

WALKING DOWN WALL STREET WITH A TABLET: A SURVEY OF STOCK MARKET PREDICTIONS USING THE WEB

Michela Nardo✉, Marco Petracco-Giudici, Minás Naltsidis

First published: 16 January 2015

<https://doi.org/10.1111/joes.12102>

Citations: 66

The views expressed in this paper are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

Abstract

'A blindfolded chimpanzee throwing darts at The Wall Street Journal could select a portfolio that would do as well as the (stock market) experts' [Malkiel (2003) The efficient market hypothesis and its critics. *Journal of Economic Perspectives* 17(1): 59–82)]. However, what if this chimpanzee could browse the Internet before throwing any darts? In this paper, we ask whether online news has any influence on the financial market, and we also investigate how much influence it has. We explore the burgeoning literature on the predictability of financial movements using online information and report its mixed findings. In addition, we collate the efforts of various disciplines, including economics, text mining, sentiment analysis and machine learning, and we offer suggestions for future research.

References

Agić, Z., Ljubešić, N. and Tadić, M. (2010) Towards sentiment analysis of financial texts in Croatian. LREC *Proceedings of the European Language Resource Association*. Available at: <http://www.lrec-conf.org/proceedings/lrec2010/> (Last accessed 26 November 2013).

[Google Scholar](#)

Ahmad, K., Cheng, D. and Almas, Y. (2006) Multi-lingual sentiment analysis of financial news streams. *Proceeding of the Conference: Grid Technology for Financial Modelling and Simulation*, Palermo, Italy.

[Google Scholar](#)

Akhtar, S., Faff, R., Oliver, B. and Subrahmanyam, A. (2012) Stock salience and the asymmetric market effect of consumer sentiment news. *Journal of Banking and Finance* 36: 3289–3301.

[Web of Science®](#) | [Google Scholar](#)

Alanyali, M., Moat, H.S. and Preis, T. (2013) Quantifying relationship between financial news and stock market. *Scientific Reports* 3: 3578, <http://www.nature.com/srep/2013/131220/srep03578/full/srep03578.html> (Last accessed 16 March 2014).

[PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Antweiler, W. and Frank, M.Z. (2004) Is all that talk just noise? The information content of internet Stock Message Boards. *Journal of Finance* 59(3): 1259–1294.

[Web of Science®](#) | [Google Scholar](#)

Asur, S. and Huberman, B.A. (2010) Predicting the future with social media. Available at <http://www.hpl.hp.com/research/scl/papers/socialmedia/socialmedia.pdf> (Last accessed 16 November 2013).

[Google Scholar](#)

Baker, M. and Wurgler, J. (2006) Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61(4): 1645–1680.

[Web of Science®](#) | [Google Scholar](#)

Baker, M. and Wurgler, J. (2007) Investor sentiment in the stock market. NBER Working paper series, 13189. Available at: <http://www.nber.org/papers/w13189> (Last accessed 16 November 2013).

[Google Scholar](#)

Barber, B. and Odean, T. (2001) The internet and the investor. *Journal of Economic Perspectives* 15(1): 41–54.

[Web of Science®](#) | [Google Scholar](#)

Bautin, M., Vijayarenu, L. and Skiena, S. (2008) International sentiment analysis for news and blogs. *Proceedings of the Second International Conference on Weblogs and Social Media*, Seattle, Washington, USA.

[Google Scholar](#)

Bennet, E. and Selvam, M. (2013) The influence of stock specific factors on the sentiment of equity investors: evidence from Indian stock market. *Proceedings of ASBBS (American Society of Business and Behavioral Sciences) Annual Conference*, Las Vegas, USA.

[Google Scholar](#)

Bennet, E., Selvam, M. and Shalin, E.E. (2011) Investors' sentiment measures. *Global Conference on Innovations in Management*, London UK.

[Google Scholar](#)

Bentley, R.A., O'Brien, M.J. and Brock, W.A. (2014) Mapping collective behavior in the Big-Data era. *Behavioral and Brain Sciences* 37: 63–119.

[PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Black, F. (1986) Noise. *Journal of Finance* 41(3): 528–543.

[Web of Science®](#) | [Google Scholar](#)

Bollen, J., Mao, H. and Zeng, X. (2011) Twitter mood predicts the stock market. *Journal of Computational Science* 2: 1–8.

