Dynamics of Trading Volume and Stocks Return: An Empirical Study Based on CNX Nifty of National Stock Exchange

*Pushpendra Singh**

ABSTRACT

Return-volume relationships are of common interest as they may unearth dependencies that can form the basis of profitable trading strategies, and this has implications for market efficiency. This research paper intends to study the relationship between stock returns and their trading volume and to test the causality effects. It focuses on the 50 stocks of CNX Nifty which is a value-weighted stock index of National Stock Exchange of India. Three proxies of trading volume namely, numbers of transactions, total traded quantity (volume) and total Rupee value of the traded quantity (turnover) have been taken and the asymmetry in the relationship of returns and volume is tested through regression. The study also tries to find the best proxy for volume through granger causality. The results indicate that there is asymmetry in the relation between returns and volume and the best proxy of the volume is the turnover or the value of shares traded.

Key words: Financial Markets, Market Capitalization, Trading Volume, granger causality.

1. Introduction

In financial market, it is important to understand the relationship between price and volume as it helps in understanding the competing theories of dissemination of information flow and improving the construction of test and its validity in to the market as argued by Karpoff (1986, 1987). First, the empirical relation between returns and volume helps discriminate between competing theories on how information is disseminated in financial markets. Second, the return volume relationship is critical in assessing the distribution of returns themselves. Third, a better understanding of the statistical structure of volume and return can help explain technical analysis. The price-volume relation can also be used to validate two well-known Wall Street adages: (i) volume is relatively heavy in bull markets and light in bear markets, and (ii) it takes volume to make prices move.This study focuses on the 50 stocks of the S&P CNX Nifty and test the asymmetry in the relationship of returns and different proxies of volume.

The CNX Nifty, also called the Nifty 50 or simply the Nifty, is a stock market index and benchmark index for Indian equity market and is owned and managed by India Index Services and Products Ltd. (IISL), which is a joint venture between NSE and CRISIL (Credit Rating and Information Services of India Ltd).

IISL is India's first specialized company focused upon the index as a core product. IISL has marketing and licensing agreement with Standard & Poor's for co-branding equity indices. 'CNX' in its name stands for 'CRISIL NSE Index'.

The CNX Nifty stocks represent about 70.14% of the free float market capitalization of the stocks listed at National Stock Exchange (NSE) as on March 31, 2014.

The CNX Nifty index is a free float market capitalization weighted index. It is calculated by taking the equity's price and multiplying it by the number of shares readily available in the market. Instead of using all of the shares outstanding like the full-market capitalization method, the free-float method excludes locked-in shares such as those held by promoters and governments. The free-float method is seen as a better way of calculating market capitalization because it provides a more accurate

* Associate Professor, Finance, GL Bajaj Institute of Management & Research, Greater Noida.

reflection of market movements. When using a free-float methodology, the resulting market capitalization is smaller than what would result from a full-market capitalization method. The index was initially calculated on full market capitalization methodology. From June 26, 2009, the computation was changed to free float methodology. The base period for the CNX Nifty index is November 3, 1995, which marked the completion of one year of operations of NSE's Capital Market Segment. The base value of the index has been set at 1000 with a base capital of Rs 2.06 trillion.

Table 1: Portfolio Characteristics of CNX Nifty

Methodology:	Free Float Market Capitalization
No. of Constituents:	50
Launch Date:	April 01, 1996
Base Date:	November 03, 1995
Base Value:	1000
Calculation Frequency:	Real-time Daily
Index Rebalancing:	Semi-Annually

Table 2: Top 10 constituents by weightage

Company's Name	Weight (%)
ITC	9.29
INFOSYS	7.64
HDFC LTD.	6.99
RELIANCE INDUSTRIES	6.96
ICICI BANK	6.61
HDFC BANK	6.28
TCS	4.41
L&T	4.05
ONGC	3.02
SBI	2.93

