

2.451 A +1.54%

USING BIG DATA TO IMPROVE FINANCIAL FORECASTING

Economists, business leaders and policymakers need good data to make accurate predictions about the future. New sources of powerful data are emerging from social media websites, search engines and other companies, which accumulate vast amounts of information every second. Alongside his colleagues, **Professor Francesco Audrino** of the University of St. Gallen in Switzerland is transforming this real-time data into economic indicators that will help predict levels of risk in a financial system.

The Importance of Accurate Forecasts

When people think about the stock market, it often conjures up an image of a room crammed with frantic traders. Central to an investor's trade, is their feeling about how risky a given investment is. There is a level of risk involved in most transactions, because an investor cannot know with absolute certainty what will happen tomorrow, next month, or next year.

However, investors still want to feel somewhat comfortable when putting their wealth at risk. They pay close attention to changes in the price of a financial product and economic conditions, as well as what might happen to them in the future. For example, investors may feel anxious buying cheap stocks in a company with yearly profit losses, because the company is at risk of bankruptcy. On the other hand, investors may feel comfortable investing a large sum of money in a technology company, if current forecasts project growth in the sector.

Investors keep a watchful eye on the many forecasts produced on products, sectors and the economy to guide their investment decisions. However, forecasts do not describe the future – they simply predict it. Therefore, they have the potential to be right or to be very wrong.

It is now a decade since the Global Financial Crash in 2008, which plunged advanced economies into economic turmoil. Although many eminent voices foresaw the financial crisis, the predictions coming from the financial figures and economic models that were used at the time failed to create an accurate forecast of the future. Since the crash, economists, business leaders and policymakers have been busy finding ways to improve the accuracy of forecasts, so that events leading to the next crash can be foreseen and prevented before they come to pass.



61%

65%

The Role of Big Data in Forecasting

Often, financial crises are collective phenomena in which investor emotions (or 'sentiment') and herding behaviour play a key role. For instance, if an investor hears enough good news stories from friends who are making successful investments by entering a market, then their fear of entering that market turns into a fear of missing out. The stories of good returns diminish the investor's perceived risk of a future loss, despite the fact that the market now has the furthest distance to fall.

14.700 V - 0.235 V - 0.15%

'We found that sentiment and attention variables constructed from social media and search engine data, respectively, contain additional predictive information for future volatility after controlling for a wide range of financial and economic variables.'



The accuracy of forecasting the future of a financial market or product relies on analysts being able to quickly extract emerging trends from data. Social media websites, search engines and other companies collect vast amounts of data every second that can be used to create measures of attention and sentiment. For example, if there is growing uncertainty about the legitimacy of a financial product, we can expect to see more Google searches on the product and more tweets mentioning the issue, as investors seek information and discuss strategies with each other. Although it is possible to collate and transform this data in real time, the question remains as to whether the resulting data will actually improve financial prediction tools.

The SentiVol Project

In recognition of this, Professor Francesco Audrino, in the Faculty of Mathematics and Statistics at the University of St. Gallen in Switzerland, set out to investigate whether the recent availability of new sources of huge amounts of data from social media and from online users' web queries can improve financial predicting tools. 'Thanks to the proliferation of the internet and social media, it is, arguably for the first time in human history, possible to broadly monitor public opinions and moods,' explains Professor Audrino.

Specifically, he and his research team wanted to explore whether they could use this data to assess how investor sentiment and attention impacts volatility in the financial markets. To do so, the team measured investor sentiment using social media data from platforms including Twitter, and evaluated investors' attention, using web queries in Google and Wikipedia.

A further goal of their 'SentiVol' project was to understand the mechanisms through which investors' opinions and activities on the Internet influence financial prices and volatilities. The team hopes that this new approach to volatility estimation will lead to a better understanding of the risk behaviour inherent in a given financial product.

WWW.SCIENTIA.GLOBAL

Transforming Web Data into Financial Data

Professor Audrino and his colleagues built a novel and extensive dataset of investor sentiment and activity in the financial market everyday over a fiveyear period from 2012 to 2016. They started by gathering a large amount of online data from social media, financial news and web queries.

Firstly, the team gathered posts from Twitter and StockTwits – a social media website exclusively for investors. Secondly, they retrieved sentiment scores for a large number of stocks and for 200 economies from RavenPack News Analytics. Finally, they collected data on information consumption from Wikipedia and search volume data from Google Trends – for example, the number of Wikipedia page views and the number of search queries on Google.