[Web of Science®](#) | [Google Scholar](#)

Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A. and Weber, I. (2012) Web search queries can predict stock market volumes? *PLoS One* 7(7): e40014.

[CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Bradley, M. and Lang, P. (1999) Affective norms for English Words (anew): stimuli, instruction manual and affective ratings. *Technical Report c-1*, Gainesville, University of Florida, USA.

[Google Scholar](#)

Chalothorn, T. and Ellman, J. (2013) Sentiment analysis: state of the art. *Proceedings of the International Conference on Advances in Computer and Electronics Technology - ACET 2013*, Hong Kong.

[Google Scholar](#)

Chang, E.C., Cheng, J.W. and Khorana, A. (2000) An examination of herd behavior in equity markets: an international perspective. *Journal of Banking & Finance* 24: 1651–1679.

[Web of Science®](#) | [Google Scholar](#)

Chen, R. and Lazer, M. (2011) Sentiment analysis of Twitter feeds for the prediction of stock market movements. CS 229 Machine Learning Final Projects, Autumn 2011. Available at: <http://cs229.stanford.edu/projects2011.html> (Last accessed 27 November 2013).

[Google Scholar](#)

Choi, H. and Varian, H. (2012) Predicting the present with Google Trends. *Economic Record* 88: 2–9.

[Web of Science®](#) | [Google Scholar](#)

Cornell, B. (2013) What moves stock prices? Another look. *Journal of Portfolio Management* 39(3): 32–38.

[Web of Science®](#) | [Google Scholar](#)

Curme, C., Preis, T., Stanley, H.E. and Moat, H.S. (2014) Quantifying the semantics of search behavior before stock market moves. *Proceedings of the National Academy of Sciences of the United States of America* 111(32): 11600–11605.

[CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Cutler, D., Poterba, J. and Summers, L. (1989) What moves stock prices? *Journal of Portfolio Management* 15: 4–12.

[Web of Science®](#) | [Google Scholar](#)

Das, S.R. and Chen, M. (2001) Yahoo! for Amazon: sentiment parsing from small talk on the web (August 5, 2001). *EFA 2001 Barcelona Meetings*. Available at SSRN: <http://ssrn.com/abstract=276189> or <http://dx.doi.org/10.2139/ssrn.276189> (Last accessed 22 November 2013).

[Google Scholar](#)

DeBond, W.F.M. and Thaler, R. (1985) Does the stock market overreact? *Journal of Finance* 40: 793–805.

[Google Scholar](#)

De Choudhury, M., Sundaram, H., John, A. and Seligmann, D.D. (2008) Can blog communication dynamics be correlated with stock market activity? *Proceedings of the Nineteenth ACM Conference on Hypertext and Hypermedia*, Pittsburgh, Pennsylvania, USA.

[Google Scholar](#)

DellaVigna, S. (2009) Psychology and economics: evidence from the field. *Journal of Economic Literature* 47(2): 315–372.

[Web of Science®](#) | [Google Scholar](#)

Denecke, K. (2008) Using SentiWordNet for multilingual sentiment analysis. *Data Engineering Workshop, ICDEW 2008, IEEE 24th International Conference*, Cancun, Mexico.

[Google Scholar](#)

Devitt, A. and Ahmad, K. (2007) Sentiment polarity identification in financial news: a cohesion-based approach. *Proceedings of the 45th Annual Meeting of the Association of Computer Linguistics* (pp. 984–991), Prague, Czech Republic.

[Google Scholar](#)

Dodds, P.S. and Danforth, C.M. (2009) Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *Journal of Happiness Studies* doi:10.1007/s10902-009-9150-9.

[Google Scholar](#)

Doshi, L., Krauss, J., Nann, S. and Gloor, P. (2009) Predicting movie prices through dynamic social network analysis. *Proceeding Collaborative Innovations Network Conference (COINs 2009)*, Savannah, USA.

[Google Scholar](#)

Engle, R. and Ng, V.K. (1993) Measuring and testing the impact of news on volatility. *Journal of Finance* 48(5): 1749-1777.

[Web of Science®](#) | [Google Scholar](#)

Ettredge, M., Gerdes, J. and Karuga, G. (2005) Using web-based search data to predict macroeconomic statistics. *Communications of the ACM* 48(2): 87-92.

