

A Thousand Words Tell More Than Just Numbers: Financial Crises and Historical Headlines



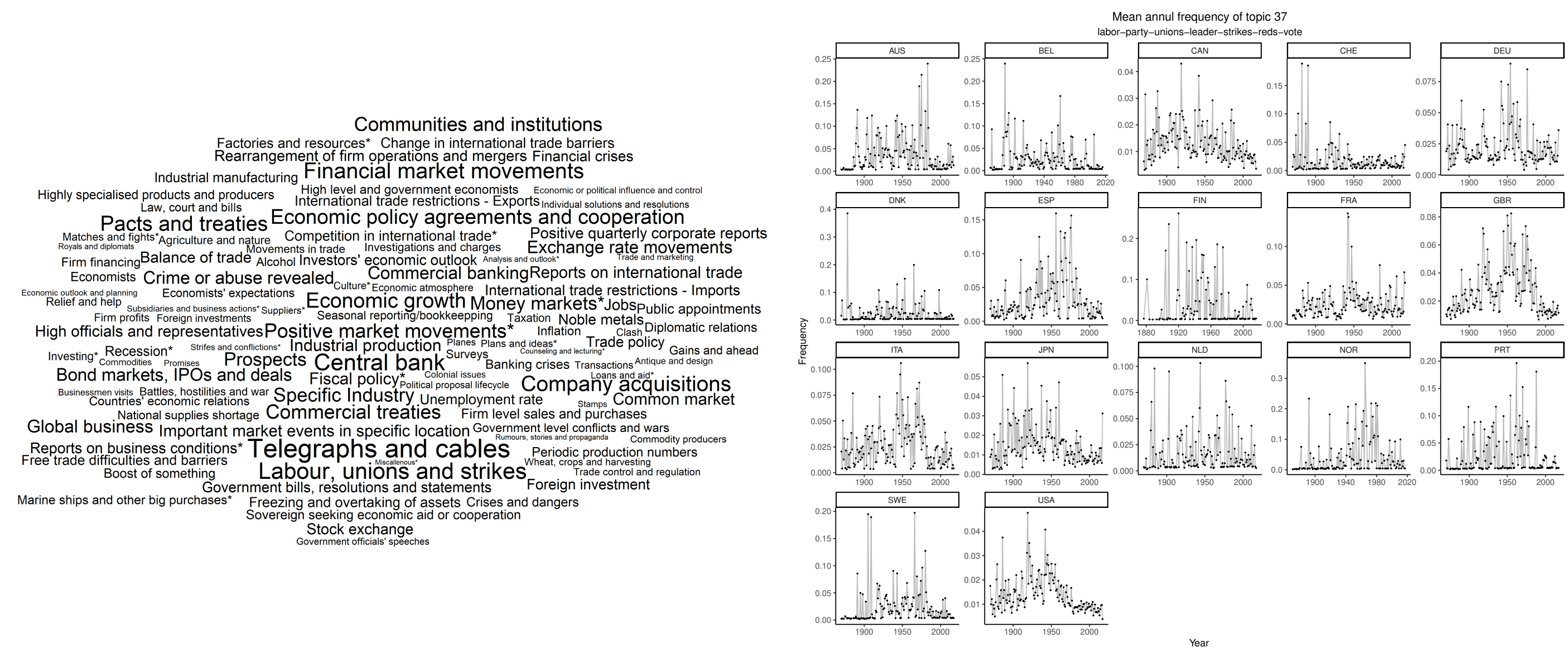
Kim Ristolainen^a, Tomi Roukka^a, Henri Nyberg^a

^aUniversity of Turku
e-mail: kim.ristolainen@utu.fi

Introduction

- We use a large corpus of global newspaper articles titles to extract latent topics and their annual frequencies for 17 developed countries from 1870 to 2016
- We use the topic frequencies to discover if newspaper information can outperform and/or improve the prediction ability of financial crisis prediction models utilizing macroeconomic - and financial variables.
- We utilize case studies for specific historical crisis episodes to identify why the text topics possibly signal a crisis better than the usual economic variables.

