10-2020

# Discovering the Relationship Between Investor Attention and Trade Volume in the Philippine Stock Exchange Using Google Search Analysis

Patrick Allan G. Morando

Isabel Louise D. Banico

Martin Stevens C. Ng

Riggs Mattieu M. Villardo

# Discovering the Relationship Between Investor Attention and Trade Volume in the Philippine Stock Exchange Using Google Search Analysis

DLSU-AKI Working Paper Series 2020-10-060

By: Patrick Allan G. Morando
Isabel Louise D. Banico
Martin Stevens C. Ng
Riggs Mattieu M. Villardo
*School of Economics*
*De La Salle University*

# Discovering the Relationship Between Investor Attention and Trade Volume in the Philippine Stock Exchange Using Google Search Analysis


**Patrick Allan G. Morando**
*De La Salle University,*
*2401 Taft Avenue, Manila, Philippines*
*E-mail: patrick_morando@dlsu.edu.ph*

**Isabel Louise D. Banico**
*De La Salle University,*
*2401 Taft Avenue, Manila, Philippines*
*E-mail: isabel_banico@dlsu.edu.ph*

**Martin Stevens C. Ng**
*De La Salle University,*
*2401 Taft Avenue, Manila, Philippines*
*E-mail: martin_ng@dlsu.edu.ph*

**Riggs Mattieu M. Villardo**
*De La Salle University,*
*2401 Taft Avenue, Manila, Philippines*
*E-mail: riggs_villardo@dlsu.edu.ph*



This study investigates the relationship between daily Google search volume and trade volume in the Philippine Stock Exchange. This is completed by employing a sample consisting of listed stocks in the Philippine Stock Exchange from 2015 to 2019 and testing the veracity of three different keyword categories in proxying investor attention. The empirical results indicate that when using company names and stock tickers as Google search keywords, there is a strong positive relationship between Google search volume and trade volume. This implies that the more investors search for a stock's company name or ticker, the more likely that there will be an increase in that stock's trade volume the following day. Additionally, this study finds evidence that when using stock tickers as Google search keywords, aggregate investor sentiment can influence this relationship, such that it is more significant during periods of high consumer confidence. The findings support the notion that daily Google search volume can measure investor attention and potentially serve as a supplementary stock indicator for retail investors in the Philippine Stock Exchange.


## 1. INTRODUCTION

Investors regard trade volume as a crucial technical indicator that can confirm trends, predict trend reversals, and determine liquidity (Westerhoff, 2006; Mahender et al., 2014). Trade volume is reported throughout the current trading day as often as once an hour; however, the reported hourly and end-of-the-day trade volumes are merely estimates. Actual and final trade volumes are reported the following day. Due to the relative inconsistency of reported trade volumes and uncertainty of future trade volumes, investors miss the opportunity to better reinforce their trading decisions through the trade volume indicator.

Given that trade volume is heavily influenced by investors' reaction towards their exposure to information (Barber & Odean, 2011), an alternative proxy to investor attention, such as internet search data, may have the capacity to help explain trade volume (Da et al., 2011; Bank et al., 2011; Bordino et al., 2012; Takeda & Wakao, 2014; Nguyen et al., 2020). Because there is already a large amount of literature on this topic that applies to developed financial markets, there is a dearth of its application in emerging financial markets, which are less informationally efficient and may possibly have different investor behavior than its developed counterparts. Therefore, there is a need to test further the veracity of internet search data as an accessible supplementary indicator that can help investors verify trade volume estimates and predict future trade volumes in emerging financial markets, such as the Philippines, where the majority of investor accounts are represented by retail investors and where there is rapid growth in the internet penetration rate and uniform use of a single search engine.

This study aims to provide retail investors in the Philippine Stock Exchange (PSE) with a supplementary indicator when trading in the stock market by testing the veracity of three different search keyword categories to verify the use of internet search data as a proxy for investor attention.