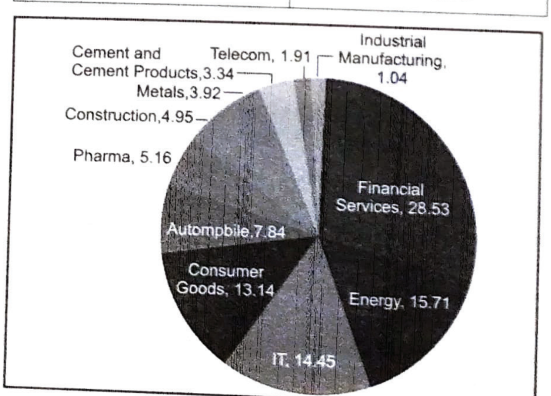

Figure 1: Sector Representation

2. Review of Literature

According to Karpoff (1986), price-volume relationship is important as it helps in understanding the dissemination of flow of information into the market. Using a vector auto-regression analysis on weekly data, we present a strong evidence of bi-directional relationship between volume and price change.There have been extensive empirical studies which support the positive relationship between price and volume.Crouch (1970) studied the relationship between daily trading volume and daily absolute changes of market index and individual stocks and found positive correlation between them. Harris (1987) used the number of transactions as a measure of volume and found a positive correlation between changes in volume and changes in squared returns for individual NYSE stocks.Hiemstra and Jones (1994) analyzed the bidirectional causality between trading volume and returns and found support for positive bidirectional causality between them for NYSE while Bhagat and Bhatiya, 1996 found strong one-directional causality from price to trading volume. Moosa and Al-Loughani (1995) examined the dynamic relationship between price and volume for four Asian stock markets excluding India and found a strong evidence for bi-directional causality for Malaysia, Singapore, and Thailand.Similar attention has also been given to the relationship between volatility and trading volume of a security. Many studies report the existence of ARCH effects in the time series of returns. Trading volume is usually considered as a proxy of information flow into the market. Any unexpected information affects both volatility and volume and so volatility and volume are expected to be positively related. Otavio and Bernardus (2006) investigated empirical relationship between returns and volatility and trading volume on Brazilian stock market and came out with a dynamic relationship between return volatility and trading volume. They also found that return volatility contains information about upcoming trading volume and vice versa. Darrat, Rahman and Zhong (2001) examined the correlation and lead-lag relationship between trading volume and volatility for all stocks in Dow Jones Industrial Average(DJIA). They used 5-minute intraday data and measured return volatility by using EGARCH model and found evidence of lead-lag relations between volume and volatility in DJIA stocks which is in accordance with the sequential information arrival hypothesis.Guillerm, Roni, Gideon, Jiang (2000) analyzed the relation between daily volume and first-order return autocorrelation for individual stocks listed on NYSE and AMEX and found that the cross-sectional variation in the relation between volume and return autocorrelation is related to the extent of informed trading in the manner consistent with the theoretical prediction.Joseph, Harrison and Jeremy (1999) analyzed the determinants of asymmetries in stock returns and developed a cross-sectional series of regression specifications and attempted to forecast the skewness in the daily returns of individual stocks. They found that negative skewness is more pronounced in the stocks that have experienced an increase in trading volume relative to trend over the past six months and positive returns over the prior thirty-six months.Brajesh and Priyanka() carried out an empirical study examining the relationship between returns and trading volume and volatility. They found that the number of transactions is a better proxy of information than the number of shares traded (volume) and value of shares traded (turnover).

3. Objective of the Study

This project is based on the similar lines of literature review and thus the objective of the study is defined as follows:

(i) To test the asymmetry in the relationship of stock returns and all three measures of volume for the 50 companies in the CNX Nifty Index.

(ii) To find the best proxy of volume from the three measures of volume taken namely, number of trades (transactions), volume (number of shares traded) and turnover (value of shares traded).

