

Trading Volume and Stock Indices: A Test of Technical Analysis

Paul Abbondante

College of Business and Public Management, University of La Verne,
1950 Third Street La Verne, CA, 91750

Abstract: Problem statement: Technical analysis and its emphasis on trading volume has been used to analyze movements in individual stock prices and make investment recommendations to either buy or sell that stock. Little attention has been paid to investigating the relationship between trading volume and various stock indices. **Approach:** Since stock indices track overall stock market movements, trends in trading volume could be used to forecast future stock market trends. Instead of focusing only on individual stocks, this study will examine movements in major stock markets as a whole. Regression analysis was used to investigate the relationship between trading volume and five popular stock indices using daily data from January, 2000 to June, 2010. A lag of 5 days was used because this represents the prior week of trading volume. The total sample size ranges from 1,534-2,638 to observations. Smaller samples were used to test the investment horizon that explains movements of the indices more completely. **Results:** The F statistics were significant for samples using 6 and 16 months of data. The F statistic was not significant using a sample of 1 month of data. This is surprising given the short term focus of technical analysis. The results indicate that above-average returns can be achieved using futures, options and exchange traded funds which track these indices. **Conclusion:** Future research efforts will include out-of-sample forecasting to determine if above-average returns can be achieved. Additional research can be conducted to determine the optimal number of lags for each index.

Key words: Stock indices, trading volume, technical analysis, efficient markets hypothesis

INTRODUCTION

“Volume is an important indicator in technical analysis as it is used to measure the worth of a market move. If the markets have made strong price moves either up or down, the perceived strength of that move depends on the volume for that period. The higher the volume during that price move, the more significant the move” (Investopedia, 2010).

Trading volume is defined as the number shares bought and sold each day. It is an important indicator in technical analysis and is used to measure the strength of stock price movements either up or down. Technical analysis attempts to identify trends in stock price movements and then make investment recommendations for either going long or short for that particular stock.

The purpose of this study is to test this aspect of technical analysis using stock indices instead of individual stocks and using recent data. The events of the last ten years provide a comprehensive history of the US business cycle. In addition, smaller sample sizes can be selected within this period to be consistent with the short term investment horizon of technical analysis.

The emphasis in technical analysis and in previous academic research efforts has been on examining the relationship between trading volume and individual stock prices. The difference here is that five stock indices will be examined for the effect of trading volume on the level of the index.

Using trading volume to predict these stock indices can be profitable because futures, options and Exchange Traded Funds (ETFs) track these indices. Therefore technical analysis can be expanded to analyze movement in the stock markets as a whole.

The regression results will be a test of technical analysis and the Efficient Markets Hypothesis (EMH). In its weak form, the EMH states that previous stock prices cannot be used to forecast current and future stock prices. In its semi-strong form, the EMH states that public information cannot be used to forecast stock prices. Information about stock trading volume is available on a daily basis, so these results will either confirm the technical analysis view of trading volume or the EMH’s view on public information.

The five indices are the Dow Jones Industrial Average, the NASDAQ Composite Index, the NYSE Composite Index, the Standard and Poor’s 500 index and the Wilshire 5000 index.

The study is organized as follows. Previous research provides a review of prior research efforts. Definitions for the five indices are presented next. The specific equation to be tested is presented next. The data and regression results are presented following that. Conclusions and suggestions for further research conclude the paper.

Previous research: The majority of the previous research efforts have concentrated on the effect of stock prices on trading volume. In addition, these efforts have focused on individual stock prices or the stock prices of related companies in similar industries. There has been research on the effect of trading volume on stock prices. There has been less research on the effect of trading volume on stock indices.

Sabri (2008) analyze the relationship of trading volume and stock indices for eight of 15 Arab stock markets in the Arab Monetary Fund database. The study used monthly data from 1994-2006 for a total of 144 observations. The study consisted of using correlation coefficients between the volume and index for each stock market. The highest correlations were found in Saudi, Muscat, Ammann and Kuwait stock markets. In addition, the correlation coefficients were higher in the oil Arab states compared to the non-oil Arab states.

The concern for this study is the use of correlation coefficients instead of regression analysis. Correlation coefficients do not measure causality. The study would have more impact if more sophisticated statistical techniques have been used.

Ravindra and Wang (2006) examine the relationship of trading volume to stock indices in Asian markets. Stock market indices from six developing markets in Asia are analyzed over the 34 month period ending in October 2005. In the South Korean market, the causality extends from the stock indices to trading volume while the causality is the opposite in the Taiwanese market.

