



Facebook drives behavior of passive households in stock markets

Citation

Siikanen, M., Baltakys, K., Kanninen, J., Vatrappu, R., Mukkamala, R., & Hussain, A. (2018). Facebook drives behavior of passive households in stock markets. *Finance Research Letters*, 27. <https://doi.org/10.1016/j.frl.2018.03.020>

Year

2018

Version

Publisher's PDF (version of record)

Link to publication

[TUTCRIS Portal \(http://www.tut.fi/tutcris\)](http://www.tut.fi/tutcris)

Published in

Finance Research Letters

DOI

[10.1016/j.frl.2018.03.020](https://doi.org/10.1016/j.frl.2018.03.020)

Copyright

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

License

CC BY-NC-ND

Take down policy

If you believe that this document breaches copyright, please contact cris.tau@tuni.fi, and we will remove access to the work immediately and investigate your claim.

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Facebook drives behavior of passive households in stock markets

Milla Siikanen^{*,a}, Kęstutis Baltakys^a, Juho Kanninen^a, Ravi Vratapu^{b,c},
Raghava Mukkamala^{b,c}, Abid Hussain^b

^a DARE Business Data Research Group, Laboratory of Industrial and Information Management, Tampere University of Technology, Finland

^b Centre for Business Data Analytics, Copenhagen Business School, Denmark

^c Westerdals Oslo School of Arts, Communication and Technology, Norway



ARTICLE INFO

Keywords:

Investor behavior
Social media
Stock markets
Investor sophistication
Decision making

JEL classification:

G10
G11

ABSTRACT

Recent studies using data on social media and stock markets have mainly focused on predicting stock returns. Instead of predicting stock price movements, we examine the relation between Facebook data and investors' decision making in stock markets with a unique data on investors' transactions on Nokia. We find that the decisions to buy versus sell are associated with Facebook data especially for passive households and for nonprofit organizations. At the same time, it seems that more sophisticated investors—financial and insurance institutions—are behaving independently from Facebook activities.

1. Introduction

Social media sites, such as Facebook and Twitter, create various opportunities for companies to improve their internal and external communications and to collaborate and communicate with their customers, partners, and other stakeholders, such as investors. Given the importance of social media in external communications, it is not surprising that social media data have been used recently to predict real-world outcomes (see e.g. [Asur and Huberman, 2010](#)). In the financial market research, numerous scholars have used Facebook data ([Karabulut, 2013](#); [Siganos et al., 2014](#); [Bukovina et al., 2015](#)) and data from other social media sites ([Bollen et al., 2011](#); [Zhang et al., 2011](#); [Zheludev et al., 2014](#); [Chen et al., 2014](#); [Nofer and Hinz, 2015](#); [Zhang et al., 2017](#); [You et al., 2017](#)).¹ The primary aim of such research has been to predict market-wide stock movements, yet there is scant research on how social media data relate to the behavior of *individual investors*, perhaps because of the lack of availability of investor account level data.

In this paper, we examine the extent to which investors' trading decisions are driven by Facebook posts and activity. To this end, we use a unique investor-level shareholding registration data set that includes the trading of all Finnish investors over multiple years. In particular, given that an investor trades, we study how Facebook data relate to investors' decisions to increase or decrease their positions. This question is addressed for different investor groups, including financial institutions, nonprofit organizations, and households, and their trades in Nokia stock. As Nokia was one of the most liquid stocks on the Finnish stock market, this unique data has been studied in several articles,² and here we combine it with social media data. Paper by [Lillo et al. \(2015\)](#) is the most closely related study to ours. It also investigates the trading behavior of different investor groups with Nokia stock, but with Thomson Reuters news articles—which are not social media data per se.

* Corresponding author.

E-mail address: milla.siikanen@tut.fi (M. Siikanen).

¹ See also [Bukovina \(2016\)](#) for an overview of research related to a link between social media and capital markets.

² See for example [Westerholm \(2009\)](#); [Tumminello et al. \(2012\)](#); [Lillo et al. \(2015\)](#) and [Ranganathan et al. \(2017\)](#).

