Google trends search query and islamic stock indices: an analysis of their lead-lag relationship based on the Malaysian data

Tew, Li Mei and Masih, Mansur

INCEIF, Malaysia, Business School, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia

28 May 2018

Google trends search query and islamic stock indices: an analysis of their lead-lag relationship based on the Malaysian data

Tew Li Mei[1] and Mansur Masih[2]

Abstract

This paper seeks to analyze the lead-lag relationship between Google trends and Islamic Stock Index prices for the case of Malaysia. In recent years, huge data uploaded and shared in the internet everyday has made it a valuable information to understand the behavior of its users and it has been proven worthy by previous literature. The lag-behind of relevant empirical analysis on Islamic stock market is the gap that this paper aims to fill in. We adopt time series analysis to examine the relationships between two Shariah index prices (FTSE BM EMAS Shariah Index and FTSE BM HIJRAH Shariah index) with Google query search volume. In the end, we identify a two-way causality relationship between the Google Trends search query and Islamic stock Index Price and this paper also reveals immediate negative effect on two Shariah Index prices (EMAS, HIJRAH) in response to one standard deviation shock of Google trends search volume about International finance. Therefore, the finding of this paper suggests that Google trends constitute an important trading signal in Shariah stock investment.

Keywords: Islamic stock index prices, Google trends, lead-lag, Malaysia

---

[1] INCEIF, Lorong Universiti A, 59100 Kuala Lumpur, Malaysia.

2 **Corresponding author**, Senior Professor, UniKL Business School, 50300, Kuala Lumpur, Malaysia.

Email: mansurmasih@unikl.edu.my

Introduction

Recent decade has seen how big data disruption has innovated unprecedented new transformation across different social functions. Data Scientist reveals an inspiring hidden value of the Google query that enables the prediction of social trends from various areas and therefore attracts considerable empirical analysis; prediction of influenza cases (Ginsberg, 2009), prediction on outbreak of disease (Pelat et.al, 2009), link between the performance of financial market with the Google query volume (Preis and Stanley, 2010, Kristoufek, 2013), and prediction of consumer behavior (Goel .et al, 2010). Google query is selected in most of the investigation studies because Google contains the largest search query data among all search engines, and Google Trends (website created by Google in 2006) disclose publicly the aggregate search volume for different query and the changes of these search volume over the time.

If Google search query is so significant important for organization especially the financial-related institution to understand market trends and make better decision in allocation of resource, the lag-behind of relevant empirical analysis on Islamic stock market is the gap that this paper aims to fill in this paper. In this paper, we will examine the possible lead-lag relationship between Google Trends search query on Islamic Finance and Islamic Stock Index for the case of Malaysia. Google Trends allows the user to narrow and concentrate the query search volume based on particular geographical zone and time span. We choose Malaysia to be our focus country given its well-developed global center for Islamic Finance and account for the largest asset holding in term of global market share. Our study yields an interesting result where we find all three variables are endogenous, implies a two-way causality between the Google Trends search query and Islamic stock Index Price. As evidenced in previous empirical study, Google search query indeed gives an impact on the stock index price, but a new finding of this paper shows that in return stock index price movement affects the query search volume too. Furthermore, another contribution of this paper is in the finding of immediate negative effect on two Shariah Index prices (EMAS, HIJRAH) in response to one standard deviation shock of Google trends search volume about International finance. This finding constitutes a crucial signal to the investors who implement tactical investment strategy to re-balance their portfolio to reduce loss. On the other hand, there are limitations in this paper which could be further addressed in future analysis.

This paper is organized in the following order; Session two will review the previous literatures on the theoretical and empirical study on Islamic stock index and the utility function of internet query data in financial market with an objective to identify the research gap for Islamic stock market. Session three will describe the methodology used in this paper and the underlying rationale. Session four will discuss about the analysis results and state the limitation of this paper. The last session will conclude this paper.