The remaining studies cited here analyze the relationship between trading volume and individual stock prices. The list here is to be considered representative rather than comprehensive.

Simon *et al.* (2001) found that stocks experiencing unusually high trading volume over a day or a week tend to appreciate over the course of the following month. It is argued that this high-volume return premium is consistent with the idea that shocks in the trading activity of a stock affect its visibility and in turn the subsequent demand and price for that stock.

Stickel and Verrecchia (1994) present the hypothesis that stock price changes are more likely to

reverse following weak volume support than strong volume support because price changes reflect demand for a stock and higher volume reflects a greater likelihood that the demand originates from informed rather than uninformed trades. Consequently, as volume increases, the probability that the price change is information driven increases. The evidence indicates that large price changes on days with weak volume support tend to reverse, at least partially, the next day. This volume effect is reinforced by, but is independent of, a bid-ask bounce effect. Returns do not reverse following days of strong volume support. In fact, large price increase with strong volume support tends to be followed by another price increase the next day (from the abstract).

You and Zhang (2009) examine trading volumes and stock price movement's days before and after the publication of the 10-K statements. Abnormal price movements were found to be indicators of future profitability based on the information in the 10-K report. Based on the results, they concluded that the 10-K reports can be used as indicators of future profitability

Majdosz *et al.* (2006) studied the effects of dividend announcements on both stock prices and trading volume on the German stock market. Stock prices react in the same direction as the dividend announcement. The increased trading volume around dividend announcement dates is attributed to portfolio adjustment by investors.

Michael and Starks (1988) also examine the relationship between stock price and trading volume using the Granger causality technique. Using data from June to August, 1981, they examine the effect of stock price movements on volume for a sample of firms listed on the NYSE. The sample consisted of transactions for 300 firms during a 49 day period which amounted to a sample of 2,989 observations.

Their conclusion is that information is processed by investors sequentially or in stages rather than simultaneously or all at once. This is especially true for quarterly earnings announcements.

MATERIALS AND METHODS

The indices: The Dow Jones Industrial Average consists of 30 stocks and is a price-weighted index. The NASDAQ Composite Index consists of all common stocks listed on the NASDAQ exchange. It does not include derivatives, exchange-traded funds and corporate bonds.

The Standard and Poor's 500 index consists of stocks selected by market value, liquidity and

industry. The index is market value-weighted. The individual companies are selected by a committee at Standard and Poor's.

The Wilshire 5000 Index now consists of more than 6,700 companies traded on various US stock exchanges. The index is calculated using a market capitalization approach (Investopedia, 2010).

Yahoo Finance (2001) provides a comprehensive list of indices on its web site.

The equation to be tested: The equation to be tested for each index is of the following form:

$$\text{Stock Index } t = a + b \text{ Volume } t-1 + c \text{ Volume } t-2 + d \text{ Volume } t-3 + e \text{ Volume } t-4 + f \text{ Volume } t-5$$

The theory to be tested is that the stock index at time t may or may not be influenced by the trading volume in the five previous days. The total of 5 days was selected because it represents the prior week of trading activity.

The equation is open-ended and the data will determine the outcome of the test. If the coefficients b to f are significant, the results would support the technical analysis view that trading volume conveys information about stock index momentum. If the coefficients are not significant, this would provide support for the Efficient Markets Hypothesis.

Multicollinearity: The independent variables are obviously highly correlated with each other since they are derived from the same data series. Multicollinearity does not affect the estimate of the coefficients b to f ; however it can result in exceptionally high standard errors. As a result, the t statistics for each coefficient will be low and the temptation will be to reject the independent variables as having any explanatory power. The overall fit of the model will not be affected so the F statistic becomes critical in assessing the explanatory power of the model.

The usual adjustments for multicollinearity include the following: delete one variable, use more data, or standardize the independent variables. Deleting one or more independent variables is not feasible because the intention is to examine the movement of a stock index and its relationship to trading volumes. Increasing the data set would not solve the problem since there are over 2,600 observations in the full sample. Standardizing the independent variables by selecting a starting date is, again, not consistent with the intention of testing the relationship of a stock index to trading volume. If there is a relationship, then the test of that relationship should be based on the daily movements of all variables.

