Is trading volume an appropriate indicator to forecast the stock index?

Dr.T.Viswanathan, *SIBM, Bengaluru* , **Prof.Kannadas S,** *SDMIMD, Mysuru*

***Abstract*—Forecasting the share price is one of the most common decision to be taken by traders and investors. Investors use tools such as fundamental and technical analysis available to forecast the share price. However, these tools are arbitrary to some extent and cannot be tested for its forecasting accuracy. The accuracy of trading may be improved by using a model that provides unbiased estimates of future price. We apply predictive analytics tools to forecast Sensex by using trading volume. We apply Bai and Perron (1998) multiple break point test to determine the structural stability of stock volume and price for the period of observation. Both the models of pure and partial structural breaks are considered and run linear regression model to forecast the market index and checked for model fitness. The nature of joint association of all variables on the price is interpreted by using F-Statistic. We then apply ARIMA model to forecast the index. Residual diagnostics is applied to examine the model fitness. We found ARIMA model provides best forecasting results.**

Index Terms— Trading volume, Share price, Forecasting, ARIMA, Causal relationship

I. Introduction

Stock trading volume is not merely a quantity of shares or contracts that is traded in the market. Volume is expressed directly beneath the price chart of the respective stock illustrating the corpus of shares have been traded per period. Volume is also analyzed to show the trend of fluctuations over a period of time. Volume is commonly used by technical analysts to endorse the trends and chart patterns. The grip of any given price movement is measured primarily by volume. In fact, a geometric progression movement in stock price may not be pertinent if the volume doesn't progress in that pace. A predominant notion is that price oscillation is preceded by volume fluctuation. Technical analysts keenly observe stock volume trading whether the respective movements reverse which means that the volume changes can be predecessor to price changes. Volume is an indicator of the given financial assets have been traded. Being an appropriate tool but is often overlooked because it is perceived a nominal indicator. But the smart market participants aware how to utilize this indicator to increase their return. A mounting market will see rising volume in the beginning during that buyers require volume to keep pushing the prices higher. But increasing price and diminishing volume shows lack of preference in the market and it acts a warning signal. Volume has two faces i.e. when the prices increases or decreases, an increase in trading volume acts a confirmed rise or fall in price is real and that the price movement has strength. When price increases or decreases and there is a decline in volume, there might be a weak price move and less interest from traders. In the booming stock market, it is crucial to look at the relationship between returns/price and volume of trading.

be noisy but the ground work behind that includes the stock volume crusts and troughs.

The relationship between the volume and the returns/volatility in the stock markets are of most common interest as they pave way for lucrative arbitraging and speculating strategies which have insinuations for market efficiency (Yu, 2004). By closely monitoring the facets of the stock market, it can be perceived that the intertemporal causality relationship between the volume and the returns shacks the illumination on the informational efficiency of the market irrespective of the source of information (Ahmet E. Kocagil, 1998). A Uni-directional causal relationship strongly illuminates that the information diffuses from volume to return volatility, return to volume volatility thus suggests that the lagged values of a measure provide the predictability component of the other measure. Thus, the feedback relationship between return, volume and volatility suggests that the variables encompass domineering material, which helps traders to price the value content of new information which create strong impact on the mindset of stock market players. A positive causal relation from returns to volume is consistent with the positive response trading strategies of noisy traders, for whom the trading decision is based on historical stock price movements (Lokman, 2005).

There are four major reasons for having awareness about the price-volume relationship. Primarily it provides insight into the structure of the financial markets of the state. Secondly, the return-volume relation is crucial for keeping track of events for future reference during similar happenings and to incorporate the similar recovery measure in case of fall in the market stability. Thirdly, the bi-directional price-volume relation is precarious to the argument over the empirical distribution of speculative/arbitraging strategies. Finally, the price-volume relationship has a significant implication for the in-depth research into stock market for arriving at the stability i.e. to sustain the efficient market hypothesis (Karpoff, 1987).

We explicitly focus on the background related to relationship between returns, and trading volume in this paper i.e. the kind of relationship between trading volume and returns, the nature of relationship, if it does exist, then the direction and degree of relationship between them, whether there exist ARCH effects in the stock returns.