4. Research Methodology

4.1 Data Collection and Analysis tools

The daily data of adjusted prices and all three measures of volume of all 50 companies of Nifty 50 are collected from Capital market database. The data period is from April 2001 till March 2014. Companies which are listed after April 2001, their data period is taken from the listing date till March 2014.

1. Daily returns of all 50 companies are calculated. Descriptive statistics like Average, Skewness* and Kurtosis* of returns are calculated for each company. These descriptive statistics are calculated for each measure of volume also, namely number of trades, volume of trades and total value of trades. Skewness is an indicator of asymmetry in the distribution of any variable. It indicates the difference between the manner in which items are distributed in a particular distribution compared with symmetrical (or normal) distribution (Figure 2). The concept of skewness gains importance from the fact that statistical theory is often based upon the assumption of the normal distribution. A measure of skewness is, therefore necessary in order to guard against the consequence of this assumption. Excel uses the following formula for estimating skewness-

$$\frac{n}{(n-1)(n-2)}\sum\left(\frac{x_j - \bar{x}}{s}\right)^3$$

Skewness can be positive or negative-

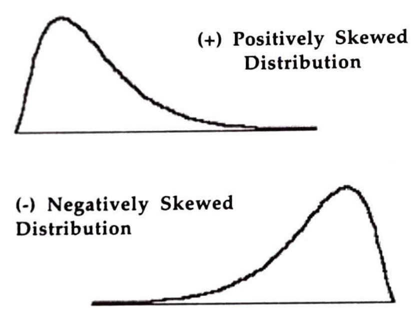

Figure 2

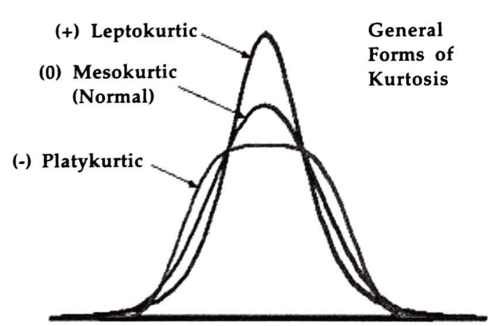

Figure 3

Kurtosis in Greek means *'bulginess.'* It refers to the degree of flatness or peakedness in the region about the mode of a frequency curve. It tells us the extent to which a distribution is more peaked or flat-topped than the normal distribution.If a curve is more peaked, it is leptokurtic and is said to have fatter tails and if it is more flat it is known as platykurtic and it is said to have thinner tails (Figure 3).

It is calculated through the following formula in excel-

$$\left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)} \qquad \left(\frac{Xj-\bar{X}}{s}\right)^4\right\}$$

$$-\frac{3(n-1)^2}{(n-2)(n-3)}$$

The more peaked curve is said to have a leptokurtic distribution and have fatter tails. The fatter tails are an indicator of ARCH effect*.

ARCH is Auto Regressive Conditional Heteroscedasticity which explains that the effect of the variable's own lags of the overall effect on the variable. Heteroscedasticity means that the residuals from the regression analysis are not equal and it can be improved by introducing autoregressive terms. This also explains the effect that returns in time series plots often show that "large changes tend to be followed by large changes, and small changes tend to be followed by small changes" which is known as Volatility Clustering.

Correlation statistics are then calculated between all three measures of volume for each company. The three measures of volume are then standardized* for further analysis. All the above work is done in MS Excel.

By standardizing a series, that series is converted to a series which has a mean of 0 and standard deviation of 1. This is done by the following formula-

$$Z = \frac{X - X}{\sigma}$$

This will help in an easy comparison of the different measures of volume and will reduce the units of measurement as the number of trades and volume are in numbers while the turnover is in money terms (Rs. Thousand).