In the next step, the researchers transformed the vast amounts of collected data to create measures of n litte

Tube



investor sentiment and activity, with the goal of constructing a dataset that is analysable using statistical regression. They constructed several variables from the dataset, including the total number of messages per day about a certain company and the number of messages expressing positive and negative investment sentiments about a given company.

Analysis and Initial Findings

In the first round of analysis in the SentiVol project, the research team applied a regularised statistical regression to investigate the effect of investor sentiment and attention on volatility in the stock market. For stock market measures, they included US companies that are listed on either on the New York Stock Exchange or on the Nasdaq Stock Market, such as Microsoft, Coca-Cola and General Electric, as well as using the Dow Jones Industrial Average Index.

'We found that sentiment and attention variables constructed from social media and search engine data, respectively, contain additional predictive information for future volatility after controlling for a wide range of financial and economic variables,' says Professor Audrino. Such variables include the exchange rate between the US dollar and Euro.

Professor Audrino and his colleagues found that the effects of investor sentiment and attention on volatility generally last one or two days. However, they also revealed that the predictive power of micro-blogging, search engine and news article data is particularly large during periods of market turbulence. The team found that both general and company-specific investor attention are the most relevant factors – for example, references to a specific company on Twitter and general mentions of the stock market online. The results confirm that the new approach developed by Professor Audrino and his team represents a significant improvement in financial forecasting.

The Broader Impact

Economists, business leaders, and policymakers are always striving to improve the accuracy of volatility predictions in the financial market. The consequences of financial market volatilities impact individual investors but can also affect the entire global economy, as evidenced by the 2008 Global Financial Crisis.

'Understanding and anticipating such high volatility phases is therefore crucial, not just for financial market participants but also for the society as a whole,' Professor Audrino explains. It is his earnest hope that his team's new approach to volatility estimation provides novel insights and tools for achieving this. Indeed, a model capable of providing reliable volatility prediction holds the potential to prevent financial institutions from entering bankruptcy and avoid a global financial crisis.

Unlocking the Potential of the Team's Approach

Having collected and analysed vast amounts of useful data, Professor Audrino and his team have an ambitious future planned out. They are now reasoning about the feasibility to set up a website called 'Open Observatory for Sentiment in Volatility' that performs real-time volatility prediction based on their results. The team's observatory will serve as an early warning device for market turbulence and crises, thus providing an invaluable source of information to investors, financial institutions and economists.

Furthermore, the team is educating students by putting their findings into the curriculum at the University of St. Gallen. They are also considering the possibility for a follow-up project with industry participants, financed either by a potential partner or the Commission for Technology and Innovation, which will allow them to see their results transferred from science into practice.

You

WWW.SCIENTIA.GLOBAL



Meet the researcher

Professor Francesco Audrino Faculty of Mathematics and Statistics University of St. Gallen St. Gallen Switzerland

Professor Francesco Audrino is Professor of Statistics at the University of St. Gallen in Switzerland, where is also in charge of the PhD programme in Economics and Finance. He earned a diploma in Mathematics and a PhD in Statistics and Finance at ETH Zürich. After his postgraduate studies, he moved to the University of Lugano, where he served as a scientific researcher, and later as an Assistant Professor in the Department of Economics. In St. Gallen, he constantly receives awards for outstanding publications. The Swiss Society of Economics and Statistics also awarded him the young economist award in 2007. Professor Audrino's current research interests include the development of new models for the analysis of univariate and multivariate financial time series, with a particular focus on approaches to handle large quantities of heterogeneous data.

CONTACT

You

You

E: Francesco.Audrino@unisg.ch
W: http://www.mathstat.unisg.ch/People.aspx
W: http://ssrn.com/author=333418
W: http://www.researchgate.net/profile/Francesco_Audrino



FONDS NATIONAL SUISSE SCHWEIZERISCHER NATIONALFONDS FONDO NAZIONALE SVIZZERO SWISS NATIONAL SCIENCE FOUNDATION

Foundation.

Daniele Ballinari, University of St. Gallen

FURTHER READING

FUNDING

KEY COLLABORATORS

F Audrino, F Sigrist, and D Ballinari, The Impact of Sentiment and Attention Measures on Stock Market Volatility, 2018, Available at SSRN: <u>https://ssrn.com/abstract=3188941</u>.

Dr Fabio Sigrist, Lucerne University of Applied Sciences and Arts

The research is partially founded by the Swiss National Science

V.SCIENTIA.GLOBAL