II. REVIEW OF LITERATURE

(Chandrapala, 2011) The study discovered the stick returns are positively correlated to the contemporary change in stock trading volume. Further, trading volume changes negatively to the stock returns. Paired sample t-test is applied in the study. It is concluded that stocks with low trading volume change outperform the stocks with high trading volume change in the subsequent period.

(Singh, 2015) This paper investigates the empirical relationship between return, volume and volatility dynamics of stock market by using data of the NIFTY index of NSE during the period from Jan 2007 to March 2014. Augmented Dickey Fuller (ADF) test, ARCH LM test, GARCH, EGARCH, TARCH, Granger Causality test are applied to find the returns of NIFTY granger cause volumes and volatility granger causes volumes in short run. The strong form of market efficiency does not hold since some private information exists that is not reflected in stock prices. This study also detects one-way causality from return to volume that is indicative of noise trading model of return volume interaction in this market.

(Singh G. S., 2016) This article examines the causal relationship between intraday return and volume by using 1-minute intraday data of 35 stocks of S&P CNX Nifty index during the period from April 2007 to March 2011. Schwarz information criterion (SIC), vector autoregression (VAR) test, Augmented Dickey Fuller (ADF) tests are applied. This study finds evidence of significant causal and lead–lag relations between the intraday return–volume associations for selected stocks. These findings reveal strong indication of unidirectional and bi-directional causality, thus supporting the sequential information arrival hypothesis (SIAH) which suggests that lagged values of volume provide the predictability component of current return and vice versa.

(Vasishth, 2015) Data are employed from January 1998 to December 2011 for Brazil, India, South Africa and South Korea, Indonesia and China. Portfolios are formed on the basis of past information on prices and/or volumes. Unrestricted and risk adjusted returns for sample portfolios are analyzed. The risk models employed in this study are Capital Asset Pricing Model (CAPM), Fama-French (F-F) Model and Fama-French augmented models. Price momentum patterns are observed for Brazil, India, South Africa and South Korea, while there are reversals in Indonesia and China. Low-volume stocks outperform high-volume stocks for all sample countries except China. Further, volume and price based bivariate strategies do a better job than univariate strategies in case of India, South Africa and South Korea. The past price and volume patterns in stock returns are not fully explained by CAPM as well as the F-F Model. Price and volume momentum factors do play a role in explaining some of these return patterns.

(Chandra, 2012) The paper contributes some empirical evidence using three different measures of FII trading volume as proxy of FII trading behaviour, and its bi-directional relationship with Indian stock market returns. Bi-directional causality between net FII investment and Indian stock market return is observed. In general, the FIIs seem to be chasing the Indian stock market returns. It is found that FII trading behaviour resulting in heavy trading volumes may cause variations in stock market returns only in the very short-term, but afterwards, it is the stock market returns which cause changes in FII trading behaviour.

(Rajib, 2010) This paper discovers that the Equity derivatives are relatively new phenomena in Indian capital market. This paper extends and updates the existing empirical research on the relationship between futures price volatility and volume in the emerging Indian capital market using improved methodology and recent data set. ARMA-generalized autoregressive conditional heteroscedastic (GARCH) and ARMA-EGARCH models with generalized error distribution have been used. The paper finds evidence of leverage effect, which indicates that negative shocks increase the futures market volatility more than positive shocks of the same magnitude.

(Hsiao, 2004) This paper examines the causal relationship between returns and volume for four Taiwan related futures contracts and the co-movement between returns and volume. First, past returns could predict the volume for the SIMEX-TW contract; however, a significant lagged volume Granger-caused returns for TF contracts. Second, regarding trading period returns and volume, it was found that volume could be used to forecast returns for TE contracts. There exists a feedback relationship for TF contracts. Third, the study found evidence of significant transitory and permanent covariances except in the case of TX contracts when the close-to-close returns were considered. Fourth, results concerning the covariance between day-time returns and volume corresponded to rejection of the null hypotheses of a zero unconditional covariance (ZUC) and a zero-conditional covariance (ZCC); but only in the case of the TE contracts did the author find evidence of ZUC.