Literature Review

Due to the scarcity of previous research that conducts analysis on the causal relationship of the big data phenomena with Islamic Stock Market, this review session will categories the literature into three strands: the correlation of Islamic Stock Indices and conventional counterparts market, the macroeconomic drivers which affect the movement of Islamic Stock Index, and the possible link of internet search behavior to the financial related decision making process.

Debate on the resemblance of Islamic Finance and its conventional counterpart has never failed to capture attention, hence the related studies became the main interest of researchers. Area of debate stretches from the argument on the difference of principles and value to the empirical analysis on the performance of both markets. Some of the distinct contrasts of Islamic Finance from its counterparty are Riba prohibition and Asset-based financial system. All these are asserted in Usmani (2002), Iqbal and Mirakhor (2007) and Abd Rahman (2010).

Theoretical debate is unconvincing without the quantitative evidence. A wide range of empirical analysis conducted to examine the possible causal relationships between Islamic Stock Market and conventional financial market. One of the earliest investigation is done by Hakin and Rashidan in 2002 to scrutinize the possible dynamic relationship (short-run) and cointegration (long-run) between Dow Jone Islamic Stock Index (DJIMI)and Wilshire 500 (W500). They test the cointegration with Johansen's method and find no long run relationship between both indices. In contrast, Dania and Malhotra (2013) find significant positive relationship of return and volatility between conventional stock indices in Northern America , European Union , Far East and Pacific with corresponding Islamic stock indices. Sukmana and Kolid (2012) find evidence that investing in Islamic stock index is less risky than conventional stock index in the case of Jakarta. They apply the GARCH method to investigate the risk performance of Jakarta Islamic Index and Jakarta Composite Index.

The second strand of literatures give more emphasize on the determinant of the risk and return of Islamic Stock Index from the perspective of macroeconomic indicators. Krasicka and Nowak (2012) suggested that both conventional stock price and Islamic stock price respond to common factors like bond rate and tax incentive for case of Malaysia. Most of the empirical results are done for the Malaysia market. Majib and Yusuf (2009) provides policy implication to government to focus on effective exchange rate, money supply and treasury bill rate in order to stabilize Islamic Stock market. Three years later, Husin (2012) extends the research to include other macroeconomic variables: Industrial Production Index(IPI), Consumer Price Index(CPI), and Islamic interbank rate(IIR). The result shows that stock price is related positively with IPI and CPI but has an insignificant relationship with IIR.

Despite all these analysis, study on the impact of big data revolution to the Islamic stock market still lags-behind, and hence this paper aims to address this question. However, in recent years, huge data uploaded and shared in the internet everyday has made it a valuable information to understand the behavior of its users and it has been proven worthy. Preis and Stanley (2010) find a correlation between the weekly trading volume of S&P 500 and the weekly search volume of the corresponding company names. Goel .et al (2010) shows that by using search query, they can predict the sale of a product weeks in advance. These search queries data is now provided free by search engine like Google in their Google Trends platform. Google Trends is a free web tool that do data collection and statistic computation continuously for all search terms input into Google search engine. Kristoufek (2013) shows that the search queries on Google Trends can be successfully utilized for portfolio diversification. The rationale of the diversification is based on the idea that the more the search on particular stock, the higher risk the stock and thus should be assigned lower weight. The same author provides empirical evidence to show the link of google trend to the Bitcoin price. Not limited to economy finding, Google trend is also useful in predicting the outbreak of disease (Pelat et.al,2009), and technology trends(Rech,2007).

As can see from the review above, the cost-free human interaction through the internet has created a massive pool of data in return, the hidden pattern in the data is capable of forecast the near future. Nevertheless, the meaning of Google Trends is unclear for Islamic stock market, in which this paper is going to explore their relationship.

Methodology

This paper adopts time series approach that is based on Johansen cointegration to justify the long -term relationship between Google trends and Islamic Stock Index and Vector Error Correction method to test for Granger causality to determine endogeneity and exogeneity of the variables.

