Reading the press to read housing markets

Christophe André, Yann Dorville, Emmanuelle Guidetti, Capucine Leroy and Annabelle Mourougane

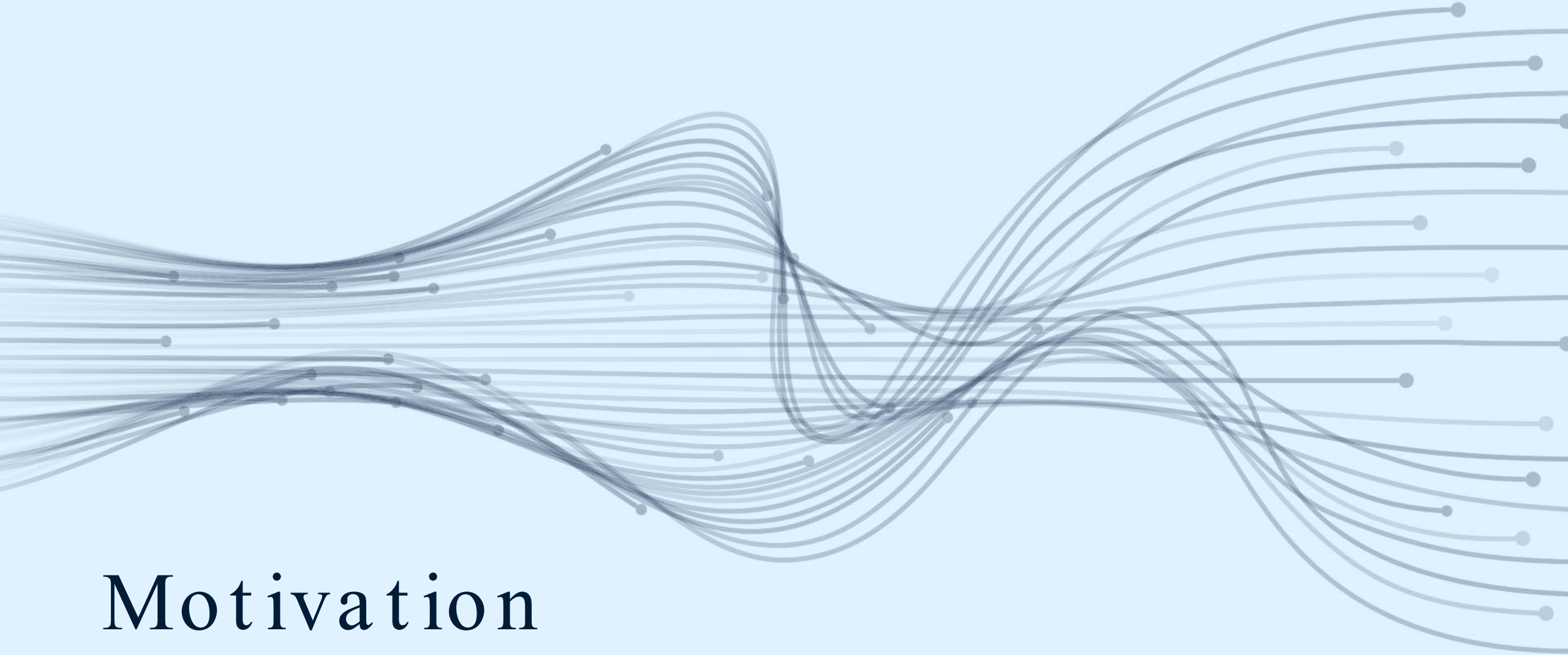
OECD

International Conference on Real Estate Statistics

Tokyo, 19th-20th February 2026

Outline

1. Motivation
2. Approach
3. Preliminary findings
4. Concluding remarks



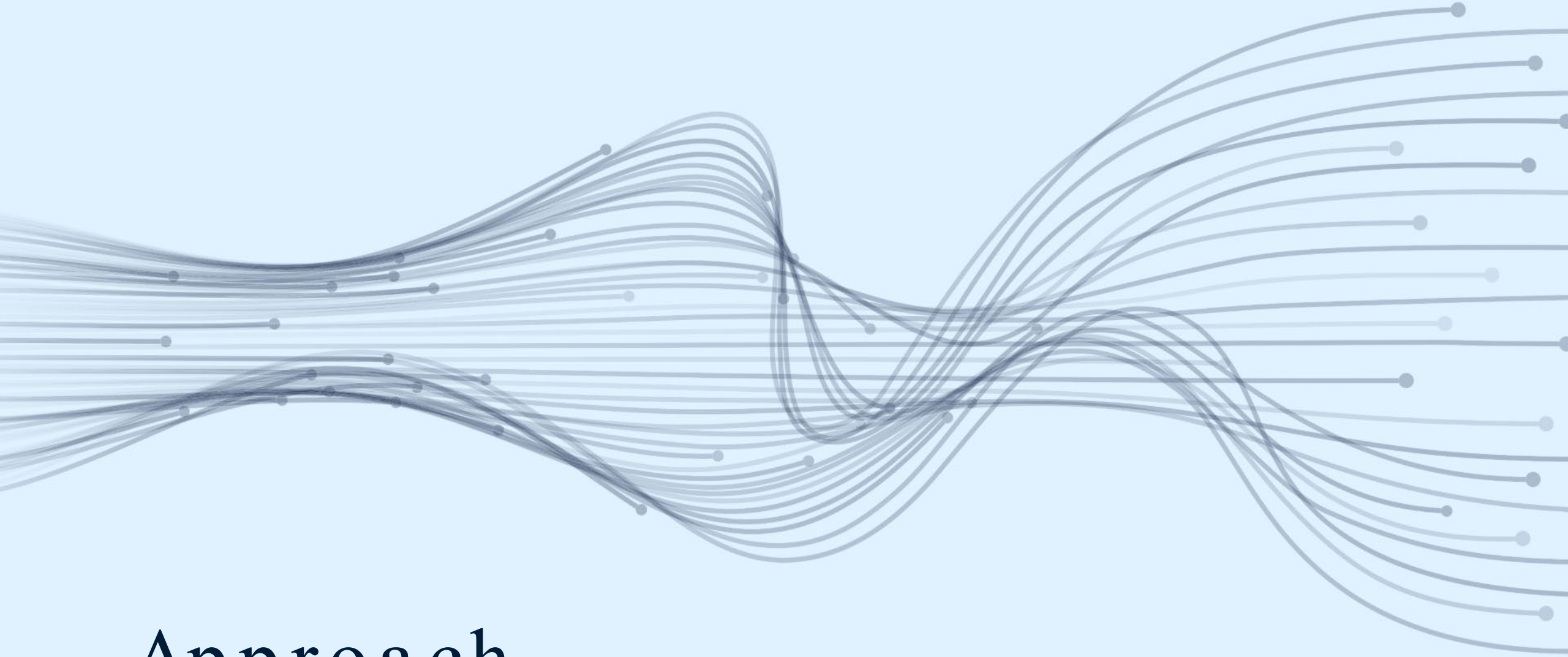
Motivation

Monitoring housing markets

- Macro-economists monitor housing markets closely, given the importance of those markets to understand economic developments and assess financial stability.
- But standard statistics (including house prices) are published with a lag.
- Sentiment indicators, based on surveys, have been developed to help forecast house prices.
- More recently the focus has turned to news-based sentiment indicators.

Economic sentiment is found to bring useful information on housing market developments

- Survey-based economic sentiments have been found to be correlated with some dimensions of housing markets, and at times leading house prices developments (Wilcox, 2015; Lee et al., 2015;Bork et al., 2020, Gupta et al., 2019 and 2020).
- News-based economic sentiments also appear to have predictive power in explaining housing market developments (Cepni et al., 2025; Biktimirov et al., 2024)
- But there is limited evidence in markets outside the United States (Canada, Australia, China).