(Bakhtiar Moazzami, 2013)This paper examines the dynamic relationship between stock market trading volume and returns for four major stock markets: New York, Tokyo, London and Toronto using daily data covering March 1, 2003 to Nov. 1, 2012 period. The authors investigate the information content of volume for the stock returns. We find a positive contemporaneous relation between volume and absolute value of return in all markets. In addition, we find support for the proposition that lagged volume has predictive power for future absolute returns. We also investigate whether the 2008 market crash has had a significant impact on the relationship between the trading volume and return on all markets. The authors also found significant support for the proposition that trading volume has predictive power for stock returns for the full sample and the period before the 2008 market crash.

(Oral, 2012) This paper focuses on this relationship by assuming the Student's t and the Stable distributions for innovations. In this paper, GARCH and Threshold GARCH (TGARCH) models are applied on the Istanbul Stock Exchange National-100 Index with the purpose of analyzing the relationships between the volatility of stock returns and the trading volume. The results support strong leverage effects, where negative shocks have larger effect on volatility of ISE National-100 returns. The results also indicate that the trading volume significantly contributes to the volatility and indicate the strong leverage effects on volatility.

(Hui-Ching, 2014)The real estate markets in Asia have attracted significant investor attention as they have grown rapidly in recent years. Both local and foreign investors continue to display a strong appetite for Asian real estate investment projects. Given the different characteristics of listed real estate stocks, the purpose of this paper is to focus on the causal relations between the financial variables of these stocks. This financial analysis can help investors to understand the characteristics of listed real estate companies, provide implications for optimal asset allocation decisions, and also increase the predictability of portfolio returns. Autoregressive Conditional Heteroskedasticity (GARCH) model, Granger causality test is applied. The paper finds that there are positive contemporaneous relations between trading volume and both returns and absolute returns. The paper examines the causal relations between the financial variables, the evidence implies that current trading volume helps to explain the returns indirectly by leading return volatility; however, trading volume does not help to explain future returns directly. Extending the causality test to international markets, the listed real estate portfolios of the four Southeast Asian countries are found to be more closely correlated than those of the other three countries studied in the paper. Among the four Southeast Asian countries, Singapore, the only developed country, is found to play an influential role, its current financial variables having predictive power for the other countries.

III. RESEARCH METHODOLOGY

Data Source. The study is based on secondary data that includes trading volume and price of Bombay Stock Exchange for the period between April 2000 and March 2017.

Tools applied

Augmented Dickey Fuller Test

The time series data of money supply, WPI, CPI and GDP Inflator is taken for unit root test. The testing procedure for the ADF test is applied to the following model:

$$\Delta y_t = \alpha + \beta_1 t + \beta_2 t^2 + \gamma y_{t-1} + \phi_1 \Delta y_{t-1} + \cdots + \phi_{p-1} \Delta y_{t-p+1} + \varepsilon_t \quad (1)$$

Where:

- $\Delta$is the first different operator
- α is a constant
- β_1 is the coefficient on a time trend
- β_2 is the coefficient on a squared time trend

Granger Causality Test

The Granger Causality test is a statistical hypothesis test for determining whether one-time series is useful in forecasting another series. This was proposed by Granger (1969) and popularized by Sims (1972)

Steps involved in Granger Causality Test

Regress the first orders of spot price of Pepper with the future price for the period of observation. Assume a particular autoregressive lag length *p*, and estimate the following unrestricted equation by ordinary least squares (OLS):

$$x_t = c_1 + \sum_{i=1}^{p} \alpha_i x_{t-i} + \sum_{i=1}^{p} \beta_i y_{t-i} + u_t$$

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_p = 0$$

Conduct an *F*-test of the null hypothesis by estimating the following restricted equation by OLS

$$x_t = c_t + \sum_{i=1}^{p} \gamma_i x_{t-i} + e_t$$

Compare their respective sum of squared residuals.

$$RSS_1 = \sum_{t=1}^{T} \hat{u}_t^2 \qquad RSS_0 = \sum_{t=1}^{T} \hat{e}_t^2$$

If the test statistic

$$S_1 = \frac{(RSS_0 - RSS_1)/p}{RSS_1/(T-2p-1)} \sim F_{p,\,T-2p-1}$$

is greater than the specified critical value, then reject the null hypothesis that *Y* does not Granger-cause *X*.