[Web of Science®](#) | [Google Scholar](#)

Fama, E. (1965) The behavior of stock market prices. *The Journal of Business* 38(1): 34-105.

[Web of Science®](#) | [Google Scholar](#)

Fawcett, T. and Provost, F. (1999) Activity monitoring: noticing interesting changes in behavior. *Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, USA.

[Google Scholar](#)

Gidófalvi, G. (2001) *Using news articles to predict stock market movement*. Mimeo, Department of Computer Science and Engineering, University of California San Diego, USA.

[Google Scholar](#)

Gidófalvi, G. and Elkan, C. (2003) Using news articles to predict stock price movements. Technical Report, Department of Computer Science and Engineering, University of California San Diego, USA.

[Google Scholar](#)

Gilbert, E. and Karahalios, K. (2010) Widespread worry and the stock market. *Fourth International AAAI Conference on Weblogs and Social Media, (ICWSM)*, Washington (DC), USA.

[Google Scholar](#)

Gloor, P., Krauss, J., Nann, S., Fischbach, K. and Schroder, D. (2009) *Web Science 2.0: identifying trends through semantic social network analysis*. 2009 International Conference on Computational Science and Engineering, Vancouver, Canada.

[Google Scholar](#)

Godbole, N., Srinivasaiah, M. and Skiena, S. (2007) Large scale sentiment analysis for news and blogs. *Proceeding of the First International AAAI Conference on Weblogs and Social Media (ICWSM) 2007*, Boulder, Colorado, USA.

[Google Scholar](#)

Goel, S., Hofman, J., Lahaie, S., Pennock, D. and Watts, D. (2010) Predicting consumer behavior with web search. *Proceedings of the National Academy of Sciences of the United States of America* **107**(41): 17486–17490.

[CAS](#) | [PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Gruhl, D., Guha, R., Kumar, R., Novak, J. and Tomkins, A. (2005) The predictive power of online chatter. *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining* (pp. 78–87), New York, USA.

[Google Scholar](#)

Gunn, S.R. (1998) Support Vector Machines for classification and regression. *Technical Report, Faculty of Engineering Science and Mathematics, School of Electronics and Computer Science, University of Southampton, UK.*

[Google Scholar](#)

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica* **47**(2): 263–291.

[Web of Science®](#) | [Google Scholar](#)

Karabulut, Y. (2013) Can Facebook predict stock market activity? *Mimeo*, Goethe University, Frankfurt, Germany.

[Google Scholar](#)

Koppel, M. and Shtrimberg, I. (2006) Good news or bad news? Let the market decide. In J.G. Shanahan, Y. Qu, J. Wiebe (eds.) *Computing Affect and Attitude In Text: Theory and Applications, The Information Retrieval Series*, Vol. **20**, Chapter 22, Springer, Dordrecht, NL.

[Google Scholar](#)

Lavrenko, V.M., Schmill, M., Lawrie, D. and Ogilvie, P. (2000a) Mining of concurrent text and time series. *6th ACM SIGKDD International Knowledge Discovery and Data Mining*, Boston, USA.

[Google Scholar](#)

Lavrenko, V.M., Schmill, M., Lawrie, D., Ogilvie, P., Jensen, D. and Allan, J. (2000b) Language models for financial news recommendations. *Proceedings of the 9th International Conference on Information and Knowledge Management*, Washington (DC), USA.

[Google Scholar](#)

Lemmon, M. and Portniaguina, E. (2006) Consumer confidence and asset prices: some empirical evidence. *Review of Financial Studies* 19(4): 1499–1529.

[Web of Science®](#) | [Google Scholar](#)

Liu, B. (2012) *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies* 5(1): 1–167.

[Google Scholar](#)

Loughran, T. and McDonald, B. (2011) When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance* 66: 67–97.

[Web of Science®](#) | [Google Scholar](#)

Lucey, B.M. and Dowling, M. (2005) The role of feelings in investor decision-making. *Journal of Economic Surveys* 19(2): 211–237.

[Web of Science®](#) | [Google Scholar](#)

Malkiel, B.G. (2003) The efficient market hypothesis and its critics. *Journal of Economic Perspectives* 17(1): 59–82.

