NLP for Nowcasting and Economic Tracking

Papers by

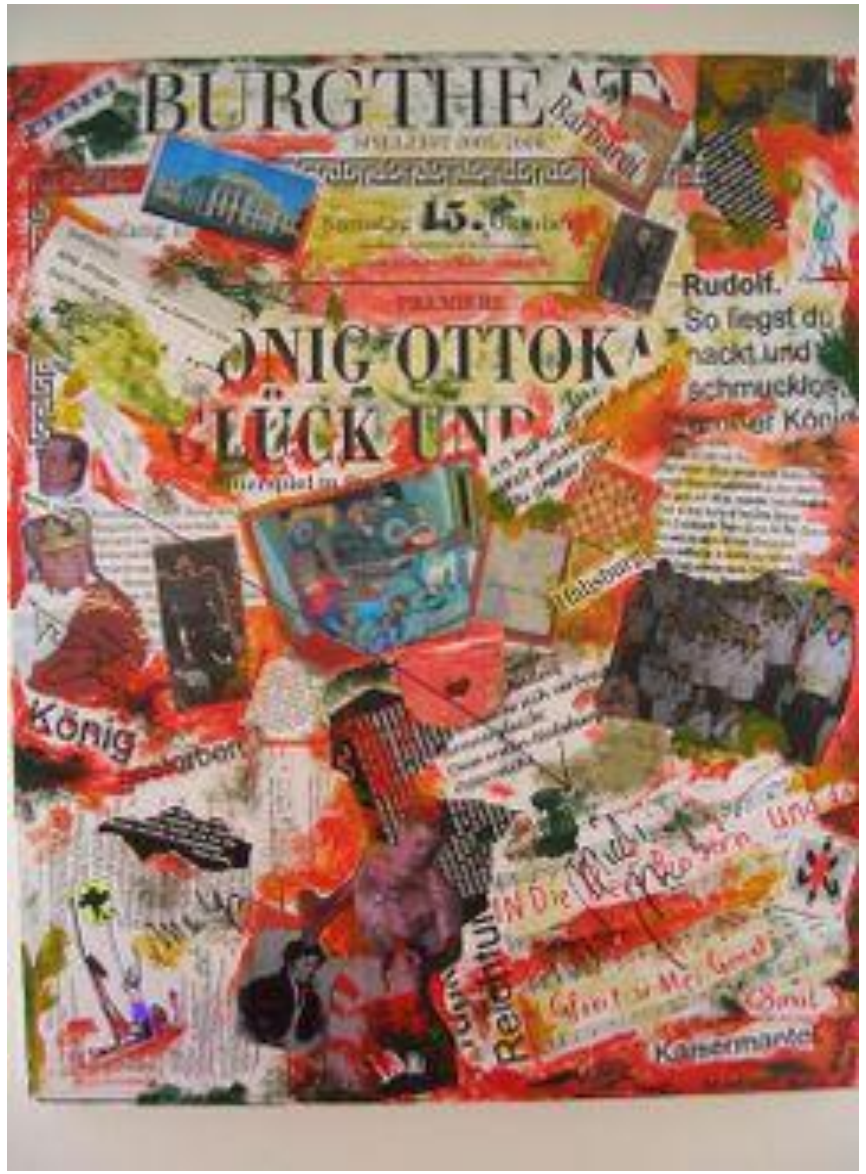
Barbaglia, Consoli, Manzan
Aprigliano, Emiliozzi, Guaitoli, Luciani, Marcucci, Monteforte