2. Stationarity* of simple returns and the standardized measures of volume is tested through Augmented Dickey Fuller test (ADF test) in EViews. All the series are not stationary. A stationary time series is one whose statistical properties such as mean, variance, autocorrelation, etc. are all constant over time. The classical regression model i.e. ordinary least square (OLS) requires that the dependent and independent variables in a regression be stationary in order to avoid the problem of what Granger and Newbold (1974) called 'spurious regression' characterized by a high R^2, significant t-statistics but results that are without economic meaning. A stationary series is relatively easy to predict: we simply predict that its statistical properties will be the same in the future as they have been in the past. Another reason for trying to make stationary a time series is to be able to obtain meaningful sample statistics such as means, variances, and correlations with other variables. Such statistics are useful as descriptors of future behavior only if the series is stationary. For example, if the series is consistently increasing over time, the sample mean and variance will grow with the size of the sample, and they will always underestimate the mean and variance in future periods. So, if the mean and variance of a series are not well-defined, then neither is its correlation with other variables.A series is made stationary through a Unit Root Test. This is done by estimating the following OLS equations-

$$\Delta y_t = \alpha + \gamma y_{(t-i)} + \theta t + \sum_{i=2}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \alpha + \gamma y_{(t-i)} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \gamma y_{(t-i)} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t$$

The first equation includes both a drift term and a deterministic trend; the second excludes the deterministic trend; and the third does not contain an intercept or a trend term. In all three equations, the parameter of interest is ã. Here, null hypothesis is that ã=0 i.e. there is a unit root (the series is non-stationary) while the alternate being that there is no unit root (the series is stationary).

3. The relationship between trading volume and returns and their asymmetric nature is investigated from the following OLS equation-

$$V_t = \alpha + \beta_i 1 r_t 1 + \beta_2 D_t 1 r_t 1$$

Where, V_t is the standardized trading volume at time t, r_t is the return at time t and $D_t=1$ when $r<0$ and $D_t=0$ when $r>=0$. The parameter β_1 measures the relationship between returns and volume irrespective of the direction of return. The parameter $\hat{a}_2$ measures the asymmetry in the relationship. A statistically significant* value of β_2 will indicate that the relation between return and trading volume for negative returns is smaller than for positive returns. So, for each company there are three OLS equations for each measure of volume.

4.2 *Statistical Significance*

The level of significance is taken to be 5%. Under the null hypothesis, it is assumed that the OLS estimators are not significant while the alternate hypothesis being that the OLS estimators are significant. So, if the probability of an estimator is more than 0.05, it would imply that the OLS estimators are insignificant.

Fourthly, to find the best proxy of volume, the causality effects are tested between returns and volume through granger causality between returns and each measure of volume for each company. The number of lags is chosen on the basis of Schwarz Information Criterion (SIC) through EViews. This measure is chosen as it is seen to be the best for large sample size (Zahid Asghar, Irum Abid). (Performance of Lag Length Selection Criteria In Three Different Situations). AIC and FPE are better for sample sizes less than 60.

3.3 *Granger Causality*

The Granger causality test is a statistical hypothesis test for determining whether one time series is useful in forecasting another. Ordinarily, regressions reflect "mere" correlations, but Clive Granger, who won a Nobel Prize in Economics, argued that certain set of tests reveal something about causality.

A time series X is said to Granger-cause Y if it can be shown, usually through a series of t-tests and F-tests on lagged values of X (and with lagged values of Y also included), that those X values provide statistically significant information about future values of Y. It is based on the philosophy that if event A happens before event B, then it is possible that A is causing B. In other words, events in the past can cause events to happen today, future events cannot.

It is estimated using the following equations-

$$y_t = \beta_{1.0} + \sum_{i-1}^{P} \beta_{1,i} y_{t-i} + \sum_{j-1}^{P} \beta_{1.p+j} x_{t-i} + e_{1t} \quad (1)$$

$$x_t = \beta_{2.0} + \sum_{i-1}^{P} \beta_{2,i} y_{t-i} + \sum_{j-1}^{P} \beta_{2.p+j} x_{t-i} + e_{2t} \quad (2)$$

There can exist bi-directional causality between x and y, uni-directional causality from x to y or from y to x or there can be no causality. The null hypothesis in this case is that 'x doesn't granger cause y' against the alternate of 'x granger cause y.' The significance level is taken to be 5% again.