Time series method is favored over OLS regression due to several shortcomings of the latter method. First of all, OLS regression is not appropriate to test for lead-lag relationship since the lead or independent variables and lag or dependent variable are pre-assumed intuitively before the test in OLS and reversal causality among variable is not allowed in the model. Vector Error Correction method in time series approach on the other way allows the data to speak for its own and provides more realistic and convincing result on the causality relationship.

OLS regression method ignores the testing for the stationarity of data and which means resulted significant coefficient from the regression could be spurious, because non-stationary data could result in significant coefficient even there isn't any economic meaning relationship among the variables. Variables can be turned stationary by taking the first-differenced form, yet the estimated coefficient from OLS failed to justify a theoretical relationship as the long-term information has been removed by the transformation of level-form data(non-stationary) into differenced form data (stationary). The estimated coefficient only tells the short-run relationship between variables. As a consequence, no policy implication could be derived from the empirical result from OLS.

Detachment of OLS from the economic reality rendered this paper to select more recent time series approach.

This paper conducts the following steps in Microfit 5 software to test for the Granger causality between Google trends search term and Islamic stock index in Malaysia.

First, stationarity of data is tested by using ADF (Augmented Dicker Fuller) test and KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test. The null hypothesis of ADF is that there should be unit root in the level form of data while KPSS asserted stationary as the null hypothesis, the contracting null hypothesis provides a cross-check on the test result. Then VAR is ran to determine the lag order of variable based on AIC and SBC criteria. After lag order is determined, Johansen cointegration tests is applied to justify the existence of a long-run theoretical relationship between the two variables. However the test is not completed yet at this step as the objective of this paper is to determine which variable is leading and which variable is lagging. This is done with the vector error correction model (VECM). After lead and lag variables are identified, we extend the analysis to examine the relative exogeneity or endogeneity via variance decomposition as this is crucial for making investment decision. At the final stage, this paper will perform generalized impulse function analysis to provide overview on the direction of change and dynamic relationship among the variables when a standard deviation shock on Google search volume.

Data, Result and Discussion

This paper uses data collected from two sources span from 2008-2017, for a total of ten years, in monthly term. There are two variables used which are the Google trends search query (IF) volume about “Islamic Finance” and two Malaysia Islamic Stock Index Price. The search query volume data is take from Google Trend website and the two Islamic Stock Index price; FTSE Bursa Malaysia EMAS index (E)and FTSE Bursa Malaysia Hijrah Shariah Index (H) are from Thomson Reuter Datastream. Google trends volume data is not presented in absolute frequency, it divides each frequency by the total searches of the same geography and time range it represents. Then each data is scaled on a range of 0 to 100. All variables are transformed into logarithms to make the variance stationary.

Variable	Implication
FTSE Bursa Malaysia EMAS (log form)- LE	Variable is Non -Stationary
FTSE Bursa Malaysia Hijrah Shariah Index (log form)-LH	Variable is Non -Stationary
Google Trends (log form) -LIF	Variable is Non -Stationary

Table 1a. Unit Root Test Result for variables in log form

Variable	Implication
FTSE Bursa Malaysia EMAS (Differenced form)-DE	Variable is Stationary
FTSE Bursa Malaysia Hijrah Shariah Index (Differenced form)-DH	Variable is Stationary
Google Trends (Differenced form) -DIF	Variable is Stationary

Table 1b. Unit Root Test Result for variables in first-differenced form

The result of unit roots test shows that all variables are non-stationary in the log form but are stationary in their differenced form (Table1a and 1b), hence all variables are I(1) based on ADF and KPSS tests. Hence Johansen Cointegration method is applicable in this case as all variables fulfill the criteria of I(1) . Johansen method is used to allow the detection of more than one cointegration.

The result for the selection of VAR order is presented in the table 2. below based on AIC (Akaike Information Criterion)and SBC(Schwarz Bayesian Criterion).As per the table has shown, AIC recommends order of 5 and SBC suggests zero lag.

	Selection Criterion	
	AIC	SBC
VAR optimal order	5	0

Table 2. Lag order selection for VAR

Given the conflict between recommendation of AIC and SBC, we check for the serial correlation for each variable and the results obtained are shown in the table below.

