The power of text-based indicators in forecasting the Italian economic activity

Valentina Aprigliano, a Simone Emiliozzi, a Gabriele Guaitoli b, Andrea Luciani, a Juri Marcucci, a Libero Monteforte a

a Directorate General for Economics, Statistics, and Research, Bank of Italy
b Warwick University

2021 OECD and Banque de France Conference
June 8, 2021

1 The opinions expressed are those of the authors and do not reflect the views of the Bank of Italy or the Eurosystem.
Motivation

Forecasting faces new hard challenges

- Macroeconomic facts have been changing rapidly (Ng, and Wright, 2013)
- Legacies of the latest deep recessions

But...

- Big data availability: high-dimensional, high-frequency and timely
- *Unstructured* information

Our contribution

- **Sentiment Indicators (TESI)** and **Uncertainty indicators (TEPU)** for Italy based on newspaper articles for different sectors and topics
- Use TESI and TEPU to track the short-term evolution of the Italian economic activity
- Benefits in forecasting with both *monthly* and *weekly* models
Motivation: Survey on most relevant source of information

Bank of Italy’s Survey on Inflation and Growth Expectations

Sample size: 1199 respondents.
Growing literature exploring the *media-economy-opinion nexus*

- Shapiro et al. (2018), Gentzkow et al. (2019), Thorsrud (JBES, 2020), Kalamara et al. (BoE WP 2020), Ardia et al. (IJF, 2019), Algaba et al. (JES, 2020), Nguyen and La Cava (RBA WP 2020), Garboden (2019), Rogers and Xu (FRB WP 2020)

- Economic perceptions affect policy preferences but these perceptions are oftenly driven by factors other than the economy, including media [Soroka et al. (2015)]
- Newspapers catch the mood (and to a certain extent they amplify and propagate pessimism or optimism)

- Unstructured information: transforming text into numerical data (tokenization)
We downloaded approximately 2 million newspaper articles in the Italian language related to economic news from September 1996 to December 2019.
The News Corpus
Number of articles by year and source and share of articles for each source in each year
Data treatment - Sentiment & Uncertainty

- Pre-processing (removal of stop-words, non-meaningful punctuation, etc.)

Examples

*Production fell by 1.2% overall between January and October.* → *Production fall overall January October*

- Building a meaningful **dictionary** related to economic topics in Italian (unigrams + n-grams)
  - Polarity (+/-) & weight (#)
  - Valence Shifters tailored to newspapers’ jargon

- Constructing sentiment score for article \( j \) as

\[
\text{SENT}_{jt} = \frac{\sum_{i=1}^{\text{No words}} \text{polarity}_{ijt} \times \text{shifter}_{ijt}}{\text{No words}_{jt}}
\]

Examples

*Gross Domestic Product has fallen* → *SENT* \(_t\) = \(-1.0\)

*Istat’s projections, GDP grew in 2019. Expansion is set to strengthen in 2020* → *SENT* \(_t\) = \(0.25\)

- Constructing also Economic Policy Uncertainty (EPU) Indicators as the share of articles containing at least an “Uncertainty” word
Data treatment - Sentiment & Uncertainty

- Pre-processing (removal of stop-words, non-meaningful punctuation, etc.)

Examples

Production fell by 1.2% overall between January and October. → Production fall overall January October

- Building a meaningful **dictionary** related to economic topics in Italian (unigrams + n-grams)
  - Polarity (+/-) & weight (#)
  - Valence Shifters tailored to newspapers’ jargon

- Constructing sentiment score for article \( j \) as \( \Rightarrow SENT_{jt} = \sum_{i=1}^{\text{No words}} \frac{\text{polarity}_{ijt} \times \text{shifter}_{ijt}}{\text{No words}_{jt}} \)

Examples

*Gross Domestic Product has fallen* → *SENT*\(_t\) = \(-1.0\)

*Istat’s projections, GDP grew in 2019. Expansion is set to strengthen in 2020* → *SENT*\(_t\) = \(0.25\)

- Constructing also Economic Policy Uncertainty (EPU) Indicators as the share of articles containing at least an “Uncertainty” word
Data treatment - Sentiment & Uncertainty

- Pre-processing (removal of stop-words, non-meaningful punctuation, etc.)

