Textual analysis of international financial markets

1. Introduction

Textual analysis, also known as text mining or text analytic, is the process of deriving useful and applicable information from the text content (Guo, Shi and Tu, 2016). In finance area, corporations' specific news with regard to mergers and acquisitions, industry forecasting of specified foreign exchange rates and numerous banks' filings, all provide abundant resources for putting this technique into practice. Because of the emergence of textual analysis application in financial markets, algorithms behind are not well defined and developed (Guo, Shi and Tu, 2016). In this essay, readability and dictionary-based measures are discussed to help readers understand how textual analysis techniques are implemented to analyse geopolitical risks and investor sentiment concerning foreign exchange markets.

This essay is structured as follows: section 2 briefly describes foreign exchange markets and currency exchange rates. Section 3 outlines two factors impact financial markets, including geopolitical risks and investor sentiment followed by two textual analysis techniques, including readability measure and dictionary-based measure in section 4, and conclusions.

2. Foreign exchange market and exchange rates

The foreign exchange market is an international decentralized market for the trading of currencies. Major financial centres in the world like London, New York, Tokyo, Hong Kong and Sydney are main trading centres. The majority of traders on the market is global institutional investment firms, but the barriers for individual retail trades do not exist nonetheless as a decentralized market. With more than $5. 3 trillion US dollars traded every

single day, the foreign exchange market ranks first in terms of liquidity when compared with other financial instrument markets (Formadvertising. co. uk, 2018).

Exchange rates between currencies are fluctuated and influenced by numerous highly correlated economic, political and even psychological factors and remain one of the most important economic indices in the international monetary market. It is imperative for international corporations that conduct currency transactions across borders to accurately forecast exchange fluctuations, which can result in considerable improvement in the firm’s overall profitability. When it comes to individuals, psychological factors, including investor sentiment plays an important role in directing investment behaviours. These factors interact in a very complicated nonlinear fashion that cannot be easily forecasted using traditional forecasting models, thus exchange rate series exhibiting high volatility, complexity and noise that result from the nature of the market.

3. Geopolitical risks and investor sentiment

Before moving on to discuss textual techniques utilised to analyse geopolitical risks and investor sentiment in foreign exchange markets and other financial markets, it is important to define and have a better understanding of these two terms.

Geopolitical risks

Geopolitics is defined as the analysis of geographic influences on power relationships in international relations (Deudney, 2018). For instance, the

worsening trade war between US and China, two largest economies in the world, would cost the world much more than US $430 billion of lost GDP (Yao, 2018). Geopolitical risk captures both the risk that these events, including wars, terrorist acts, and tensions between states that affect the normal and peaceful course of international relations, and the new risks associated with an escalation of existing events. (Caldara and Iacoviello, 2018)

It is crucial to comprehend geopolitical risk in a world where advancement and development of communication and globalisation has made information available instantly. It also can have a direct relationship to the amount of risk investors are willing to take.

Investor sentiment

Market sentiment, also known as investor attention, is the prevailing attitude of investors to the anticipated price movements in the financial market (Rao and Ramachandran, 2014). This attitude is the combination of numerous fundamental and technical factors, including but not limited to market price history, exchange rate forecasts, reports and filings of financial organisations, occasional elements, and national and worldwide events. Geopolitical risks also play an important role in directing investor attitude. For instance, if most investors share the feeling of upward price development in the market, the sentiment is said to be bullish (Chen, 2018). On the contrary, if the market sentiment is bearish, most investors hold the expectation that price trend is downward (Chen, 2018).

4. Textual analysis techniques

There have been growing interest in the adoption of the state-of-the-art textual analysis techniques to help extract information from various social media platform, news, journal articles to analyse geopolitical risks and investor sentiment. In this section, two approaches widely used in this area, including readability measure and dictionary-based measure are discussed in detail.

1) Readability Measure

Readability means the degree that a given group of people find certain reading matter compelling and comprehensive (McLaughlin, 1969). Researchers discover that companies with more readable analysts' reports often have more trade volume over three days after report date (De Franco et al., 2012). Biddle (2009) finds that firms which have high reporting quality often have greater capital investment inflow and function more efficiently. Based on this readability measure, Li (2008) finds that companies with lower reported earnings tend to have annual reports that are harder to read. This finding is consistent with the point that firms with poorer earnings are likely to have more words and longer sentences to explain their financial situation to existing and prospective investors.

Manela and Moreira (2017) conducted a textual analysis to measure uncertainty starting in 1890 utilising Wall Street Journal front-page articles. Based on this method, they find news coverage related to wars between countries and intergovernmental policies explains most of the time variation in risk, and this finding is consistent with the fact that time variation in catastrophic disasters that rarely happened is the main source of aggregate

asset prices fluctuations. As can be seen from the finding above, readability measure of textual analysis has its own advantage in analysing geopolitical risks in financial markets.

However, there are some drawbacks that are worth mentioning explicitly under this measure. First, one of the key issues in readability is how “ readability” is defined. For example, words like “ finance”, “ corporations”, “ applications”, “ management” and “ consumers” are usually well understood by investors but they are defined as relatively complicated words according to readability measure (Loughran and McDonald, 2014). Second, unclear market communication deliberately made by some financial institutions may cause market participants to delay important investment decisions, generating more uncertainty. According to Alloway (2018), the temptation of federal reserve is to write longer and with potentially more complexity as the subject matter becomes more sensitive and the outlook more uncertain. As a result, readability measure is required for investors to understand different types of documents.