[Web of Science®](#) | [Google Scholar](#)

Mao, H., Counts, S. and Bollen, J. (2014) Quantifying the effects of online bullishness on international financial markets. *ECB Workshop on Using Big Data for Forecasting and Statistics*, Frankfurt, Germany.

[Google Scholar](#)

McLaren, N. and Shanbhogue, R. (2011) Using internet search data as economic indicators. *Bank of England Quarterly Bulletin* 51(2): 134–140.

[Google Scholar](#)

Mishne, G. and Glance, N. (2006) Predicting movie sales from blogger sentiment. In *AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs*, Stanford University, California, USA.

[Google Scholar](#)

Mishne, G. and Rijke, M. (2006) Capturing global mood levels using blog posts. In *AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs*, Stanford University, California, USA.

[Google Scholar](#)

Mittermayer, M.A. (2004) Forecasting intraday stock price trends with text mining techniques. *Proceeding of the 37th Hawaii International Conference on System Sciences*, Hawaii.

[Google Scholar](#)

Moat, H.S., Curme, C., Avakian, A., Kenett, D.Y., Stanley, H.E. and Preis, T. (2013) Quantifying Wikipedia usage patterns before stock market moves. *Scientific Reports*, 3: 1801, <http://www.nature.com/srep/2013/130508/srep01801/full/srep01801.html> (Last accessed 6 March 2014).

[CAS](#) | [Web of Science®](#) | [Google Scholar](#)

Moat, H.S., Preis, T., Olivola, C., Liu, C. and Chater, N. (2014) Using Big Data to predict collective behavior in the real world. *Behavioral and Brain Sciences* 37(1): 92–93.

[PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Morstatter, F., Pfeffer, J., Liu, H. and Carley, K. (2013) Is the sample good enough? Comparing data from Twitter's Streaming API with Twitter's Firehose. *Association for the Advancement of Artificial Intelligence* (www.aaai.org), <http://www.public.asu.edu/~fmorstat/paperpdfs/icwsm2013.pdf> (Last accessed August 2014).

[Web of Science®](#) | [Google Scholar](#)

Niederhoffer, V. (1971) The analysis of world events and stock prices. *Journal of Business* 44(2): 193–219.

[Web of Science®](#) | [Google Scholar](#)

Odean, T. (1998) Volume, volatility, price and profit when all traders are above average. *Journal of Finance* 53: 1887–1934.

[Web of Science®](#) | [Google Scholar](#)

O'Hare, N., Davy, M., Bermingham, A., Ferguson, P., Sheridan, P., Gurrin, C. and Smeaton, A. (2009) Topic-dependent sentiment analysis of financial blogs. *TSA '09, 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion Measurement*, Hong Kong.

[Google Scholar](#)

Pang, B. and Lee, L. (2008) Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2(1–2): 1–135.

[Google Scholar](#)

Preis, T., Reith, D. and 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.

[PubMed](#) | [Web of Science®](#) | [Google Scholar](#)

Preis, T., Moat, H.S., Stanley, H.E. and Bishop, S.R. (2012) Quantifying the advantage of looking forward. *Scientific Reports* 2: 350, <http://www.nature.com/srep/2012/120405/srep00350/full/srep00350.html> (Last accessed 16 May 2014).

Preis, T., Moat, H.S. and Stanley, H.E. (2013) Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports* 3: 1684, <http://www.nature.com/srep/2013/130425/srep01684/full/srep01684.html> (Last accessed 16 May 2014).

Radford, W., Hachey, B., Curran, J.R. and Milosavljevic, M. (2009) Tracking information flow in financial text. *Proceedings of the Australasian Language Technology Workshop, ALTW (2009)*, University of New South Wales, Sydney, Australia.

Remus, R., Heyer, G. and Ahmad, K. (2009) Sentiment in German language news and blogs and the DAX. *Text Mining Services 2009* (pp. 149–158). Leipzig: Leipziger Beitrage zur Informatik, Germany.

Ren, J., Wang, W., Wang, J. and Shaoyi Liao, S. (2013) Exploring the contribution of unlabelled data in financial sentiment analysis. *Proceedings of the XXVII AAAI Conference on Artificial Intelligence*, Bellevue, Washington, USA.

