

# Previous Thesis Report on Saltzman-ML-Uncertainty-2018

Yuhao Zhao

2024-01-16

## 1. Research Question

This paper is focusing on how to use machine learning and specifically natural language processing (NLP) method to extract and measure/quantify various areas of uncertainty from original text data source and then assess their effects on real economy indicators.

## 2. Motivation

Uncertainty plays a crucial role in influencing investment decisions, consumer spending, and labor market outcomes, with potential cross-border spillover effects. Despite its significance, measuring uncertainty accurately is challenging. This study is motivated by the need to develop a more detailed understanding of uncertainty by leveraging machine learning tools, specifically NLP. I find this research is compelling because the research aims to provide insights into how different types of uncertainty impact the economy, distinguishing between economic/business uncertainty and political/governmental uncertainty.

## 3. Contribution related

This paper contributes to the existing body of literature on uncertainty measurements. It clarify that general uncertainty is not a good indicators of economic performances but we should use subsetting or more focused definition of uncertainty measures. And it addresses the advantages of using ML-based methods of extracting, quantifying and differentiating the groups of uncertainty by building upon previous paper that use NLP algorithms on topic modeling.

## 4. Data

### Source

The paper used text data of Federal Reserve Biege Books published between 1970 and 2018.

### Pre-processing

The data was pre-processed heavily by using a text-mining engine from Amenity Analytics. The readable text was first tokenized and classified as events of interests by built-in event and sentiment classifiers. Following event generation, they use customized rule for more accurately extracting of “uncertainty” events by considering context of sentences.

## 5. Approach

After grouping and defining the main index of measuring uncertainty, the authors first use Principal Component Analysis to group 13 general areas of uncertainties into two chunks: Business & Economics (B&E) and Politics & Government (P&G) uncertainty. Then, they use vector-autoregressive (VAR) model to explore the relationship between uncertainty and economy performance.

The intuition behind their approach is to see how/whether B&E uncertainty have different effects on real economy performance measures (Unemployment rate, GDP, FFR, S&P 500 index) across years compared to P&G uncertainty. And

The major assumptions here are: 1) we agree with the definition of uncertainty count the author used at beginning 2) we assumed the two group of uncertainty are independent of each other.

## **6. Results**

The results show that economic response to shocks in B&E uncertainty is more statistically significant, persistent and stronger compared to P&G uncertainty.

## **7. Extensions**

The next step of the research can be to manipulate the algorithms in classifying “uncertainty” since there could be possible improvement in extracting features or groups of the measures.