How to forecast employment figures by reading the newspaper

Clément Bortoli
Département de la conjoncture
Stéphanie Combes
Département des méthodes statistiques
Thomas Renault
Université Paris 1 Panthéon-Sorbonne
IÉSEG School of Management

Factory closures, recruitment drives, huge orders, publication of financial results: the press is a treasure trove of macroeconomic and microeconomic information on the current state of affairs in the business world. This is particularly true when it comes to employment: the media regularly report on decisions taken by businesses which have a direct impact on the labour market. The development of new big data techniques, especially in textual analysis, as well as the fact that some daily newspapers now provide online archives of articles stretching over a long period of time, allow people to make use of this data to create indicators of media sentiment about the current state of the economy.

Among French media, Le Monde *newspaper was selected for the purposes of this study because the content available on its website covers a period of time that is particularly long for France, including many articles printed in the hard copy edition before the advent of the internet. The resulting database contains over a million articles published in* Le Monde *between 1990 and the present day. By combining statistical models and techniques of textual analysis, articles dealing with the economic situation in France can be singled out, leaving a sample of around 200,000 texts. These articles can be classified by tone, positive or negative, based on a list of key words, and thus a monthly indicator of media sentiment regarding employment or the economic situation in general can be calculated.*

An indicator of this type would provide a rapid, pertinent and easy-to-read signal of short-term fluctuations in the economy, displaying similarities to the French business climate indicator derived from the business tendency surveys, published regularly by INSEE. For one thing, both indicators are available rapidly, almost in real time. For another, this indicator of media sentiment regarding the general economic situation in France is also closely correlated with the level of salaried employment, even more so than the monthly indicator of media sentiment specifically regarding employment.

The predictive power of such an indicator can then be assessed and compared to the performance of the business climate indicator. When introduced into a very short-term forecasting model for payroll employment, this media sentiment indicator generally provides real information: from the second month of the quarter onwards, it can significantly boost the accuracy of predictions compared to a simple model based purely on previous fluctuations in employment and economic activity. However, when this indicator is used alone in forecasts, it remains less effective than the business climate indicator. Finally, a model using both the business climate indicator and the media sentiment indicator yields slightly more accurate predictions than a model based solely on INSEE's business climate indicator, even though the improvement in performance is small and non-significant. Therefore, to some extent, it would appear that media sentiment contains certain residual information that is not captured by the business tendency surveys. To put it another way, when it comes to making economic forecasts, particularly for the Conjoncture in France *report, media information can be a useful addition to the INSEE's business tendency surveys but is in no way an adequate substitute.*

How to forecast employment figures by reading the newspaper

In theory, information from the media may be useful in forecasting employment in the market sector

Systematic analysis of the tone of texts published by the media can provide qualitative indications of fluctuations in economic activity in real time. Media information shares certain properties with the information derived from the business tendency surveys: it is available rapidly, several weeks before the quantitative outlook indicators; it can be summarised by a single indicator, called the "media sentiment indicator", which could, in theory, be of use to the forecaster. For example, Thorsrud (2016) uses the published content of certain Norwegian media outlets to generate an advance indicator of economic activity in the country. The rise of websites operated by major media groups, along with techniques linked to big data, make it easier to exploit the content of this wealth of information. Indicators of this kind are also facilitated by the development of open data tools such as Google Trends (Bortoli & Combes, 2015).