Examples
Production fell by 1.2% overall between January and October. → Production fall overall January October

- Building a meaningful dictionary related to economic topics in Italian (unigrams + n-grams)
  - Polarity (+/-) & weight (#)
  - Valence Shifters tailored to newspapers’ jargon

- Constructing sentiment score for article $j$ as $\Rightarrow SENT_{jt} = \sum_{i=1}^{No\ words_{jt}} \text{polarity}_{ijt} \times \text{shifter}_{ijt} / \text{No \ words}_{jt}$

Examples
Gross Domestic Product has fallen $\Rightarrow SENT_t = -1.0$
Istat’s projections, GDP grew in 2019. Expansion is set to strengthen in 2020 $\Rightarrow SENT_t = 0.25$

- Constructing also Economic Policy Uncertainty (EPU) Indicators as the share of articles containing at least an “Uncertainty” word
Typically sentiment indices provided by statistical Institutes (NSI) or PMIs are based on information collected up to the mid of the reference month.

Sentiment based on newspapers’ articles accounts for facts and events occurring daily.
TESI and Economic Activity

Iraq disarmament crisis
Dot.com bubble crash
Euro circulation in IT
Eurozone (11 countries)
Twin Towers
IT elections
Iraq War
Italian Budget Law
tesI
Italy GDP q-o-q growth rates
-5
-4
-3
-2
-1
0
1
2
3
US Stock Mkt collapse
Lehman Brothers collapse
Greek default risk
EA spread crisis
Government crisis
Brexit
Trump’s election
IT Banking sector tensions
IT Elections
IT Budget law
tesI
Italy GDP q-o-q growth rates
-5
-4
-3
-2
-1
0
1
2
3
TESI and PMI

Aprigliano et al. Sentiment & Economic Activity Forecasting OECD-BdF - June 8, 2021 10 / 32
1 **Sentiment by topics (**# 15**), grouping > 300 article pre-labeled categories**
   - Fiscal Policy/Government
   - Monetary policy
   - Labor Markets
   - Economic conditions
   - Prices
   - Foreign Policy
   - ...

2 **Sentiment by sector (**# 21**)**
   - Manufacturing
   - Services
   - Retail
   - ...
Sentiment index - by topics (Fiscal/Government)
Sentiment index - by topics (Monetary Policy)
Economic Policy Uncertainty (EPU)
Empirical application #1 - Bayesian Model Averaging (BMA)

Short-term forecasting with **monthly data**

- **Target.** q-o-q growth of the Italian **GDP** and of its main **demand/supply components**
- **Model.** Bayesian Model Averaging (BMA) (Bencivelli, Marcellino, and Moretti. EE, 2017)
- **Data.**
  - Baseline model: soft indicators (NSI surveys and PMIs), industrial production index;
  - Augmented model: baseline + sentiment from newspapers’ articles (economic conditions, manufacturing, service sectors)
- Pseudo real-time forecasting exercise
Empirical application - BMA Results on point forecasts

Short-term forecasting - Relative RMSFE for nowcasting and forecasting

- T-model tends to lower the RMSFE during the most turbulent period (2011-2014), in particular for HHC.

Table: Relative RMSFE for nowcasting (n) and forecasting (f)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>f</td>
<td>n</td>
</tr>
<tr>
<td>GDP</td>
<td>0.93</td>
<td>0.91</td>
<td>1.17</td>
</tr>
<tr>
<td>VAS</td>
<td>0.97</td>
<td>1.21</td>
<td>1.08</td>
</tr>
<tr>
<td>GFI</td>
<td>1.03</td>
<td>0.94</td>
<td>1.13</td>
</tr>
<tr>
<td>HHC</td>
<td>0.83</td>
<td>0.79</td>
<td>1.46</td>
</tr>
</tbody>
</table>
Empirical application - BMA Results on density forecasts

Short-term forecasting - Average log score based on WLRT (Amisano & Giacomini, 2007)

- T-model definitely outperforms the benchmark overall, and in particular during the sovereign debt crisis. Text-based indicators squeeze the uncertainty around nowcasts.

