

RELATIONSHIP BETWEEN STOCK MARKET INDICES AND GOOGLE TRENDS

Aysun Bozanta, Mustafa Coskun, Birgul Kutlu, Meltem Ozturan

Department of Management Information Systems, Bogazici University, Istanbul, Turkey

aysun.bozanta@boun.edu.tr, mustafa.coskun@boun.edu.tr, birgul.kutlu@boun.edu.tr,
meltem.ozturan@boun.edu.tr

Abstract: This study analyzes the relationship between stock market indices and Google trends. For this purpose, 12 stock market indices whose domestic market capitalization values are more than 1.5 billion USD were chosen. Their historical adjusted close values were collected in weekly basis for 2014 and 2015. Additionally, for the same period, weekly Google search data for those indices were collected. Results showed strong correlations between number of searches and rates of stock market indices. Correlation coefficients for relations are mostly positive which implies that the stock market indices are searched more while stock market values are increasing.

Keywords: stock market, stock market index, Google trends, search query

Introduction

Recent technological developments have connected people all around the world more than ever before and provide latest information about any progress instantly. People are likely to surf the web for information retrieval using search engines. Search engines have also been used for financial issues. Stock market movements are one of the popular topics that people especially investors want to monitor.

In June 2016, 70.16 percent of worldwide Internet users searched the web without bots using online search engine Google (NetMarketShare, 2016). It is claimed in literature that large scale online data can be the predictor of public perception on various contexts such as presidential elections (Tumasjan et al., 2010), commercial sales (Mishne & Glance, 2006; Choi & Varian, 2012), and influenza epidemics (Culotta, 2010). Thus, each rise or fall in stock markets may be reflected in Google searches by investors. The volume of queries in search engine can be used as a measure of investors' attention on the stock market (Da, Engelberg & Gao, 2011).

The volume of search queries for Dow Jones reached to eleven times greater than average search volumes when its volatility has become an almost record high of over 150% annualized on October 10, 2008 (Dimpfl & Jank, 2016). When Lehman Brothers disclosed bankruptcy in September 2008, the volume of Google searches on "Lehman" was 24 times higher than the average (Ramos, Veiga & Latoeiro, 2013). In November 2009, the number of submitted searches for "Dubai World" increased six-fold, when Dubai World rescheduled a debt due to accumulated losses (Ramos, Veiga & Latoeiro, 2013). After a year, when British Petroleum caused an environmental disaster in the Gulf of Mexico, the frequency of Google searches on "BP" became three-fold during this period (Ramos, Veiga & Latoeiro, 2013).

These facts show that financial actions may cause collective behavior in online environments. Therefore, the relationship between financial events and online behaviors has become an attractive topic for researchers. Various online data sources such as social media platforms and search engines were used to analyze different relationships. Twitter is one of the mostly used social media platform for this purpose. For instance, public mood states which were measured from Twitter were found as the predictive of changes in Dow Jones Industrial Average closing values (Bollen, Mao & Zeng, 2011). Zhang, Fuehres and Gloor (2011) collected six-month Twitter feeds and analyzed collective hope and fear for each day. It is found that emotional tweet percentage significantly negatively correlated with Dow Jones, NASDAQ and S&P 500, but displayed significant positive correlation to VIX (Zhang, Fuehres & Gloor, 2011).

Online news sources are also analyzed whether they are the predictor of financial metrics. Gilbert and Karahalios (2010) used a blog dataset consisting a stream of all LiveJournal posts and shown that a broad index of mood from an online community has novel predictive information about the stock market. In another study, a positive correlation was found between the daily number of mentions of a company in the Financial Times and the daily transaction volume of a company's stock both on the day before the news is released, and on the same day as the news is released (Alanyali, Moat & Preis, 2013).

Search engine queries are other sources of information that can be used as the predictor of fluctuations in stock market indices. It was found that weekly transaction volumes of S&P 500 companies were correlated with weekly search volume of corresponding company names (Preis, Reith & Stanley, 2010). In a sample of S&P 500 firms over the period 2005-2008, it was found that, over a weekly horizon, online search intensity reliably predicts abnormal stock returns and trading volume, and that the sensitivity of returns to search intensity is positively related to the difficulty with which a stock can be arbitrated (Joseph, Wintoki & Zhang, 2011). Bordino and his colleagues (2012) shown that daily trading volumes of stocks traded in NASDAQ-100 were also correlated with daily volumes of queries related to the same stocks. In another study, the volume of Google search queries were found as the predictor of foreign currency rates (Smith, 2012). Another interesting study in 2013 showed that there is a relationship between subsequent stock market moves and the historic data relating to politics or business from the search engine Google and online encyclopedia Wikipedia, as well as judgments from the service Amazon Mechanical Turk (Curme et al., 2014). In particular, they found that increase in search volume of the topics tends to precede stock market falls. Arditi, Yechiam, and Zahavi (2015) analyzed the relationship between stock-specific events and associated Google searches using daily data of 13 stocks from the Dow-Jones and NASDAQ100 indices, over a period of 4 trading years. They found a correlation between the magnitude of stock returns at the beginning of the period and the volume, peak, and duration of search generated during that period. That study revealed that losses improved the fit between people's search behavior and the extent of real-world events triggering the search (Arditi, Yechiam & Zahavi, 2015).