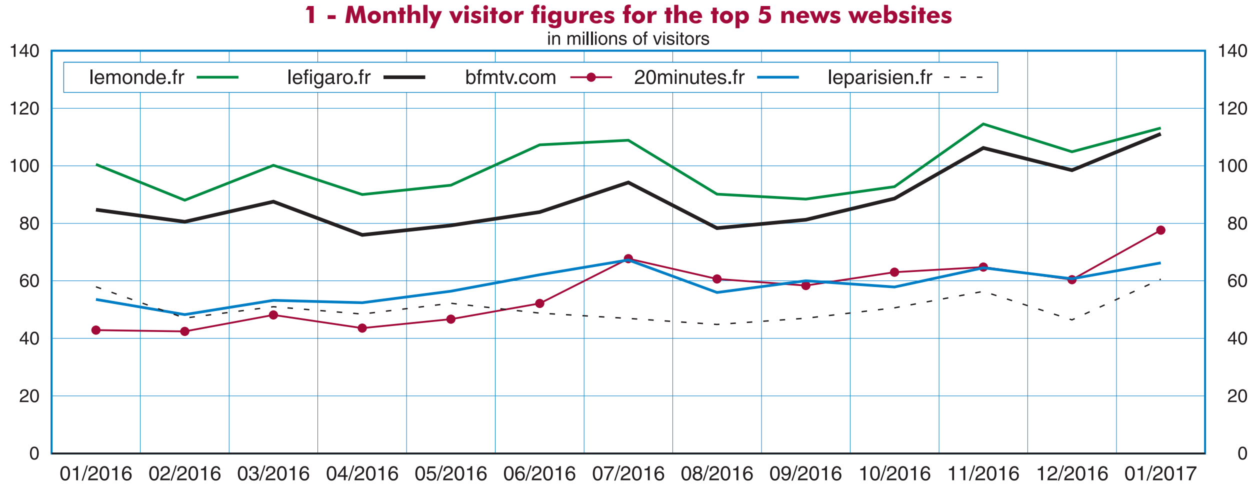
A media sentiment indicator could prove to be particularly useful for estimating employment in the very short term. There is a steady stream of articles dealing with news items which have a direct impact on the labour market, such as the announcement of recruitment drives, the opening of new facilities or, conversely, redundancy programmes. A media sentiment indicator based on these articles can provide information on current and future employment trends, long before publication of the first quantitative data. On the one hand, this indicator could sum up a number of signals which take a certain amount of time to find their way into the statistical system (the first "flash" estimates of payroll employment in the market sector are published 45 days after the end of the quarter in question). On the other hand, the media climate may itself exert a certain influence over business leaders' decisions to create or shed jobs.

Media sentiment indicators are constructed using a database of over a million articles published by *Le Monde* since 1990

A database of over one million articles published since 1990 on lemonde.fr, of which 200,000 are related to the French economy

Among the various French media sources whose published content could be used to construct a media sentiment indicator, *Le Monde* offers some particularly useful characteristics. It is one of France's leading publications: in its paper format, it is currently the second best-selling national daily paper behind *Le Figaro* (circulation of approx. 260,000 copies per day), and its website lemonde.fr is France's most-visited news site, just ahead of the *Figaro* website (*Graph 1*). Furthermore, the content available on the website covers a time period which is particularly long for France, including many articles printed in hard copy before the rise of the internet. This provides a database of 1.4 million articles published between 1990 and the present day.

1 - Monthly visitor figures for the top 5 news websites



Source: ACPM

In order to construct a media sentiment indicator on the current economic situation in France, this initial database was trimmed down in order to single out only those articles dealing with the economy, and whose content was directly related to the situation in France. Various filters and algorithms were used to select these articles, ending up with a working database of around 200,000 articles (*Box 1*).

Scores are given to articles according to their tone, positive or negative

The positive or negative tone of these articles was measured using a "sentiment dictionary", a list of recurring terms tracked in the articles and classified on the basis of their positive or negative connotations. There are already a number of English-language dictionaries designed for textual analysis: the *Harvard IV-4 Psychological Dictionary* is the best-known, but other dictionaries are used in specific fields of research, such as the *Loughran-McDonald* glossary for the financial sector.

Box 1 - How to pick out articles dealing with the state of economy in France from a raw database of 1.3 million published articles

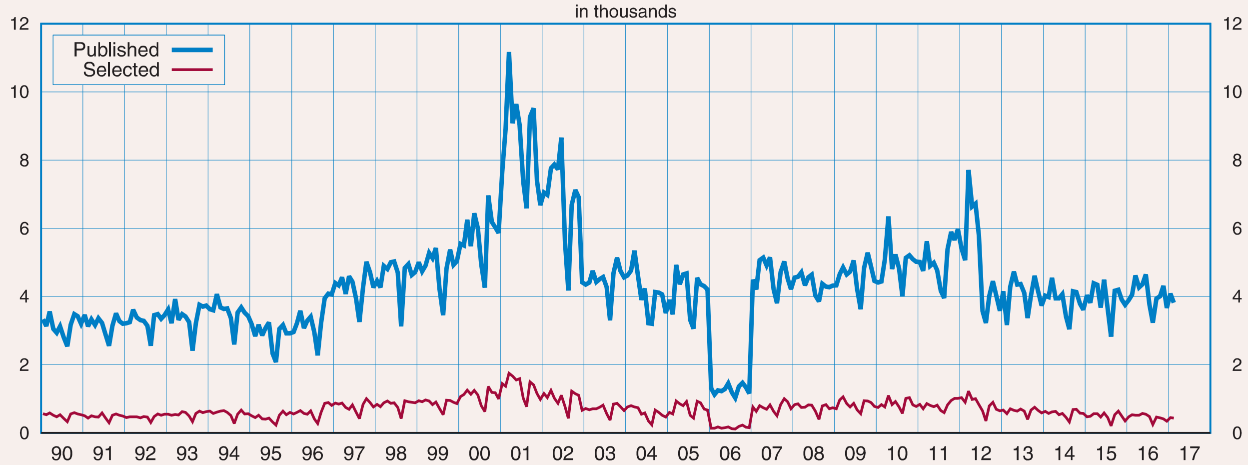
The initial database contains 1,405,038 articles published since 1990 and uploaded to the *Le Monde* website. Access to some of these articles is reserved for paying subscribers: in these cases, only the title, the first few lines and certain information regarding the article (date of publication, author's name and category) are available for free. The more recent articles (published since 2005) are assigned to different categories by the journalists at *Le Monde*: economy, international, politics, sport, etc.