It is worth noting that with lagged dependent variables, as in Granger-causality regressions, the test is valid only asymptotically. An asymptotically equivalent test is given by

$$S_1 = \frac{T(RSS_0 - RSS_1)}{RSS_1} \sim \chi^2(p)$$

IV. EMPIRICAL RESULTS

Bai Perron multiple break point test

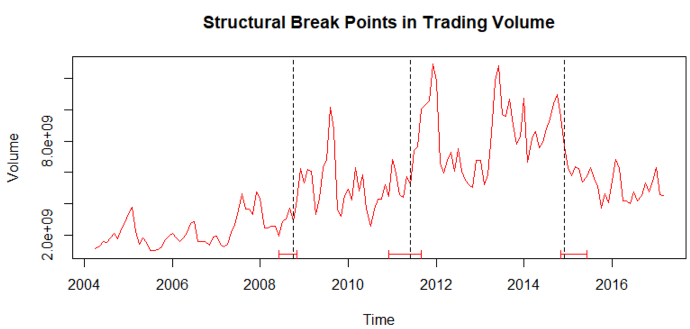
domestic economic factors and market sentiments causing bull or bear trend. These changes cause trend reversal in the series with unknown priori.

Use of linear regression model to predict the future requires time invariant coefficients. Any variation in coefficients provides only biased forecasts of the regression model. We apply Bai and Perron (1998) multiple break point test to determine the structural stability of stock volume and price for the period of observation. Both the models of pure and partial structural breaks are considered.

Table Bai Perron Multiple breakpoint test

Break test options: Trimming 0.15, Max. breaks 5, Sig. level 0.05				
			Break Dates	
Break Test	Scaled-F Statistic	Critical Value	Sequential	Repartition
0 vs.1 *	917.3500	8.58	2005M11	2003M05
1 vs.2 *	348.0382	10.13	2010M12	2005M11
2 vs.3 *	152.0860	11.14	2013M06	2010M12
3 vs.4 *	14.33311	11.83	2003M05	2013M06
4 vs.5 *	2.932939	12.25		
* Significant at 0.05 level.				

Period	Coefficient	Std Error	t-Statistic	Probability
Apr 2000 – Feb 2003	0.383727	0.001131	339.1932	0.0000
March 2003 – Oct 2005	0.390911	0.001210	323.1204	0.0000
Nov 2005 – Nov 2009	0.422096	0.000828	509.6514	0.0000
Dec 2009– March 2014	0.438577	0.001196	366.7264	0.0000
April 2014 – March 2017	0.451183	0.000961	469.5968	0.0000
R - Squared	0.961625			
Adjusted R Squared	0.960853			
Akaike Info criterion	-0.980547			



We run the linear regression model by taking stock volume as the regressor and price as regressand. The stability of the model is checked by applying Bai Perron (2003) sequential break point test. The model is applied to determine the number of structural breaks in the linear regression model. Sub Wald statistic is applied to determine the number of unknown breaks and break points. The null hypothesis is set as there are L number of breaks as against the alternate hypothesis of L+ 1 breaks. The stability of the model is tested at 5 % significance level. The results of multiple break point test is shown in Table. The model indicates 4 break points in stock volume during the period 2003–M02, 2005-M10, 2009-M11 and 2014-M03.

Table: Parameter Stability test for multiple break points

Application of linear regression model to forecast the stock price using volume requires the model to be best fit. The fitness of the model is determined based on its forecasting ability. We run five sets of regression by accommodating the following break points 2003M02, 2005M10, 2009M11 and 2014M03. The results indicate the coefficients are significant @ 5 % confidence level. The stability of the model is inferred based on adjusted R–squared that shows best fitness of the model.

Stability Diagnostics Test of Model Fitness

Linear Regression Model				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	2.553517	0.642341	3.975332	0.0001
Volume	0.264029	0.030138	8.760662	0.0000
DM1	0.208451	0.043088	4.837789	0.0000
DM2	0.761603	0.038510	19.77684	0.0000
DM3	0.372002	0.030583	12.16363	0.0000
DM4	0.337718	0.033072	10.21165	0.0000
R-squared	0.958545			
Adjusted R-squared	0.957498			
F-Statistic	915.6578			
Prob(F-Statistic)	0.0000			

Understanding the relationship between price and volume would facilitate the decision-making process of an investor. The change of volume is associated with price change. However, the nature of association exhibit various patterns like increase in volume and price, increase in volume but decrease in price and vice versa. These patterns go through various trends such as bull, bear and volatile markets. The predicting power of the volume to forecast the price is estimated by running a linear regression model. The issue of structural breaks in the series are addressed by introducing dummy variables for the multiple break periods. The equation is estimated by considering price as the regressand and volume & dummies as regressors.

