PROJECT BRIEF

News-based Economic Sentiment Index

Pablo Ruiz Carrizosa
Economist · Bogotá
2026 · version 1.1

Can the tone of thousands of news articles anticipate economic activity?

PURPOSE

Activity indicators arrive with lag, but the public conversation does not stop: every day thousands of news articles register optimism or pessimism about the economy. The project turns that tone into a quantitative indicator, with three deliverables: **(i)** a monthly sentiment index for 8 economies in the region since 2015, built with the same rules for all of them; **(ii)** a PPP-GDP-weighted regional aggregate, comparable with the monthly activity indicators; and **(iii)** a battery of validation against each country's official statistics.

DATA AND CONSTRUCTION

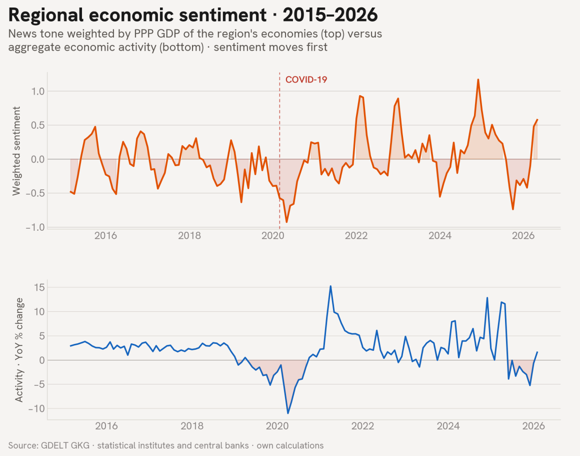
- GDELT, an open database that indexes more than 140,000 press sources; each article about a country's economy receives a tone score.
- Aggregation by country and month, distinguishing economic from political news, with minimum coverage thresholds: a country-month with few articles does not enter the index.
- Standardization of each series against its own history: the index reads in deviations from each country's normal tone.
- PPP-GDP-weighted regional aggregate, with weight redistribution when a country lacks sufficient coverage.
- Homogeneous 2015–2026 series: *scraping* tests with Google News were run, but the index uses GDELT exclusively to avoid an artificial break between sources.

TOOLS

Python for the whole pipeline: bulk download of GDELT (GKG and Doc API), tone scoring with specialized dictionaries, pandas for the panels and statsmodels for the statistical validation.

LIMITATIONS

Press sentiment is a complementary signal, not a substitute for the official figure; media coverage varies across countries and over time, and the country-by-country validation is still preliminary.



PPP-GDP-weighted regional sentiment (top) and aggregate economic activity of the same economies (bottom), 2015–2026. Source: GDELT GKG · statistical institutes and central banks · own calculations.

MAIN RESULT

A higher-frequency, lower-lag signal than official statistics: in some economies the index **anticipates the movements of the official activity indicator**, and the signal is better where press coverage is denser.

FUTURE WORK

Extend the index's coverage to more countries in the region, taking advantage of the own per-country indices that are already built and tested.

pablo.ruiz2000@gmail.com