The first step consisted of identifying articles dealing with the economy from among the older texts, i.e. those not already tagged with category information by the journalists. A machine learning algorithm was trained on a sample of 10,000 articles from the economy category and 10,000 from other categories: the algorithm calculated the probability that an article belonged (or did not belong) to the "economy" category based on the frequency with which certain words from both samples cropped up. So the presence of the word "employment" in an article would increase the probability that this text belonged in the "economy" category, because in the training sample this term is more frequently found in economic articles than in articles from other categories. This algorithm, which can be described as a "naive Bayes" classifier (Kotsiantis *et al.*, 2006), served to categorise all of the older texts found in the database.

In parallel, another process was used to identify those articles concerned primarily with France. Two lists containing the names of geographical entities were used: one list containing French terms (names of towns, départements, regions) and another containing international terms (names of countries and their capital cities). The selection process retained only those articles containing at least as many references to geographical locations in France as to locations overseas. Articles containing the names of certain statistical institutions (INSEE, DARES, Pôle Emploi etc.) could potentially have been removed, to avoid creating a media sentiment indicator which would actually be heavily dependent on the publications of these bodies: in practice, this filter was not significant on account of the small proportion of articles concerned each month (no more than 5% of all economic articles).

At the end of this process, only those articles dealing with the economic situation in France remained. The final sample contains 226,493 articles, equivalent to around 700 per month: the proportion of articles retained each month generally fell somewhere between 10% and 20% (see *Graph*). ■

Number of articles published each month, and retained in the final sample



Note: significantly fewer articles were published on lemonde.fr in 2006 than in 2005 or 2007.
Sources: lemonde.fr, INSEE calculations

In French, however, pre-existing glossaries of this kind are much rarer. For the purposes of this study, two sentiment dictionaries were devised, based on an initial list of terms compiled by hand, then enlarged in successive phases with the expressions most frequently associated with the words found in the initial list:

• The first dictionary contains expressions specific to the labour market. It contains 53 positive terms ("création d'emplois / job creation", "plan d'embauche / recruitment programme", "hausse de l'activité / increase in activity" etc.) and 121 negative terms ("destruction d'emplois / job destructions", "plan social / redundancy scheme", "liquidation judiciaire / entering administration" etc.).

• The second dictionary contains more general terms. It comprises 485 words with positive connotations ("amélioration / improvement", "favorable / favourable" etc.) and 1507 words with negative connotations ("instabilité / instability", "affaiblissement / weakening" etc.).

Negative terms are clearly predominant. This imbalance is not unusual (see in particular Schrauf & Sanchez, 2004). A similar disparity can also be observed in many of the English-language dictionaries, for example *Loughran-McDonald* (the *Harvard IV-4 Psychological Dictionary* is far more balanced, however).

Two media sentiment indicators were calculated, one for employment and the other for the economic situation in general

Each one of these dictionaries was used to assign a "sentiment score" to each article, based on the number of positive and negative terms found in the text. Two scoring systems were possible: "continuous coding" or "discrete coding". The indicators obtained via these two methods have similar statistical properties, but continuous coding increases the accuracy of forecasts, so this option was selected (*Box 2*).