5. Results/Findings

After calculating the returns of all the companies and their descriptive statistics, it is found that only in case of 8 companies, the distribution of returns in negatively skewed although the skewness statistics are not very large. Rest 42 companies have their returns distribution as positively skewed which means there is more probability of having positive returns and is inconsistent with the results of Kumar and Singh because of their assumption of lognormal distribution of the returns. The negative skewness is only for ACC ltd., Bajaj Auto Ltd., Cairn India, Cipla, HCL Technologies, Infosys, JP Associates and Tata Steel.

In case of only three companies i.e. Ambuja Cements Ltd., Maruti Suzuki, and Sesa Goa, the kurtosis is less than 3 meaning the distribution of returns is platykurtic. Rest all companies have a leptokurtic distribution of returns. Leptokurtic distributions have fat tails implying there is an ARCH effect in the distribution of returns. It is also shown through the Durbin Watson statistic which is explained later in this section(Table A-1).

For the number of trades, the kurtosis value is below 3 for Axis Bank and Tata Steel, implying Platykurtic distribution. While for HDFC, JP Associates and Kotak Mahindra Bank, the distribution of number of trades are near normally distributed (Table A-2).

For the turnover (total value of trades) the kurtosis is less than 3 only for Tata Steel, implying platykurtic distribution. Rest all companies have kurtosis greater than 3. It implies that turnover is more volatile as it is only in one company that these are normally distributed (Table A-3).

For the volume (number of shares traded), only in case of JP Associates, the kurtosis is less than 3 while for Tata Steel and State Bank of India, the kurtosis is around 3 meaning the volumes are more or less normally distributed (mesokurtic). For every other company, the kurtosis value is very high meaning the distributions are Leptokurtic with flat tails (Table A- 4).

None of the companies have their volumes negatively skewed; all companies have their distribution of volume, no. of trades and turnover positively skewed.

The Unit Root test of the simple returns and the modulus of simple returns indicate that the returns are stationary at level [Integrated of order 0 i.e. I(0)]. As discussed in the methodology, ADF test is used to test the stationarity of the above variables (Table A-.6 to A-7).As discussed in the previous section, the volumes measures are standardized for further analysis. Even after standardizing the measures of volume, all measures for each company are not stationary at level. The following companies are non-stationary at level but stationary at first difference i.e. they are integrated of order 1 [I(1)] (refer appendix Table A-8 to A-10).

Table 3

Standardized Trades	Standardized Turnover	Standardized Volume
Axis Bank	Coal India	Coal India
Coal India	DLF Ltd.	IDFC
HDFC Ltd.	HDFC Bank	Maruti Suzuki
NTPC Ltd.	Power Grid	NTPC Ltd.
Power Grid		Power Grid

Recalling the earlier mentioned OLS Regression-

$$V_t = \alpha + \beta_1 |r_t| + \beta_2 D_t |r_t|$$

Where, V_t is the standardized trading volume at time t, r_t is the return at time t and D_t=1 when r<0 and D_t=0 when r>=0. The parameter β_1 measures the relationship between returns and volume irrespective of the direction of return. The parameter β_2 measures the asymmetry in the relationship. A statistically significant value of β_2 will indicate that the relation between return and trading volume for negative returns is smaller than for positive returns.

Analyzing company wise, the results indicate thatfor 28% of the companies (14 companies), dummy variables for each measure of volume are insignificant out of which Lupin Ltd. is one such company where all the variables i.e. constant, $\hat{a}_1$ and dummy for each measure of volume are insignificant. In 18% of the companies (9 companies), only one dummy is significant. In 30% of the companies (15 companies), all dummies are significant while in 24% of the companies (12 companies), only two dummies are significant. So, overall in 54% of the companies 2 or more dummies are significant while for 46% of the companies, the dummy variable is not significant.

