Public Sentiments & Performance of Industrial Indices: Evidence from Twitter Happiness Index

Mohammad Shameem Jawed, Archit V. Tapar & Amol S. Dhaigude

In this study we analyzed sentiments of the people in terms of Happiness Index calculated based on twitter discussions. We used most liquid composite stock indices from India - Nifty50, CNX Midcap, CNX Smallcap and many sectoral indices. The study was conducted for the period – 2014-2018, on daily data. We used bivariate Vector Auto-Regressive (VAR) framework to test Granger causality of the Sentiments measured as Happiness Index with the next day's market returns. Key findings of the study suggest that SmallCap and Nifty50 show marginal relationship with previous day's happiness scores, while amongst sectoral indices only Financial Services and Automotive sector are significantly prone to get impacted by sentiments.

Mohammad Shameem Jawed (E-mail: msjawed@iimv.ac.in) is Assistant Professor (Finance & Accounting), IIM Visakhapatnam 530 003. ***Archit V. Tapar*** (E-mail: archit.tapar@iimrohtak.ac.in) is Asistant Professor (Marketing), IIM Rohtak,124010. ***Amol S. Dhaigude*** **(E-**mail: amold@tapmi.edu.in**) is** Associate Professor (Operations & Information Science), T A Pai Management Institute (TAPMI), Manipal 576104.

Introduction

The conventional financial theory suggests that the investor demonstrates rational behavior, leaving no room for any irrational behavior in influencing pricing of assets. With the introduction of behavioral finance, the view with regard to investor's behavior has been changing. Behavioral finance scholars defined investor sentiment as "a belief about future cash flows and investment risks that is not justified by the facts at hand" (Baker & Wurgler, 2007: 129). Li et al. (2017:497) referred sentiment as "the extent of investors' expectations diverge from the norm, either manifested as excessive optimism or pessimism". Researchers of behavioral finance have been actively exploring and substituting proxies for investor sentiments that may have an effect on pricing of assets (Baker & Wurgler, 2006; Kim & Kim, 2014; Shen, Liu & Zhang, 2018). Kim & Kim (2014) identified that scholars have used four broad sources of proxies for investor sentiments namely: a) investors survey, b) several market variables, c) news and social media, and d) popular internet message boards.

Extant literature has used social media like Twitter (Shen, Liu & Zhang, 2018; You, Guo & Peng, 2017; Zhang et al., 2018), Facebook (Siganos et al., 2014; Siikanen, 2018), etc. to operationalize investor sentiments and examine its effect on stock markets. Shen, Liu & Zhang (2018) used event study methodology to examine the impact of happiness sentiments extracted from the online social media, Twitter, on 26 international stock market returns. The study identified the skewness of stock returns to be significantly greater in the highest happiness days as compared to lowest happiness days. Examining 10 international stock markets You, Guo & Peng (2017) identified daily happiness as a powerful predicting variable for stock returns. Although analysis using the Granger non-causality test showcased the varying effect in different quantiles but the effect was found to be significant. Zhang et al. (2016) emphasized the crucial role played by the social media in stock markets. Examining 11 international stock markets Zhang et al. (2016) identified positive impact of happiness sentiments on the performance of stock markets. The happiness sentiments in this study were extracted from Twitter and were further grouped into quantiles of least to the most happiness days. Zhang et al. (2017), further extended the studies and identified linear and non-linear relationship among different international markets. Yet another study conducted by Li et al. (2017) identified bi-directional relationship between daily happiness sentiment and stocks return, excess trading volume and range-based volatility. Unlike other studies on investor sentiment analysis, the Li et al. (2017) study examined the local investor sentiment on the international stock market.

Based on the extant literature we found that very few studies looked at the emerging market context. Moreover, none have explored the causal relationship of various sectoral stock market indices with Happiness Index, as they may show a very different behavior than market in general – depending upon their nature of business and hence could give new evidences and cues for the investors who take market positions based on sentiments.

Data

Investor sentiment is measured by the daily happiness index extracted from Twitter. Data for daily happiness index include observations spanning the period 09-09-2008 to 02-14-2019. Daily time series data of happiness index was collected from site http://hedonometer.org/data. The index is derived using Natural Language Processing technique applied on a random sample of about 10% of all tweets posted in Twitter. Returns of Nifty50, Nifty500, Nifty MidCap were captured for period 01-02-2014 to 12-31-2018 from NSE website. Return of Nifty SmallCap was captured for period 04-01-2016 to 12-31-2018 (as the index was introduced in the year 2016 only). Returns of different sectoral indices viz. Bank, Automotive, Information Technology, Financial Services, Metal, Media, FMCG, Pharmaceutical and Infrastructure were captured from NSE site for period 01-02-2014 to 12-31-2018. Data obtained from NSE site were cleaned by removing the blank columns. R programming was used.

Methodology Used

In this study, the key objective is to find the causal (Granger-Causal) relationship between people's sentiment and stock market return – especially, to understand if it varies in large verses small firms' indices or is there a variance in their causal behavior within sectors.