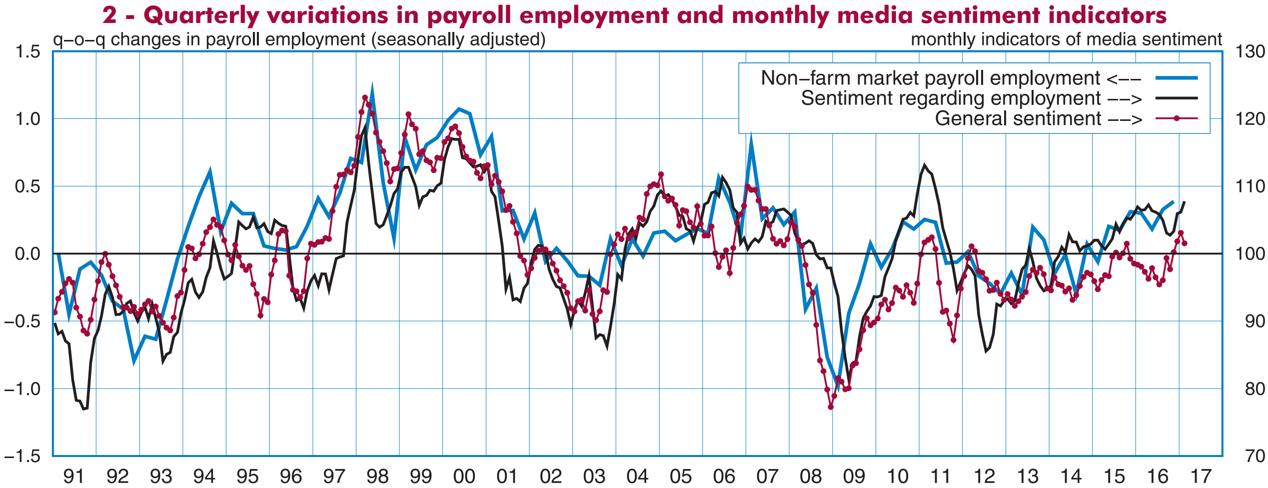
Media sentiment for a given month is assessed by calculating the mean score for all articles published in the month of interest. Using this protocol, two indicators are derived from the dictionaries: an indicator reflecting media sentiment on the topic of "employment", using the glossary of terms and expressions specific to the labour market, and a "general" indicator of media sentiment concerning the economic situation in France, using the generic glossary.

For forecasting purposes, general media sentiment is less effective than the business tendency surveys, but can complement them

These indicators appear to show a strong correlation with variations in employment

The media sentiment indicators obtained in this manner appear to be closely correlated with the quarterly variations in employment since 1990, with a coefficient of 0.7 for sentiment regarding "employment" and 0.8 for the "general" sentiment (*Graph 2*). The general sentiment indicator also appears to

2 - Quarterly variations in payroll employment and monthly media sentiment indicators



Note: a moving average of order 5 has been applied to the media sentiment indicators to make them easier to read.
Source: INSEE

be slightly ahead, timewise: turnarounds in the labour market appear to affect the general indicator before payroll employment, while sentiment regarding employment appears to react to short-term fluctuations on the labour market only after a slight time lag. *Ex-ante*, the general sentiment indicator would thus appear to be a more useful forecasting tool than the employment sentiment indicator.

Modelling confirms that the general sentiment indicator has predictive power

These media sentiment indicators may be used for forecasting purposes. Forecasts for payroll employment (quarterly variable) are made using the monthly indicators, by means of a commonly used approach known as blocking. This involves constructing a different forecasting model (or "calibration") for each month in the quarter, in order to make use of all of the information available at a given date (see for example Bec & Mogliani, 2013). Forecasts for the variation in payroll employment in the market sector during the current quarter therefore become more reliable as the quarter goes on (*Appendix*).

Box 2 - Scoring the articles and calculating the media sentiment indicators

To start with, the dictionaries were used to assign a "sentiment score" to each article, based on the number of positive and negative terms it contains. Two calculation conventions were tested. First of all (the "continuous coding" model), the sentiment score assigned to article *i* published in month *t* was calculated on the basis of the number of positive terms it contained (p_{it}), along with the number of negative terms (n_{it}) and the total word count (m_{it}), as per Baker *et al.* (2016):