Currently, Facebook is clearly the most widely used social media platform, with 2.2 billion monthly active users worldwide (Statista, 2018). As of January 2013, social media sites such as Facebook and Twitter are used by about 45% of S&P1500 firms to communicate externally formal and informal information about their business (Jung et al., 2017). Specifically, companies communicate both corporate disclosures and other information via social media (Zhou et al., 2014). Yang et al. (2017) show that social media, and mass media in general, influences investor's trading decisions. Snow and Rasso (2017) argue that less sophisticated investors potentially benefit most from disclosures communicated via social media, because, on social media platforms, the information is essentially "pushed" to them, which makes this information easier to access. In addition, Snow and Rasso (2017) show that less sophisticated investors process financial information received from social media differently from information received via company's investor relations website.

It is important to remember that typically companies use official exchange-routed company announcements as a primary communication channel (see e.g. Jung et al., 2017), followed by other channels, including newspapers and social media.³ Additionally, communicating information via social media is voluntary, while some company announcement releases are mandatory. Furthermore, Jung et al. (2017) show that companies disseminate strategically, i.e. companies are less likely to disseminate information in Twitter when the news is bad. In this regard, we wish to determine how the investment decisions of, for example, less sophisticated and professional investors, among other investor groups, correlate with potentially biased Facebook information. We note that the relationship between Facebook data and trading can also be related to the attention grabbing behavior of investors, especially households (see Barber and Odean, 2007).

2. Data

2.1. Shareholding registration record data

To identify the trading of different investor categories, we use shareholding registration record data including all domestic investors from June 7, 2010 to the end of 2016, obtained from Euroclear Ltd.⁴ Each record in the data contains detailed information about the investor and the change in his/her holdings. During our analysis period, 282,269 distinct Finnish investors traded Nokia stock. We divide them into five groups according to their sector codes: nonfinancial corporations, financial and insurance corporations, general governmental organizations, nonprofit organizations, and households. Household investors are further divided into four investor activity groups. Investor's activity group is defined based on the number of days the investor traded during the past eight weeks, including the analyzed week. If the number of active days in the past 8 weeks is equal to 1, the investor is considered inactive; if it is between 2 and 5, the investor is passive; 6–20 means moderate; and 21–40 means active. Notably, this is a dynamic group, as one investor might appear in several groups throughout the analysis period.

For the purposes of our analysis, we calculate the number of investors in each group who changed their holdings during a week and the number of investors who increased their holdings (bought more than sold) during that week. Table 1 gives the descriptive statistics of the investor groups and their weekly trading in our data sample. We see that financial and governmental institutions are on average most active sector groups, where as households and nonprofit organizations are least active.

2.2. Facebook data

We collect daily numbers of posts and related comments, likes, and shares from Nokia's Facebook wall⁵ between June 2010 and December 2016 using the Social Data Analytics Tool (SODATO) (see Hussain et al., 2014; Hussain and Vatrappu, 2014a, 2014b). The comments, likes, and shares are always related to a specific post, i.e. the post is the main action. Therefore, we assign the numbers of comments, likes, and shares to the date of the original post—that is, not the date when the actual comment, like, or share was made. In effect, the numbers of comments, likes, and shares quantify the attention the posts released on a particular day received.

We aggregate the daily Facebook data to weekly by summing the numbers of posts, comments, likes, and shares during a week. We take the week beginning on Saturday and ending on Friday, since trading does not occur on weekends. This way, we relate the Facebook activity on weekends to the week in which they can actually affect investors' trading decisions. In total, our sample comprises of 342 weekly observations for posts, comments, likes, and shares. Table 2 gives descriptive statistics of these time series. We can see that on average, there is more than one post made per day, and calculate that one post got on average 274 comments, 4379 likes, and 7 shares.

³ See Siikanen et al. (2017b,a), and references therein, for effects of company announcements in stock markets.