## 2. LITERATURE REVIEW

To many investors, trade volume is a crucial technical indicator that reinforces their trading decisions because of the indicator's evident correlation with stock returns (Campbell et al., 1993). Because of the attention-driven nature of trade volume, it has been shown that this indicator is closely correlated with investor attention, to the point that it is even considered a proxy for investor attention. However, it must be noted that investor attention typically refers to the attention given by retail investors due to their limited allocation of attention; thus, retail investors are the ones that are greatly prone to investment decisions that are attention-based.

In recent literature, investor attention proxies have been classified as either indirect or direct (Han et al., 2017). Indirect proxies of investor attention, such as trade volume, are a more traditional way of measuring investor attention than direct proxies. They are acceptable measures because they can explain market dynamics, and thus, they can help investors form and reinforce their investment decisions (Kaniel et al., 2012; Alanyali et al., 2013). However, indirect proxies fail to accurately and consistently indicate whether investors actually show interest in corresponding financial instruments or whether news presented by the media is truly noticed by investors (Barber & Odean, 2008; Gervais et al., 2001; Kaniel et al., 2012). This means that while investors can use indirect proxies to supplement their trading decisions, the information that indirect proxies give has the possibility of being inaccurate and untimely.

Direct proxies of investor attention, such as internet search data, aim to address the issues presented by indirect proxies. Direct proxies assume that investors interested in a certain security will also use internet search engines (e.g. Google Search) to gather information related to it. This gives direct proxies the advantage of encompassing investor attention in a more accurate and timely way. Previous studies have revealed that direct proxies are correlated with indirect proxies of investor attention, which opens the opportunity to use direct proxies as a supplementary indicator alongside indirect proxies to help investors better reinforce their investment decisions (Bank et al., 2011; Preis et al., 2013). However, direct proxies require a careful selection of keywords to properly capture investor attention, such that popular keyword categories are stock tickers, company names, and terms related to financial markets.

## 3. FRAMEWORK

***Overreaction and Underreaction Hypothesis***

The overreaction hypothesis and underreaction hypothesis reveal that market dynamics are heavily affected by reactions to new information and that there are different determinants that can affect how new information is reacted to (Stefanescu et al., 2012). This shows that different kinds of information can cause different kinds of investor reactions and that factors outside the actual information can also influence how the information is interpreted. In accordance with direct proxies of investor attention, it is assumed that investors will search for relevant information on companies that they receive news on. Therefore, the overreaction hypothesis and underreaction hypothesis can explain how the release of new information can trigger various magnitudes of internet search activity and movements in the market.

***Search Intent***

By analyzing the nature of the search intent, triggered internet search activity can then be classified to be informational, navigational, or transactional in nature. In the case of our field of study, the keywords used are all informational in nature (Broder, 2002). This means that search intent identification can be used to explain how investors will use the internet to obtain further relevant information on a certain stock that they initially received new information on.

***Signaling Theory***

Bordino et al. (2012) states that the internet search activity of individuals interested in certain stocks can be used as signals to anticipate the trading volume movements of the same stock. Therefore, when applying the signaling game in the context of direct proxies of investor attention, the retail investors who search for information on the internet are the signalers, the signal is the aggregate search volume, and the receivers are investors that decide to view the aggregate search volume. However, unlike the classical theory of signaling, the retail investors do not send signals intentionally, and there is no cost component to the signal. This shows how the signaling theory can explain the dynamics of retail investors through the examination of internet search activity.

***Sentiment Analysis***

To further add nuance to the dynamics of retail investors, we use sentiment analysis to examine whether sentiment affects internet search signals and its corresponding relationship with market movement. This can be done directly through keyword analysis or indirectly through the consumer confidence index (CCI) (Statman & Fisher, 2002). As previously mentioned, the field of our study uses keywords that are informational in nature; thus, they are neutral, meaning that they do not display positive or negative subjectivity. Because of this, the sentiment of our keywords cannot be directly analyzed. Therefore, we use the CCI to examine how sentiment affects the dynamics between internet search signals and market movements.

***Prospect Theory***

The prospect theory (Kahneman & Tversky, 1979) states that investment decisions are heavily influenced by the loss aversion phenomenon; thus, the theory may be used to explain how investors may become more cautious and be less likely to commit to trades when investor sentiment is low. Applying the prospect theory to direct proxies of investor attention, it can be suggested that internet searches are more likely to be translated to actual trades when investor sentiment is high, which suggests that internet search signals may possibly happen more frequently and have more significance when there are higher levels of investor sentiment.