Approach

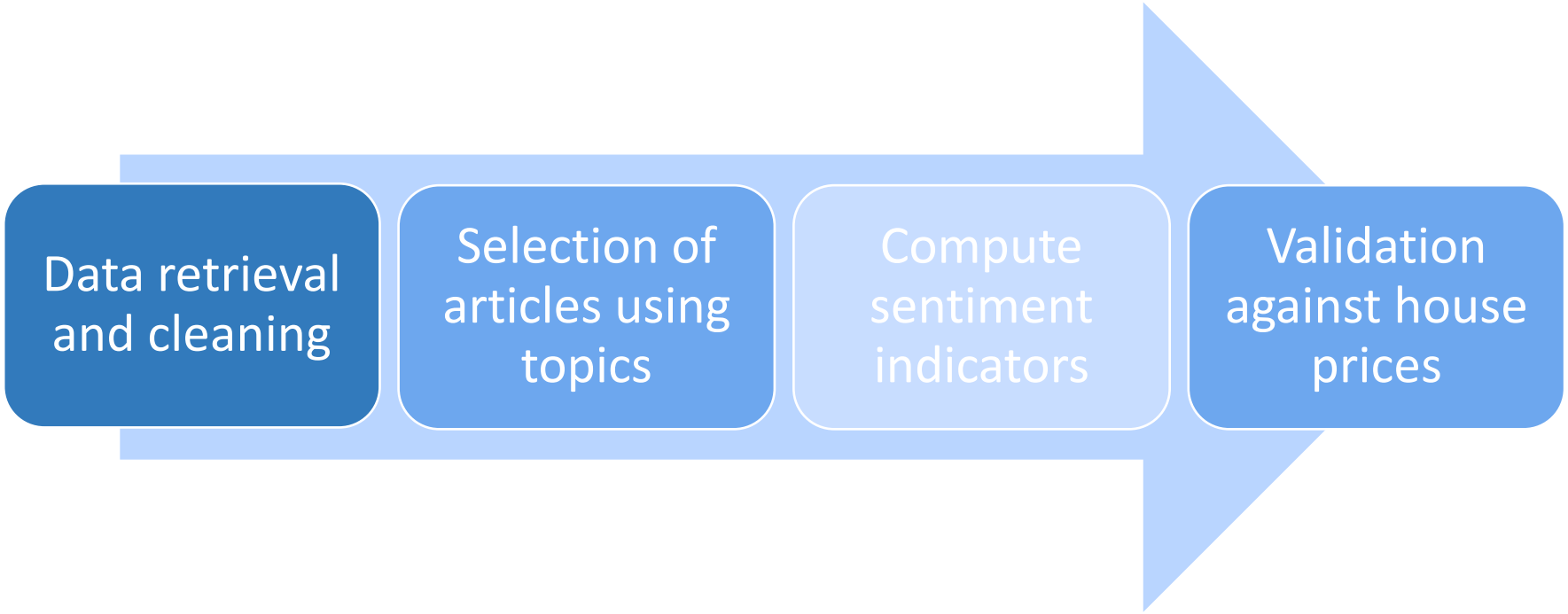
Objective and scope

- Objective : Explore the usefulness of news-based sentiments to get more timely information on house prices.
- 10 OECD Countries:
  - 4 large economies (the United States, Japan, the United Kingdom, Canada)
  - 6 smaller economies (Finland, Ireland, Israel, Korea, the Netherlands and Portugal)
- Period: 2017-25
- Monthly data

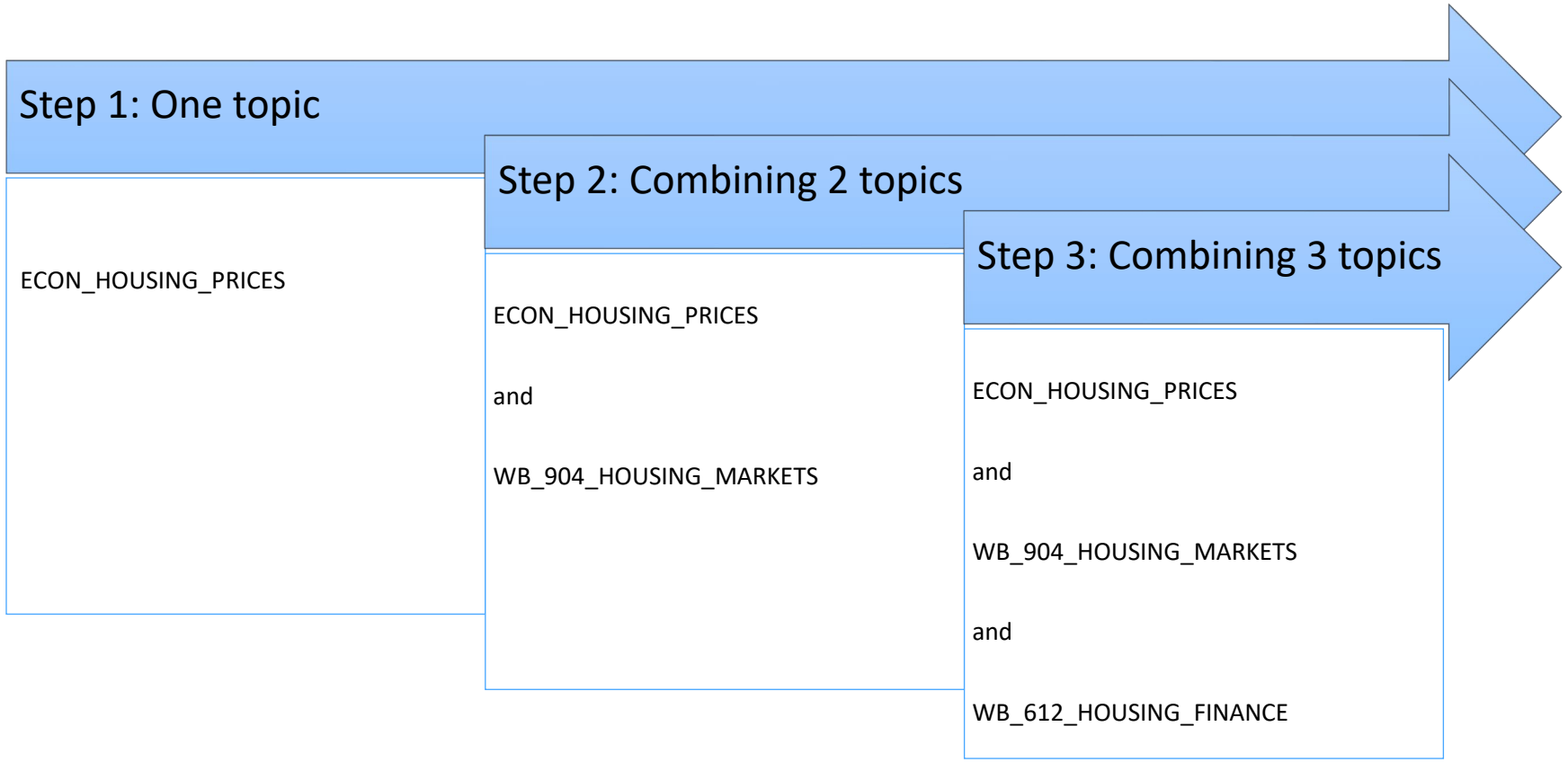
What is GDELT?

- Stands for Global Database of Events, Language and Tone (GDELT)
- Open-source database, collects information on news articles, very timely (updated every 15 minutes).
- We use the Global Entity Graph from GDELT, which includes around 4.5 millions articles per month.
- About 50% of the articles are in English.
- Information : the URL, the title, the language of origin and a range of indicators for each article.
- Articles can be selected using topics.
- Some evidence on the link between GDELT-based economic sentiments and macro-economic variables (GDP, inflation), but not for housing markets.

