

# Summary of “The Colour of Finance Words”

Diego Garcia, Xiaowen Hu, Maximilian Rohrer(JFE, 2023)

Guowei Zhao summarised on Dec. 2023

## 1 What are the research questions?

- Can a dictionary built through machine recognition more accurately measure sentiment and influence stock price movements compared to a manually constructed dictionary?

## 2 Why are the research questions interesting?

- Tetlock (2007) used the Harvard-IV financial dictionary to study textual information in news media, after which research on textual data in finance and accounting became very popular.
- The primary method currently used to measure textual sentiment involves calculating the frequency of emotional words, known as the 'bag-of-words' approach, based on the LM financial dictionary.
- Gentzkow (2019) points out that the aforementioned approach is, to some extent, weaker than more complex machine learning model methods.

## 3 What is the paper’s contribution?

- It provides a new approach for the literature on how to measure textual sentiment.
- Using machine learning algorithms, a new dictionary was constructed that includes both common unigrams and multi-grams, effectively supplementing unigrams with ambiguous sentiment.

## 4 What hypotheses are tested in the paper?

- A new financial dictionary (ML dictionary) was constructed, which is more capable of fully measuring the sentiment in financial texts compared to the LM dictionary.
- An examination of the breadth of the ML dictionary, the advantages of the ML dictionary over the LM dictionary, and the role of bigrams in the dictionary.

## 5 Do these hypotheses follow from and answer the research questions?

- Yes, all hypotheses are derived from the newly constructed ML dictionary and compared to other dictionaries for measuring textual sentiment effects

## 6 Do these hypotheses follow from theory? Explain logic of the hypotheses.

- Existing literature and theories have shown that sentiment analysis of financial texts can influence and predict the volatility of stock prices, so a new ML lexicon constructed on this basis can also predict the impact of textual sentiment on stock prices through the release of new information.

## **7 Sample: comment on the appropriateness of the sample selection procedures.**

- The dataset for text messages was selected from: conference calls, annual statements, and WSJ news reports. The main reasons for selecting conference call transcripts as the primary study were their higher signal-to-noise ratio and the fact that conference calls were temporally prioritized over annual statements and WSJ reports. Moreover, conference call texts contain more colloquial vocabulary; annual statement texts contain more specialized terminology; and WSJ presentation is more storytelling.

## **8 Dependent and independent variables: comment on the appropriateness of variable definition and measurement.**

- The dependent variable is the stock return, and the independent variable  $S_j$  is a sentiment indicator constructed from the proportion of the number of document words accounted for by the sentiment subscripts in the document obtained by machine learning text analysis.

$$S_j = \sum_{i \in D_i} \left( \frac{tf_{ij}}{N_j} \right)$$

## **9 Regression/prediction model specification: comment on the appropriateness of the regress/predict model specification.**

- The model uses a simple linear regression, illustrated by the regression coefficients of the sentiment indicators, which show that the new ML dictionary is a more accurate and adequate measure of the sentiment of newly published textual messages compared to other dictionaries such as LM.

## **10 What difficulties arise in drawing inferences from the empirical work?**

- Even if there is a correlation between the sentiment of a text message and stock returns, it is difficult to infer that it is the release of the message that causes stock returns to move.

## **11 Describe at least one publishable and feasible extension of this research.**

- Firstly, different word vector models can be compared to analyze the text to construct the prediction effect of different dictionaries.
- Secondly, consider the problem of the number of lexemes, one-word and two-word may not be sufficient, and a larger number of lexemes can better avoid the ambiguity problem.