## 4. METHODOLOGY

This study uses the Google search volume data of PSE listed stocks, obtained from Google Trends, as a proxy for investor attention from January 2015 to December 2019. The relationship between search volume and trade volume in the PSE is examined through the use of three search keyword categories, namely, company names (e.g. Jollibee Foods Corporation), stock tickers (e.g. JFC), and stock tickers—with the subsequently added word "stock" (e.g. JFC stock) which represent the first, second, and third keyword categories, respectively. Each keyword category utilizes a unique keyword selection process, which provides a final sample size of 134, 95, and 49 stocks for the first, second, and third keyword categories, respectively.

The two main variables utilized in this study are search intensity and abnormal trade volume, wherein search intensity is quantified by three different indicators. The first indicator, *lnSVI*, represents the natural logarithm of search interest. Using the search interest obtained from Google Trends, we define the *lnSVI* at period *t* as follows:

$$\ln SVI_t = \ln(SVI_t + 1)$$

where $SVI_t$ is the search interest at period $t$ obtained from Google Trends. The second indicator, $\Delta\ln SVI$, measures the change of search intensity; we define $\Delta\ln SVI$ at period $t$ as follows:

$$\Delta\ln SVI_t = \ln SVI_t - \ln SVI_{t-1}$$

The third indicator, $ASVI$, measures the abnormal level of search intensity; we define $ASVI$ at period $t$ as follows:

$$ASVI_t = \ln SVI_t - \text{median}(\ln SVI_{t-1}, \dots , \ln SVI_{t-n})$$

where $\text{median}(\ln SVI_{t-1}, \dots , \ln SVI_{t-n})$ is the median of $lnSVI$ for $n$ periods. Correspondingly, abnormal trade volume ($ATV_{i,t}$) of stock $i$ at period $t$ is defined as follows:

$$ATV_{i,t} = \frac{TV_{i,t} - \frac{\sum_{t=1}^{L} TV_{i,t}}{L}}{\frac{\sum_{t=1}^{L} TV_{i,t}}{L}}$$

where $TV_{i,t}$ is the trade volume of stock $i$ at period $t$ and $L$ is the number of days we have in our examination.

Accordingly, a multivariate regression analysis is performed on each sample by using the three variables of search intensity to estimate the following model:

$$ATV_{i,t} = \beta_0 + \beta_1 SI_{i,t-1} + \beta_2 \frac{\sum_{t=1}^{N} ATV_{i,t}}{N} + \epsilon_{i,t}$$

where *SI* represents the three search intensity indicators, *ATV* is the abnormal trade of the stock at a certain period, and *N* is the number of stocks included in each sample.

Furthermore, this study examines whether investor sentiment would affect the relationship between search volume and trade volume by sorting the samples based on its investor sentiment. In this case, aggregate sentiment based on the current state of the economy is captured by using the current quarter confidence levels from the Bangko Sentral ng Pilipinas' Consumer Expectation Report. We divide our sample into two: (1) a period with confidence levels above the full sample median and (2) a period with confidence levels below the full sample median, such that the former and the latter represents a period with high confidence and low confidence, respectively.

## 5. RESULTS AND DISCUSSION

As provided by our regression results from examining the relationship between search intensity and trade volume, Table 1 shows the coefficients of search intensity for each keyword category, wherein models 1, 2, and 3 use *lnSVI, ΔlnSVI* and, *ASVI,* respectively, as measures of search intensity.

