

## Information gleaned from press articles can help predict economic activity in real time

Most of the short-term economic activity indicators commonly used in the *Notes de Conjoncture* are monthly or quarterly, only becoming available at the end of the month or the quarter in question. However, it can be crucial to be aware of any change in activity even earlier, especially during the Covid-19 health crisis which has given rise to some sudden and far-reaching economic trends. The qualitative information in articles in the French economic press can be used for this purpose. In particular, it could be used to calculate an indicator that measures the tone of media opinion on economic activity. This indicator provides real-time information on the French economy which is often consistent with change in GDP measured *ex post*.

### The aim of the media sentiment indicator is to summarise the qualitative information collected in press articles

Press articles contain a wealth of qualitative information on the current economic context, they deal with a range of topics and are available two or three weeks before the usual short-term indicators, such as the monthly outlook surveys (available only at the end of the month). Using text analysis techniques, automated online data collection (*web scraping*) and *machine learning*, a media sentiment indicator of the French economy was produced based on online articles in *Les Échos*, a French daily newspaper. To do this, the words appearing in an article are categorised and classified as “positive” (or “negative”), according to whether they reflect an “optimistic” (or “pessimistic”) opinion on French economic activity on the day the article was published.<sup>1</sup> An indicator can then be calculated which compares the occurrence of “positive” words against “negative” words: it therefore measures the general tone of the paper regarding the economic situation in the country on a given day. According to the number of positive and negative terms in the

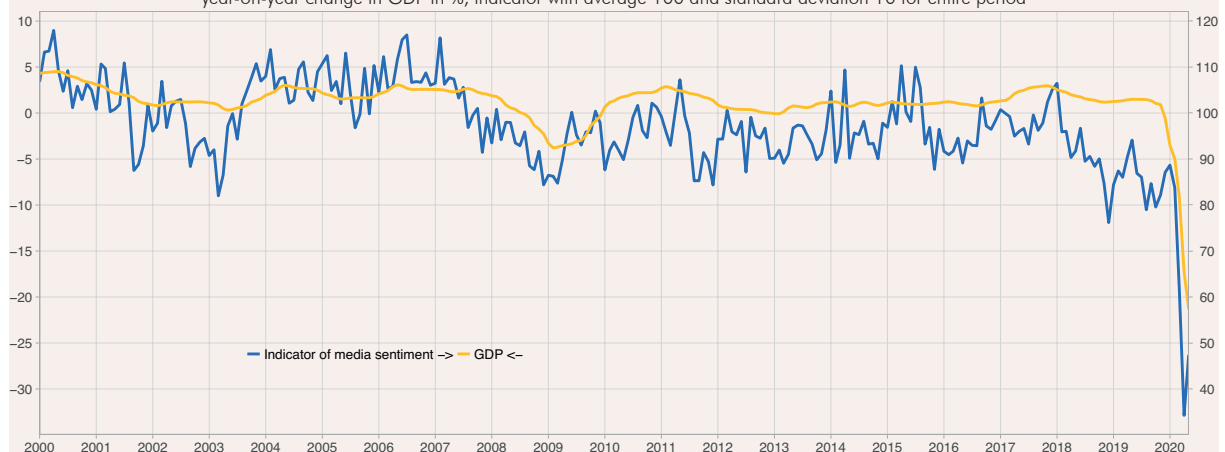
article, a “sentiment score” can be attributed: the value of the media sentiment indicator in a given month is the average of the “scores” from that month’s articles. The indicator was then centred around an average of 100 and reduced to a standard deviation of 10 for the entire period since 2000. Articles containing indicators from economic institutes like INSEE were not taken into account in the analysis in order to avoid any circularity bias. This restriction does not entirely eliminate the possibility of circularity, however, because INSEE publications can also influence the tone of articles even if they are not specifically quoted.