$$sentiment_{it} = \frac{p_{it} - n_{it}}{m_{it}}$$

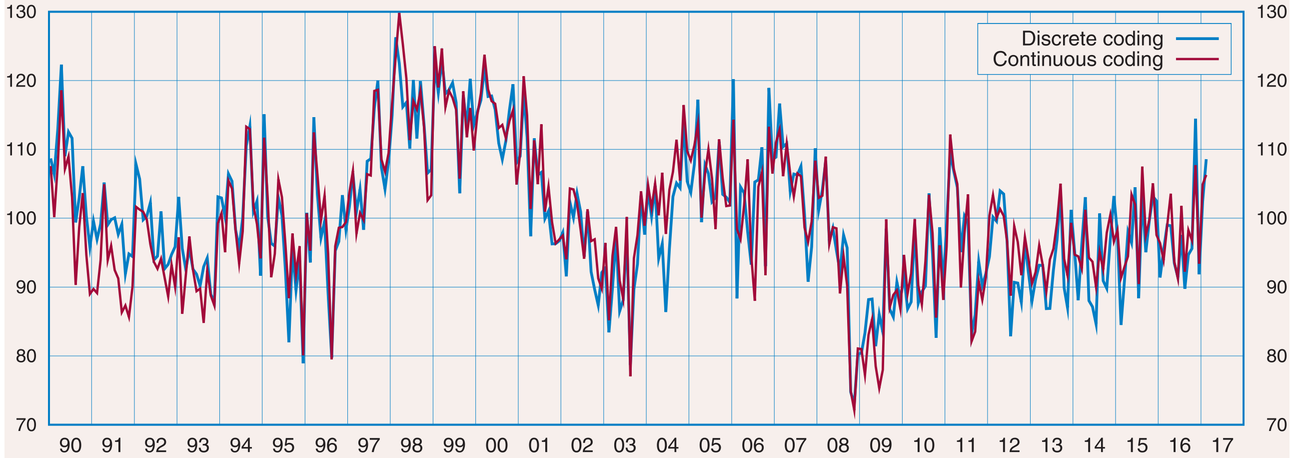
This formula presents certain disadvantages, bearing in mind the database used here: the database contains a number of very short texts, particularly those entries for which the whole article is reserved for paying subscribers. If the non-neutral terms are concentrated towards the start of the article, this calculation method risks over-estimating their proportional significance in the article as a whole. In order to get around this problem, "discrete coding" was also tested. This is simply a matter of comparing the respective numbers of positive terms and negative terms, without taking the length of the text into account. The score assigned to the article is determined as follows:

$$sentiment_{it} = \begin{cases} 1 \text{ if } p_{it} > n_{it} \\ 0 \text{ if } p_{it} = n_{it} \\ -1 \text{ if } p_{it} < n_{it} \end{cases}$$

Once a sentiment score had been assigned to each article, a monthly media sentiment indicator was calculated in the form of a simple mean value. The time series was then normalised to give a mean of 100 and a standard deviation of 10; this is a standard normalisation process, similar to that used to create the composite business climate indicators derived from the business tendency surveys.

In practice, the indicators obtained via continuous coding were preferred because they offered slightly better forecasting capacities when applied to the calibration models included in this report, compared to the discrete-coded indicators. Nonetheless, both types of indicator achieved a correlation score of over 0.90 (*Graph*), ensuring that the approach adopted to summarise media sentiment was robust to the choice of method used to score the articles. ■

Two different types of coding for the general media sentiment indicator



Source: INSEE

The general media sentiment indicator contains information which may be useful for forecasting purposes from the second month of the quarter onwards. When the model using this indicator as its only source of exogenous information is estimated for the whole period studied here (1990-2016), the indicator proves to be significant in every month of the quarter. Moreover, the adjusted R^2 value is higher than that observed in a model based solely on lags in activity and employment, a sign of a closer fit with the data. Finally, a "real time" forecast simulation indicates that, from the second month of the quarter onwards, the root mean square forecasting error (RMSE out of sample, or RMSFE) decreases significantly when the media sentiment indicator is added to the model based solely on lags in activity and employment (*Table*). However, the media sentiment indicator for employment is less accurate than the general indicator, despite the fact that it is supposed to be more focused on this topic. It does not even add any new information to the past trends for employment and activity.

The business climate indicator derived from the business tendency surveys offers greater predictive power

In spite of the information it provides, the general media sentiment indicator cannot be considered as a substitute for the composite indicators currently constructed on the basis of the business tendency surveys. These surveys, conducted each month by INSEE on a sample of 15,000-20,000 enterprises in the market sectors, are used to assess the business climate. They are also used to produce the forecasting scenario set out in *Conjoncture in France* (see *Special analysis* "Forecasting employment based on business tendency survey responses", p. 19). The media sentiment indicators are in no way a replacement for this tool, because INSEE needs to have an independent, controlled source of information for assessing the business climate: the long-term stability of the media sentiment indicator could be affected by external events (changes in publishing strategy, changes in editorial line etc.). But, above all, the indicators derived from the business tendency surveys retain their superior predictive power. A model based solely on the business climate indicator for the French economy yields more accurate employment forecasts than a model based solely on the media sentiment indicator, regardless of the month of the quarter in which the forecast is made: the former model more closely matches the data when estimated for the period as a whole (higher coefficient of determination, or R^2) and yields significantly fewer forecasting errors when producing "real time" simulations (lower RMSFE).

