A Big Data approach to understand Central Banks

Big Data Spain 2018

November 2018
Summary

01 Why is the use of NLP important in economics and Monetary policy?

The data and methodology

02 Understanding Central Banks: “What”, “How” and “Who” is talking (or writing) about?
Why is the use of NLP important in economics and Monetary policy?

The data and methodology
Why is the use of NLP important in economics and Monetary policy?

**Text as a key source of information to enrich economic analysis**

- 80% of the total amount of web pages on the internet is given by **textual or unstructured data**

- The potential use of textual information and text sources **improves the understanding of economic and financial systems**

- **Text mining** to extract meaning from strings of letters

- It helps us to **understand what drives monetary policy decisions**
A Big Data approach to understand Central Banks

80% of available data

20% of used data
## The data and methodology

### From Extraction to Sentiment Analysis

<table>
<thead>
<tr>
<th>Information extraction</th>
<th>Pre-Processing and text parsing</th>
<th>Transformation</th>
<th>Text mining and NPL</th>
<th>Sentiment analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents</td>
<td>Extract words</td>
<td>Text filtering</td>
<td>Analysis and Machine learning</td>
<td>Apply sentiment dictionaries</td>
</tr>
<tr>
<td>Web pages</td>
<td>Identify parts of speech</td>
<td>Indexing to quantify text in lists of term counts</td>
<td>Topics extraction (LDA)</td>
<td>Semantic analysis and classification</td>
</tr>
<tr>
<td></td>
<td>Tokenization and multi-word tokens</td>
<td>Create the Document-term matrix</td>
<td>Clustering</td>
<td>Clustering</td>
</tr>
<tr>
<td></td>
<td>Stopword Removal</td>
<td>Weighting matrix</td>
<td>Modelling (STM and DTM)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stemming</td>
<td>Factorization (SVD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Case-folding</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A Big Data approach to understand Central Banks / 6
The data and methodology

Analyzing central banks’ communication: Examined documents

Statements / Press Releases
Immediately after the meeting on monetary policy, a short report about the decision on interest rates is released. If there’s a press conference, the president of the CB explains the decision and answer questions from journalists.

Minutes
A more detailed document explaining the monetary policy decision containing an overview of financial market, economic and monetary developments.

Speeches
Collection of speeches and articles by senior central bank officials published in the central bank websites.
Analyzing central banks’ communication: Cleaning and transforming the text

Extracting and organizing the data:
- Extract words
- Identify parts of speech
- Stopword Removal
- Case-folding

Preparing it for the analysis:
- Text filtering
- Indexing to quantify text in lists of term counts

Working with text in numbers:
- Create the Document-term matrix
- Weighting matrix
- Factorization

Converting it into numbers:
- Stopword Removal
- Stemming
- Tokenization and multi-word tokens
The data and methodology

Analyzing central banks’ communication: Dynamic topic models

Latent Dirichlet Allocation (LDA) and Dynamic Topic Model (DTM)
The data and methodology

Analyzing central banks’ communication: Sentiment analysis

Loughran and McDonald (2011)

<table>
<thead>
<tr>
<th>Positive words</th>
<th>Negative words</th>
</tr>
</thead>
<tbody>
<tr>
<td>achieve</td>
<td>bankruptcy</td>
</tr>
<tr>
<td>benefit</td>
<td>fallout</td>
</tr>
<tr>
<td>efficiency</td>
<td>imbalance</td>
</tr>
<tr>
<td>outperform</td>
<td>migrant</td>
</tr>
<tr>
<td>progress</td>
<td>adverse</td>
</tr>
<tr>
<td>stabilize</td>
<td>escalate</td>
</tr>
<tr>
<td>strength</td>
<td>vulnerability</td>
</tr>
<tr>
<td>versatility</td>
<td>worsening</td>
</tr>
</tbody>
</table>

Positive words

Negative words

FED Financial Stability dictionary (2017)