Historical Newspaper Topics 1870–2016

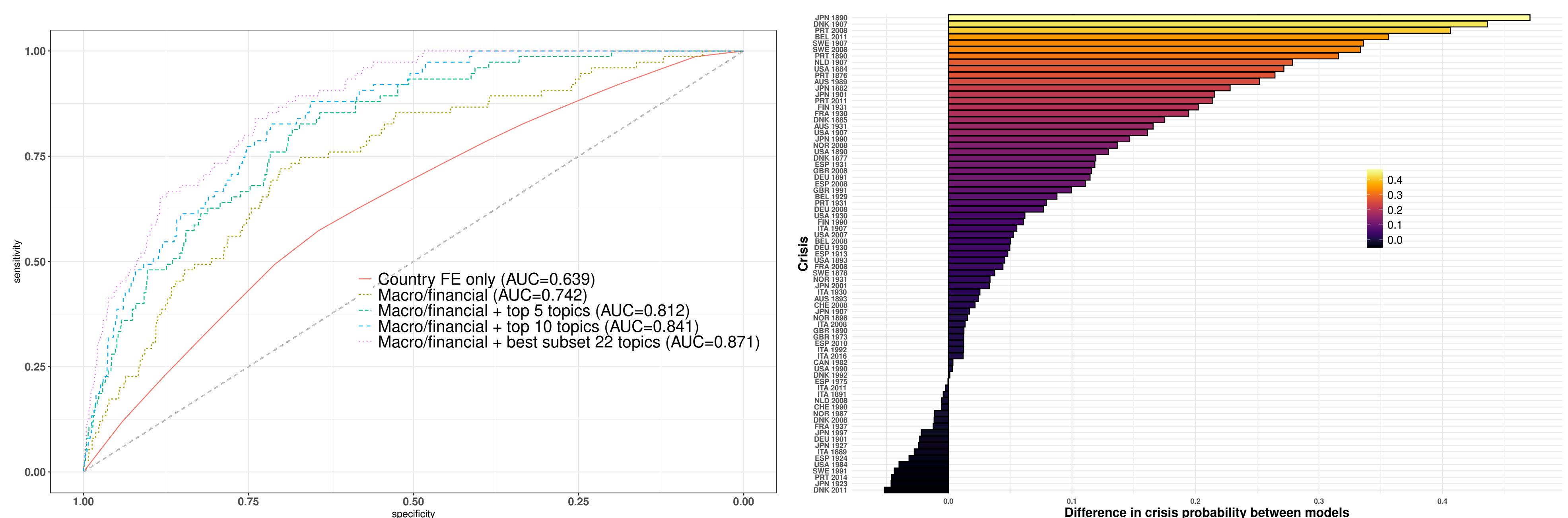


We use the latent Dirichlet Allocation (LDA) model to extract 110 latent topics and their annual frequencies in economy related articles in six global newspapers from 1870 to 2016. The peak years of the topic frequencies for a panel of 17 developed countries capture important topic-country related historical events.

Top Leading Indicators

- Int. trade restrictions - Imports** (-)
- Int. trade restrictions - Exports*** (-)
- Commodities** (+)
- Industrial production*** (-)
- Change in int. trade barriers* (+)
- Relief and help** (-)
- Money markets* (-)
- Political proposal lifecycle*** (+)
- Crises and dangers* (-)
- Foreign investment*** (+)