The approach



Incremental topic selection



How did we measure news-based sentiment?

- The key measure is the score, which captures the average tone of an article.
- A second measure is the magnitude, which measures the emotional content of an article.
- We also test combinations of the score with other metrics (e.g. magnitude) and variants of the score.

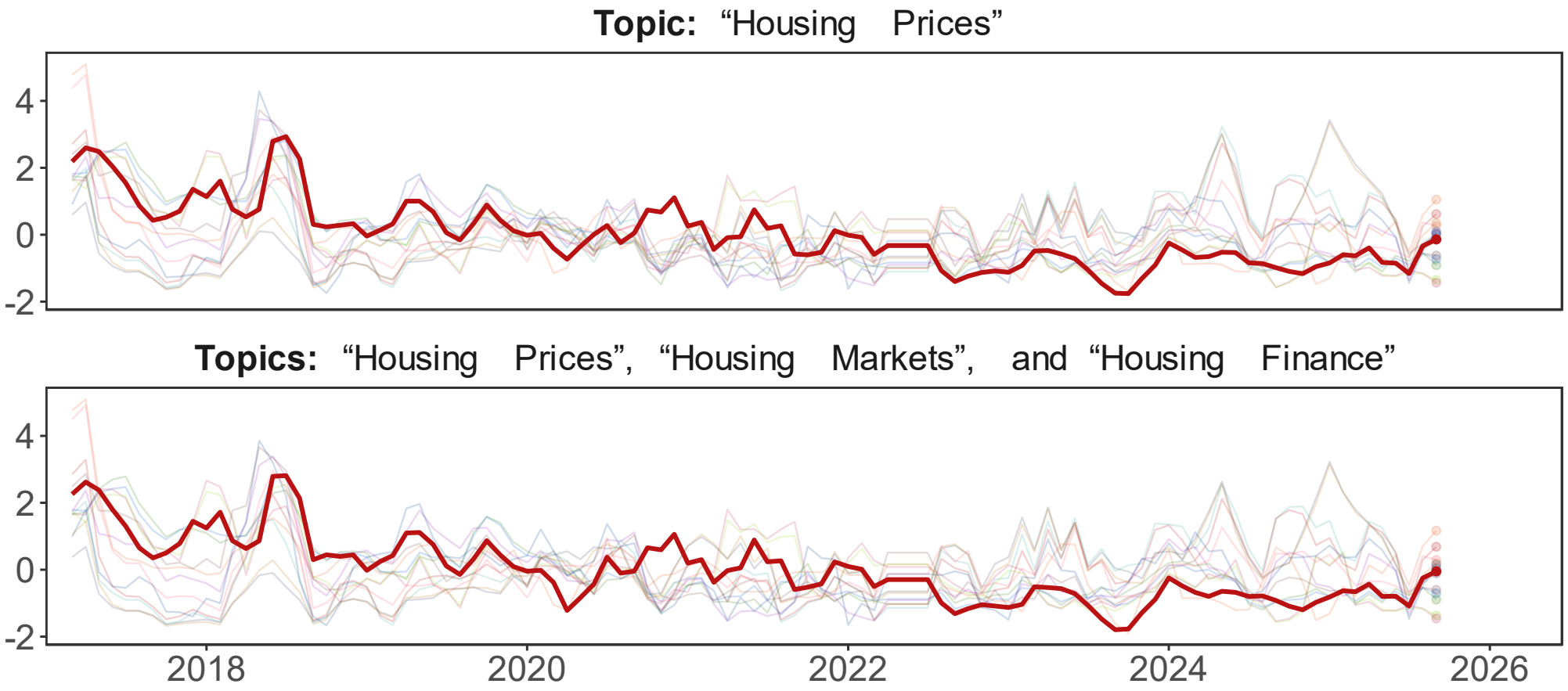
→ In total we test 13 different sentiments (per combination of topics)

Validation

- To assess the plausibility of the indicators, we compare them with real house prices (taken from BIS and deflated by CPI, OECD and FRED) using standard metrics (RMSE, directional accuracy, correlations).
- To assess the usefulness of the sentiment indicators, we examine whether they bring additional information compared to a standard house price model, whereby real house price inflation is explained by fundamentals.
  - Panel estimations using a linear model and GBM
  - Expanding window and quasi-real time test
  - RMSE relative to the benchmark and Clark-West tests