⁴ Grinblatt and Keloharju (2000, 2001); Tumminello et al. (2012); Lillo et al. (2015) and Baltakys et al. (2018) use data sets from the same source, and provide descriptions of the data. However, they use data from before 2009, when all transactions were reported separately with exact trading dates. After moving to Central Counterparty Clearing in late 2009, the Euroclear research data set contains only aggregated daily trades without specifying the actual trading dates—instead a registration date is reported for each record. Thus, we reverse engineer the trading dates from the registration dates. We use the official T + 3 settlement convention for data before and on October 8, 2014 and T + 2 afterwards (see Euroclear, 2014). Using the derived trading dates, we aggregate transactions on a weekly basis, and this reduced the possible noise of inaccurate trading date derivation.

⁵ <https://www.facebook.com/nokia>.

Table 1

Descriptive statistics on investor groups. *N* gives the total number of investors per group. Mean, median and standard deviation (st.Dev) relate to the weekly observations on numbers of investors in each group that changed their net holdings during a week. In Panel B, household investors are categorized into activeness groups on the basis of their trading in the past eight weeks (40 trading days).

Panel A: Investor categories				
Sector	<i>N</i>	Mean (%of all)	Median	st.Dev
Companies	12,213	271 (2.2%)	230	166
Financial	427	28 (6.6%)	27	9
Governmental	89	7 (7.9%)	7	4
Nonprofit	1177	18 (1.5%)	16	12
Households	268,363	4640 (1.7%)	3694	3179
Total	282,269			
Panel B: Activity groups of household investors				
Activeness; # of active days	<i>N</i>	Mean (%of all)	Median	st.Dev
Active; (20, 40]	1228	54 (4.4%)	51	22
Moderate; (5, 20]	16,019	502 (3.1%)	450	227
Passive; (1, 5]	120,906	1856 (1.5%)	1402	1422
Inactive; 1	264,942	2228 (0.8%)	1670	1897

Table 2

Descriptive statistics on Facebook data. *N* gives the total number of each Facebook activity in our sample. Mean, median and standard deviation (st.Dev) relate to the weekly observations on numbers of each Facebook activity.

Activity	<i>N</i>	Mean	Median	st.Dev
Post	2906	8	8	6
Comment	797,586	2332	1585	2808
Like	12,725,171	37,208	11,977	43,500
Share	919,380	2688	461	4525

2.3. Company announcement data

The announcement data is collected from NASDAQ OMX Nordic's website.⁶ The data set includes all the announcements that Nokia filed with Nasdaq between June 2010 and December 2016. Altogether, we have 507 company announcements in the sample. We aggregate the announcement data into weekly by summing the number of announcements from Saturday to Friday, i.e. in similar way as the Facebook data. In the regressions, we use a dummy variable to indicate whether there was at least one announcement release during a week. Our sample includes 187 weeks with at least one announcement release (out of total 342 weeks).

2.4. Weekly return data

The daily adjusted closing price data used to calculate the returns is collected from NASDAQ OMX Nordic's website.⁷ For each week, we calculate the log return as $Ret_t = \ln[P_t/P_{t-1}]$, where P_t is the closing price from the last trading day on the week (usually Friday), and P_{t-1} is the closing price from last trading day on the previous week $t - 1$ (usually previous week's Friday). The average weekly return for Nokia during the sample period was -0.16% .

3. Framework of the empirical analysis

Our analysis is based on logistic regressions to explain how Facebook activity relates to an increase versus a decrease in Nokia shares in investors' portfolios.⁸ To identify the groups of investors whose trading behavior is related to Facebook data, we run separate regressions for each investor group with each Facebook variable.

⁶ <http://www.nasdaqomxnordic.com/news/companynews>, see the page also for detailed information.

⁷ <http://www.nasdaqomxnordic.com/shares/microsite?Instrument=HEX24311>.

