	0.7566
θ	0.1820	0.1900
Ljung-Box Q-22	33.048 (0.0619)	24.442 (0.3245)

Ljung-Box Q test is used to test whether the residual series are independently distributed. Large values for Ljung-Box Q -22 test for the price range indicate that there is a significant persistence in the volatility. However the residual series for the fitted CCARR model demonstrate a significance reduction in Ljung-Box Q-22 statistics with p-values exceeding 0.05, suggesting the absence of serial autocorrelation up to 22 trading days.

4.4 Comparison between LNCARR Model and CCARR Model

In this section we test the in-sample prediction and out of sample forecasting ability of proposed CCARR model. To test the differences in prediction and forecasting power between CCARR and LNCARR, we conduct in-sample prediction and out of sample forecasting .In order to test how well the proposed CCARR models perform in extreme situations such as a recession period, we conduct the analysis for the period from December 2007 to June 2009 for S&P500 and April 2008 to June 2009 for FTSE 100.

In-sample prediction for the LNCARR is its conditional mean range and that of the CCARR is the product of estimated long term and short term volatility components. In order to compare the in-sample prediction and out of sample forecasting ability, we calculate AIC and mean-absolute error (MAE) statistics for both CCARR and LNCARR. MAE is calculated as follows:

$$\mathrm{MAE} = \left(\frac{\sum_{\forall i,t} \left(\left| MV_{i,t} - PV_{i,t} \right| \right)}{N} \right) \tag{8}$$

The unobserved real volatility is represented by $MV_{i,t}$ and here we use price range $(R_{i,t})$ as proxy variable for real volatility. Predicted values $PV_{i,t}$ are the fitted values for price range $(\widehat{R}_{i,t})$.

Table 5 presents the model comparison between LNCARR (1, 1) model and CCARR model. The CCARR model shows a better performance over the LNCARR model for all periods. During the full period of in-sample and the time of recession CCARR models have smaller MAE values and lower AIC values when compare to the LNCARR models for both stock indices. Further, diagnostic test results for the residuals indicates that they are independently and identically distributed in the CCARR model. However there is a clear rejection of the null hypothesis in LNCAR model where errors show high persistence in the residual of price range data. We calculated the one step out of sample forecasted values for both stock indices. Based on the out of sample statistics CCARR model dominates over LNCARR model with respect to MAE statistics.

Table 5: Model Comparison between LNCARR (1, 1) and CCARR for S&P500 and FTSE 100

Sample Period		S&P500		FTSE 100	
		LNCARR (1,1)	CCARR	LNCARR (1,1)	CCARR
In-Sample	MAE	0.422	0.420	0.415	0.410
	Standardized Residuals Q(22)	51.724 (0.0003)	33.048 (0.0612)	50.084 (0.0050)	24.442 (0.3245)
	AIC	8507.395	8466.001	9424.89	9375.982
Recession	MAE	0.88	0.87	0.93	0.92
Out of Sample	MAE	0.47	0.46	0.34	0.33