Preliminary findings

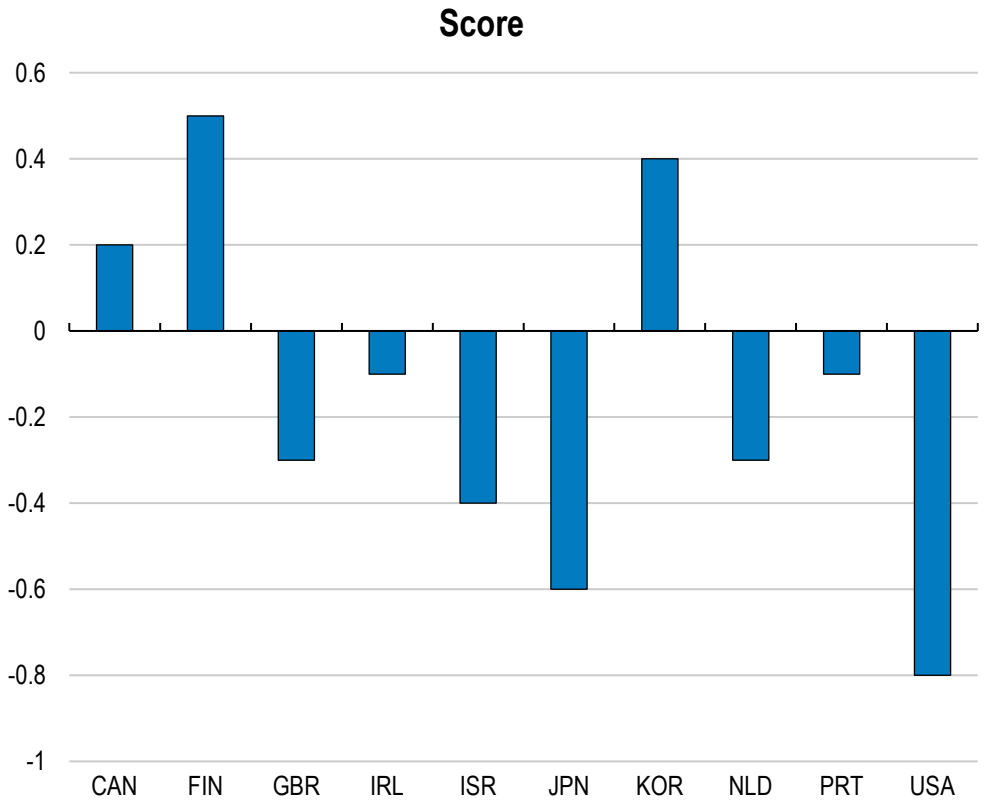
News-based sentiments differ in variability and are dominated by one topic

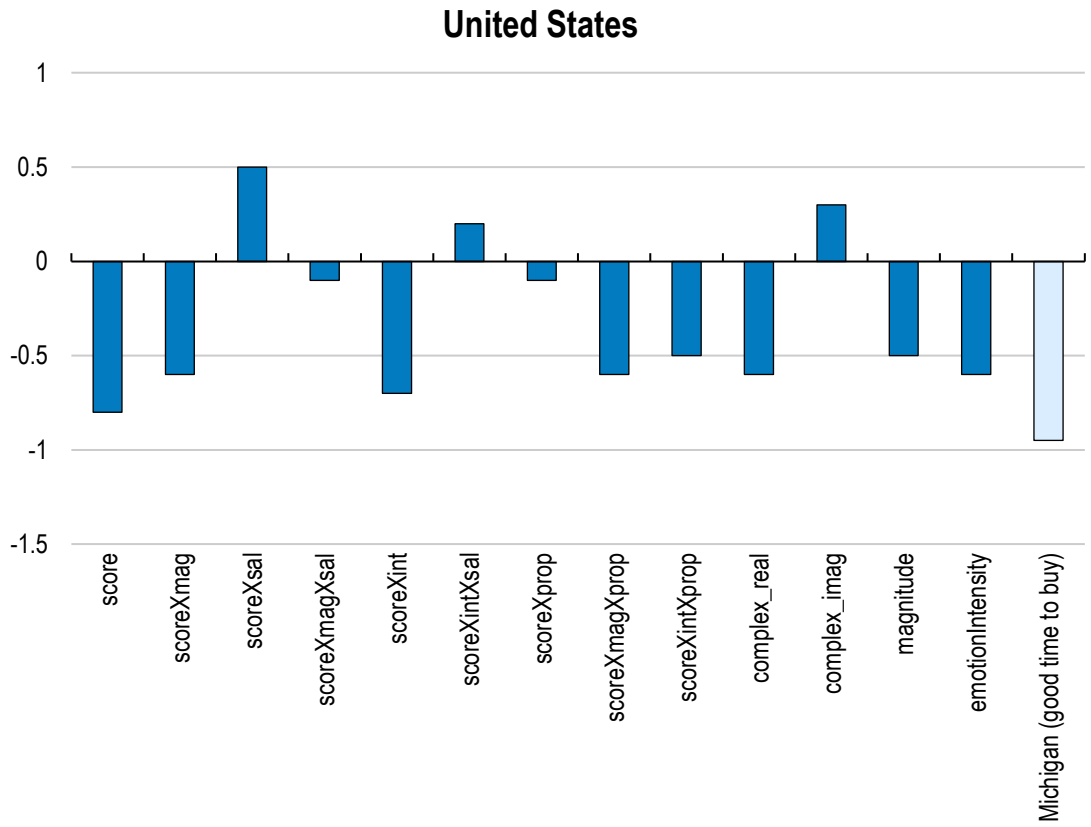


News-based sentiment indices in Japan (Score in red)

