

# Discussion of Leif A. Thorsrud: “Nowcasting using news topics. Big Data versus big bank”

Georg Strasser

European Central Bank

Advances in Short-Term Forecasting, 29 September 2017

The opinions expressed in this presentation are those of the author and do not necessarily reflect the views of the European Central Bank or the Eurosystem.

# Exploring Untapped Data Sources

- Aggregate(d) Information
- Disaggregate (but structured) Information
- Unstructured Information

Still a barely charted territory:

- Relevance, potentially tapping previously hidden information
- Effort vs. benefit of analysing raw data
- Noise?

# Analysis of Textual Data

varies with amount of structure imposed

- 1 Timing a known event (of known size)
- 2 Quantifying and timing a pre-specified quantity
- 3 Extraction of unspecified quantities, quantifying and timing them

Examples

- Shock identification (e.g. Romer and Romer 2010 for exogenous tax changes, Ramey 2011 for government spending shocks)
- Measure prominence of a given topic over time
- Proxy for information flow (information, absolute and relative intensity, timing)

# Approach

- **Textual data** from a leading business newspaper in Norway: daily tone-adjusted topic frequencies, 1988-2014
- Application: Nowcasting **quarterly GDP** growth
- Latent Dirichlet Allocation model to **cluster words into topics**
- Mixed-frequency, time-varying **dynamic factor model**, **threshold** for factor loadings

# Key Findings

- News-based nowcast of GDP (“NEWS”) **predicts revisions** (significant in Mincer-Zarnowitz regression of revisions on first releases)
- NEWS (nowcasts based on textual data) performs **as good as Norges bank forecast (NB)**, and outperforms it in 2007/2009, i.e. at business cycle turning point
- **Tone-adjustment is key** for nowcasting performance

# Network of News Topics

Long sample (1988-2014), thus large set of random news events on which articles are based, 20 of 80 identified topics used

- Some topics have a neutral message (even with given direction): Example topic “Funding” (topic 42) includes loan, competition, creditor, loss, bankruptcy, leverage
- Tone adjustment crucial: Does classification in English fit 1:1 to Norwegian?
- Peculiar top keywords in some topics. Examples: Topic “Nordic countries” (topic 37) rarely covers Finland, “Europe” (topic 56) focuses on France-Germany-Russia, “EU” (topic 25) features “no” as a keyword, “foreign” (topic 77) focuses on Japan, immigration, and games ...

# Topic Variable Dynamics

Modelling approach: daily news process (and output process with missing observations)

Might consider:

- Noisy raw daily news data ... noise removed by backward-looking MA filter
- How persistent is effect of a news article? Of a news topic?
- What kind of news should we capture in model? Headline of the day? Grinding debates?

Other (extreme) approach: quarterly processes, declining uncertainty about most recent observation?

# Latent Topics and Nowcasting

- Latent Dirichlet Allocation model to cluster words into topics: Requires a set of topics from which the them of the article is drawn ( $\theta_m \tilde{Dir}(\alpha)$ )
- Only 20 of 80 topics identified in news dataset used (“truncated”)

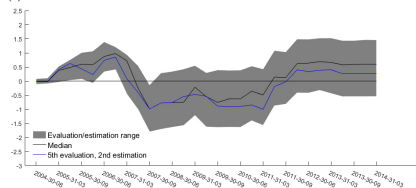
How are latent topics determined? Are they determined ex-ante, iteratively, or ex-post? Likewise: topic truncation real-time or ex-post?

(regime changes might coincide with new topics; compare to latent threshold mechanism for time-varying factor loadings)

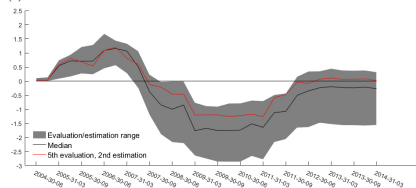


# Cumulative Difference in Squared Prediction Error against NEWS

(a) News relative to *NB*



(b) News relative to *SAM*



# Performance of NEWS vs. NB/SAM

Performance of NEWS relative to judgement (NB) and model-averaging (SAM) is balanced time-wise:

- **Better: 27 months** (01/2007-03/2009, relative to NB only 01/2007-09/2007)
- **Same: 51 months** (04/2009-03/2011, 01/2012-03/2014)
- **Worse: 27 months** (07/2004-12/2006, 4/2011-12/2011)

# Strengths and Weaknesses of NEWS

Judgemental forecasts stem from newspaper *readers*, model averages are based on past performance

- Possible causes of outperformance during 2007/2008: NEWS less dependent on history, more robust to mis-judgement and overfitting
- Possible causes of underperformance during 2011: Blank spots, changes in phrasing, change in perception (e.g. asymmetric inflation risk)?

Direct measure NEWS less prone to perception bias, but maybe struggling with detecting changes of context?

# Forecasts Tailored to Purpose

- Example: Evaluation release
  - Judgement (NB) might be optimized to first release (incentives?)
  - Model-averaging (SAM) might closely resemble procedures of statistical agency for early releases
  - NEWS less focussed on initial release
- Desired forecasting target determines choice of evaluation release
- Predicting the final (5th) release using the second release is not an obvious choice. It is the setup where NEWS shows its relative strengths.

# How to make most of the NEWS forecast

Overall similar performance as NB and SAM, but periods of outperformance and underperformance. What to do?

- Forecast combination
- Find (fix?) cause for NEWS predictor's underperformance in 2011, rendering it the best predictor overall?
- Predictable pattern of periods of NEWS out-/underperformance? (e.g. switching by topic with largest contribution?)

# This paper ...

- Great paper, relevant, timely
- Sophisticated methodology to exploit textual data in forecasting
- Test of practical usefulness might require application to other countries

## Going forward ...

- Quest for “best practice” in nowcasting with textual data: More on what works and what doesn't and exploring why. (Examples: tone adjustment, topic selection)
- Potential to exploit the time-variation in the relevance of “topics” as well ...
- ... and analysis of “higher-order” information in textual data, such a variability of shifts in tone, persistence of news topics ...