⁸ Another option would be (instead of restricting the analysis to a binary outcome) to use linear regressions with continuous dependent variable (i.e. how much an investor changed the position). However, in order to use continuous dependent variable, proportional changes in investors' positions would have to be calculated, which, in turn, requires information on investors' holdings. In contrast to changes in holdings, the levels of holdings, however, were not accurately available. The use of "changes in holdings" as a non-proportional variable is problematic, because investors are trading by very different amounts of shares. These problems are addressed by using logistic regression.

The dependent variable in our regressions is a dummy variable with value 1 if an investor increased his/her holdings in Nokia stock during a given week (bought more than sold) and 0 if the investor decreased the holdings ($D_t^{\text{increased}}$). In a given week, only investors whose net position for Nokia changed are included. The explanatory variable of main interest is the number of posts, comments, likes, or shares depending on the regression (FB). We control for company announcement releases with company announcement dummy (NEWS_t), which is 1 if there was an announcement released during week t and 0 otherwise. Additionally, we use the number of investors in the group who increased their holdings during the previous week scaled by the total number of investors who changed their holdings during the previous week. This is depicted as follows:

$$\text{scD}_{t-1}^{\text{increased}} = \frac{1}{n_{t-1}} \sum_{i=1}^{n_{t-1}} D_{i,t-1}^{\text{increased}}$$

where n_{t-1} is the number of investors who changed (increased or decreased) their holdings in Nokia during week $t - 1$. We also add control variables for the return on present week (Ret_t) and the previous week (Ret_{t-1}). Lastly, we include monthly (M_t) and yearly (Y_t) dummy variables. The monthly dummies control for the potential yearly seasonality in the trading (for example, realizing the losses in December for tax purposes, see e.g. Grinblatt and Keloharju, 2001), and the yearly dummies accommodate the analysis for example to possible changes due to the abandonment of Nokia's mobile business (in 2014, Nokia's mobile business was acquired by Microsoft, changing the focus of the company to a telecommunications infrastructure business). To summarize, the regressions we run are of the following form:

$$g(D_t^{\text{increased}}) = \alpha_1 + \alpha_2 \cdot \text{FB}_t + \alpha_3 \cdot \text{NEWS}_t + \alpha_4 \cdot \text{scD}_{t-1}^{\text{increased}} + \alpha_5 \cdot \text{Ret}_t + \alpha_6 \cdot \text{Ret}_{t-1} + \sum_{j=1}^{11} \alpha_{j+6} \cdot M_j + \sum_{j=1}^6 \alpha_{j+17} \cdot Y_j \quad (1)$$

where g is the logit function.

4. Results

Panel A in Table 3 shows that for households and nonprofit institutions, all the regression estimates are statistically significant. The results indicate that the decisions of investors in these groups to buy vs. sell have a clear association with the Facebook data. For nonprofit institutions, the economic significance is relatively high: the odds of a nonprofit institution buying rather than selling range from 1.111 to 1.212 when the amount of Facebook activity increases by one standard deviation. For financial institutions, Panel A in Table 3 shows no association between the buy vs. sell decisions and the Facebook data. The results for companies and governmental institutions are something between those of financial institutions and households and nonprofit institutions, as half or less of the estimates are statistically significant.

To take a closer look at the effect of Facebook on the trading of households, Panel B in Table 3 presents the estimated regression results for individual investors in different activity groups. We observe that, in general, the more active a household is, the weaker is the association between Facebook data and buying/selling behavior. The odds ratios for passive and inactive investors are more modest than those of nonprofit institutions, though for posts they are still relatively high (1.088 and 1.072). For brevity, we do not report the regression estimates for interception and control variables here, but they are available in Online appendix. In general, most of the estimates for control variables are statistically significant.

Grinblatt and Keloharju (2000) argue that, roughly speaking, finance and insurance institutions, as well as companies, can be viewed as the most sophisticated investor groups, as they generally take larger positions, have more resources to spend on research, and in many cases view investment as full-time career. In light of this, our findings indicate that more sophisticated investors are more independent of Facebook activities, as there is clearly no association between Facebook activities and decisions of financial institutions. Assuming that an investor's activeness is related to his/her sophistication, our findings on household activity groups supports the result that more sophisticated investors behave more independently of Facebook data.