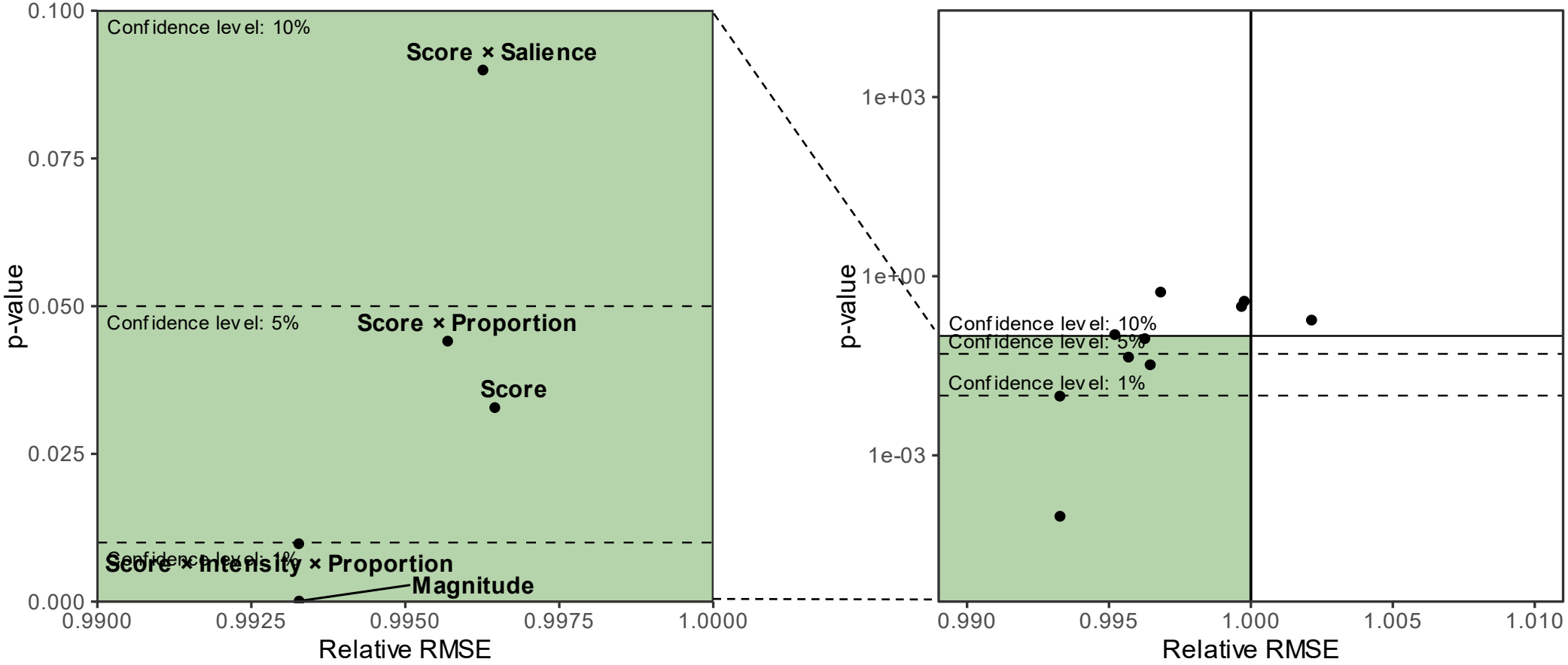
No clear signal from the descriptive analysis

Correlation with real house prices (average 2017-25)