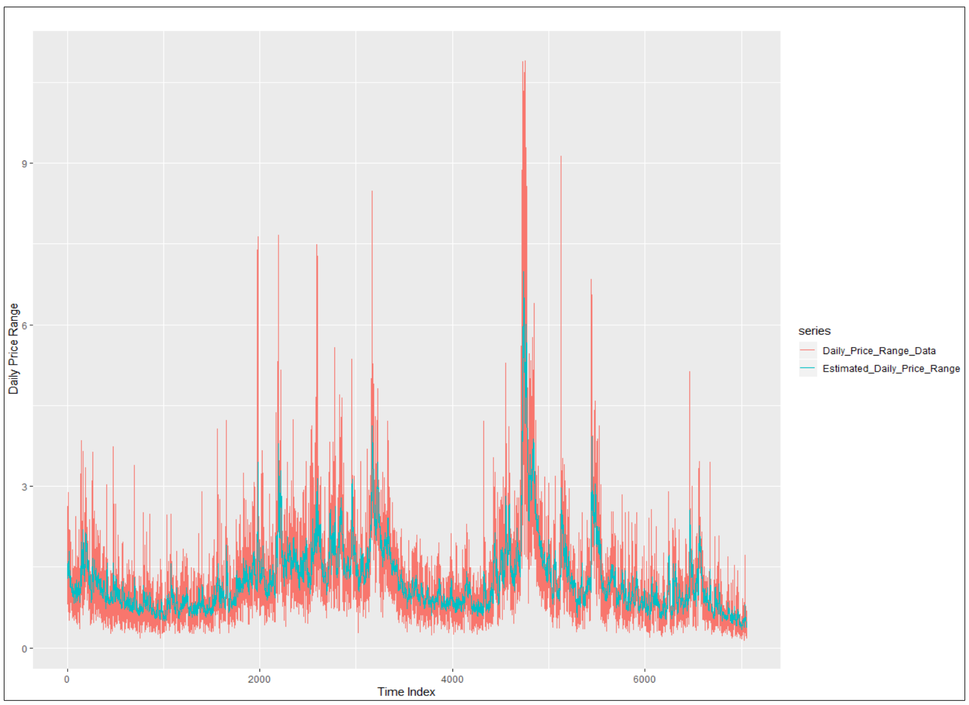

Figure 5: In-Sample Prediction by CCARR Model for S&P500

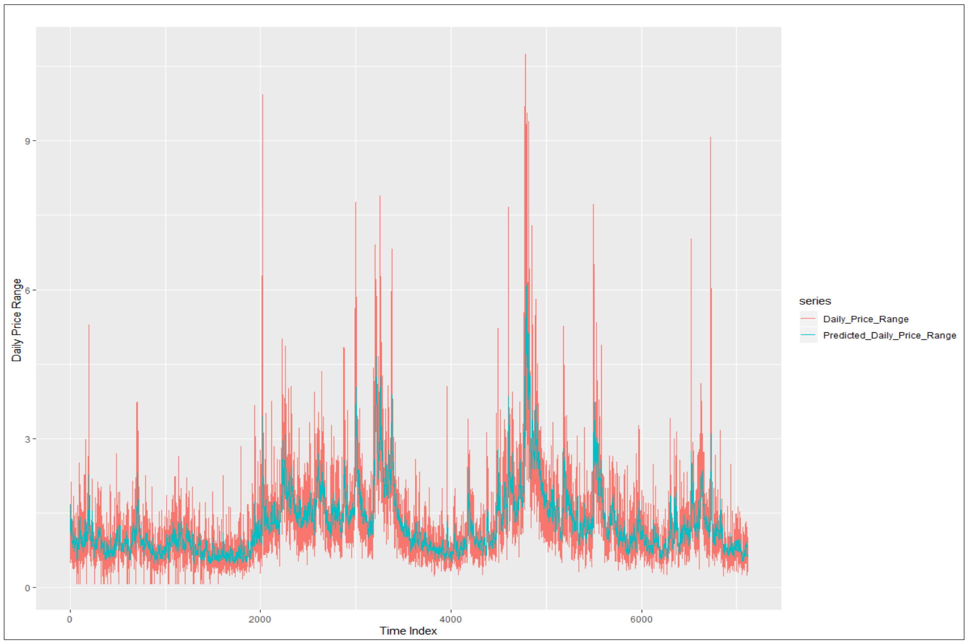

Figure 6: In-Sample Prediction by CCARR Model for FTSE 100

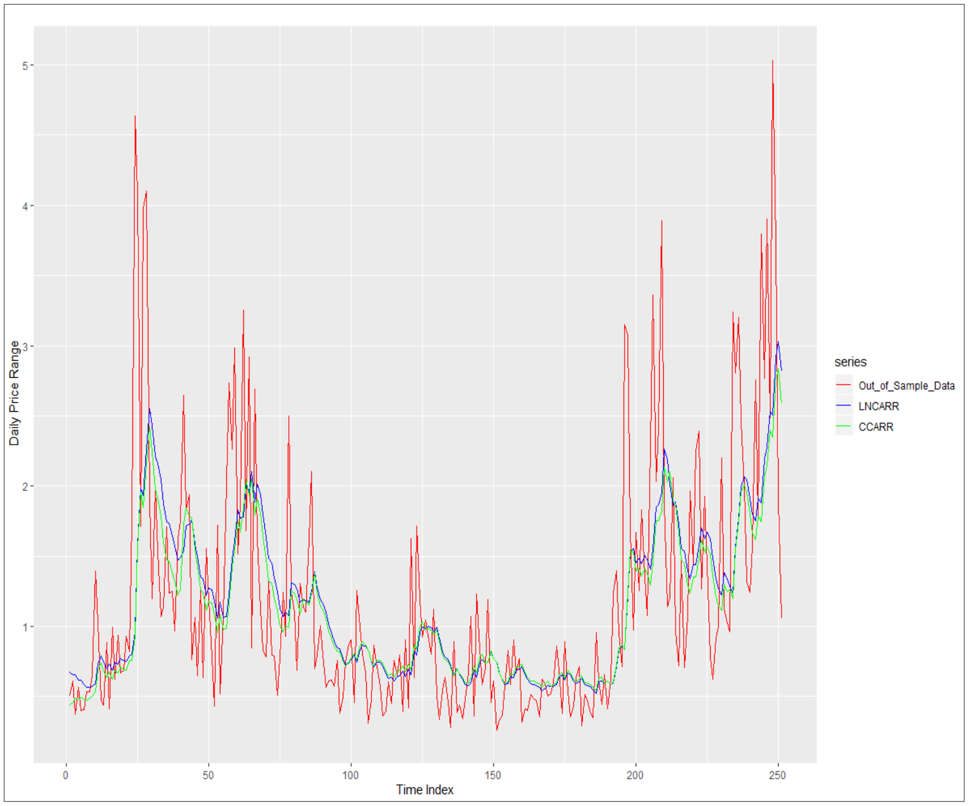

Figure 7: 1-step ahead forecasted value Comparison between LNCARR (1, 1) and CCARR for S&P500