Variable	Chi-Sq P-value	Implication (10%)
DE	0.168	There is no serial correlation
DH	0.265	There is no serial correlation
IF	0.122	There is no serial correlation

Table 3. Serial Correlation test results on all variables

Test results confirm the non-existence of serial correlation among the variables, therefore give way to the adoption of a lower order of VAR. However, we resort to higher VAR order (5) in our paper with consideration that Microfit software does not allow a zero-lag order and we have reasonably big time series (115 observation). Once the VAR order of lag is determined, we proceed to test for cointegration by adopting Johansen method. Both Eigenvalue and Trace criteria indicate one cointegration. The finding of cointegration in this step is crucial in order to proceed to succeeding steps. The existence of cointegration assets there is at least a meaningful theoretical long-term relationship between the variables that this paper has proposed in the beginning. To interpret the statistical meaning of cointegrated variables, it simply indicates that all three variables produce cointegration equation that result in a stationary error term. In economic sense, we inclined to believe intuitively that when there is more search about Islamic Finance, it is an indication of increased interest and attention among investors on Islamic market that subsequent turn into a rise of demand for Islamic Stocks, which eventually boost up the index price for Islamic Stocks. To quantify this intuitive relationship we adopt Long Run Structural Modelling (LRSM) to test our theoretical expectation against the statistical findings. We first normalize our first variable of interest (FTSE Bursa Shariah EMAS Index-LE) and obtain result as depicted in the table 4.

Variable	Coefficient	Standard Error	Implication
LH	-2.1247	1.2577	Variable is Not significant
LIF	-0.68877	0.73122	Variable is Not significant
LE	1	-	-

Table 4. LRSM with exact -identifying restriction on EMAS index

As it contradicts with what we find in the cointegration test earlier, we decided to verify the significance of the variables by subjecting the estimates to over-identifying restriction.	Coefficient	Standard Error	Implication

We set the LH to be zero and get the finding as in the table 5. Variable			
LH	0	0	-
LIF	0.54604	0.057319	Variable is Significant
LE	1	-	-

Table 5. LRSM with over-identifying restriction on Hijrah index

With the finding, we have a clearer understanding that Google search volume is significant variable explaining the change of FTSE BM Shariah EMAS index price. Since we are checking for two Shariah Stock Index, we repeat the LRSM for second variable of interest which is the LH (FTSE BM Hijrah Shariah Index) and the result is shown in the table 6.

Variable	Coefficient	Standard Error	Implication
LH	1	-	-
LIF	0.32417	0.15381	Variable is significant
LE	-0.47065	0.27860	Variable is not significant

Table 6. LRSM with exact -identifying restriction on Hijrah index

As predicted, the results resemble previous test result with EMAS Shariah Index, and so to justify the insignificance of LE, we made it the over-identifying restriction and get the expected result as the previous Shariah Index as displayed in the table 7.

Variable	Coefficient	Standard Error	Implication
LH	0	0	-
LIF	0.58055	0.052782	Variable is significant
LE	1	-	-

Table 7. LRSM with over -identifying restriction on Emas index

Thus far, cointegration result only tell that the Google search query volume and Shariah Index price are cointegrated, yet the cointegration reveals nothing about causality, that which variable is leading and which variable is lagging. By knowing which is the lead variable, in the case if the Google query volume is the leading variable, investors could monitor the volume of search in helping to decide on buy or sell decision for the Shariah stocks in Malaysia. In light of this purpose, we test the Granger causality using Vector Error Correction Model (VECM). This step will test the extent variable changes is caused by the change of other variable in the previous period for short-term components and long-term component. Therefore the equation of VECM will include short term components and long term component, where a variable is endogenous when the t-ratio of error term rejects the null hypothesis or exogenous if the t-ratio of error term fails to reject the null hypothesis.

