Listening to the buzz: social media sentiment and retail depositors' trust

Workshop Harnessing Big Data & Machine Learning Technologies for Central Banks

Discussant: Ralph Verhoeks

DeNederlandscheBank
EUROSYSTEEM

Added value

- The paper makes use of Twitter data to better describe the behavior of retail depositors

- Besides measuring sentiment towards banks, it introduces an original method to capture informational spillovers across banks

- It shows the impact of contagion channels and its relevance for policy

Suggestions for further improvements

1. Time horizon: April 2015 - April 2016

- From a period of 13 months, it is difficult to draw strong conclusions

- A short period does not cover many swings in attention (compare with business cycle).

- Until November '15 no significant variation in sentiment

- The period after the time horizon is interesting with distress at Italian banks

- Currently, the study is more a specific case study for a few months with some banks being in distress in that period

Suggestions for further improvements

2. Definition of distressed banks

- You show there is an important difference between distressed and other banks

- This distinction is being made on the basis of "public interventions in banks".

- Being in distress may not be a black/white issue, but a gradual scale.

- Alternative: measure distress more based on financial indicators.

- Objectify the measurement of distressed banks

- → This is important, as the results are mainly interesting for distressed banks with only 13 x 8 = 104 observations

Suggestions for further improvements

3. Dictionary to classify tweets

- A self-made dictionary is employed with a list of 130 words

- An external word list is preferable, as it is less subjective

- See for instance Loughran and McDonald (2011), who developed an alternative for the *Harvard Psychological Dictionary*, better reflecting the tone of financial text

- Alternative: Italian translations of an external dictionary

- Include your list of words with English translation

- Sensitivity analysis: to what extent are outcomes sensitive to the dictionary used?

DeNederlandscheBank
EUROSYSTEEM

Suggestions for further improvements

4. Forecasting vs. 'nowcasting'

$$D_{i,t} = \alpha_i + \beta \overline{S_{i,t}} + \gamma \overline{I_{i,t}} + \delta \overline{S_{i,t}} * T1R_{i,t-1} + \sum_k \zeta_k C_{k,i,t-1} + \eta\, Y_{i,t-1}$$
$$+ \theta D_{i,t-1} + \varepsilon_{i,t}$$

- 'Forecasting' and 'nowcasting' are being mixed up sometimes

- E.g.: 'the forecasting power of the benchmark market discipline model is improved by Twitter information'

Suggestions for further improvements

- This raises some interesting questions:

  1. Is sentiment leading and can it be used a predictor?

  2. Does it capture 'animal spirits' not yet captured by fundamentals?

  3. Is there any causality? What is leading what?

DeNederlandscheBank
EUROSYSTEEM

Thank you!