To meet the aforementioned objective, we start with Augmented Dickey Fuller test to check stationarity of the time series. For the stationary series, a bivariate Vector-Auto-Regression (VAR) test was done to capture the linear interdependencies among happiness index time series and all the Index return time series respectively. Finally, Granger Causality test was done on the fitted VAR models to find the casualty between Happiness Index and Index Returns with null hypothesis as

H_0: Happiness Index Value do not Granger-cause Index Returns.

Augmented Dickey Fuller (ADF) Test

The ADF test is used for checking stationarity using unit root. Unit roots can cause unpredictable results in time series analysis. The basic objective is to test the null hypothesis that β_2=1 against the alternative β_2<1.

$$X_t = \beta_1 + \beta_2 X_{t-1} + \varepsilon_t$$

If β_2= 1, the process variance increases with t and it is non-stationary. If β_2 lies between 1 and -1, the series is stationary as the variance is fixed.

Vector Auto-Regressive Models (VAR)

A VAR is in a sense a systems regression model that there is more than one dependent variable.

Simplest case is a bivariate VAR

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \ldots + \beta_{1k} y_{1t-k} + \alpha_{11} y_{2t-1} + \ldots + \alpha_{1k} y_{2t-k} + u_{it}$$

$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \ldots + \beta_{2k} y_{2t-k} + \alpha_{21} y_{1t-1} + \ldots + \alpha_{2k} y_{1t-k} + u_{it}$$

where uit is an iid disturbance term with E (uit) =0, i=1,2; E (u1t u2t) =0.

Granger Causality

Granger causality is a way to test causality between two variables in a time series. Granger causality method is a probabilistic account of causality which uses empirical data sets to find patterns of correlation. Granger causality test finds answer to questions such as "Do changes in y_1 cause changes in y_2?" If y_1 causes y_2, lags of y_1 should be significant in the equation for y_2. If this is the case, we say that y_1 "Granger-causes" y_2.

- If y_2 causes y_1, lags of y_2 should be significant in the equation for y_1.

- If both sets of lags are significant, there is "bi-directional causality".
- If both sets of lags are insignificant, there is "no relation".

Granger Causality test also states instantaneous causality. It states that the prediction of unobserved current variable y_2 can be improved by including the current information of y_1.

Results & Discussion

Augmented Dickey fuller test was used to check stationarity of Indices. At significance level of 0.05, the hypothesis test came to be statistically significant i.e. we reject the null hypothesis and accept the alternative hypothesis that returns of indices are stationary (Table 1).

Table 1 Augmented Dickey Fuller Test Results

Composite Indices	Dickey-Fuller Value	p-value
Nifty-50	-10.008	0.01
Nifty-500	-10.612	0.01
Nifty-Midcap	-9.4683	0.01
Nifty-Small cap	-9.641	0.01
Sectoral/industrial Indices		
Bank	-8.4244	0.01
Automotive	-11.386	0.01
Information Technology	-9.5167	0.01
Financial Sector	-10.756	0.01
Metal	-10.216	0.01
Media	-10.821	0.01
Pharmaceutical	-10.78	0.01
Infrastructure	-10.297	0.01

Happiness Index values depends on last day value of Happiness Index.

Vector Auto-Regression test was done to capture the linear interdependencies among Happiness Index time series and Index Return time series. A slight negative co-relation was observed between happiness index and Index returns for all indices i.e. if Happiness Index decreases Index Returns increases but the relationship is very weak. Happiness Index values depends on last day value of Happiness Index.

For Nifty-fifty and Nifty-500, current day return does not depend on past day return. But for Nifty-500, Index Returns depend on past day value of happiness index at a significance level of 0.2, which means people's level of happiness to some extend influences returns of Nifty-500. For Nifty-Mid Cap and Nifty- Small Cap, Index Returns depend on past day returns at a significance level of 0.01 but do not depend on past day value of Happiness Index. In sectoral Indices, Returns depended on past value of Happiness Index at a significance level of 0.1 for Automotive and Media sector.

Returns depended on past day returns for Media, Infrastructure and Pharmaceutical sector indices. For Bank, FMCG and Financial Service sectors returns were not dependent on past values of returns or happiness index. To find the causality between Returns and Happiness Index, Granger causality test was done with null hypothesis - Happiness Index do not Granger-cause Index Returns. The results of the test are reported in Table 2. The results indicated that Nifty-500 showed Happiness Index Granger cause returns at significance level of 0.2 and Automotive and Media showed Happiness Index Granger cause returns at significance level of 0.1. One reason behind this could be that common people are direct consumers of automotive and media industry and these are not life essentials. So, when people's sentiments remain low they do not spend much on luxury and entertainment. Remaining sectors are either of essential goods and services or where common people are not the direct and mass consumer. This is justified by the fact that the emotion does not influence business decisions but impacts common people consumer behavior.

Though people's sentiment instantaneously affects the Returns but in long run people's sentiment does not create much impact.