Forecasting error over the period 2000-2016, depending on the model used and the month in the quarter

As a % - Explained variable: quarterly variations in employment in the non-farm market sector (standard deviation: 0.4%)

	Model (1): past employment and activity only	Model (2): media sentiment only	Model (3): French climate only	Model (4): media sentiment and French climate
1st month of the quarter	0.213	0.211	0.194*	0.193
2nd month of the quarter	0.216	0.194*	0.170*	0.168
3rd month of the quarter	0.216	0.194*	0.164*	0.161

How to read it: all the models contain employment and activity lags. For each month within the quarter, the stars indicate that, according to the test devised by Harvey, Leybourne and Newbold (1997), the root mean square forecasting error (RMSFE) of the model is significantly lower at the 10% threshold than the RMSFE of the "previous" model. As such, in the first month of the quarter, the RMSFE in Model 2 (media sentiment only) is not significantly lower than the RMSFE in Model 1 (past employment and activity only), but the RMSFE in Model 3 (business climate only) is significantly lower than the RMSFE in Model 2.
Source: INSEE

Simultaneous use of these two indicators yields slightly more accurate predictions, but the improvement in performance is not significant.

Nevertheless, media sentiment and the business climate appear to complement one another, at least to a certain extent. Whichever month of the quarter under consideration, the model using both the media sentiment and business climate indicators outperforms the model which uses the business climate as its only exogenous information: the media sentiment indicator is significant, the "sample" performances are better (higher R^2) and forecasting errors "in real time" are slightly lower. However, the difference in RMSFE is too small to conclude that the increase in predictive power is significant. The media sentiment thus appears to contain residual information which is not fully captured by INSEE's business tendency surveys. This residual information could to some extent be useful in more accurately predicting variations in market sector employment. This indicator could thus be a useful addition to the tools already at forecasters' disposal when producing a diagnosis of the short-term outlook. The slight upturn observed in this indicator since mid-2016 complements the recent improvement in the business climate since late 2016. This bodes well for job creation, which should remain solid in early 2017. ■

Bibliography

Baker S. R., **Bloom N.** and **Davis S. J.** (2016), "Measuring economic policy uncertainty", *The Quarterly Journal of Economics*, vol. 131 issue 4, p. 1593-1636.

Bec F. and **Mogliani M.** (2013), "Nowcasting French GDP in Real-Time from Survey Opinions: Information or Forecast Combinations ?", Banque de France, *Document de travail* n°436.

Bortoli C. and **Combes S.** (2015), "Contribution from Google Trends for forecasting the short-term economic outlook in France: limited avenues", *Conjoncture in France*, INSEE, March, p. 43-56.

Harvey D., **Leybourne S.** and **Newbold P.** (1997), "Testing the equality of prediction mean squared errors", *International Journal of forecasting*, vol. 13 issue 2, p. 281-291.

Kotsiantis S. B., **Zaharakis I. D.** and **Pintelas P. E.** (2006), "Machine learning: a review of classification and combining techniques", *Artificial Intelligence Review*, vol. 26 issue 3, p.159-190.

Schrauf R. W. and **Sanchez J.** (2004), "The preponderance of negative emotion words in the emotion lexicon: a cross-generational and cross-linguistic study", *Journal of multilingual and multicultural development*, vol. 25 issue 2-3, p. 266-284.

Thorsrud L. A. (2016), "Nowcasting using news topics, Big Data versus big bank", Norges Bank, *Working Paper*, n° 2016-20. ■

Appendix - Short-term prediction models to test the performances of the different advance composite indicators of market sector employment

Various models could potentially be used to forecast variations in payroll employment in the market sectors during the current quarter, based on the one hand on past variations in the variable of interest and in GDP, and on the other hand on short-term outlook indicators such as the French business climate indicators published by INSEE and on the general media sentiment indicator explored in this article. Various techniques can theoretically be used to manage the difference in frequency between the variable to be forecast (quarterly) and the explanatory variables (monthly): the approach adopted here is known as "blocking", commonly used by forecasters and involving a different calibration process for each month of the quarter, making use of all information available at the time of calculation. The calibrations for "month 1", "month 2" and "month 3" use all of the information available at the end of the first, second and third month of the quarter, respectively.