Table 1: The data set

Index	Dates	Number of observations
Dow Jones	January 3, 2000-June 29, 2010	2,638
Industrial average		
NASDAQ	January 3, 2000-June 29, 2010	2,638
NYSE	January 3, 2001-June 29, 2010	2,385
S and P 500	January 3, 2000-June 29, 2010	2,638
Wilshire 5000	May 26, 2004-June 29, 2010	1,534

The data: The daily data for the five indices were retrieved from the Yahoo Finance web site. The web site provides a list of indices from the Dow Jones company, the New York Stock Exchange, the NASDAQ exchange, Standard and Poor's and others (Yahoo Finance, 2010). The data include time series for the following variables: date, open, high, low, close, volume and adjusted close. The adjusted close takes into account any dividends and stock splits. Since the emphasis in technical analysis is on short-term investing, daily results were used. Further details are included in Table 1.

Data for the Dow Jones Industrial Average, NASDAQ index and Sand P 500 index are easily available. Data for the NYSE Index do not begin until a year later. Data for the Wilshire 5000 Index began in October, 2001, however the volume data were not reported reliably until May, 2004.

Also, the Wilshire 5000 Index volume reported for January 18, 2010 was 49,000 shares. The volumes for the days before and after this date were 1,831,549,400 shares and 1,443,292,800 shares respectively. These amounts were verified at other financial web sites such as MSN Money and Market Watch. A review of the business news for this date did not indicate any unusual events; therefore no adjustment was made in this data series.

The data are available upon request.

RESULTS

There were a total of four regressions estimated for each index. They are differentiated by the sample size used for each. The sample sizes are the following.

Full sample-details are presented in Table 1: Recent trend-after examining graphs of the five indices, a trend upward began in March, 2009. Since technical analysis places greater emphasis on short-term investing, this month was chosen as the beginning of the next sample to avoid charges of bias against technical analysis. As a result, these results could be open to a charge of bias in favor of technical analysis.

Table 2: Full Sample: January 3, 2000-June 29, 2010

Index	Dow Jones	Nasdaq	NYSE	S and P 500	Wilshire 5000
Intercept	10.462.118*	2.079.0718*	6.668.692*	1.235.3172*	9.910.8595*
Volume					
t-1	-7.33E-08	-1.08E-08	3.7E-09	-1.44E-08*	1.798E-07
t-2	-9.04E-09	6.03E-09	1.4E-08	-4.32E-09	1.738E-07
t-3	4.405E-09	1.42E-08	2.29E-08	-1.94E-09	1.945E-07
t-4	1.331E-08	1.717E-08	3.86E-08	-1.33E-09	2.315E-07
t-5	6.99E-.08	3.627E-08	6.55E-08	2.505E-09	3.639E-07*
n	2.633	2.633	2.380	2.633	1.527
R ²	0.0016937	0.0014585	0.031806	0.0312613	0.0456943
F statistic	0.8914	0.7674	15.5976*	16.9547*	14.5658*

Note: *: Significant at the 95% confidence level

Table 3: March 2, 2009-June 29, 2010

Index	Dow Jones	NASDAQ	NYSE	S and P 500	Wilshire 5000
Intercept	12.398.404*	1.670.1312*	8.719.146*	1.336.4163*	13.517.245*
Volume					
t-1	-2.05E-07*	1.936E-08	-1.5E-07*	-2.26E-09*	-4.948E-07*
t-2	-6.95E-08	2.437E-08	-5.8E-08	-7.5E-09	-3.511E-07*
t-3	-6.46E-08	2.577E-08	-5.5E-08	-7.2E-09	-2.832E-07
t-4	-6.70E-08	2.847E-08	-5.1E-08	-7.17E-09	-2.797E-07
t-5	-1.24E-07	6.957E-08*	-8.9E-08*	-1.36E-08*	-2.461E-07
n	331.000	331.000	331.00	331.00	331.000
R squared	0.2668174	0.0341547	0.303528	0.261484	0.200942
F statistic	23.6546*	2.2985*	28.3275*	23.0143*	16.3458*

Note: *: Significant at the 95% confidence level

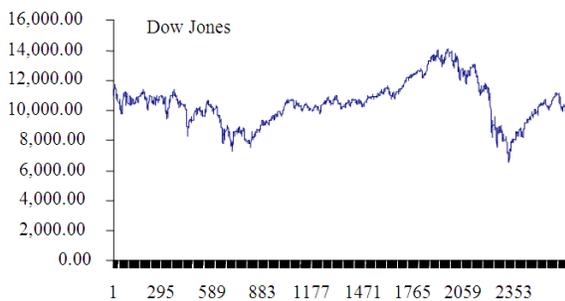


Fig. 1: The Dow Jones industrial average

The Fig. 1 of the Dow Jones Industrial Average is shown as an example for the decision to analyze the recent trends. Due to the large data set, individual dates are not graphed on the x axis (Fig. 1). The numbers represent the observation count. The point where the index begins its recent increase occurs in early March, 2009. Figure 1 show for the other similar results.