Predicting Financial Crisis

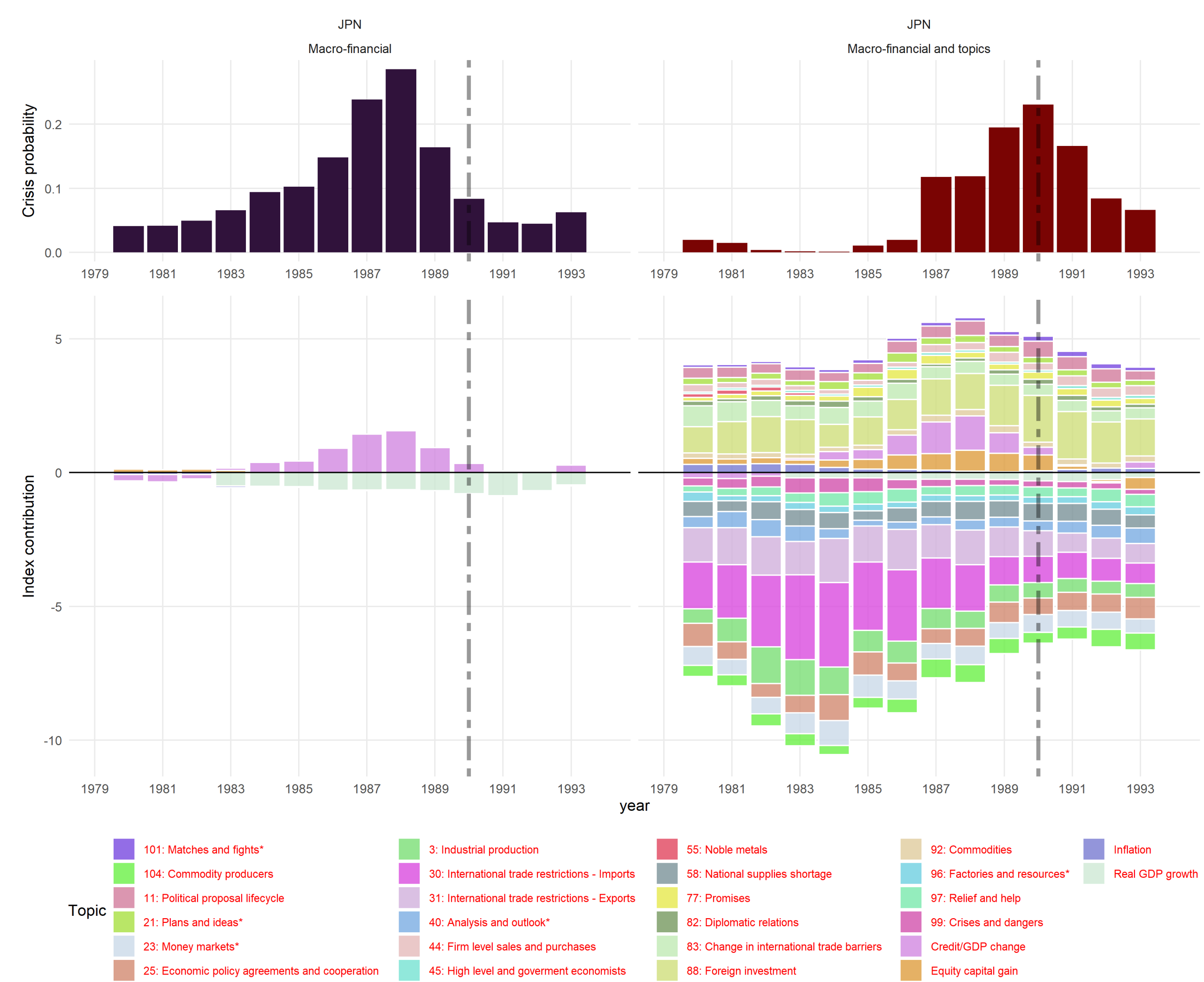


We use a panel logit regression with a forward stepwise selection process to estimate models using the topic frequency variables with and without macroeconomic - and financial variables to predict financial crisis. The models using text-based indicators outperform the baseline model using only the macroeconomic - and financial variables statistically significantly. The results hold in a large set of robustness tests including an out-of-sample prediction exercise.

Conclusion

- We form historical economic topic frequency series that capture topic related historical events for a panel of countries
- We show that these newspaper text based variables outperform the usual crisis indicators by themselves and also improve the prediction accuracy significantly when both types of indicators are included in the models
- Crisis case studies reveal that the best performing topic frequency leading indicators capture phenomena like important changes in economic policies, external shocks, detailed information on production or output and economic atmosphere
- These variables capture events that have been discussed in public and individual studies to have at least partly caused the specific crises episodes, but these variables have not been implemented to crisis prediction models until now

Crisis Case Studies



We perform a series of case studies for specific historical crisis episodes to see what events the topic based indicators capture and why they outperform the prediction ability of the models relative to the baseline models. For example the 1990 financial crisis in Japan was preceded by the USA-Japan trade war in the first half of the 1980s, which came to an end when the Plaza Accord was signed in 1985. This resulted in to the appreciation of the yen w.r.t the dollar, which later negatively affected the competitiveness of Japanese exporters. The Bank of Japan reacted to the appreciation of yen by raising interest rates, which has been argued to be the trigger of the collapse of asset prices. The biggest contributor to the crisis probability during the peak year relative to the preceding years is the decreased frequency or the absence of the international trade restrictions on imports topic.