Discussion by

Roberta Zizza
Banca d'Italia

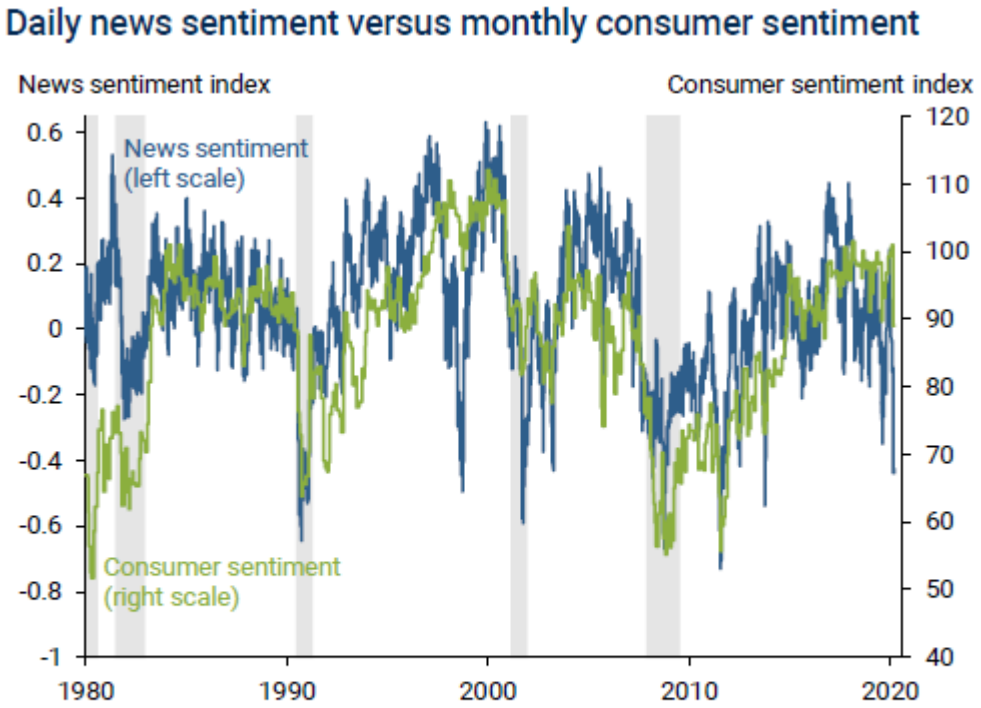
2020 Banca d'Italia and FED Board joint Conference on
«Nontraditional Data & Statistical Learning with Applications to Macroeconomics»
11-13 November 2020

Clipping newspapers used to be a hobby for kids…



… or a way to preserve anonimity…

… but since Baker et al. (2016) a vast literature has flourished showing that clipping newspapers can also serve for building news-based indicators of uncertainty and sentiment, that in turn can help tracking the state of the economy and predicting where the economy is headed



Note: Moving average of daily news sentiment; see Shapiro, Sudhof, and Wilson (2020) for methodology. Gray bars indicate NBER recession dates.
Source: Daily News Sentiment Index and Michigan survey.

Framing in the literature

Two general approaches in the field of computational linguistics to quantify the emotional content in a set of texts:

- *lexical approach*, which relies on a predefined list of words associated with an emotion, referred to as lexicons. Sentiment lexicons typically classify words into three categories: negative, neutral, or positive. <<BOTH PAPERS IN THIS SESSION>>

- *machine-learning* (ML) techniques: learn sentiment weights for words and even entire phrases (based on large "training" datasets), then use those weights to measure the sentiment of the given textual passage.

A bird’s eye view

Barbaglia et al.

- US (6 newspapers)
- Sentiment
- Aspect-based and fine-grained
- Use of different tenses
- GDP, CPI, Empl, IP (q & m)
- Also quantile analysis

Aprigliano et al.

- Italy (4 newspapers)
- Sentiment and uncertainty
- By sector and topic
- By journal
- GDP, C, I, VAServ (q)
- Bayesian MA techniques

Barbaglia et al.

Points of strength:

- *aspect-basedness*: sentiment is calculated only on the basis of the words in a neighborhood of and that are semantically dependent on 6 tokens of interest (economy, financial sector, inflation, output, monetary policy, unemployment)
- *fine-grainedness*: terms are not only positive, neutral or negative, but given a value between -1 and 1
- use of verbal tenses (past, present, future)
- Role of sentiment in the tails of the distribution

Barbaglia et al.

My comments:

- more terms related to *Inflation* could help (deflation, prices, oil, fuel, sales, cost-of-living, bills; see Angelico et al. 2020). Why not 'unconventional policies' in the list for *Monetary policy*?
- select only articles that do not explicitly mention a nation other than the US, since your targets are US aggregates. Get rid of articles on the US-China war trade, on the Iraqi conflict, on the US-North Korea relationship, possibly inducing a bias (more bad than good news)?
- comparison with standard survey-based sentiment measures (e.g. Michigan Consumer Sentiment Index)
- comparison of the indicators based on the three verbal tenses within each token. Is S(*future*) at time t a good proxy of S(*present*) some periods ahead? Could indicators based on future tense be used to extend the forecast horizon?

Aprigliano et al.

Points of strength:

- both sentiment and uncertainty
- develop indicators by topic (15) , sector (21) and journal (bias)
- calculated only on the basis of the words in a neighborhood (within 5 words) of the tokens of interest; use of valence shifters
- horse-race with standard survey-based sentiment measures
- most appreciable results from density forecasts

Aprigliano et al.

My comments:

- extend polarity beyond the dichotomy -1/1
- evaluate verbal tenses
- look at forecasting properties of higher frequency variables (e.g. industrial production)

Q&A and General Discussion

2020 Banca d'Italia and FED Board joint Conference on
«Nontraditional Data & Statistical Learning with Applications to Macroeconomics»
11-13 November 2020