

IFC-Bank of Italy Workshop on "Data science in central banking: enhancing the access to and sharing of data"

17-19 October 2023

Central bank communication on economic activity¹

Sercan Eraslan and Eniko Gabor-Toth,
Deutsche Bundesbank

¹ This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Italy, the BIS, the IFC or the other central banks and institutions represented at the event.

Central bank communication on economic activity

3rd IFC Workshop on Data Science in Central Banking

Eniko Gabor-Toth¹ Sercan Eraslan¹

¹Deutsche Bundesbank

October 19, 2023

This presentation represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem.

Central bank communication is important for shaping expectations on...

- monetary policy
- inflation
- ... and last but not least on economic activity

This paper:

- Measure the information content of the central bank narrative on underlying economic conditions and risks
- Generate sentiment indices for topics/sections in the monthly reports of the Bundesbank (German and English)
- Forecasting exercise: do the MR sentiment and uncertainty indices improve the forecasting accuracy of key macro indicators in real-time?

Review of the relevant literature

So far most studies using text-based data in business cycle analysis and finance make use of newspaper/press articles...

- Tetlock (2007), Baker et al. (2016), Ulbricht et al. (2017), Ardia et al. (2019), Larsen & Thorsrud (2019), Ellingsen et al. (2020), Kalamara et al. (2020), Thorsrud (2020), Aguilar et al. (2021), Larsen et al. (2021) etc.

Studies on central bank communication have mostly focused on monetary policy/inflation...

- Gürkaynak et al. (2005), Ehrmann & Fratzscher (2007), Haldane & McMahon (2018), Enders et al. (2019), Coibion et al. (2020), Ehrmann & Talmi (2020), Swanson (2021) etc.

... growing literature on incorporation of text-based data obtained from central banks and research institutes into macroeconomic forecasting...

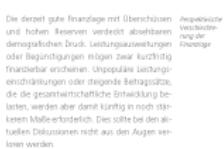
- Sharpe et al. (2020): Fed's Green Book for the US; Clements & Reade (2020), Jones et al. (2020): BoE's IR for the UK; Müller (2020): Bundesbank and research institutes forecast reports for DE

Textual database

Deutsche Bundesbank
Monatsbericht
März 2019
17



Beginnt werden die Forderungen nach Beitragsmildigungen mit dem Argument, dass es doch zu einer „Doppelbelastung“ kommt: Beiträge würden sowohl in der Alters- als auch in der Rentenphase fallen. Hierzu kommt es jedoch lediglich in engen Ausnahmefällen.² In der Regel unterliegen Zulassungen zu Vermögenspräferenzungen oder Aufzulassungen von Pensionsrückstellungen keinen GKV-Bertragen.³ Insoweit erscheint es folgerichtig, Beitragsmildigungen und Beitragsmildigungen zu verhindern: ggf. erhöhte Arbeitsmobilisierung zu verbieten und keine Beitragsmildigung vorzunehmen.



Soziale Pflegeversicherung

Das Defizit der sozialen Pflegeversicherung stieg im Jahr 2018 um 1 Mrd. € auf 375 Mrd. €.⁴ Die Einnahmen wuchsen um 4%, dagegen die Beiträge der Beschäftigten um gut 41% und die de-

beitigungen der Arbeitgeber um 10%.

abschließend

5. Mit dem 1.1.2019 in Kraft getretene Änderungen im Pflegeversicherungsgesetz wird seit 2004 der GKV-Bertrag erhoben.

6. Vgl. „Bundeslig-Checkliste 19/1401“ vom 21. September 2018.

7. So etwa in Form von Empfangsverhandlungen. Kurzfristige Abweichungen von den tatsächlichen Beitragsmildigungen sind zulässig, wenn es sich um eine geistige Verunsichertheit (§ 1 Absatz 1 Nr. 1 Buchstabe a) handelt.

8. Vgl. „Deutscher Rückzug, Wissenscheckliste: Derzeit, Verwendung des sog. „Doppelbelastung“ von Beitragsmildigungen und Beitragsmildigungen in der Pflegeversicherung“ in der gesetzlichen Kranken- und Pflegeversicherung, 2018, S. 15.

9. Viele der Heftungspraktiken gegenwärtig abgesetztes Pflegeversicherungsgesetz sind nicht mehr im Pflegeversicherungsgesetz beschrieben, der zur Verhörgesetzeszeit 2010 parallel übernommen wurde. Zur weiteren Bedeutung der Konkurrenz siehe Kapitel 3 und 4.