### The media sentiment indicator can anticipate fluctuations in economic activity

Calculated over a long period, the media sentiment indicator often reflects variations in year-on-year changes in GDP, especially during major economic contraction episodes (*Graph 1*). For example, between 2007 and 2008, the value of the indicator fell substantially, whereas in 2008 real GDP declined by 2.2%. Since February 2020, the media sentiment indicator has fallen sharply, well below its previous

1. The methodology used is similar to that described in the article by C. Bortoli, S. Combes, T. Renault, *Prévoir la croissance du PIB en lisant le journal* [Predicting GDP growth by reading the paper], *Économie et Statistiques* no.505-506, 2018.

**1 – Media sentiment indicator and year-on-year change in French GDP**  
 year-on-year change in GDP in %, indicator with average 100 and standard deviation 10 for entire period



How to read it: in May, the indicator was more than 5 standard deviations lower than its long-term average (100 between 2000 and 2020).

Source: INSEE, *Les Échos* daily newspaper. INSEE calculations. Last point: May 2020 for media sentiment indicator, Q2 forecast for GDP

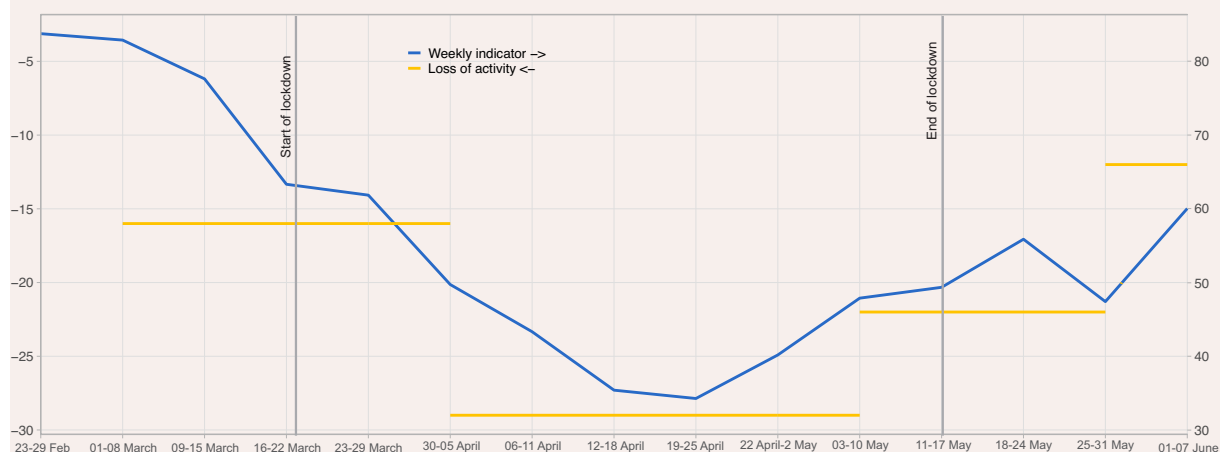
worst levels. This strong deterioration in media opinion on the short-term economic situation is consistent with estimates of economic activity losses in *Points de Conjoncture* since the start of the current crisis.

**The indicator suggests a decline in economic activity even before lockdown was introduced, but this was then accentuated when it actually came into force**

The indicator can be calculated on a daily basis, but because of its volatility, it is not easy to produce a good interpretation. A trade-off between frequency and volatility shows that it is preferable to calculate the indicator on a weekly basis. The resulting indicator shows a sharp drop in activity from the

second week of March, after which the general lockdown of the population was announced then introduced (*Graph 2*). This fall then intensified until the end of April, when the indicator rebounded, reflecting the less pessimistic tone of press articles as the end of lockdown approached. This trend is consistent with the estimated levels of current activity in previous *Points de Conjoncture*. However, the media indicator was a little slow in reaching its low point given that there was a much more sudden decline in activity during the second half of March. Since mid-May, the economic recovery appears to be established: the indicator seems to be returning to the levels observed just before lockdown, but nevertheless remains below its February level. ■

**2 – Weekly media sentiment indicator and monthly loss of activity**



How to read it: at the beginning of March, the indicator was 15% lower than its long-term average (100 between 2000 and 2020). During lockdown, it tumbled to 35, more than 6 standard deviations below this average.

Source: INSEE, *Les Échos* daily newspaper