Log Price = C(1) + C(2)*LOGV + C(3)*DM1 + C(4)*DM2 + C(5)*DM3 + C(6)*DM4

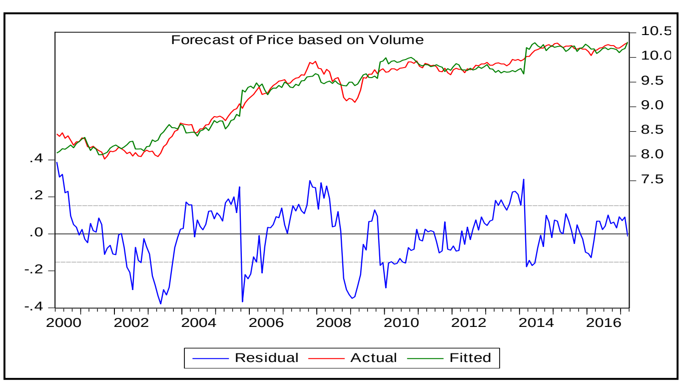
Log Price = 2.55 + 0.264*LOGV + 0.208*DM1 + 0.76*DM2 + 0.372*DM3 + 0.337*DM4

The model diagnostic is done by examining whether the independent dummies are individually significant to influence the price. . The nature of joint association of all variables on the price is interpreted by using F-Statistic. As the result of F-Statistic is significant @ 5% level, it in inferred that all variables jointly influence the stock price. All coefficient for the dummies are significant @ 5 % level. It shows running the regression with structural breaks provides a better forecast. The obtained adjusted R square of 95.74 % shows the best fitness of the model. All coefficient parameters are significant @ 5%. It indicates all variables independently influence the price

Residual Diagnostics

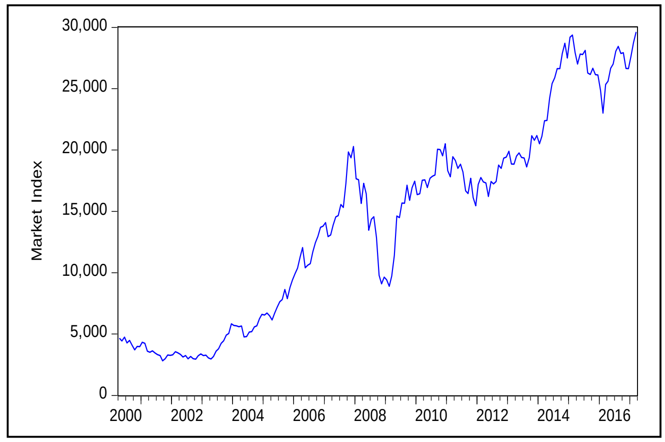
The historical time series of volume and price between 2000 and 2017 is considered to forecast the stock price. The residuals are the difference between the forecasted and actual price. The model is considered to be a best fit upon meeting certain futures. The figure shows the forecasted vs actual price and the obtained residuals.

Figure: Forecast of Price and Residual Diagnostics



The most important determinants of model fitness are the residuals should normally be distributed, no auto correlation and the variance of the residuals should be homoscedastic.

Table: Residual Diagnostics test of Volume and Price



Residual Diagnostics	Normality Jarque Bera test	Serial Correlation Breush Godfrey LM Test	Homoscedasticity Breusch Pegan Godfrey test
Model Parameter	3.232	167.488**	5.687**
Probability	0.198	0.0000*	0.0004*

** F- Statistic, * significant @ 5% confidence level

Application of linear regression and testing of the hypothesis requires the series to be normally distributed. The normality is tested using Jarque Bera test under the null hypothesis of normal distribution vs. non-normal distribution. Since the obtained probability is more than 5 %, the test is not statistically significant @ 5% level. It shows the residuals of liner regression are not normally distributed. Breush Godfrey LM Test is applied to examine whether the residuals are correlated with the lagged value of itself causing auto correlation. Presence of auto correlation may produce biased coefficients and inflated R^2 value. The LM test shows there is no auto correlation in the residuals. Breusch Pegan Godfrey test is used to test whether the variance of residuals is time invariant or homoscedastic. For a regression model is considered as best fit when $\mu=0$, σ=constant (Homoscedastic). The result of Breush Pegan test is significant @ 5 %

showing the variance of residuals is homoscedastic. Therefore it is inferred that the stock volume has the power to influence the share price in the market.