Facebook can be seen as a secondary information channel compared to first-hand official company announcements published on the exchange, and companies are likely to strategically select information disseminated in Facebook (Jung et al., 2017). Nonprofit organizations and households, as arguably less sophisticated investors (Grinblatt and Keloharju, 2000), may allow their trading decision to be affected by Facebook posts and activity, especially if they have no access to professional data sources. In line with this view, Ammann and Schaub (2017) find that the trading decisions of unsophisticated investors are affected by postings that do not contain value-relevant information on a social trading platform.

As our question is if the decisions of different investors are associated with the Facebook data, we are mostly interested in whether the regression estimates for the Facebook variables are statistically and economically significant, while the signs of the coefficients are not in the main focus.⁹ However, a couple of words about the signs of the estimates in Table 3. In Panel A, the signs for posts, comments, and likes are consistently positive, except comments for households. The signs for shares are both positive (governmental and nonprofit) and negative (companies, financial, households), though not all of them are statistically significant, which can explain

⁹ The number of data points in the regression analysis is 332–341, which does not automatically lead to significant estimates as very large data samples do.

Table 3

Regression estimates: Trading of investor groups and Facebook data. The estimates related to Facebook variables of logistic regressions described in Section 3 (Eq. (1)) for all the investor categories. The dependent variable is a dummy variable getting value of 1 if an investor increased his/her holdings during the week, and 0 if the investor decreased the holdings. In addition to the Facebook related variables (for which we report the estimates here), we control for company announcement releases, number of investors who changed their position during previous week (scaled), current and previous weeks returns, and in addition we have monthly and yearly dummies. The regression estimates for control variables (omitted here) are available in Online appendix. In Panel B, household investors are categorized into activeness groups on the basis of their trading in the past eight weeks (40 trading days). Number of observations (weeks in the analysis) is 341 for all the other regressions, except 332 for group governmental. *p*-values are given in parentheses (), and odds ratios (ORs) are given in curly brackets {}. ORs are calculated on the basis of one standard deviation change in the explanatory variable.

Panel A: Investor categories				
	Posts	Comments	Likes	Shares
Companies	0.011*** (3.71E–12) {1.064}	5.55E–06 (0.098) {1.016}	7.61E–07** (6.44E–03) {1.034}	–6.40E–07 (0.775) {0.997}
Financial	5.92E–03 (0.185) {1.034}	1.25E–05 (0.197) {1.036}	9.42E–07 (0.226) {1.042}	–1.94E–06 (0.758) {0.991}
Governmental	0.015 (0.091) {1.086}	5.45E–05** (7.20E–03) {1.165}	2.62E–06 (0.112) {1.121}	2.19E–05 (0.087) {1.104}
Nonprofit	0.033*** (2.27E–07) {1.203}	3.76E–05** (4.79E–03) {1.111}	4.43E–06*** (2.23E–04) {1.212}	3.03E–05** (1.84E–03) {1.147}
Households	0.011*** (5.74E–83) {1.064}	–4.46E–06*** (7.64E–07) {0.988}	1.92E–07** (6.89E–03) {1.008}	–9.04E–06*** (1.23E–41) {0.960}
Panel B: Activity groups of household investors				
Active	2.41E–04 (0.942) {1.001}	–5.82E–06 (0.386) {0.984}	–9.74E–07 (0.065) {0.959}	–1.36E–05** (1.32E–03) {0.940}
Moderate	1.85E–03 (0.083) {1.011}	4.11E–06 (0.071) {1.012}	2.67E–07 (0.140) {1.012}	–1.71E–06 (0.225) {0.992}
Passive	0.012*** (8.55E–55) {1.072}	–9.33E–06*** (2.37E–10) {0.974}	–7.33E–07*** (4.53E–10) {0.969}	–1.60E–05*** (1.73E–49) {0.930}
Inactive	0.015*** (5.12E–67) {1.088}	1.01E–05*** (5.19E–12) {1.029}	1.84E–06*** (1.24E–42) {1.083}	1.09E–07 (0.911) {1.000}