Four models were produced for the period as a whole: Model 1 is based exclusively on past variations in payroll employment and economic activity; Model 2 incorporates the general media sentiment indicator; Model 3, in addition to past variations in payroll employment and economic activity, incorporates the French business climate indicator; finally, Model 4 incorporates these two indicators simultaneously. To save time, only the significant variables were retained in each model. All the models were estimated for the period 1990 - 2016.

How to forecast employment figures by reading the newspaper

Calibration in Month 1

At the end of the first month of the quarter, figures for employment in the previous quarter are not yet available (the first publication comes in the middle of the second month) and thus cannot be used for calibration purposes. However, GDP for the previous quarter is available, since the new publication calendar for the quarterly accounts (which came into force in 2016) includes a "30 day" estimate.

Model 1 (past employment and activity only)

$$emploi_t = \underset{(-2.5)}{-0.08} + \underset{(6.5)}{0.45}.emploi_{t-2} + \underset{(6.6)}{0.39}.pib_{t-1} + u_t$$

adjusted $R^2 = 0.64$
DW = 1.50

Model 2 (+ media sentiment only)

$$emploi_t = \underset{(-3.4)}{-1.08} + \underset{(4.1)}{0.32}.emploi_{t-2} + \underset{(5.8)}{0.34}.pib_{t-1} + \underset{(3.2)}{0.01}.sentiment\_mediatique_{t,m1} + u_t$$

adjusted $R^2 = 0.67$
DW = 1.42

Model 3 (+ French climate only)

$$emploi_t = \underset{(-1.2)}{-0.03} + \underset{(8.5)}{0.59}.emploi_{t-2} + \underset{(3.8)}{0.23}.pib_{t-1} + \underset{(4.7)}{0.03}.(climat\_france_{t,m1} - climat\_france_{t-1,m1}) + u_t$$

adjusted $R^2 = 0.70$
DW = 1.87

Model 4 (+media sentiment and French climate)

$$emploi_t = \underset{(-2.3)}{-0.71} + \underset{(5.8)}{0.49}.emploi_{t-2} + \underset{(3.6)}{0.22}.pib_{t-1} + \underset{(4.0)}{0.02}.(climat\_france_{t,m1} - climat\_france_{t-1,m1}) + \underset{(2.2)}{0.01}.sentiment\_mediatique_{t,m1} + u_t$$

adjusted $R^2 = 0.71$
DW = 1.82

Calibration in Month 2

At the end of the second month in the quarter, the first figures for payroll employment can be used.

Model 1 (past employment and activity only)

$$emploi_t = \underset{(-2.0)}{-0.06} + \underset{(4.1)}{0.40}.emploi_{t-1} + \underset{(2.4)}{0.21}.emploi_{t-2} + \underset{(4.6)}{0.28}.pib_{t-1} + u_t$$

adjusted $R^2 = 0.69$
DW = 1.96

Model 2 (media sentiment only)

$$emploi_t = \underset{(-5.2)}{-1.72} + \underset{(5.0)}{0.38}.emploi_{t-1} + \underset{(2.9)}{0.17}.pib_{t-1} + \underset{(3.3)}{0.01}.sentiment\_mediatique_{t,m2} + \underset{(3.2)}{0.01}.sentiment\_mediatique_{t,m1} + u_t$$

adjusted $R^2 = 0.73$
DW = 1.98

Model 3 (French climate only)

$$emploi_t = \underset{(-2.4)}{-0.90} + \underset{(3.1)}{0.30}.emploi_{t-1} + \underset{(3.9)}{0.37}.emploi_{t-2} + \underset{(2.4)}{0.01}.climat\_france_{t,m2} + \underset{(5.2)}{0.06}.(climat\_france_{t,m2} - climat\_france_{t,m1})$$
$$+ \underset{(2.3)}{0.01}.(climat\_france_{t,m1} - climat\_france_{t-1,m1}) + u_t$$

adjusted $R^2 = 0.78$
DW = 2.15

Model 4 (media sentiment and French)

$$emploi_t = \underset{(-3.1)}{-1.35} + \underset{(3.2)}{0.30}.emploi_{t-1} + \underset{(2.8)}{0.29}.emploi_{t-2} + \underset{(2.2)}{0.01}.climat\_france_{t,m2} + \underset{(5.1)}{0.05}.(climat\_france_{t,m2} - climat\_france_{t,m1})$$
$$+ \underset{(1.9)}{0.01}.(climat\_france_{t,m1} - climat\_france_{t-1,m1}) + \underset{(1.9)}{0.01}.sentiment\_mediatique_{t,m1} + u_t$$

adjusted $R^2 = 0.79$
DW = 2.19

Calibration in Month 3

At the end of the third month in the quarter, the available quantitative indicators (GDP and employment) are the same as the previous month. Furthermore, media sentiment in the final month of the quarter does not seem to contain any information which is useful for forecasting. So models 1 and 2 are identical in the second and third months.