In this study, the relationship between stock market indices and Google trends is analyzed. Since previous studies have mostly focused on stock market indices in USA such as NASDAQ or S&P500, this study aims to analyze stock indices from different countries.

Materials and Methods

As it is mentioned before, the main aim for this exploratory study is to find out possible correlations between stock market indices and Google Trends for a period of two years, 2014 and 2015. For this purpose, the biggest stock markets were chosen according to their market capitalization. Thus, December 2015 market capitalization value of 1.5 million USD was chosen as the cut off point for market capitalization. Then, based on the World Federation of Exchanges Members statistics, 12 stock market indices from America, Europe and Asia were chosen with domestic market capitalizations listed in Table 1 (WFE, 2016).

Table 1: Chosen Stock Market Indices and Countries

Stock Market Index	Country	Domestic Market Capitalization Value (Dec, 2015)
NYSE	USA	17,786,787.40
NASDAQ	USA	7,280,752.17
Nikkei225	Japan	4,894,919.10
SSE Composite Index	China	4,549,288.00
LSE.L	England	3,878,774.20
Shenzhen Stock Exchange	China	3,638,731.30
Euronext 100	European Union	3,305,901.40
HKEX	Hong Kong	3,184,874.20
Deutsche Boerse AG	Germany	1,715,800.50
TMX Group Limited	Canada	1,591,928.60
SIX Swiss Exchange	Switzerland	1,519,323.50
S&P BSE SENSEX	India	1,516,216.70

After stating the stock market indices for the study, the weekly adjusted close values of them were collected for a chosen time interval of 104 weeks, from first week of 2014 to last week of 2015. Since the trending search data for terms can be collected from Google Trends web site (GoogleTrends, 2016) in weekly time interval, the stock market data was collected in same proportion.

There are several sources in the World Wide Web for collecting finance data but for being consistent and reliable in the analysis, the most popular finance data source Yahoo Finance (YahooFinance, 2016) was chosen. The historical prices of the chosen stock markets were collected for 104 weeks from first week of 2014 to last week of 2015.

After collecting weekly stock market index data of those 12 biggest markets, the trending search data of them were gathered from Google Trend web site. In this perspective, the keywords were stated carefully for those stock markets. Accordingly, first of all, the official web sites of those stock markets were examined. Then, the official abbreviations of those stock market indices were used as trending data keyword. In fact, Google also informs users about the type of keyword while searching. Therefore, Google trend keyword type “stock exchange” and related country were chosen during trend data search.

Moreover, it should be stated that Google gives the search trend data in 1-100 scale. This means Google standardizes the search trend of the chosen term for the chosen time interval into a scale of 1 to 100. The most popular (trending) week is given a value of “100” and the others are declared accordingly.

Lastly, the main purpose of this study is to examine the possible predictive effect of Google trend on stock market indices. For this purpose, Google trend and following week’s adjusted close values of stock market indices were gathered.

Results

Results showed the existence of correlation between adjusted closed values of stock market indices and their Google search volumes. Out of 12 separate correlations of different indices, 6 of them were positively correlated, 3 of them were negatively correlated and remaining 3 were not correlated with Google search volume data.

Table 2: Google Searching Trend vs. Adjusted Closed Values of Stock Market Indices

Stock Market Index	Country	Correlation Results		
		Pearson Correlation Coefficient	Significance (2 Tailed)	Direction of Relation
NYSE	USA	-0.076	0.445	-
NASDAQ	USA	-0.789	0.000**	Negative
Nikkei225	Japan	0.358	0.001**	Positive
SSE Composite Index	China	0.661	0.000**	Positive
LSE.L	England	-0.508	0.000**	Negative
Shenzhen Stock Exchange	China	0.211	0.033*	Positive
Euronext 100	Europe	0.452	0.000**	Positive
HKEX	Hong Kong	0.498	0.000**	Positive
Deutsche Boerse AG	Germany	-0.125	0.209	-
TMX Group Limited	Canada	-0.197	0.047*	Negative
SIX Swiss Exchange	Switzerland	-0.159	0.107	-
S&P BSE SENSEX	India	0.360	0.000**	Positive

* significant at 5% alpha level; ** significant at 1% alpha level

When domestic results sorted in terms of their correlation values, the most significant and the correlated one appears as NASDAQ. This result shows that NASDAQ index has the most impact on Google in terms of peak and trough points. Another finding for NASDAQ is that the correlation is negative which explains that when people search NASDAQ more, the adjusted close value decreases.