$$\Delta LIF = a + \Sigma\beta\Delta LE_{t-i} + \Sigma\beta\Delta LH_{t-i} + \Sigma\phi LE_{t-i} + \Sigma\phi LH_{t-i} + ecm_{t-1}$$

$$\Delta LE = a + \Sigma\beta\Delta LIF_{t-i} + \Sigma\beta\Delta LH_{t-i} + \Sigma\phi LIF_{t-i} + \Sigma\phi LH_{t-i} + ecm_{t-1}$$

$$\Delta LH = a + \Sigma\beta\Delta LIF_{t-i} + \Sigma\beta\Delta LE_{t-i} + \Sigma\phi LIF_{t-i} + \Sigma\phi LE_{t-i} + ecm_{t-1}$$

We have an appalling result that all our three variables are endogenous, which implies that neither Google search query nor Islamic stock index is independent from each other. It articulates an important lesson that both variables are interrelated or a two-way causality relationship. Investors should be conscious because search query volume is no longer an independent variable that affect the performance of Shariah stocks index, as the result shows that stock price index give an impact to the query search volume as well. Nevertheless, we are unable to know which variable has more impact on the other via VECM , therefore we turn to Variance Decomposition (VDC) to find out the relative endogeneity of Shariah index price and Google query search volume. In other word, we are interest to know among all the endogenous variable, which is the most leaded and which is the most lagged variable.

VDC decomposes the variance of the forecast error of each variable into proportion attributable to the shock from each variable, and the most lead variable is identify through variable that has the highest variance of forecast error attributed to its own lagged. both orthogonalized and generalized methods for VDC produce same result. The results are presented in the table (8a-9b) for six months and a year.

	EMAS	HIJRAH	GOOGLE TRENDS
EMAS	39%	43%	18%
HIJRAH	43%	45%	17%
GOOGLE TRENDS	48%	22%	56%

Table 8a. Generalized VDC , Forecast at Horizon : 12 months

	EMAS	HIJRAH	GOOGLE TRENDS
EMAS	45%	48%	8%
HIJRAH	43%	50%	7%
GOOGLE TRENDS	14%	13%	73%

Table 8b. Generalized VDC , Forecast at Horizon : 6 months

	EMAS	HIJRAH	GOOGLE TRENDS
EMAS	74%	8%	17%
HIJRAH	65%	19%	16%
GOOGLE TRENDS	16%	2%	82%

Table 9a. Orthogonalized VDC Forecast at Horizon: 6 months

	EMAS	HIJRAH	GOOGLE TRENDS
EMAS	59%	8%	33%
HIJRAH	51%	18%	31%
GOOGLE TRENDS	27%	3%	70%

Table 9b. Orthogonalized VDC Forecast at Horizon: 12 months

From the table, the highlighted diagonal percentage represents the extent of variation of the variable that is explained by its own past variation which implies the relative exogeneity .

By referring to the table, we can rank the variables as follow for generalized and orthogonalized method:

