

Research Project Notebook

Wiktor Owczarz

July 9, 2018

1 Problem Statement

Rationale

Hypothesis: central bank researchers face higher **reputation risk**, compared to university researchers, when publishing banking competition studies. Economists employed by central banks are placed under **higher public scrutiny**, hence their publications can have stronger **signalling effect**.

As a result: central bank researchers are more prone to making **Type II Error** (failing to reject false null hypothesis). The cost of making **Type I Error** is high and might lead to financial instability.

Define reputation/signalling in context of economic theory

Define role of a central bank (supervisor/regulator) + literature

Objectives

Main Goal: Quantify if there are significant, measurable differences between banking competition papers published by the two groups of researchers.

Thesis Purpose: Create **methodology** that can be applied to analyse and measure the differences in competition papers.

Short: write methodology paper, define **heterogeneity** and to measure it in a robust way

Possible "sources" of heterogeneity:

1. **Publication Bias:** Apply publication bias methodology but in reverse, check if CB papers have larger confidence levels, are they minimising risk of Type I error by further decreasing α
2. **Text Analysis:** Looking at the two groups we find significant **differences in analysed texts**. As a result, we would be able to classify papers to either group based on the text structure.
 - (a) Apply Latent Semantic Analysis to find **proximity** between different papers based on **bag of words approach**.
 - (b)
 - (c)

2 Methodology

Objective is to **create number of tools** that will allow us to analyse and compare competition texts and quantify if any differences.

1. Initial Analysis

- (a) Create corpus from the collected articles
- (b) Transform data into text vectors
- (c) Use tf-idf to weight the words
- (d) Use SVDTruncated to reduce dimensions
- (e) Run simple KMeans

Very simplistic approach and does not fully capture the propose of the task

Rui: focus on understanding the text, clean stop words (not only 'the') but also remember the context. Visualize the data, herarchical clustering

2. Features engineering:

- (a) word density: avg. length of words
- (b) punctuation cound
- (c) part of speech distributions
- (d) lexical diversity
- (e) **Uncertainty measure**: % share of the words that introduce ambiguity in the text

Features add further complexity but ignore the data, calculations

3. Ambivalent significance:

- (a) We cannot reject vs we accept nul hypothesis (*strength of the phrase*)
- (b) What are the confidence levels used 1%, 5% or 10%?
- (c) Limitations of p-value, look at t-stat distributions
- (d) In general **publication bias analysis** (Harvey 2017)

This is second part of the methodology, focus on differences in t-stats/p-values do we see any differences there?

Daily Notes:

Harvey 2017: *The Scientific Outlook in Financial Economics (05.07.2018)*

- Points to out the problem with too high reliance on p-values and t-statistics, which are poorly understood
- P-hacking: cherry picking the most significant results. **This should not be an issue since design is the same across papers, same model and data is fully accessible**
- **T-ratio/P-value distributions** if not uniform (skewed to right) show that there is actually bias.

Use p-value distributions to visualise problem

- Harvey: “complexity stems from the challenge in interpreting results when the human researcher interacts with the results” == results interpretation is crucial in social science (soft science)
- published articles find “support” for the idea being tested because: priori beliefs, narrow hypothesis, confirmation bias, manipulation of the data (

We are interested in the **“opposite” of the publication bias** since we want to see if the results of the central bankers fail to reject the null more often than other researchers.

References: (work-in-progress)

1. Stanley et. al. *Meta-analysis of economics research reporting guidelines*
2. A. P. Field, R. Gillett *How to do a meta-analysis*, British Journal of Mathematical and Statistical Psychology
3. C. R. Harvey, *Presidential Address: The Scientific Outlook in Financial Economics*