Figure 8: 1-step ahead forecasted value Comparison between LNCARR (1, 1) and CCARR for FTSE 100

Figure 7 and Figure 8 show how well the proposed model performs in 1-step prediction. It can be seen in the figures that the CCARR model picks high volatility periods (high spikes) as LNCARR does however CCARR quickly capture the low volatile periods (short spikes) while LNCARR does not have the flexibly to adapt to such situations.

5. Conclusion

In this study we propose a composite range based model to estimate long term and short term volatility components. The proposed methodology models the long term volatility by using a stochastic process which itself exhibits conditional volatility. Further, both short term and long term volatility components are driven by the past realization of price range data. The empirical results based on MAE and AIC values show that the CCARR model is dominates the LNCARR model (which as selected based on performance out of other CARR models) in performance, especially during the recession periods. The proposed CCARR model does better than the single component LNCARR model with respect to the residual diagnostics as well.

Reference

Akgiray, V. (1989). Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts. *The Journal of Business*, 62(1), 55-80.

Bauwens, L., and Giot, P. (2000). The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks. *Annales D'Économie Et De Statistique*, (60), 117-149.

Beckers, Stan (1983). Variance of Security Price Return Based on High, Low and Closing Prices. *Journal of Business* 56, 97–112.

Ben Taieb, S., Bontempi, G., Atiya, A. F., & Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Systems with Applications*, 39(8), 7067-7083.

Bollerslev, Tim (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics* 31, 307–327.

Brandt, M.W. and C.S. Jones (2006), "Volatility Forecasting with Range-Based EGARCH Models". *Journal of Business and Economic Statistics*, 24, 470-486.

Chiang, M., Chou, R. Y., & Wang, L (2016). Outlier detection in the lognormal logarithmic conditional autoregressive range model. *Oxford Bulletin of Economics and Statistics*, 78(1), 126-144.

Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money, Credit and Banking*, Vol. 37, pp. 561-582.

Conrad, Christian, Custovic, Anessa and Ghysels, Eric, (2018), Long- and Short-Term Cryptocurrency Volatility Components: A GARCH-MIDAS Analysis, *Journal of Risk and Financial Management*, 11, issue 2, p. 1-12.

Engle, Robert (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50, 987–1008.

Engle, R. F., Ghysels, E., & Sohn, B (2013). Stock market volatility and macroeconomic fundamentals, *Review of Economics and Statistics*, 95(3), 776–797.

Engle, R. & Lee, G. (1999) .A long-run and short-run component model of stock return volatility. In Engle, R. & White, H. (eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, pp. 475–497. Oxford University Press.

Engle, R. & Rangel, J. (2008).The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies* 21, 1187–1222.

Engle, Robert, & Jeffrey Russell (1998). Autoregressive Conditional Duration: A New Model for Irregular Spaced Transaction Data. *Econometrica* 66, 1127–1162.

Fernandes, M., Mota, B. and G. Rocha (2005), “A multivariate conditional autoregressive range model,” Economics Letters, 86, 435-440.

Ghysels, E., Sinko, A., & Valkanov, R. (2006) MIDAS regressions: Further results and new directions. *Econometric Reviews* **26**, 53–90.

Hautsch, Nikolaus (2002) .Modelling Intraday Trading Activity Using Box-Cox-ACD Models, CoFE Discussion Paper, No. 02/05, University of Konstanz, Center of Finance and Econometrics (CoFE), Konstanz, http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-7845.

Nguyen, Duc Khuong & Walther, Thomas (2017): *Modeling and forecasting commodity market volatility with long-term economic and financial variables.*

Parkinson, Michael (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business* 53, 61–65.

Quirós, J.L., & Izquierdo, J.D. (2011). Volatility forecasting with range models: An evaluation of new alternatives to the CARR model.

Swanson, S. (2017). Range-Based Component Models for Conditional Volatility and Dynamic Correlations (Master’s thesis). Retrieved from https://digitalcollections.trentu.ca/islandora/object/etd:549

Zhang, M.Y., Russell, J. R., & Tsay, R.S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. J. *Econometrics*, 104(1):179–207.

Zou, W. & Guo, M. (2014).Using CARR model and GARCH model to forecast volatility of the stock index: Evidence from China's Shanghai stock market. *2014 International Conference on Management Science & Engineering 21th Annual Conference Proceedings*, Helsinki, pp. 1106-1112.