Model 3 (French climate only)

$$emploi_t = \underset{(-3.9)}{-1.30} + \underset{(3.0)}{0.26}.emploi_{t-1} + \underset{(3.8)}{0.30}.emploi_{t-2} + \underset{(3.9)}{0.01}.climat\_france_{t,m3} + \underset{(5.9)}{0.04}.(climat\_france_{t,m3} - climat\_france_{t,m1}) + u_t$$

adjusted $R^2 = 0.80$
DW = 2.19

Model 4 (media sentiment and French)

$$emploi_t = \underset{(-4.8)}{-1.77} + \underset{(3.1)}{0.26}.emploi_{t-1} + \underset{(2.8)}{0.23}.emploi_{t-2} + \underset{(3.5)}{0.01}.climat\_france_{t,m3} + \underset{(5.8)}{0.04}.(climat\_france_{t,m3} - climat\_france_{t,m1})$$
$$+ \underset{(2.6)}{0.01}.sentiment\_mediatique_{t,m1} + u_t$$

adjusted $R^2 = 0.81$
DW = 2.23

Regardless of the month in which the forecast is made, media sentiment is significant in Model 2. Moreover, the adjusted R^2 value for this model is higher than that for Model 1, indicating that it is a better match for the data. Model 3, which uses the French business climate indicator, is a better fit than Model 2. Finally, when the two indicators are used simultaneously in Model 4, the media sentiment indicator is still significant and the adjusted R^2 value shows a slight improvement on Model 3.

These first conclusions are formulated on the basis of estimates which make use of the available sample in its entirety. However, strong "sample" performances can sometimes be attributed to the phenomenon of "overfitting" (Bortoli & Combes, 2015). The media sentiment indicator could thus prove to be ineffective for forecasting purposes. In order to test this hypothesis, which seems unlikely given the parsimony of the models in question, their performance "out of sample" were tested. This involved running a forecasting simulation "in real time".[1]

For each model, the choice of explanatory variables is made once and for all. A first version of the models is then estimated for the period stretching from Q1 1990 to Q4 1999, then used to forecast the variation in employment in Q1 2000: the result thus obtained is then compared to the actual variation recorded in that quarter. A new version of the model is then calculated for the period up to Q1 2000, and used to forecast variation in employment in Q2. Gradually, forecasts are generated for every quarter from 2000 to 2016. This "out of sample" performance can then be evaluated by calculating its root mean square forecasting error. The results obtained for each model in the different months of the quarter are summarised in the table on page 40.

In Model 1, the forecasting errors observed at the end of the quarter are no smaller than those observed at the end of the first month: this means that market-sector employment forecasts are not improved by the information provided by the first lag in employment variation, given the information already contained in the second lag in this variable and the first lag in GDP growth. In the other models, however, the forecasts improve in quality throughout the quarter: the media sentiment and business climate indicators which become available as the quarter progresses serve to boost the predictive capacity of the models (with the exception of the media sentiment indicator during the third month of the quarter, which does not appear to offer any significant new information).

Regardless of the month in question, the root mean square forecasting error (outside of the sample) of Model 4 (media sentiment and business climate affaires) is lower than that observed in Model 3 (business climate only), which in turn is lower than RMSFE in Model 2 (media sentiment only), itself lower than in Model 1 (past employment and activity only): this would appear to corroborate the conclusions derived from the respective "in complete sample" performances of these two indicators. Nevertheless, the differences in RMSFE can be very slight: the test devised by Harvey, Leybourne & Newbold (1997) can be used to determine whether or not these differences are significant. In the first month of the quarter, Model 2 (media sentiment only) is not significantly better than Model 1 (past employment and activity only) in terms of predictive power. It does become significant at the 10% level from the second month of the quarter onwards. Meanwhile, Model 3 (business climate only) is always better than Model 2 at the 10% level. Finally, even though the forecasting errors observed in Model 4 (media sentiment and business climate) are systematically lower than those of Model 3, the difference in terms of RMSFE is never significant. ■

1. The simulation is generated using the historical series for payroll employment in its current published format, not the series for variations in employment as measured in the initial publications: strictly speaking this is "virtual real time" rather than actual "real time".