Ruiz, E.J., Hristidis, V., Castillo, C., Gionis, A. and Jaimes, A. (2012) Correlating financial time series with micro-blogging activity. In E. Adar, J. Teevan, E. Agichtein and Y. Maarek (eds.). *Proceedings of the Fifth International Conference on Web Search and Web Data Mining* (pp. 513–522). Seattle (WA), USA. ACM 2012 ISBN 978-1-4503-0747-5.

Ruiz-Martínez, J.M., Valencia-García, R. and García-Sánchez, F. (2012) Semantic-based sentiment analysis in financial news. *Proceedings of the 1st International Workshop on Finance and Economics on the Semantic Web* (pp. 38–51). Available at: <http://ceur-ws.org/Vol-862/FEOSWp4.pdf> (Last accessed 21 October 2013).

Saavedra, S., Hagerty, K. and Uzzi, B. (2011) Synchronicity, instant messaging, and performance among financial traders. *PNAS Early Edition*, S. L. Levin (ed.), Princeton: Princeton University Press.

Scharfstein, D. and Stein, J. (1990) Herd behavior and investment. *The American Economic Review* 80(3): 465–479.

Schumaker, R. and Chen, H. (2006) Textual analysis of stock market prediction using breaking financial news: the AZFinText system. *12th Americas Conference on Information Systems (AMCIS-2006)*, Acapulco, Mexico.

[Google Scholar](#)

Schumaker, R. and Chen, H. (2009) A quantitative stock prediction system based on financial news. *Information Processing and Management* 45(5): 571–583.

[Web of Science®](#) | [Google Scholar](#)

Shleifer, A. and Summers, L.H. (1990) The noise trader approach to finance. *The Journal of Economic Perspectives* 4(2): 19–33.

[Web of Science®](#) | [Google Scholar](#)

Tetlock, P. (2007) Giving content to investor sentiment: the role of media in the stock market. *The Journal of Finance* 62(3): 1139–1168.

[Web of Science®](#) | [Google Scholar](#)

Tobback, E., Daelemans, W., Junqué de Fortuny, E., Naudts, H. and Martens, D. (2014) Belgian economic policy uncertainty index: improvement through text mining. *ECB Workshop on Using Big Data for Forecasting and Statistics*, Frankfurt, Germany.

[Google Scholar](#)

Tulankar, S., Athale, R. and Bhujbal, S. (2013) Sentiment analysis of equities using data techniques and visualizing the trends. *International Journal of Computer Science Issues* 10(4): 265–269.

[Google Scholar](#)

Tumarkin, R. and Whitelaw, R.F. (2001) News or noise? Internet message board activity and stock prices. *Financial Analysts Journal* 57: 41–51.

[Web of Science®](#) | [Google Scholar](#)

Tumasjan, A., Sprenger, T.O., Sandner, P.G. and Welpe, I.M. (2010) Predicting elections with Twitter: what 140 characters reveal about political sentiment. *Fourth International AAAI Conference on Weblogs and Social Media (ICWSM)*, Washington (DC), USA.

[Google Scholar](#)

Wysocki, P.D. (1998) Cheap talk on the Wweb: the determinants of posting on stock message boards. Working Paper n. 98025, University of Michigan Business School.

[Google Scholar](#)

Zhang, X., Fuehres, H. and Gloor, P. (2010) Predicting stock market indicators through Twitter ‘I hope it is not as bad as I fear’. *Procedia – Social and Behavioral Science*, 2010.

[Google Scholar](#)

Zhai, J., Cohen, N. and Atreya, A. (2011) Sentiment analysis of news articles for financial signal prediction. *Mimeo*, University of Stanford, USA.

[Google Scholar](#)

Citing Literature



[Download PDF](#)

ABOUT WILEY ONLINE LIBRARY

[Privacy Policy](#)

[Terms of Use](#)

[About Cookies](#)

[Manage Cookies](#)

[Accessibility](#)

[Wiley Research DE&I Statement and Publishing Policies](#)

[Developing World Access](#)

HELP & SUPPORT

[Contact Us](#)

[Training and Support](#)

[DMCA & Reporting Piracy](#)

OPPORTUNITIES

[Subscription Agents](#)

[Advertisers & Corporate Partners](#)

CONNECT WITH WILEY

[The Wiley Network](#)

[Wiley Press Room](#)