SSE Composite Index with correlation value 0.661 follows NASDAQ as the second mostly correlated index. However, its direction of correlation is positive which means when people search this index more than the adjusted close score is also increasing.

LSE.L which is London stock market index is negatively correlated with a correlation value of -0.508. HKEX, Euronext 100, S&P BSE SENSEX, Nikkei225, and Shenzhen stock market indices are significantly correlated with related Google search volume data in a positive manner. Lastly, TMX Group Limited is negatively correlated with a correlation value of -0.197.

Discussion and Conclusion

Search engine query data provides great insights into many areas such as presidential elections, commercial sales, and influenza epidemics. Financial issues have been also studied with the data from search engines. In this study, the relationship between adjusted closed values of stock market indices and the search volume of related indices has been investigated. For this purpose, correlation analysis between adjusted close values of 12 stock market indices, of which daily domestic market capitalization values are more than 1.5 billion USD, and Google search volume data for those indices has been applied for the two year period, 2014 to 2015 on weekly basis.

The results of the correlation analysis showed that out of 12 market indices, 9 of them were significantly correlated and 3 of them were not.

One of the uncorrelated indices was NYSE. This might be because of the global structure of NYSE where the investment funds in NYSE are more about global companies and firms rather than individual players. The other two uncorrelated indices were Deutsche Boerse and SIX Swiss Exchange. The reason for these results could be the effect of Euronext that has been formed as a merger of the Amsterdam, Brussels, and Paris Bourse, in order to take advantage of the harmonization of European Union financial markets and which has become the possessor of the market capitalization value in Europe. Hence foundation of Euronext might have decreased the attractiveness of these two indices and therefore has affected their correlations with the Google search volume data.

When the significantly correlated market indices are considered, it was seen that, in line with the previous studies (Curme et al., 2014; Arditi, Yechiam & Zahavi, 2015), there were negative correlations for indices that have dominancy in the stock market. These indices are NASDAQ, LSE.L and TMX Group Limited.

For other countries rather than USA, Canada and England, positive correlations were found between stock market indices and Google search volume data which implies when there is an increase in search volume this may indicate an increase in stock market indices.

The main limitation of the study is the weekly collection of stock market index value due to the weekly retrieval ability of Google search volume data. For this reason daily correlation analysis could not be applied.

In this study, Google search volume data was used as the predictor of stock market index which was measured one week after Google trends. In future studies, this analysis may be applied for other time horizons. On the other hand, only the names of stock market indices were used as the keywords in Google search queries. Other keywords such as names of companies related to those indices can also be used and their effect can be investigated.

References

- Alanyali, M., Moat, H. S., & Preis, T. (2013). Quantifying the relationship between financial news and the stock market. *Sci. Rep*, 3, 3578.
- Arditi, E., Yechiam, E., & Zahavi, G. (2015). Association between Stock Market Gains and Losses and Google Searches. *PloS one*, 10(10), e0141354.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.
- Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., & Weber, I. (2012). Web search queries can predict stock market volumes. *PloS one*, 7(7), e40014.
- Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88(s1), 2-9.

- Culotta, A. (2010). Towards detecting influenza epidemics by analyzing twitter messages. *In Proceedings of the first Workshop on Social Media Analytics* (115-122), ACM.
- Curme, C., Preis, T., Stanley, H. E., & Moat, H. S. (2014). Quantifying the semantics of search behavior before stock market moves. *Proceedings of the National Academy of Sciences*, 111(32), 11600-11605.
- Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5), 1461-1499.
- Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility?. *European Financial Management*, 22(2), 171-192.
- Gilbert, E., & Karahalios, K. (2010). Widespread worry and the stock market. *In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, ICWSM* (59-65).
- GoogleTrends. (2016). Google Trends. Accessed at July 6, 2016 from <https://www.google.com/trends/>.
- Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, 27(4), 1116-1127.
- Mishne, G., & Glance, N. S. (2006). Predicting movie sales from blogger sentiment. In AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs (155-158).
- NetMarketShare. (2016). Desktop Search Engine Market Share. Accessed at July 6, 2016 from <https://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0>
- Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: Insights from search engine query data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1933), 5707-5719.
- Ramos, S. B., Veiga, H., & Latoeiro, P. (2013). *Predictability of stock market activity using Google search queries* (No. ws130605).
- Smith, G. P. (2012). Google Internet search activity and volatility prediction in the market for foreign currency. *Finance Research Letters*, 9(2), 103-110.
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review*, 0894439310386557.
- WFE. (2016). World Federation of Exchanges. Accessed at July 6, 2016 from <http://www.world-exchanges.org/home>.
- YahooFinance. (2016). Yahoo Finance. 2016 Accessed at July 6, 2016 from <http://finance.yahoo.com>.
- Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through twitter "I hope it is not as bad as I fear". *Procedia-Social and Behavioral Sciences*, 26, 55-62.