Year 2010-the sample begins with the first trading day of the year, January 4, 2010. The sample size is approximately 6 months.

Latest month-the sample begins with May 25, 2010 and contains 25 observations which is the minimum number for reliable results with a 5 day lag.

The results are presented in the Table 2 based on sample size.

Analysis of results in Table 2: Technical analysis focuses on short term investing. Therefore a regression

analysis using ten and a half years' of data should not show good results. This is demonstrated by the results for the Dow Jones and NASDAQ indices. The coefficients are not significant at the 95% level. The R squared is less than 1% and the F statistics are also not significant. These are strong indicators that trading volume does not explain movements in the individual index. On a preliminary level, these results do not support technical analysis.

The results for the NYSE, Sand P 500 and Wilshire 5000 indices were somewhat surprising. Given the large data set, there are a number of short-term trends included in the sample period. Despite this, the R squared results are significant as measured by the F statistic. This is supported by the low p values which were calculated for each equation.

Only two of the 15 coefficients in these three equations are significant. This is attributed to the presence of multicollinearity which was discussed above. In terms of value, the coefficients are significantly less than one. This can be attributed to the different dimensions of the independent and dependent variables. The independent variables are measured in billions and the dependent variables are measured in thousands and tens of thousands at their highest.

The fact that lagged values of trading volumes have some explanatory power for the three indices provides support for the technical analysis approach. As mentioned above, this was somewhat unexpected given the short term investment horizon of technical analysis and the long time period in the data set.

Table 4: January 4, 2010-June 29, 2010

Index	Dow Jones	NASDAQ	NYSE	S and P 500	Wilshire 5000
Intercept	9, 614.3924*	1, 881. 00297*	6,714.989*	1,020.1133*	10,808.111*
Volume					
t-1	-1.85E-08	6.712E-09	-4.2E-08	-2.28E-09	-4.274E-08
t-2	3.893E-08	3.693E-08	2.44E-08	4.552E-09	1.467E-07
t-3	2.362E-08	2.574E-08	9.11E-09	2.664E-09	7.663E-08
t-4	5.501E-08	3.766E-08	4.29E-08	6.78E-09	1.578E-07
t-5	8.107E-08	7.066E-08*	5.7E-08	1.012E-08*	2.271E-07*
n	118	118	118	118	118
R squared	0.2431447	0.2832629	0.12883	0.2374283	0.1510324
F statistic	7.1961*	8.8527*	3.3126*	6.9743*	4.0206*

Note: *: Significant at the 95% confidence level

Table 5: May 25, 2010-June 29, 2010

Index	Dow Jones	NASDAQ	NYSE	S and P 500	Wilshire 5000
Intercept	11,112.135*	2, 331.6402*	7,467.677*	1,168.7893*	11,825.884*
Volume					
t-1	-1.62E-07	-7.89E-08	-1.2E-07	-1.59E-08	-4.06E-07*
t-2	-7.34E-08	-9.6E-09	-5.7E-08	-6.96E-09	-5.83E-08
t-3	-3.54E-09	-3.352E-09	-8.7E-09	-1.16E-09	2.233E-09
t-4	5.414E-08	2.441E-08	4.92E-08	8.3E-09	9.088E-08
t-5	-1.31E-09	2.208E-08	1.03E-09	-1.29E-10	9.369E-08
n	20	20	20	20	20
R squared	0.3926761	0.346144	0.363296	0.3338897	0.4597118
F statistic	1.8104	1.4823	1.5976	1.4036	2.3824

Note: * Significant at the 95% confidence level

Analysis of results in Table 3: The shorter time period covered in these regressions would lead to the expectation that the technical analysis approach would perform better. It should be noted that the time period in this sample still extends for 16 months.

The results show that, again, the coefficients are extremely low and generally not significant. Once again, the explanation and the blame can be placed on multicollinearity. Caution must be used in interpreting the significance of the coefficients for the one period lag for four of the five indices and the five period lag for three of the five indices.