Order	Variable
1	GOOGLE TRENDS
2	EMAS
3	HIJRAH

Table 10a. Variables ranking estimated by Orthogonalized VDC

Order	Variable
1	GOOGLE TRENDS
2	HIJRAH
3	EMAS

Table 10b. Variables Ranking estimated by Generalized VDC

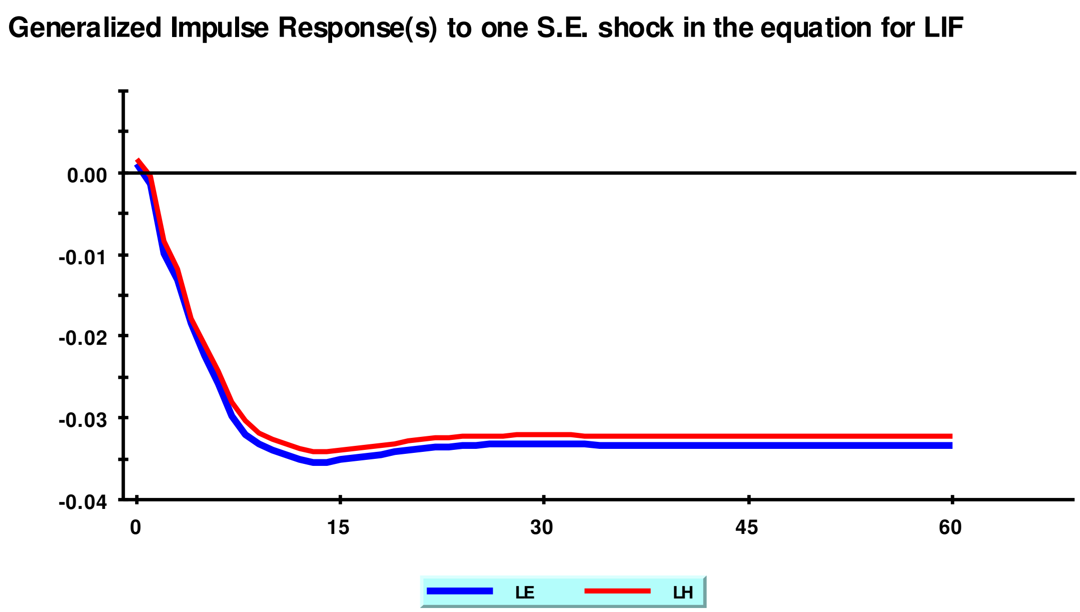
Despite the conflict on the rank between two Shariah stock index, Google trends remain the first in both test. This result implies that even all the variables are inter-related, Google trends appear to have stronger impact on the other two variables, or relative exogenous compared to other variables. Thus, in this stage of analysis, we believe that Google search query volume is still a good indicator for investors for undertaking tactical investing strategy.

This paper suggests that in order to explain the different results for the two Shariah Index price under generalized and orthogonalized method, we must recognize the limitation of orthogonalized method. Firstly, it assumes that when one variable is shocked, all other variables are "switch off"; Second, the generated percentage of VDC is subjected to the ordering of variables in the test. Therefore this paper inclined to believe that Generalized method produces more reliable result.

The decreasing percentage from 6 months to 12 months (Google trends reduce from 73% to 56%) implies that the impact of its own lag to its current variation decreases with time which is logical in the sense that past shock is temporary. From the perspective of an investor, they would be eager to know how big the response function of stock index for one standard deviation shock in

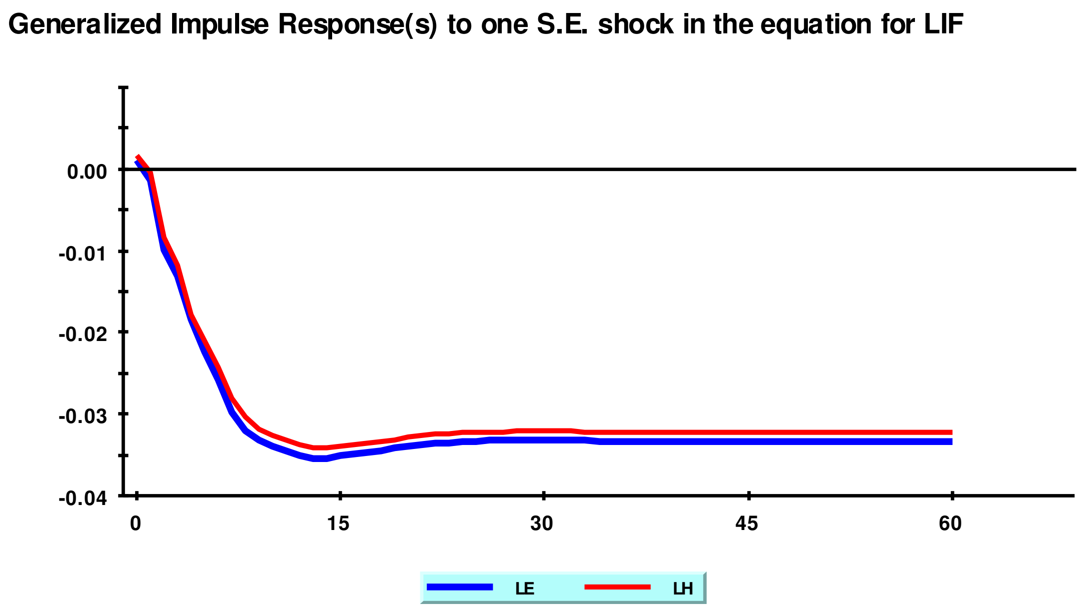
Google trends since this will affect their realized gain or loss. In view of the significant of this information to investors, this paper conducts impulse response analysis to examine the direction and dynamic relationship between Google trends and stock index and the graph below is the result we obtained.