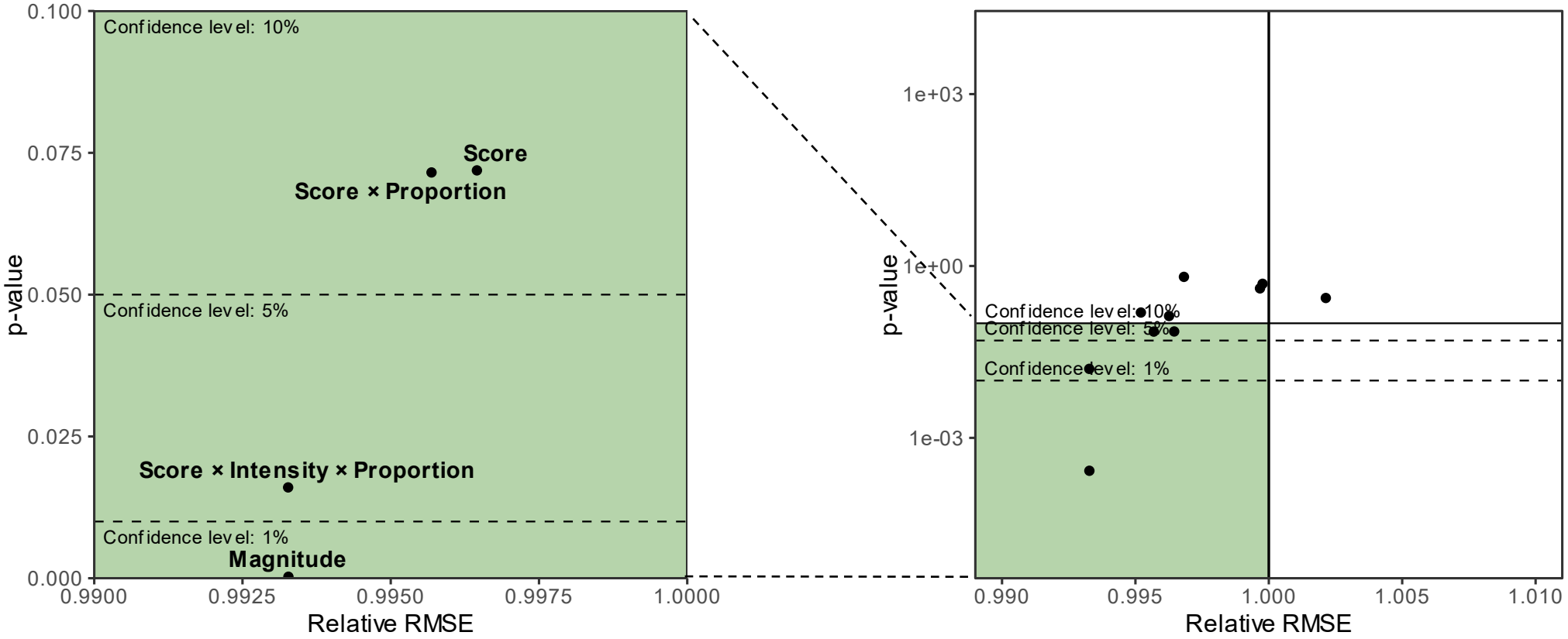


A handful of sentiment-based models outperform a benchmark



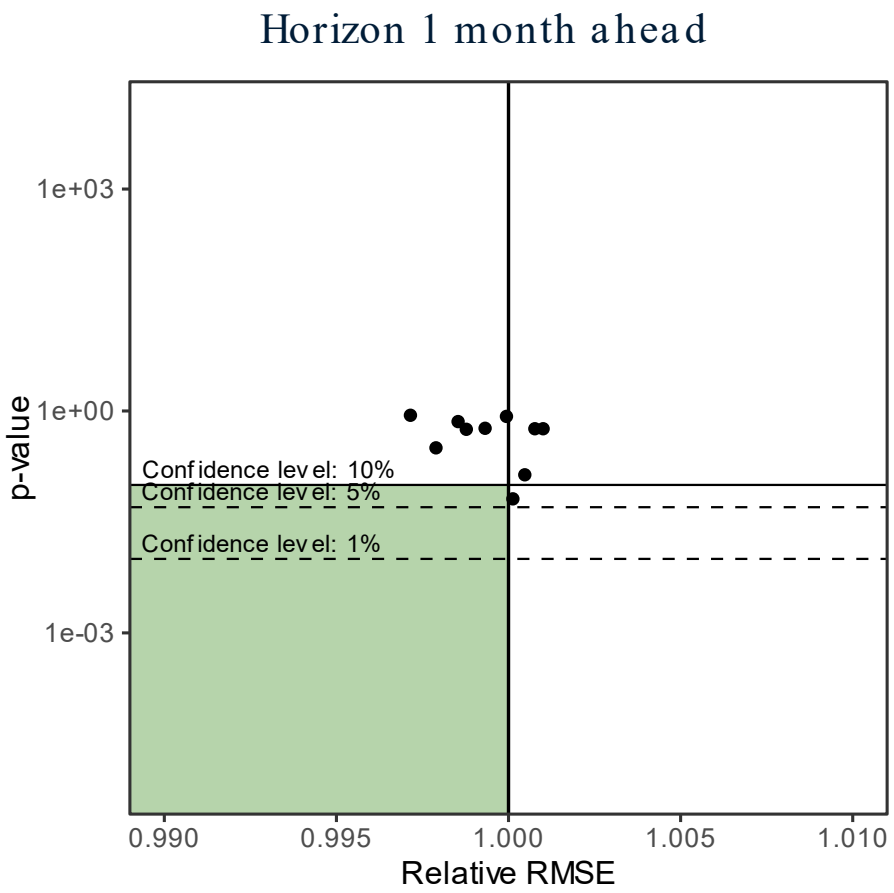
Gradient Boosted Trees at horizon h = 4 months. Indices are constructed with topic: "*Housing Prices*", "*Housing Markets*" and "*Housing Finances*". p-values are computed with a Clark-West test

