

Using Newspapers for Textual Indicators: Which and How Many?

ERIK ANDRES-ESCAIOLA, CORINNA GHIRELLI, LUIS MOLINA, JAVIER J. PÉREZ AND ELENA VIDAL

Summary of Banco de España Working Paper no. 2235

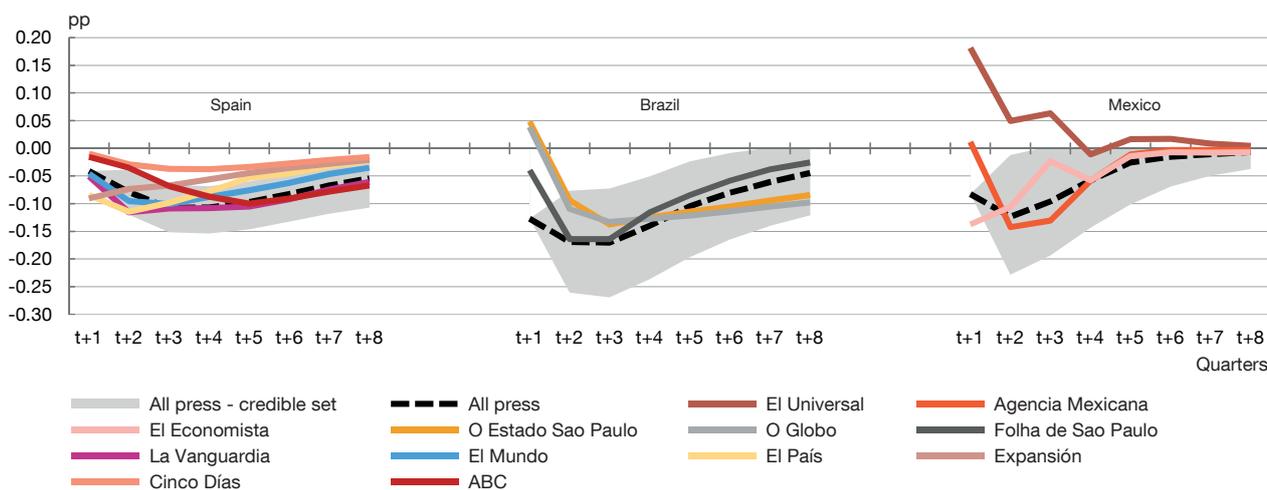
A broad and expanding literature in economics uses textual analysis to develop real-time measures of economic activity, policy uncertainty or social unrest. These measures are in turn related with economic and financial variables to show their relevance for economic and policy analysis and forecasting.

The literature typically focuses on one broad source of text: either the local press (for single-country studies) or the foreign-based press (for cross-country studies), and the results and their interpretation do not take into account

the data sources. For instance, the well-known economic policy uncertainty (EPU) index pioneered by Baker et al. (2016) is based on local newspapers. On the other hand, a very recent literature has constructed newspaper-based indicators using major international news sources (mostly leading Anglophone newspapers and networks), such as the IMF's Reported Social Unrest Index (RSUI) (Barret et al 2020).

There is no clear consensus on how many sources should be used to construct text-based indexes or which type of press should be considered. Furthermore, the literature has developed many newspaper-based economic indicators abstracting from explaining why one newspaper is selected. In our paper we tried to shed some light on this issue by answering the following questions: (i) Do sources with different levels of proximity –meaning local press versus foreign press– convey different information?; (ii) Is it better to use a broad coverage of newspapers or is it sufficient to rely on only one source? These questions are also relevant since

Figure 1
Impulse response functions of GDP to EPU shock, by newspaper (a)



SOURCES: Andres-Escayola et al (2022).

a Each panel depicts the median impulse responses (IRF) of GDP to a rise of one standard deviation in the EPU index in Spain, Brazil and Mexico, respectively. In each panel, the dashed grey line depicts the median IRF of GDP to shocks in the EPU constructed based on all available press (pooling together local and foreign sources) and the grey area depicts its corresponding 84%–16% credible set. For Spain, we consider a set of Anglophone newspapers as foreign sources. For Mexico and Brazil, foreign sources are represented by both Anglophone and Spanish newspapers. All other lines represent the IRFs of GDP to shocks in the EPU index constructed considering only one local newspaper at a time. Filled symbols indicate statistical significance within the 84%–16% credible set, while empty symbols represent not-significant estimates. The horizontal axis measures quarters since the shock.

newspaper repositories are usually not free of charge and the production of newspaper-based economic indicators is growing fast. Also, we prove that the interpretation of results is different according to the source used to build the text based index.

We focus on the well-known and broadly used EPU indicator by Baker et al. (2016) and construct this indicator for six Latin American countries (Brazil, Mexico, Chile, Colombia, Peru and Argentina) and for Spain on the basis of, alternatively, local and foreign newspapers –including Spanish newspapers for the six Latin American economies–, changing also slightly the words that measure uncertainty and economic policy.¹ Then we compare the narratives emerging from these indexes and examine their macroeconomic effects using a structural Bayesian vector autoregressive (BVAR) model.

The main result is unexpected² as it shows that the proximity of the press to the country of interest does not matter. EPU indexes based on local and foreign newspapers show very similar narratives and deliver similar macroeconomic impulse responses. This finding enables researchers to rely on foreign sources to construct EPU indexes. This option may foster the comparability of results across countries and lay the groundwork for cross-country studies of uncertainty.

Related to the second question, we show that it is often the case that macroeconomic responses to EPU indexes based on only one newspaper lie outside the credible bands of responses to the EPU index based on all sources available. Figure 1 provides some examples for the case of Spain, Mexico and Brazil. These results suggest that EPU indexes based on one single newspaper may provide biased results with respect to EPU indexes based on a broad set of newspapers, so the best option would be to rely on multiple sources. Then we advocate maximizing the breadth of the press coverage when constructing newspaper-based indicators: the larger the press coverage, the better.

REFERENCES

- Andres-Escayola E., C. Ghirelli, L. Molina, J. J. Pérez and E. Vidal. (2022). “Using newspapers for textual indicators: which and how many?”, Bank of Spain Working Paper N. 2235.
- Baker, S. R., N. Bloom, and S. J. Davis (2016). “Measuring Economic Policy Uncertainty”, *The Quarterly Journal of Economics*, 131(4), pp. 1593-1636.
- Barrett, P., M. Appendino, K. Engullen, and J. de León Miranda (2020). “Measuring Social Unrest Using Media Reports”, IMF Working Papers 2020/129, International Monetary Fund.

¹ All these indexes could be downloaded from the Banco de España web page (<https://www.bde.es/bde/en/areas/analisis-economi/enfoque/america-latina/indices-de-incertidumbre-de-la-politica-economica--epu--para-los-principales-paises-de-america-latin-dfa6ffaa8331481.html>).

² One possible hypothesis is that news with different levels of proximity convey different messages to the readers and this might affect newspaper-based indicators.