Table: Average Log Score for nowcasting (n) and forecasting (f)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>f</td>
<td>n</td>
</tr>
<tr>
<td>GDP</td>
<td>9.1</td>
<td>8.6</td>
<td>-15.6</td>
</tr>
<tr>
<td>VAS</td>
<td>5.3</td>
<td>6.6</td>
<td>-7.1</td>
</tr>
<tr>
<td>GFI</td>
<td>3.6</td>
<td>23.7</td>
<td>6.9</td>
</tr>
<tr>
<td>HHC</td>
<td>14.6</td>
<td>12.2</td>
<td>-9.4</td>
</tr>
</tbody>
</table>
Empirical application - Nowcasts of GDP qoq

T-model

Baseline
Empirical application - Nowcasts of GDP qoq

T-model

Baseline
Empirical application - Nowcasts of GDP qoq

T-model

Baseline
Empirical application - Nowcasts of GFI qoq

T-model

Baseline
Empirical application - Nowcasts of GFI qoq
Empirical application - Nowcasts of VAS qoq

T-model

Baseline
Empirical application - Nowcasts of VAS qoq

T-model

Baseline
Empirical application - Nowcasts of VAS qoq

T-model

Baseline
Empirical application - Nowcasts of HHC qoq

T-model

Baseline
Empirical application - Nowcasts of HHC qoq

T-model

Baseline
Empirical application - Nowcasts of HHC qoq

T-model

Baseline
Empirical application - Results from BMA

Posterior Inclusion Probabilities for GDP qoq Nowcasts

- PIP measures relative importance of each regressor to explain the variance of the target variable. TESI is picked more frequently than PMI or NSI confidence surveys.
Empirical application - Results from BMA

Posterior Inclusion Probabilities for VAS qoq Nowcasts

- SI for Services is picked more frequently the corresponding PMI or Istat SI.
Empirical application - Results from BMA
Posterior Inclusion Probabilities for GFI qoq Nowcasts

- EPU for Services is picked more frequently the std. dev. of expected earnings.
Empirical application - Results from BMA

Regression Coefficients and 25th-75th percentiles for GDP qoq Nowcasts
Empirical application - Results from BMA
Regression Coefficients and 25th-75th percentiles for GFI qoq Nowcasts
Empirical application #2 - A Weekly Economic Indicator

Following Stock and Watson (2002) and Lewis, Mertens and Stock (2020), we build a weekly indicator of economic activity:

- Explore the role of information timeliness

We find that:

- TESI and TEPU help nowcast the GDP (RMSFE reduced by 15 – 17% from baseline)
- Gains seem due to:
  - Better tracking than other weekly variables
  - More timely tracking than monthly indicators
- CSSED analysis shows stable gains over most of the out-of-sample period
The model

We extract the first Principal Component from two different sets of variables:

1. **Group 1 (baseline)**
   - Electric Consumption, Expected Earnings std (weekly)
   - PMI and NSI sentiment indicators (monthly)

2. **Group 2 (factiva)**
   - Baseline
   - TESI and TEPU indicators (weekly)

We use it to nowcast GDP growth yoy
Forecasting error

- Compute the nowcasting errors by using the ex-post available data on GDP as

\[ E_t^i = \Delta Y_{(yoy),t} - \Delta \hat{Y}_{(yoy),t} \]

- \( \approx 13 \) nowcasts per quarter

- We find large gains on weekly nowcasts when adding Sentiment and EPU indicators

**Table:** Relative RMSFE

<table>
<thead>
<tr>
<th></th>
<th>Expanding</th>
<th>Rolling (335 weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sample</td>
<td>0.85</td>
<td>0.83</td>
</tr>
<tr>
<td>Negative GDP</td>
<td>0.88</td>
<td>0.96</td>
</tr>
<tr>
<td>Positive GDP</td>
<td>0.82</td>
<td>0.75</td>
</tr>
</tbody>
</table>
Weekly indicator: last vintage

Expanding sample - ex-post indicator

% change GDP, YoY


Baseline
Factiva
GDP
Wrap-up

- We developed an Italian economic dictionary with polarity and shifters.
- We used Factiva newspapers data to estimate Sentiment and EPU indices at daily and weekly frequencies.
- We evaluated their properties in two short-term forecasting exercises:
  1. **Monthly**: point-forecast gains in recessions; large density forecast gains overall.
  2. **Weekly**: large point-forecast gains across all the sample.
- Further developments:
  - Forecast-maximizing dictionary weighting scheme.
  - Regional forecasting.
THANK YOU