Robustness test – Diebold Mariano test

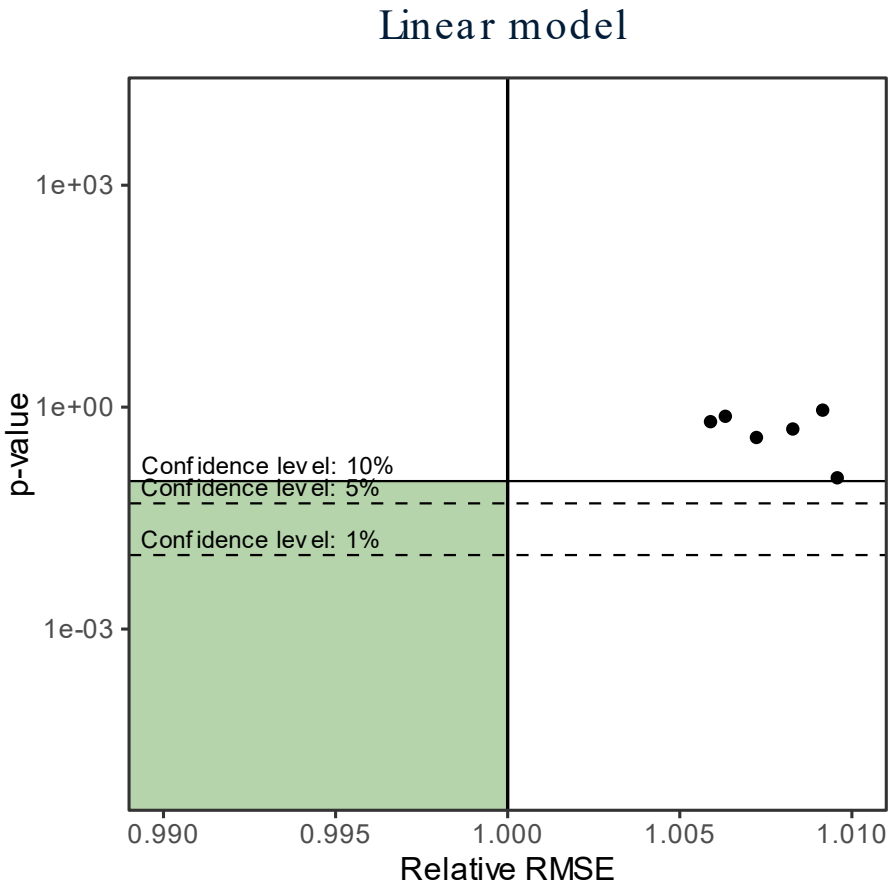


Gradient Boosted Trees at horizon h = 4 months. Indices are constructed with topic: "Housing Prices", "Housing Markets" and "Housing Finances". p-values are computed with a Diebold-Mariano test

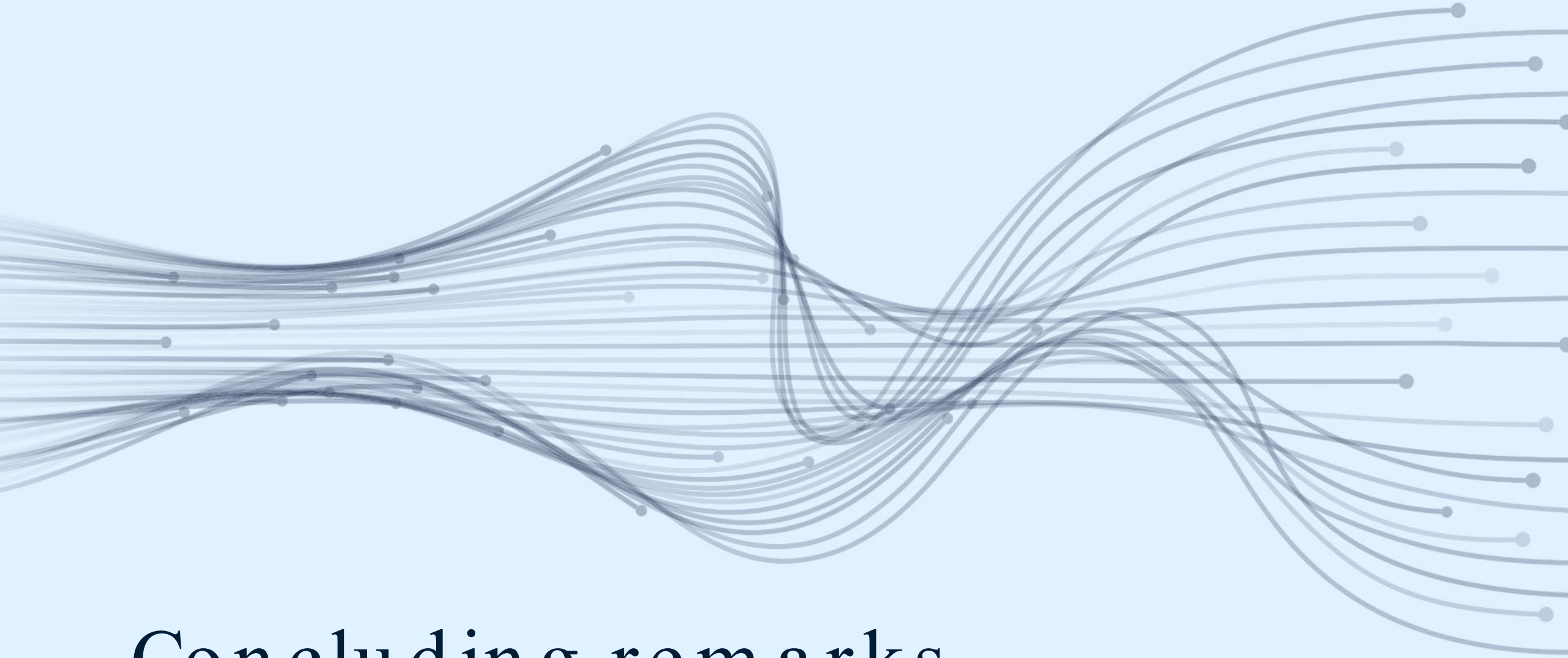
Robustness test – change in the horizon and testing a linear relation



Gradient Boosted Trees, h = 1, Clark-West test



Linear regression, h = 4, Clark-West test



Concluding remarks

Main (preliminary) findings

- A handful of news-based sentiments appear to reduce the prediction errors of house prices.
- The reduction is small but statistically significant
- But this depends on the horizon of prediction
- …and the relation appears to be non linear

Challenges and next steps

- Challenges in using GDELT to complement official statistics, including related to retrieving and processing the data

- Next steps
  - Continue the validation process
  - Run additional robustness test (e.g. adding further topics)
  - Expand the list of countries

Thank you

ありがとうございます