Table 1
*Search Intensity Coefficients*

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| First Keyword Category: Company Name | 0.0897*** | 0.0227*** | 0.0668*** |
| Second Keyword Category: Ticker | 0.0774*** | 0.0085* | 0.0337*** |
| Third Keyword Category: Ticker + "Stock" | 0.0139 | -0.0115 | 0.0179 |

Note: *** indicates statistical significance at 1% level
* indicates statistical significance at 10% level

The findings of the first and second keyword categories are consistent with the notion that investors searching for a stock's company name and ticker are likely to trade that stock the following day (Takeda and Wakao, 2014; Nguyen et al., 2020). This means that the first and second keyword categories can serve as a measure of investor attention and can potentially be used as a supplementary stock market indicator. These findings support our theory that the arrival of new information can lead to investor reactions that generate informational search activity, which retail investors can utilize as a signal that can influence their investment decisions by helping them anticipate stock market movements.

On the other hand, the findings of the third keyword category imply that there is no evidence that investors searching for a stock's ticker—with the subsequently added word "stock" are likely to trade that stock the following day. This means that we cannot confidently say that the third keyword category can be used as a measure of investor attention and a supplementary stock market indicator. This result could be due to the possibility that investors searching for more information on a particular stock are unlikely to search for the stock's ticker and append the word "stock" after it.

Tables 2 and 3 show the coefficients of search intensity for each keyword category when we sort our samples based on investor sentiment, such that tables 2 and 3 represent the samples with high and low confidence, respectively. Similarly, models 1, 2, and 3 use *lnSVI, ΔlnSVI* and, *ASVI,* respectively, as measures of search intensity.

Table 2
*Search Intensity Coefficients (High Confidence)*

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| First Keyword Category: Company Name | 0.1003*** | 0.0233*** | 0.0768*** |
| Second Keyword Category: Ticker | 0.0968*** | 0.0080 | 0.0548*** |
| Third Keyword Category: Ticker + "Stock" | 0.0083 | -0.0151 | 0.0125 |

Note: *** indicates statistical significance at 1% level

Table 3.
*Search Intensity Coefficients (Low Confidence)*

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| First Keyword Category: Company Name | 0.0705*** | 0.0220*** | 0.0571*** |
| Second Keyword Category: Ticker | 0.0481*** | 0.0087 | 0.0138 |
| Third Keyword Category: Ticker + "Stock" | 0.0324 | -0.0062 | 0.0248 |
| Note: *** indicates statistical significance at 1% level | | | |

By accounting for sentiment, the findings of the first and third keyword category in the sample with both high and low confidence show similar results with the findings when sentiment was not accounted for. Hence, it can be implied that, when using the first and third keyword category, the relationship between search intensity and trade volume is not significantly affected by sentiment.

In contrast, the findings of the second keyword category reveal that the sample with high confidence, as compared to the sample with low confidence, shows more similar results to the findings when sentiment was not accounted for. Therefore, it can be implied that, when using the second keyword category, the positive relationship between search intensity and trade volume is more prominent in periods of high confidence rather than low confidence. The findings of the second keyword category may be explained by the loss aversion phenomenon, which ultimately implies that internet search activity is more likely to be translated to actual trades during periods of high consumer confidence, causing more frequent and significant search volume-based signals.

## 6. CONCLUSION

In this study, we verified whether Google search volume can serve as a measure of investor attention and, subsequently, a supplementary stock indicator for retail investors in the PSE. This was done by examining the relationship between Google search volume and trade volume; however, different in comparison to previous literature, this examination was done using three different keyword categories to test each one's veracity in proxying investor attention. Moreover, this study utilized daily Google search volume as a proxy for investor attention, which to the best of our knowledge, has not yet been done by previous literature in the scope of emerging markets.

The empirical findings reveal that the more investors search for a stock's company name or ticker, the more likely that there will be an increase in that stock's trade volume the following day. This implies that Google searches using a stock's company name and ticker can successfully serve as proxies for investor attention and aid in predicting trade volumes in the PSE. Furthermore, the findings show that Google searches using a stock's ticker are more likely to be associated with the stock's trade volume during periods of high consumer confidence.

The findings of this study could aid in facilitating the research on the relationship between internet search data and stock market movements in emerging markets, which are less informationally efficient than its developed counterparts and have a rapid growth in internet usage. This is because this study not only reinforces the implications of previous literature but also provides new insights on the topic by testing the veracity of three different keyword categories and using daily search volume data to verify the capacity of Google search volume as a proxy for investor attention, which are things that have not yet been done by previous literature in the scope of emerging markets.