The overall results are encouraging. All five of the F statistics are significant. This is consistent support for the technical analysis approach. The interpretation is that an awareness of recent trading activity can lead to a better understanding of the movement in the overall index.

Analysis of results in Table 4: The next sample size consists of the first six months of 2010. The results are similar to those shown in Table 3. Again, the coefficients for the independent variables are low and only a few are significant.

While the F statistics are significant, they are consistently lower than those shown for the larger sample in Table 3. This calls into question the timing of the investment horizon using technical analysis. The expectation was that the results in Table 4 would be

better than those in Table 3. Since this is not the case, additional research using more detailed data sets is indicated.

Analysis of results in Table 5: Table 5 presents the results for the shortest time period, approximately one month. The results continue the trend seen in the Table 4 and 5. For the final time, the presence of multicollinearity results in higher standard errors for the coefficients and, consequently, lower t statistics. As a result, more reliance was placed on the R squared and F statistics.

Surprisingly, the model did not perform well for any of the five indices. As stated previously, the investment horizon for technical analysis is usually short term and yet, the shortest data set yielded the worst results. None of the F statistics is significant. This calls into question the timing of the investment recommendations from technical analysis.

DISCUSSION

The results presented in the tables above lead to the following conclusions. The best results were shown in Table 3 which covers the period from March, 2009 to June, 2010 which is 16 months. Knowing the previous week's volume does convey information about the level of these stock indices. As is commonly the case, trends in dependent variables are more easily subject to

explanation and forecasting once the relevant independent variables have been identified. For these indices, the longer the trend, the more predictable it becomes within a certain range.

These results indicate that technical analysis can be applied to longer investment horizons than it is usually the case. The number of observations required for reliable coefficient estimates preclude examining short term investment activities such as day trading. There are three topics for future research efforts based on the results presented here.

The first is to conduct out-of-sample forecasting and determine whether the results presented above can lead to above-average returns in the market. This was beyond the scope of this study and will be the focus of the next study in the series.

Editorial comment: if above-average returns can be achieved using this model, there will be no second study published. The author has no intentions of becoming rich and famous, just rich.

The second topic is to experiment with different data sets to determine if an optimal investment horizon can be identified. Different methods can be utilized in this effort. A popular index can be selected such as such as the Dow Jones or NASDAQ, then various models can be tested by varying the number of lags and varying the time period or number of observations. This could be accomplished by examining each of these independently and then, based on the results, selecting the optimal number of lags and the optimal number of observations.

The third topic is still selected lesser-known indices and perform the same analysis conducted here. Since these indices are not followed as closely as the ones selected here, it may be possible that a greater degree of market inefficiency exists thus leading to more accurate forecasts using lag trading volumes. The potential for above average returns is also limited here because it may not be possible to invest in derivatives or exchange traded funds linked to these lesser-known indices.

At the beginning of this project, the author's personal bias was in favor of the efficient markets hypothesis. The lesson to be learned here is to use caution in the dismissal of competing theories.

CONCLUSION

Future research efforts will include out-of-sample forecasting to determine if above-average returns can be achieved. Additional research can be conducted to determine the optimal number of lags for each index.

REFERENCES

- Investopedia, 2010. Volume description. Investopedia ULC.
<http://www.investopedia.com/terms/v/volume.asp>
- Michael, S. and L. Starks, 1988. An empirical analysis of the stock price-volume relationship. *J. Bank. Finance*, 12: 31-41.
- Majdosz, G., P. Majdosz and R. Mestel, 2006. Implications of dividend announcements for stock prices and trading volume of DAX companies. *Czech J. Econ. Finance*, 56: 58.
- Ravindra, K. and Y. Wang, 2006. The causality between stock index returns and volumes in the asian equity markets. *J. Int. Bus. Res.*, 5: 63-74.
- Sabri, N.R., 2008. The impact of trading volume on stock price volatility in the Arab economy. *J. Derivat. Hedge Funds*, 14: 285-298.
- Gervais, S., R. Kaniel and D.H. Mingelgrin, 2001. The high-volume return premium. *J. Finance*, 56: 877-919.
- Stickel, S.E. and R.E. Verrecchia, 1994. Evidence that trading volume sustains stock price changes. *Financ. Anal. J.*, 50: 57-67.
- Yahoo Finance, 2001. Major US indices. Yahoo.
http://finance.yahoo.com/indices?e=new_york
- You, H. and X.J. Zhang, 2009. Financial reporting complexity and investor under reaction to 10-K information. *Rev. Account. Stud.*, 14: 559-586.