Out of the significant coefficients' of dummies, 18 out of 23for number of trades, 27 out of 29for turnover and 23 out of 26 for volume have negative coefficients. So, overall 87% of the significant dummies have negative coefficients of $\hat{a}_2$. Only for Bharti Airtel Ltd., HDFC Ltd., Infosys Technologies, Reliance Industries Ltd. and Siemens Ltd. have their all significant coefficients positive. This shows that only for these companies, there is no asymmetry in the relationship of different proxies of volume and returns.

Table 4

Particulars	Number of Trades	Turnover	Volume
Insignificant Dummies	27	21	24
Significant Dummies	23	29	26
Negative Coefficient	18	27	23
Positive Coefficient	5	2	3

Analyzing variable wise,the dummy for the number of trades is significant for 46% of the companies (23 companies) while the dummies for the volume and the turnover are significant for 52% and 58% of the companies respectively. Overall, the dummies are significant for 52% of the companies combined. The difference of 2% lies because of different ways of analyzing the results (refer appendix Table A-11 to A-13).

The coefficient β_1 which measures the relationship of volume and returns irrespective of the direction of returns is positive for all the companies and all the proxies. The coefficients are significant in 47 out of 50 for number of trades, 47 out of 50 for turnover and 48 out of 50 for volume.

Also, the Durbin Watson statistic is much less than two for most of the companies signifying the presence of ARCH effect in the relation of returns and volumes(refer Appendix Table A-14).

The results of Granger Causality indicate mixed results.The results differ if we select the default number of lags and if we select the lags chosen on the basis of Schwarz Information Criterion (refer Appendix Table A-17).

On the basis of the default lags (2), while analyzing the results company wise, it is found that for 5 companies (Bajaj Auto, HDFC Ltd., ICICI Bank, Reliance Industries, Sun Pharma), there doesn't exist any causality between volumes and the returns while there exists bi-directional causality between returns and volumes for 9 companies. It is evident in 24 companies that the returns cause volumes (only those companies are selected where returns are causing volumes for atleast two proxies of volume) while it is only for 3 companies that the volumes are causing returns. So, overall the results indicate that in 66% of the companies (24+9=33 companies), returns are causing volume and in 24% of the companies, volumes are causing returns (3+9=12 companies) (refer appendix Table A-15).

Analyzing variable wise, simple returns cause number of trades in 68% of the companies while in 30% of the companies, number of trades cause simple returns. In 80% of the companies, simple returns cause turnover and in 52% of the companies, turnover cause simple returns. Simple returns cause volume in 66% of the companies, volume cause simple returns in 44% of the companies.

On the basis of lags specified by SIC, while analyzing the results company wise, it is found that for 4 companies (Bajaj Auto Ltd., HDFC Ltd., Reliance Industries, and Sun Pharma), there doesn't exist any causality between volumes and the returns while there exists bi-directional causality between returns and volumes for 8 companies. It is evident in 21 companies that the returns cause volumes (only those companies are selected where returns are causing volumes for atleast two proxies of volume) while it is only for 8 companies that the volumes are causing returns. So, overall the results indicate that in 58% of the companies (21+8=29 companies), returns are causing volume and in 32% of the companies, volumes are causing returns (8+8=16 companies) (refer appendix Table A-16).

Analyzing variable wise, simple returns cause the number of trades in 54% of the companies while in 38% of the companies, number of trades cause simple returns. Simple returns cause volume in 54% of the companies, volume cause simple returns in 42% of the companies. In 78% of the companies, simple returns cause turnover and in 44% of the companies, turnover cause simple returns.