****p* < 0.001; ***p* < 0.01; **p* < 0.05

the variation. Panel B with activity groups reports positive estimates for posts, but there is more variation for comments, likes, and shares as passive and inactive investors have negative estimates. Looking deeper into the reasons of these findings is out of the scope of this paper and left for the future research, as it would require semantic analysis.¹⁰

5. Summary and conclusion

This paper gives the first empirical evidence that Facebook activities affect the trading of different investors differently. We provide evidence that the decisions of arguably less sophisticated investors—that is, households and nonprofit organizations—to increase or decrease shareholdings are clearly associated with Facebook data. At the same time, the decisions of financial institutions, which are likely to be among the most sophisticated investors in the market, are not associated with Facebook activity. Moreover, less active households' decisions are related to Facebook, while the decisions of more active ones are not, which gives additional evidence that the less sophisticated the investor, the more closely related the behavior is to Facebook. Given that Facebook is not a regulated information channel compared to first-hand official exchange releases, companies are likely to strategically select what information to disseminate in Facebook (Jung et al., 2017). This suggests that less sophisticated investors, who may not have access to professional sources for financial data and news, may be driven by biased information.

In the future research we are planning to do sentimental analysis on the posts and comments to give a more comprehensive

¹⁰ Additionally, one could consider if observed associations can represent a reverse causality so that investors are not reacting to social media posts but companies are posting on Facebook in response to changes in investment behavior. However, the reverse causality seems unlikely, because the information about numbers of traders changing their position is not public.

picture of the reactions of different investors to Facebook activities. Concentrating only on Nokia may introduce some investor clientele bias, since the investors interested in Nokia may in general be more social media and technology savvy and follow the posts because of their inclination towards technology. At this point, we were only able to collect the data for Nokia, but in the future research we are planning to extend the sample to a wider variety of stocks.

Acknowledgments

We want to thank Hannu Kärkkäinen and Jari Jussila for their valuable efforts to enable this project. The research leading to these results received funding from Tampere University of Technology through Networked Big Data Science and Engineering Center and TUT's doctoral school. This project also received funding from the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement no. 675044. First author is grateful for the grants received from KAUTE Foundation and Nordea Bank Foundation sr. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary material

Online appendix associated with this article can be found, in the online version, at doi:[10.1016/j.frl.2018.03.020](https://doi.org/10.1016/j.frl.2018.03.020).