Granger Causality Test

We apply Engle Granger Causality test to examine the causal relationship between volume and price. If volume granger cause price, the past values of it should contain information to predict and vice versa. The test is used to forecast the series of one variable based on the other variable. If the price granger cause volume, the patterns of price should approximately get repeated in the volume series after some time lag. We run bivariate equation by taking volume and price as dependent and independent variable

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \ldots + \alpha_l y_{t-l} + \beta_1 x_{t-1} + \ldots + \beta_l x_{-l} + \epsilon_t$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \ldots + \alpha_l x_{t-l} + \beta_1 y_{t-1} + \ldots + \beta_l y_{-l} + u_t$$

Where y_t= Volume, x_t = Price β=coefficient

The reported F-statistics are the Wald statistics for the joint hypothesis: $\beta 1 = \beta 2 = \beta i = 0$

Null Hypothesis	**F –Statistic**	**Probability**
Causality between Price and Volume		
Price does not Granger Cause Volume	15.5884*	5.E-07
Volume does not Granger Cause Price	0.16226	0.8503

- Significant @ 5% level

The Granger causality test is applied to examine the nature of association between volume and price. The output of the test indicates whether the price is useful in forecasting volume and vice versa. The hypotheses are tested for both unidirectional and bidirectional causality. The significance of the result is examined using F Test. The causal relationship is tested pairwise between the two variables. The results of Causality test are significant at 5 % level for the Granger Cause between price and volume (15.5884). So the null hypothesis of price does not Granger cause volume is rejected. In the case of volume and price, the test is not statistically significant @ 5% level. It is concluded that volume does not Granger cause price. Therefore, it appears that Granger Causality runs from one way to price and volume and not the other way.

References

Bakhtiar Moazzami, B. D. (2013). Return and Volume and the 2008 Market Crash. The Journal of Applied Business and Economics, 33-37.

Chandra, A. (2012). Cause and effect between FII trading behaviour and stock market returns: The Indian experience. Journal of Indian Business Research, 286-300.

Chandrapala, P. (2011). The Relationship Between Trading Volume and Stock Returns. Journal of Competitiveness , 41-49.

Hsiao, J.-L. (2004). The Relationship between Returns and Trading Volume: Preliminary Evidence Concerning Taiwan Index Futures Contracts. Asia Pacific Management Review, 709-727.

Hui-Ching, S. H. (2014). The causal relationships between stock returns, trading volume, and volatility: Empirical evidence from Asian listed real estate companies. International Journal of Managerial Finance, 218-240.

Muhammad Irfan Javaid Attari, S. R. (2012). THE DYNAMIC RELATIONSHIP BETWEEN STOCK VOLATILITY AND TRADING VOLUME. Asian Economic and Financial Review, 1085-1097.

Oral, E. (2012). An empirical analysis of trading volume and return volatility relationship on Istanbul stock exchange national -100 Index. Journal of Applied Finance and Banking, 149-158.

Pramod Kumar Naik, R. G. (2018). The relationship between stock market volatility and trading volume: evidence from south Africa. Journal of Developing Areas, 99-114.

Rajib, P. C. (2010). Volatility persistence and trading volume in an emerging futures market -Evidence from NSE Nifty stock index futures. The Journal of Risk Finance, 296-309.

Sharma, R. R. (2016). Do trading volume and bid-ask spread contain information to predict stock returns? Intraday evidence from india . Asian Economic and Financial Review, 135-150.

Singh, G. (2015). The Empirical Investigation of Relationship between Return, Volume & Volatility in Indian Stock Market. IPE Journal of Management, 89-107.

Singh, G. S. (2016). The Empirical Investigation of Causal Relationship between Intraday Return and Volume in Indian Stock Market. SAGE Publications, 199-210.

Vasishth, S. S. (2015). Past price changes, trading volume and prediction of portfolio returns-Evidence from select emerging markets. Journal of Advances in Management Research, 330-356.