2) Dictionary-Based Measure

Technically, a dictionary is a tabulated collection of items, each with an associated attribute. Equipped with such lists where collections of words are created to identify a particular attribute of a document, researchers can count words associated with each attribute and provide a comparative measure (Guo, Shi and Tu, 2016). The easiness of its programming and replication has made it the most widely used approach in the finance area.

Tetlock (2007) finds that pessimistic and negative tone in Wall Street Journals can lead to lower stock return and higher stock volatility than expected previously. It has also been proved by him that when pessimism deviation increases one-standard, Dow index in the next day will decline 8. 1 basis point. MacSkassy et al (2008) document that more negative words in S&P 500 firm news are associated with lower earnings after controlling for accounting information and analyst forecasts. Based on more than 800, 000 Thomson-Reuters news, Shina and Heston (2015) point out that negative news tone is related with low stock return in the short term. Das and Chen (2007) extract small investor sentiment from websites, classify the message with a " fuzzy" method and verify that sentiment is based on stock movement. Additionally, according to Nassirtoussi, A. K. et al., (2015), the proposed layers and algorithm techniques implemented in the foreign exchange market system to detect investor sentiment can achieve a significant high directional-accuracies of approximately 84 percent.

However, the challenge of this technique remains as a result of homographs. Homographs means words with different meaning, but the same spelling (Dictionary. cambridge. org, 2018). For example, according to a research conduct by Loughran and McDoanld (2011), consider the following two sentences " The board of directors in corporation X agree to grant call options to its managers and employees" and " Management decided to call inferior products back because of customer dissatisfaction". In these two sentences, apparently the former is a good news while the latter is a bad news. The word " call" that appears twice, in one of the sentences each, has

different meanings but dictionary-based approach is not able to differentiate them.

Conclusion

Compared with traditional quantitative methods, textual analysis measures covered in this essay are less accurate in general. However, various literature has demonstrated textual analysis has become more and more popular over time among financial markets and foreign exchange markets recently due to increasing need to handle tons of texts from various information sources to analyse geopolitical risks and investor sentiment.

Reference:

- Alloway, T. (2018) *Grading the Federal Reserve's Readability* . Available from: https://www. ft. com/content/53510a62-76d3-11e4-8273-00144feabdc0 (Accessed 22 December 2018).
- Biddle, G., Hilary, G. and Verdi, R. (2009). How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics* , 48(2-3), pp. 112-131.
- Caldara, D. and Iacoviello, M. (2018). Measuring Geopolitical Risk. *International Finance Discussion Paper* , 2018(1222), pp. 1-66.
- Chen, J. (2018). *Market Sentiment* . [online] Investopedia. Available at: https://www. investopedia. com/terms/m/marketsentiment. asp [Accessed 25 Dec. 2018].

- Das, S. and Chen, M. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science* , 53(9), pp. 1375-1388.
- Deudney, D. (2018). *Geopolitics | political science* . [online] Encyclopedia Britannica. Available from: https://www. britannica. com/topic/geopolitics (Accessed 25 Dec. 2018).
- Dictionary. cambridge. org. (2018). *HOMOGRAPH | meaning in the Cambridge English Dictionary* . [online] Available at: https://dictionary. cambridge. org/dictionary/english/homograph [Accessed 26 Dec. 2018].
- Formadvertising. co. uk. (2018). *Forex Statistics and Trends for 2017 – Creative Agency for FX Providers* . [online] Available from: https://www. formadvertising. co. uk/forex-trading-industry-statistics-trends-from-2017/ (Accessed 25 Dec. 2018)
- Franco, G., Hope, O. K., Vyas, D, Zhou, Y. (2015). Analyst report readability. *Contemporary Accounting Research* , 32 (1), pp. 76-104.
- Guo, L., Shi, F. and Tu, J. (2016). Textual analysis and machine leaning: Crack unstructured data in finance and accounting. *The Journal of Finance and Data Science* , 2(3), pp. 153-170.
- Heston, S. and Sinha, N. (2016). News versus Sentiment: Predicting Stock Returns from News Stories. *Finance and Economics Discussion Series* , 2016(048), pp. 1-35.

- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* , 45(2-3), pp. 221-247.
- Loughran, T. and McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* , 66(1), pp. 35-65.
- Loughran, T. and McDonald, B. (2014). Measuring Readability in Financial Disclosures. *The Journal of Finance* , 69(4), pp. 1643-1671.
- Manela, A. and Moreira, A. (2017). News implied volatility and disaster concerns. *Journal of Financial Economics* , 123(1), pp. 137-162.
- McLaughlin, G. (1969). SMOG Grading – a New Readability Formula. *Journal of Reading* , 12(8), pp. 639-646.
- Nassirtoussi, A. K. et al., (2015). Text mining of news-headlines for FOREX market prediction: A Multi-layer Dimension Reduction Algorithm with semantics and sentiment. *Expert Systems with Applications* , 42(1), pp. 306–324.
- Rao, K. and Ramachandran, A. (2014). Exchange Rate Market Sentiment Analysis of Major Global Currencies. *Open Journal of Statistics* , 04(01), pp. 49-69.
- Tetlock, P. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance* , 62(3), pp. 1139-1168.

- Tetlock, P., Saar-tsechansky, M. and Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance* , 63(3), pp. 1437-1467.
- Yao, A. (2018). *Cost of the trade war: US$430 billion and counting* . [online] South China Morning Post. Available from: https://www. scmp. com/comment/insight-opinion/united-states/article/2165755/worsening-us-china-trade-war-might-cost-world (Accessed 25 Dec. 2018).