Furthermore, we also tested the instantaneous causality, which is reported in Table 3. The key results indicate that the prediction of unobserved current variable Returns can be improved by including the available current information on variable Happiness Index. Nifty-50, Nifty-Small cap, Financial service and Media showed instantaneous causality between Happiness Index and Returns. This implies that though people's sentiment instantaneously affects the Returns but in long run people's sentiment does not create much impact. Media is the industry which is heavily influenced by

Table 2 Granger Causality H0: Happiness Index Does Not Granger-cause Returns

Composite Indices	p-value
Nifty-50	0.4165
Nifty-500	0.1781
Nifty-Midcap	0.9478
Nifty-Small cap	0.5854
Sectoral/Industrial Indices	
Bank	0.7827
Automotive	0.0727
Information Technology	0.571
Financial Service	0.4613
Metal	0.2884
Media	0.09054
Pharmaceutical	0.6079
Infrastructure	0.2897

Table 3 Instantaneous Causality Between Happiness Index & Various Index Returns

Composite Indices	p-value
Nifty-50	0.125
Nifty-500	0.879
Nifty-Midcap	0.904
Nifty-Small cap	0.192
Sectoral/Industrial Indices	
Bank	0.394
Automotive	0.379
Information Technology	0.303
Financial Service	0.079
Metal	0.209
Media	0.094
Pharmaceutical	0.292
Infrastructure	0.419

people's sentiment both instantaneously and in the long run as people spend less on entertainment when they are not in a happy mood.

Conclusion

In our study we analyzed sentiments of the people in terms of Happiness Index calculated based on Twitter discussion provided by Hedonometer. We used the most liquid composite stock indices from India viz. Nifty-50, CNX Midcap, CNX Small-cap and many sectoral indices viz. Bank, Automotive, Information Technology, Financial Services, Metal, Media, FMCG, Pharmaceuticals, Infrastructure at NSE and Consumer Durables at BSE. The study was conducted for the period 2014-2018. We used bi-variate Vector Auto-Regressive (VAR) framework to test Granger causality of the sentiments measured as Happiness Index with the next day's market Returns.

Key findings of the study suggest that SmallCap and Most Liquid Index – Nifty50 show weak relationship with previous day's happiness scores. Also, within various sectors we found that previous day's happiness scores Granger cause only Financial Services and Automotive sector indices – making them more prone to get impacted by sentiments. While the instantaneous causality test results suggest that Media sector index also is caused by the contemporaneous changes in the happiness scores gathered from the tweets in the universe.

Our study has some limitations, which may be addressed by some further advanced studies. Happiness Index talks about state of the happiness in general but only a limited set of people trade in stock markets – thus the results may not be showing a true picture of the relationship. Secondly, a more controlled test with bifurcation of data into various quartiles based on the change in happiness from the mean happiness may give better insights.

References

Baker, M. & Wurgler, J. (2006), "Investor Sentiment and The Cross Section of Stock Returns", *The Journal of Finance*, 61(4): 1645-80.

Baker, M. & Wurgler, J. (2007), "Investor Sentiment in The Stock Market", *Journal of Economic Perspectives*, 21(2): 129-52.

Kim, S. H. & Kim, D. (2014), "Investor Sentiment from Internet Message Postings and the Predictability of Stock Returns", *Journal of Economic Behavior & Organization*, 107: 708-29.

Li, X., Shen, D., Xue, M. & Zhang, W. (2017), "Daily Happiness and Stock Returns: The Case of Chinese Company Listed in the United States", *Economic Modelling*, 64: 496-501.

Shen, D., Liu, L. & Zhang, Y. (2018), "Quantifying The Cross-sectional Relationship Between Online Sentiment and the Skewness of Stock Returns", *Physica A: Statistical Mechanics and Its Applications*, *490*: 928-34.

Siganos, A., Vagenas-Nanos, E. & Verwijmeren, P. (2014), "Facebook's Daily Sentiment and International Stock Markets", *Journal of Economic Behavior & Organization*, 107: 730-43.

Siikanen, M., Baltakys, K., Kanniainen, J., Vatrapu, R., Mukkamala, R. & Hussain, A. (2018), "Facebook Drives Behavior of Passive Households in Stock Markets", *Finance Research Letters*, 27: 208-13.

You, W., Guo, Y. & Peng, C. (2017), "Twitter's Daily Happiness Sentiment and The Predictability of Stock Returns", *Finance Research Letters*, 23: 58-64.

Zhang, W., Li, X., Shen, D. & Teglio, A. (2016), "Daily Happiness and Stock Returns: Some International Evidence", *Physica A: Statistical Mechanics and its Applications*, 460: 201-09.

Zhang, W., Wang, P., Li, X. & Shen, D. (2018), "Twitter's Daily Happiness Sentiment and International Stock Returns: Evidence from Linear and Non-linear Causality Tests", *Journal of Behavioral and Experimental Finance*, 18: 50-53.