References

- Ammann, M., Schaub, N., 2017. The impact of internet postings on individual investors. *Proceedings American Finance Association 2018 Annual Meeting*.
- Asur, S., Huberman, B.A., 2010. Predicting the future with social media. *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*. 1. IEEE, pp. 492–499.
- Baltakys, K., Kanniainen, J., Emmert-Streib, F., 2018. Multilayer aggregation with statistical validation: application to investor networks. *arXiv preprint arXiv:1708.09850*.
- Barber, B.M., Odean, T., 2007. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. *Rev. Financ. Stud.* 21 (2), 785–818.
- Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. *J. Comput. Sci.* 2 (1), 1–8.
- Bukovina, J., 2016. Social media big data and capital markets overview. *J. Behav. Exp. Finance* 11, 18–26.
- Bukovina, J., et al., 2015. Sentiment and blue-chip returns. Firm level evidence from a dynamic threshold model. *Technical Report*. Mendel University in Brno, Faculty of Business and Economics.
- Chen, H., De, P., Hu, Y., Hwang, B.-H., 2014. Wisdom of crowds: the value of stock opinions transmitted through social media. *Rev. Financ. Stud.* 27 (5), 1367–1403.
- Euroclear, 2014. T+2 Implementation Questions & Answers. https://www.euroclear.com/dam/EFI/Campaigns/T-2_cycleQuestionsAndAnswers.pdf.
- Grinblatt, M., Keloharju, M., 2000. The investment behavior and performance of various investor types: a study of finland's unique data set. *J. Financ. Econ.* 55 (1), 43–67.
- Grinblatt, M., Keloharju, M., 2001. What makes investors trade? *J. Finance* 56 (2), 589–616.
- Hussain, A., Vatrappu, R., 2014a. Social data analytics tool: design, development, and demonstrative case studies. *Enterprise Distributed Object Computing Conference Workshops and Demonstrations (EDOCW), 2014 IEEE 18th International*. IEEE, pp. 414–417.
- Hussain, A., Vatrappu, R., 2014b. Social data analytics tool (sodato). *International Conference on Design Science Research in Information Systems*. Springer, pp. 368–372.
- Hussain, A., Vatrappu, R., Hardt, D., Jaffari, Z.A., 2014. Social Data Analytics Tool: A Demonstrative Case Study of Methodology and Software. *Analyzing Social Media Data and Web Networks*. Springer, pp. 99–118.
- Jung, M.J., Naughton, J.P., Tahoun, A., Wang, C., 2017. Do firms strategically disseminate? evidence from corporate use of social media. *Acc. Rev.* in press <https://doi.org/10.2308/accr-51906>.
- Karabulut, Y., 2013. Can facebook predict stock market activity? *AFA 2013 San Diego Meetings Paper*. available at SSRN.
- Lillo, F., Micciché, S., Tumminello, M., Pilo, J., Mantegna, R.N., 2015. How news affects the trading behaviour of different categories of investors in a financial market. *Quant. Finance* 15 (2), 213–229.
- Nofer, M., Hinz, O., 2015. Using twitter to predict the stock market. *Bus. Inf. Syst. Eng.* 57 (4), 229.
- Ranganathan, S., Kivelä, M., Kanniainen, J., 2017. Dynamics of investor spanning trees around dot-com bubble. *arXiv preprint arXiv:1708.04430*.
- Siganos, A., Vagenas-Nanos, E., Verwijmeren, P., 2014. Facebook's daily sentiment and international stock markets. *J. Econ. Behav. Organ.* 107, 730–743.
- Siikanen, M., Kanniainen, J., Luoma, A., 2017a. What drives the sensitivity of limit order books to company announcement arrivals? *Econ. Lett.* 159, 65–68.
- Siikanen, M., Kanniainen, J., Valli, J., 2017b. Limit order books and liquidity around scheduled and non-scheduled announcements: empirical evidence from nasdaq nordic. *Finance Res. Lett.* 21, 264–271.
- Snow, N.M., Rasso, J., 2017. If the Tweet Fits: How Investors Process Financial Information Received via Social Media. *SSRN. working paper*
- Statista, 2018. Most Famous Social Network Sites Worldwide as of January 2018. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.
- Tumminello, M., Lillo, F., Pilo, J., Mantegna, R.N., 2012. Identification of clusters of investors from their real trading activity in a financial market. *New J. Phys.* 14 (1), 013041.
- Westerholm, P.J., 2009. Do uninformed crossed and internalized trades tap into unexpressed liquidity? The case of nokia. *Acc. Finance* 49 (2), 407–424.
- Yang, W., Lin, D., Yi, Z., 2017. Impacts of the mass media effect on investor sentiment. *Finance Res. Lett.* 22, 1–4.
- You, W., Guo, Y., Peng, C., 2017. Twitter's daily happiness sentiment and the predictability of stock returns. *Finance Res. Lett.* 23, 58–64.
- 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-Soc. Behav. Sci.* 26, 55–62.
- Zhang, Y., An, Y., Feng, X., Jin, X., 2017. Celebrities and ordinaries in social networks: who knows more information? *Finance Res. Lett.* 20, 153–161.
- Zheludev, I., Smith, R., Aste, T., 2014. When can social media lead financial markets? *Sci. Rep.* 4, 4213.
- Zhou, M., Lei, L., Wang, J., Fan, W., Wang, A.G., 2014. Social media adoption and corporate disclosure. *J. Inf. Syst.* 29 (2), 23–50.