The table 5 compares the results of the granger causality for both the categories of lags variable wise. The values show that percentage of the effects out of the total 50 companies.

From the above table 5, the results are better in the case of default lags but even after comparing both the results, it is clear that turnover cause returns in more companies than trades or volume does. Also, simple returns cause turnover in almost 80% of the companies while simple returns cause trades and volume in around 60% of the companies.

6. Conclusion

From the results of the regression analysis and granger causality, it is imperative that there is an asymmetry in the relationship of returns and different proxies of volume. It indicates that the relation between returns and different proxies of volume for negative returns is smaller than for

Table 5

Category	SIMPLE RETURNS Cause STAND TRADES	STAND TRADES Cause SIMPLE RETURNS	SIMPLE RETURNS Cause STAND TURNOVER	STAND TURNOVER Cause SIMPLE RETURNS	SIMPLE RETURNS Cause STAND VOLUME	STAND VOLUME Cause SIMPLE RETURNS
Default Lags	68%	30%	80%	52%	66%	44%
Specified Lags	54%	38%	78%	44%	54%	42%

positive returns.Although the volume is high in both bull and bear markets but the volume in bull market is more than in bear market. From the results of the granger causality, it is inferred that simple returns cause the different proxies of volume mostly(which is inconsistent with the study of Brajesh and Priyanka). For both the considerations of lag length, it can be concluded that the best proxy of volume is the total turnover or the value of shares traded in the market. This study can further be improved by applying ARCH/GARCH models and improving the results of the regression analysis.

REFERENCES

1. Ribeiro Otavio, Medeiros De(2006), The Empirical Relationship between Stock Returns, Return Volatility and Trading Volume in the Brazilian Stock Market, Working Paper
2. Darrat Ali F., Rahman Shafiqur, Zhong Maosen (2001) The Role of Futures Trading in Spot Market Fluctuations: Perpetrator of Volatility or Victim of Regret?, Journal of Financial Research
3. Otavio R. De Medeiros, Bernardus F. N. Van Doornik *"The Empirical Relationship between Stock Returns, Return Volatility and Trading Volume in the Brazilian Stock Market* (2005)".
4. Guillermo Llorente; Roni Michaely; Gideon Saar; Jiang Wang *"Dynamic Volume-Return Relation of Individual Stocks"*, The Review of Financial Studies, Vol. 15, No. 4. (Autumn, 2002), pp. 1005-1047.
5. Timothy J. Brailsford *"The empirical relationship between trading volume, returns and volatility* (1994)", Department of Accounting and Finance, University of Melbourne.
6. Joseph Chen, Harrison Hong, Jeremy C. Stein *"Forecasting crashes: trading volume, past returns and conditional skewness in stock prices* (2000)", NATIONAL BUREAU OF ECONOMIC RESEARCH, Cambridge.
7. Craig Hiemstra, Jonathan D. Jones *"Testing for Linear and Nonlinear Granger Causality in the Stock Prices-Volume Relation* (1994)", Journal of Finance.
8. C.W.J. Granger and P. Newbold *"Spurious Regressions in Econometrics* (1973)", University of Nottingham.
9. Brajesh Kumar, Priyanka Singh *"The Dynamic Relationship between Stock Returns, Trading Volume and Volatility: Evidence from Indian Stock Market."*
10. Jeff Fleming, Chris Kirby, Barbara Ostdick *"ARCH Effects and Trading Volume* (2005)", Rice University and Clemson University Working Paper.
11. Ali F. Darrat, Shafiqur Rahman, Maosen Zhong *"Intraday trading volume and return volatility of the DJIA stocks: A note"*.
12. Zahid Asghar, Irum Abid *"Performance of Lag Length Selection Criteria In Three Different Situations"*.
13. Sanjiv Bhatia, Sanjai Bhagat *"Trading Volume and Price Variability: Evidence on Lead-Lag Relations from Granger-Causality Tests"*.