<table>
<thead>
<tr>
<th>Positive words</th>
<th>Negative words</th>
</tr>
</thead>
<tbody>
<tr>
<td>benefit</td>
<td>adverse</td>
</tr>
<tr>
<td>stabilize</td>
<td>escalate</td>
</tr>
<tr>
<td>bottleneck</td>
<td>challenge</td>
</tr>
<tr>
<td>corrupt</td>
<td>deterioration</td>
</tr>
<tr>
<td>downgraded</td>
<td>vulnerability</td>
</tr>
<tr>
<td>benefit</td>
<td>improve</td>
</tr>
<tr>
<td>enhance</td>
<td>upgraded</td>
</tr>
<tr>
<td>favourable</td>
<td>smooth</td>
</tr>
<tr>
<td>stabilized</td>
<td>strengthened</td>
</tr>
<tr>
<td>favourable</td>
<td>smooth</td>
</tr>
<tr>
<td>downgraded</td>
<td>worsen</td>
</tr>
</tbody>
</table>

Positive words

Negative words

Average tone = $\frac{\sum Positive \ words - \sum Negative \ words}{\sum Total \ words}$
Main outputs

Analyzing the Central Bank of Turkey, European Central Bank and Federal Reserve
Main outputs

More than words: Getting the relation between words…

In the case of CBRT:
Main outputs

...their evolution over time...

Most frequent words by year in the analyzed documents (the case of CBRT)
Main outputs

...as well as the topical content covered in the text
Understanding Central Banks: “What”, “How” and “Who” is talking (or writing) about
We go Inside of the CB Reports to identify the topics using Machine Learning and Dynamic Topic Models. They can be different…

Source: BBVA Research
Topics are dynamic and can change over time...and the picture can change...particularly if important events hit the economy.
Networks are a useful tool to show the interconnectedness & complexity...helping us to understand “How” the Central Banks talk..

Monetary Policy in the North (ECB) and in the EM (Turkey): Complexity and interconnectedness (Networks)

Pre Lehman (1999-2007)

Financial Crisis (2007-2012)

QE & Post Crisis (2013-2018)

Source: BBVA Research
Sentiment analysis reinforces the analysis by describing “How” the Central Bank talks (“tone”)

Turkey (CBRT) : Economic Activity & Inflation Tone  
(Tone economic activity and Inflation in the MP Minutes)

Turkey (CBRT): Economic Activity & Employment Tone  
(Tone economic activity and employment in the MP Minutes)

Source: BBVA Research
Through Sentiment Analysis we can check the monetary policy stance... how “Tight” or “Ease” is the Wording of the reports

Central Bank of Turkey: Monetary Policy Sentiment
(Standardized, estimated through Big Data LDA and STM Techniques from Minutes & Statements)

Monetary Policy “Statements”

Monetary Policy “Minutes”

A more formal Statement...

More extensive and analytical...

Source: BBVA Research
And how the market rates react to the Central Bank changes in monetary policy sentiment …

Response to Short term and Long term interest rates to positive/Negative changes in Sentiment CB Turkey
(Response of interbank deposits rates and 2Y BondSwaps to mild and strong changes in sentiment. Changes relative to t−1. T=event)

Source: BBVA Research
Remember that in the case of Sentiment Analysis, we are using unsupervised methods so...always cross-check for Robustness

Monetary Policy in Turkey: Experts vs Algorithms
(Sentiments from LDA Algorithm and MP Surprises by Demiralp et Al. 1=Hawkish, 0= Neutral, -1=Dovish)

Experts vs Algorithms in Turkey: Size of Surprises & Sentiments
(Sentiments from LDA Algorithm and MP Surprises by Demiralp et Al)

Source: BBVA Research
Last… but not least … we are working on the Federal Reserve Board (FED) Topics and Stance…

Source: BBVA Research
…complementing our “What” and “How” the Central Banks talk with “who” is talking…

General and Governor FED Hawkish/Dovish index by speaker over time
(Moving average 12 months)

From a EM Crisis
Reactive and tightening…
(Mr Greenspan)
1987-2003

To a Governor
Managing the crisis…
(Mr Bernanke)
2016-2014

To a Lady managing
the Exit Strategy…
(Mrs Yellen)
2014-2018

To a Normalization
Policy
(Mr Powell)
2018-

Source: BBVA Research
You can find us at:

www.bbvaresearch.com

Alvaro Ortiz
Tomasa Rodrigo

@alvaroortiz1968
@TomasaRodrigo

Thank you!
A Big Data approach to understand Central Banks

Big Data Spain 2018

November 2018