Abstract

Intro

Method

* Lexicon
* BERT
* Midas (Forecasting)

Data

BERT vs LEX

* News
* Lexicon

Results

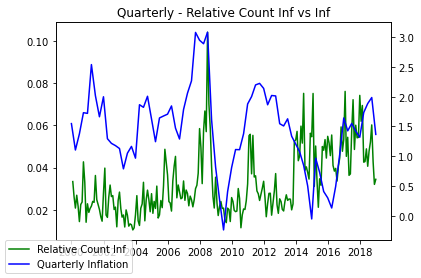
* News
* ECB
* Forecasting

Conclusion

**Introduction**

**Methodology**

**Data**

* **Newspaper data is based on dpa newspaper articles from 1991 to 2018 (subject to change)**
* **ECB data is based on all press statements from 1998 onwards. Only the press statements themselves are considered without the Q&A part.**
* **News about inflation was identified by using keywords which are related to inflation**
* **First observation: Amount of news is not uniformly. **
* **News amount peaked at the inflation peak in 2008 and fell back to its baseline immediately after. Inflation got more salient in the news over the last years.**

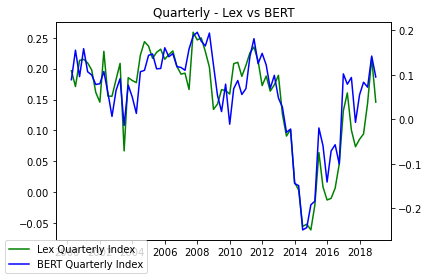
**Results**

**News**

**BERT vs Lexicon approach – News**

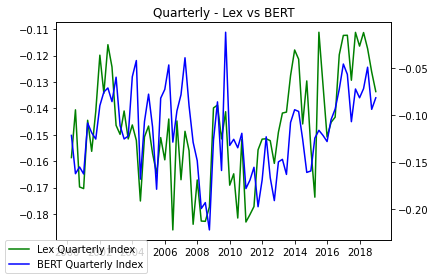
**Direction**

* **News Direction: Lexicon and BERT show similar results. Differences seem to be (more or less) random.**

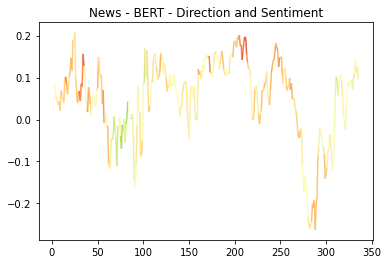
****

**Sentiment**

* **News Sentiment: Time series for sentiment in news regarding inflation is relatively noisy. Differences between BERT and Lexicon are therefore quite large. Simply increasing the size of the training data as well as the size of the dataset could help.**

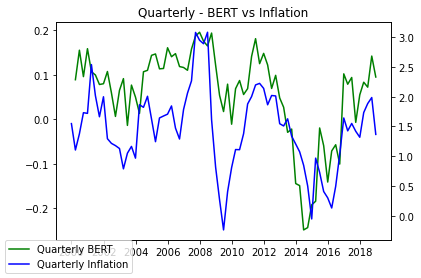
****

* **News Sentiment and Direction: Plotting the sentiment (as color) and direction (as line) in the same graph**

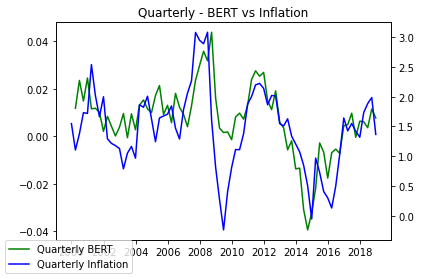
****

**News (BERT) vs real inflation**

* **News vs real inflation: The direction index follows the real inflation relatively closely except for the decreasing inflation during the great recession.**

****

* **News vs real inflation: Combining the direction index with the sentiment index by simply multiplying both with each other produces an index that matches the real inflation even closer.**

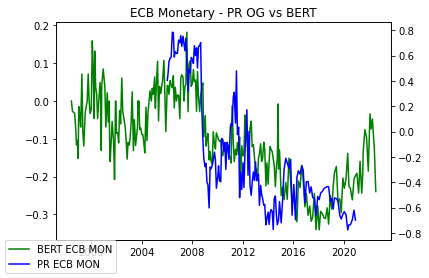
****

**ECB**

**Monetary**

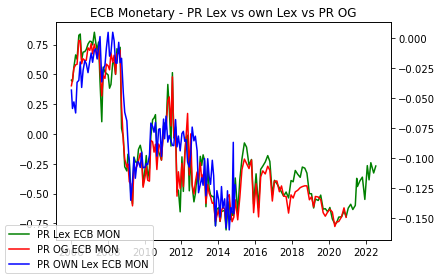
**BERT vs Lexicon approach - ECB**

* **ECB: The original PR index and BERT show the same decrease after 2008 and are overall on the same level. However, the BERT index shows a slightly different behavior after 2012 and increases in 2020. The BERT index suffers from white noise due to the binary sentence classification.**



**Monetary Lexicon comparison**

* **ECB: Comparing the original PR index with the lexicon based on my own trainings-data shows some (random?) differences but overall produce comparable results. The differences between the original PR index and my replication are small but noticeable. If I want to use these results, I must find out why.**

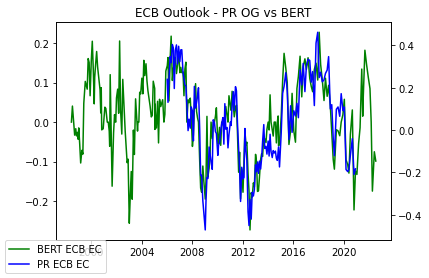
****

* **My own lexicon and BERT seem to work relatively well and hopefully can be used as an alternative to the PR index.**

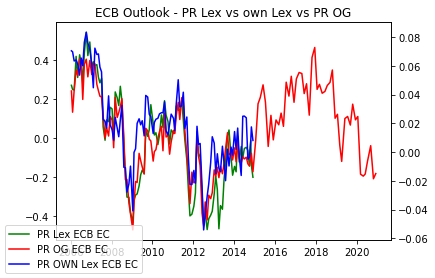
**Outlook**

**BERT vs Lexicon approach - ECB**

* **ECB Outlook: The index based on the original PR lex and my BERT index produce almost similar results. The only noticeable exception can be found during the end of 2014/beginning of 2015**

****

**Outlook Lexicon comparison**

****

* **ECB Outlook: The three lexicon approaches show similar results. There are some notable differences, but they seem to be random. All three approaches produce sensible results.**

**Forecasting**

**Conclusion**

**Appendix**