Research Project Notebook

Wiktor Owczarz

April 23, 2018

The following document is an outline and notebook concerning a review study that aims to quantify between group differences in competition studies written by two groups of economists: Central Bank vs. University.

1 Problem Statement

Rationale

Describe the rationale for the review in the context of what is already known. (literature review + hypothesis)

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

Define reputation/signalling in context of economic theory

Objective

What we do and how to do it correctly

Goal: We aim to quantify if there are significant, measurable differences between banking competition papers published by two groups of researchers.

Possible "*sources*" of heterogeneity:

1. **Meta Study:** When looking at the same market, using the same *(similar?)* methodology two groups get either contrary or close but **different** (insignificant, smaller confidence, higher variability) outcomes.

2. **Semantic Analysis:** Looking at the two groups we find significant semantic differences. As a result, we would be able to classify paper to either group based on the words they use.

   (a) First idea: Latent Semantic Analysis to find **proximity** between different papers.

   (b) If an *algorithm* would be able to easly distinguish between two groups, then

Goal is to define **heterogenity** and to measure it in a robust way

2 Methodology

2.1 Meta-Study:

Generally accepted structure required for the meta-study:

2.1.1 Save study protocol *before* actual study

The *before* protocol **has to include**:

- Problem statement
- Literature search strategy
  - Which databases
  - Key words/definition of the search strategy
  - Publication restrictions (e.g. language)
- Exclusion/inclusion criteria. Decide which of the collected papers to include/exclude.
  - Clear distinction between central bank and who university economist.

> After the selection process is done create PRISMA Flow diagram to summarise the articles exclusion/inclusion process.

2.1.2 Calculate the size effects

Def. Effect size is usually a standardised measure of the magnitude of observed effect (Clark-Carter, 2003; Field, 2005c).

2.1.3 Do the basic meta-analysis

Estimate effects in the population by combining the effect sizes from a variety of articles. The estimate is a weighted mean of the effect sizes.

> Here ideal outcome would be to find clear **heterogeneity** in the effect sizes between two groups.

2.2 Language Processing / Semantic Analysis

The initial idea is to run Latent Semantic Analysis and by looking at word frequency in each document try to classify papers based on their comparability/similarity between each other.

3 Notes:

- Affiliation based on: https://ideas.repec.org
  - How to categorize researchers that have both affiliations? 50% or more?
  - How to categorize papers that have more than one author and therefore **multiple affiliations**

4 Problems

The Effect size

According to Stanley et al. in *"Reporting Guidelines for Meta-Regression Analyses in Economics"*, in order to conduct successful meta-study:

- A precise definition of how effects are measured (the 'effect size'), accompanied by any relevant formulas.

- An explicit description about how measured effects are comparable, including any methods used to standardise or convert them to a common metric.

As I already pointed out we face significant difficulty when it comes to the second point. Also we cannot just look at the sign and/or significance since the main purpose of the meta-study is to aggregate multiple papers to average out the bias that might come from single study and therefore find the *"true" outcome*.

However, maybe we could use **Cohen's d**: $\frac{\mu_1 - \mu_2}{\sigma}$, where two means would be papers classified to central bank or university and the variance can be a pooled estimate based on both groups individual size and variances (for me it looks very similar to pooled t-static for unequal variances). We would have to do the comparison for individual models separately.

Although, just looking at Panzar-Rosse I can see problem since different papers use **different estimation methods**:

1. Revenue equation: $log(TR)$ with $log(TA)$ as a control variable

2. $log(TR/TA)$ without $log(TA)$ as a control variable

3. Price equation: $log(TR/TA)$ with $log(TA)$ as a control variable

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