Graph 1.



The graph shows that generalized impulse response to one standard deviation shock in Google trend volume lead to the negative effect on both Shariah index price and the drop persists for almost 12 months before it levels off after the 15^{th} month. This conveys an important warning signal to investors which they should be caution when there is a surge in the Google trends search about Islamic Finance. This signal might be transformed into a sell or hold strategy depend on the investment objective of the investor. For example a tactical investment strategy which aims to take profit in short run fluctuation will suggest a sell while strategic investment strategy that focuses on long term growth will suggest a hold strategy.

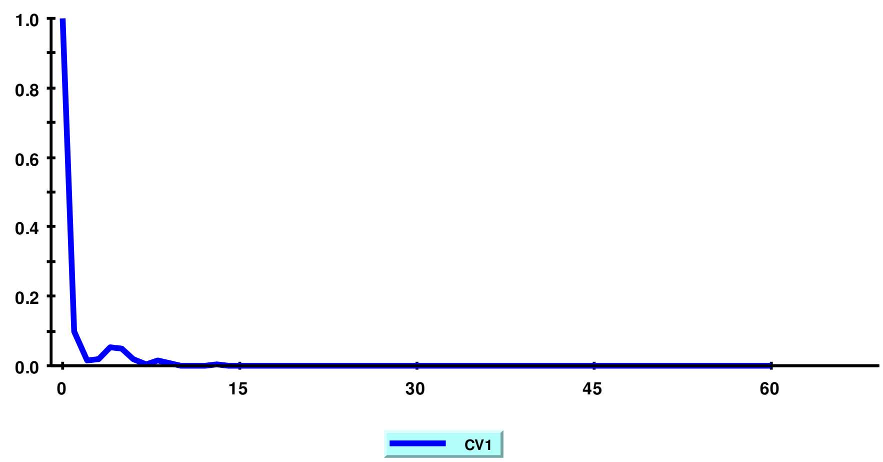
Graph 2.



For completeness , this paper also performs Persistence Profile test and the graph illustrates that in case of system wide shock to the cointegrated variables the relationship takes approximately one year to resume to equilibrium which is consistent with results from generalized impulse response.

Graph 3

Persistence Profile of the effect of a system-wide shock to CV(s)



This paper recognizes that there are limitations in this analysis. First, this analysis uses only one query term in Google Trend. Secondly, we use the term “Islamic Finance” has too broad definition, yet due to unavailable of more accurate query data like Islamic stocks, Islamic investment or the index names, we must rely on the first term. Beside this, our analysis is confined to Malaysia market only, which may not applicable to other countries. Regardless of these limitations, we believe this paper constitutes a ground for future study to extend the topic. Example of further analysis includes the finding into the underlying channel which lead to the opposite relationship between Google trends volume and Islamic stock index price.