Because this study confirms that the Google search volume of a stock's company name and ticker can verify and anticipate stock market movements, we recommend that retail investors in the PSE should consider rethinking their trading strategies by incorporating investor attention measures to these strategies. Given the rapid growth of internet penetration and retail investor participation in the Philippines, it is highly likely that investor attention will become a more significant determinant of movements in the Philippine stock market. This means that the utilization of investor attention-based trading strategies could become more relevant and practical. Moreover, given the potential usefulness of using Google search volume as a supplementary stock indicator, financial platform providers, such as financial data platform and stock screener platform providers, should consider adding an option to view the search volume relevant to a stock. Additional features that allow editing of keywords in use, time-frequency, and location are likely to fine-tune results accordingly. These features will make search volume-based indicators even more accessible to retail investors and, at the same time, promote its usage in trading and investment analysis. Lastly, given the ability of search volume to predict stock market liquidity, we further recommend that policymakers consider using measures of investor attention to monitor market liquidity and anticipate market-wide liquidity shocks. This will help policymakers act more promptly when there is a need for market intervention.

Future studies that wish to examine the relationship between Google search volume and trade volume in emerging markets may want to consider widening the scope of their study by including a cross-comparison between different emerging countries and the possibility of using an even lower time-frequency in regard to their data. This could give a more comprehensive analysis on the topic by verifying whether the findings of this study and previous literature still hold true for other emerging markets and at lower time-frequencies. Additionally, future studies may also want to consider using machine learning to analyze real-time data or non-parametric models, such as regression trees, to not only accommodate imbalanced datasets but also to clarify the structure and classification of such relationships between intention as proxied by Google search and trade activity.

## 7. REFERENCES


Alanyali, M., Moat, H., & Preis, T. (2013). Quantifying the relationship between financial news and the stock market. Scientific Reports, 3(1).

Barber, B., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. SSRN Electronic Journal, 21 (2), 785–818.

Barber, B. M., & Odean, T. (2011). The behavior of individual investors. SSRN Electronic Journal. doi:10.2139/ssrn.1872211

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. Financial Markets and Portfolio Management, 25(3), 239–264.

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). doi:10.1371/journal.pone.0040014

Broder, A. (2002). A taxonomy of web search. ACM SIGIR Forum, 36(2), 3-10.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. The Quarterly Journal of Economics, 108(4), 905–939.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. The Journal of Finance, 66(5), 1461–1499.

Gervais, S., Kaniel, R., & Mingelgrin, D. (2001). The high-volume return premium. The Journal Of Finance, 56(3), 877–919.

Han, L., Li, Z., & Yin, L. (2017). The effects of investor attention on commodity futures markets. Journal Of Futures Markets, 37(10), 1031–1049.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–292.

Kaniel, R., Ozoguz, A., & Starks, L. (2012). The high volume return premium: Cross-country evidence. *Journal Of Financial Economics*, *103*(2), 255–279.

Mahender, Aggarwal, S., & Verma, H. L. (2014). Investors' perceptions on trading volume and stock return volatility in Indian stock market. *The IUP Journal of Applied Economics, 13*(4), 52–73.

Nguyen, C., Hoang, L., Shim, J., & Truong, P. (2020). Internet search intensity, liquidity and returns in emerging markets. *Research In International Business And Finance*, *52*.

Preis, T., Moat, H., & Stanley, H. (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports*, *3*(1).

Statman, M., & Fisher, K. L. (2002). Consumer confidence and stock returns. *SSRN Electronic Journal*. doi:10.2139/ssrn.317304

Stefanescu, R., Dumitriu, R., & Nistor, C. (2012). Overreaction and underreaction on the Bucharest Stock Exchange. *SSRN Electronic Journal*. doi:10.2139/ssrn.2044459

Takeda, F., & Wakao, T. (2014). Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific-Basin Finance Journal*, *27*, 1–18.

Westerhoff, F. H. (2006). Technical analysis based on price-volume signals and the power of trading breaks. *International Journal of Theoretical and Applied Finance, 09*(02), 227–244.