Conclusion

Big data revolution speed tracks and improves the decision making of financial participants which has been evidenced in previous literatures. With Google Trends providing a free online tool for public to access to these search queries data, the query data becomes an important data source among the researchers and it has been used to forecast trend in stock markets. This paper intends to fill in the gap where we will investigate the utility of Google trends to estimate Islamic Stocks price. This paper uses data collected from two sources span from 2008-2017, for a total of ten years, in monthly term. There are two variables used which are the Google trends search query (IF) volume about “Islamic Finance” and two Malaysia Islamic Stocks Index Price. The search query volume data is take from Google Trend website and the two Islamic Stock Index price; FTSE Bursa Malaysia EMAS index (E)and FTSE Bursa Malaysia Hijrah Shariah Index (H) are from Thomson Reuter Datastream. We adopt time series approach that based on Johansen cointegration to justify the long -term relationship between Google trends and Islamic Stock Index and Vector Error Correction method to test for Granger causality to determine endogeneity and exogeneity of the variables. This paper also performs generalized impulse function analysis to examine the direction of change and dynamic relationship among the variables when a standard deviation shock on Google search volume. Our finding shows a two-way causality relationship between the Google Trends search query and Islamic stock Index Price and immediate negative effect on two Shariah Index price (EMAS, HIJRAH) in response to one standard deviation shock of Google trends search volume suggests that Google trends constitute an important trading signal for Shariah stock investment strategy. This paper recognizes that there are limitations in this analysis, yet we believe this paper constitutes a ground for future study to extend the topic. Example of further analysis includes the finding into the underlying channel which lead to the opposite relationship between Google trends volume and Islamic stock index price

References

Abd Rahman, Z., (2010). Contracts and the Products of Islamic Banking. CERT Publications, Kuala Lumpur, Malaysia.

Iqbal, Z., Mirakhor, A., (2007). An Introduction to Islamic Finance: Theory and Practice. John Wiley & Sons, Singapore

Usmani, M.T.,(2002). An Introduction to Islamic Finance. Kluwer Law International, The Netherlands.

Ajmi, A. N., Hammoudeh, S., Nguyen, D. K., & Sarafrazi, S. (2014). How strong are the causal relationships between Islamic stock markets and conventional financial systems? Evidence from linear and nonlinear tests. *Journal of International Financial Markets, Institutions and Money*, *28*, 213-227.

Hakim, S., & Rashidian, M. (2002, October). Risk and return of Islamic stock market indexes. In *9th Economic Research Forum Annual Conference in Sharjah, UAE* (26-28).

Dania, A., & Malhotra, D. K. (2013). An empirical examination of the dynamic linkages of faith-based socially responsible investing. *The Journal of Wealth Management*, *16*(1), 65.

Sukmana, R., & Kolid, M. (2012). Impact of global financial crisis on Islamic and conventional stocks in emerging market: an application of ARCH and GARCH method. *Asian Academy of Management Journal of Accounting & Finance*, *31*(2), 357-370.

Hussin, M. Y. M., Muhammad, F., Abu, M. F., & Awang, S. A. (2012). Macroeconomic variables and Malaysian Islamic stock market: a time series analysis. *Journal of Business Studies Quarterly*, *3*(4), 1.

Majid, M., & Yusof, R. M. (2009). Long-run relationship between Islamic stock returns and macroeconomic variables: An application of the autoregressive distributed lag model. *Humanomics*, *25*(2), 127-141.

Krasicka, M. O., & Nowak, S. (2012). *What's in it for Me? A Primer on Differences between Islamic and Conventional Finance in Malaysia* (No. 12-151). International Monetary Fund.

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 A, 368, 5707–5719

Goel, S., Hofman, J., Lehaie, S., Pennock, D. M.&Watts, D. J.(2010), Predicting consumer behavior with Web search. Proceedings of the National Academy of Sciences of the United States of America 7, 17486–17490.

Kristoufek, L. (2013). Can Google Trends search queries contribute to risk diversification?. *Scientific reports*, *3*, article no: 2713.

Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific reports*, *3*, article no: 3415.

Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., & Valleron, A. J. (2009). More diseases tracked by using Google Trends. *Emerging infectious diseases*, *15*(8), 1327-1328.

Rech, J. (2007). Discovering trends in software engineering with google trend. *ACM SIGSOFT Software Engineering Notes*, *32*(2), 1-2.

Ginsberg, J.(2009), Detecting influenza epidemics using search engine query data. *Nature, 457,* 1012–1014