Deutsche
Pflege-
versicherung von
derzeit
mit aktuellem
Kostensteuer-
Gesetz verändert

schätzliche Satz seit dem Jahresbeginn über nur auf 1,0%. Sehr mit den neuen Beitragsmildigungen aus dem Tarifsteuer- und Versorgungsgesetz könnten somit wieder Überzahle entstehen.
Diskutiert wird derzeit, die Beiträge auf Betriebsrenten und andere Vermögensbezüge zu erhöhen. Dies wäre eine weitere Verstärkung der finanziellen Belastung. Würden die gesetzlichen Beiträge halbiert, ließe sich 2024 der Fall, dass die Einnahmen von etwa 3 Mrd. € pro Jahr entfallen.⁵ Dies entspricht Mindestentnahmen in Höhe von etwa 0,2 Beitragsatzpunktpunkten.

- Data source: monthly and quarterly reports of the Deutsche Bundesbank: German and English in PDFs
- Corpus: 396 reports from January 1990 through December 2022
- Challenges: scanned documents, changes in PDF versions over time, unicode errors, text structure (two columns), reconstruction of words separated at line endings, side notes, footnotes

From files to database

Key Question: How can we obtain a structured, informative, quantitative representation of these texts?

- ① Remove graphs and charts, footnotes, side notes, convert scanned pdfs into editable pdfs
- ② Extract texts using: R (tabulizer, tesseract, pdftools)
- ③ Text cleaning and pre-processing, reconstruct words separated at line endings in R, experiment with ChatGPT 3.5 for pre-processing
- ④ Assign topics

Text extraction: example

Turning raw text files into structured databases is often a challenge:

- June 2000
- July 2000

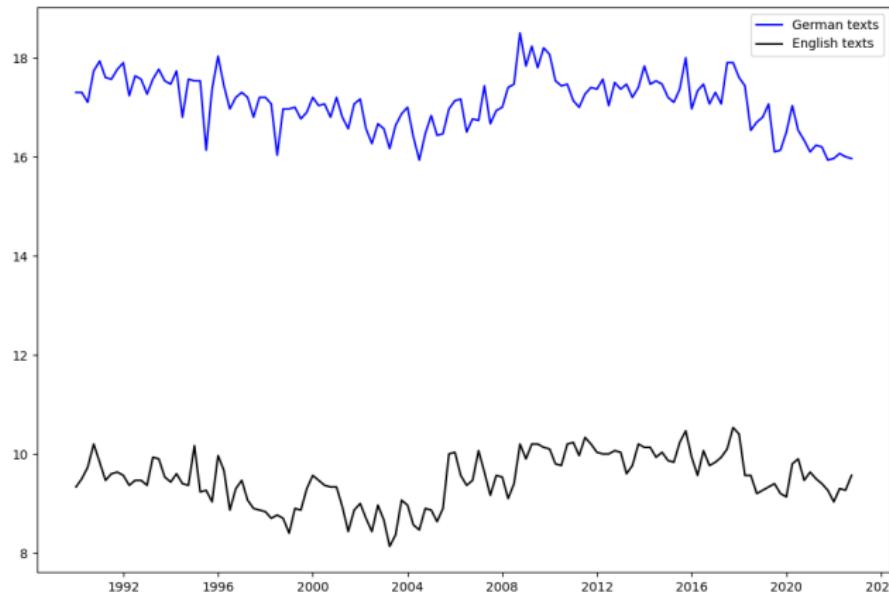
Konjunkturlage \ Gesamtwirtschaft \ Ersteren Berechnungen des statistischen Bundesamtes folge ist die gesamtwirtschaftliche Produktion auch im ersten Vierteljahr 2000 relativ kräftig gestiegen. Saison- und kalenderbereinigt erhöhte sie sich gegenüber den Herbstmonaten 1999 um 3,014 %. Auf ein Jahr hochgerechnet würde damit die deutsche Wirtschaft seit Jahresmitte 1999 mit fast 3 % den vergleichbaren Vorjahrsstand übertrafen. Das reale Bruttoinlandsprodukt nach Aus-nschaltung von Kalenderregelmäßigkeiten ist nun 2,3 %.

Main themes over time: TF-IDF

- Generate a heatmap for important keywords in German texts.

Readability: Flesch-Kincaid Grade Level

- 1st grade - 5th grade Very easy to read.
- 5th grade - 11th grade Average level.
- 11th grade - 18th grade The text is for skilled readers. For example, an academic paper.



Textual analysis

- The sentiment/tone is simply calculated as the relation between positive and negative words:

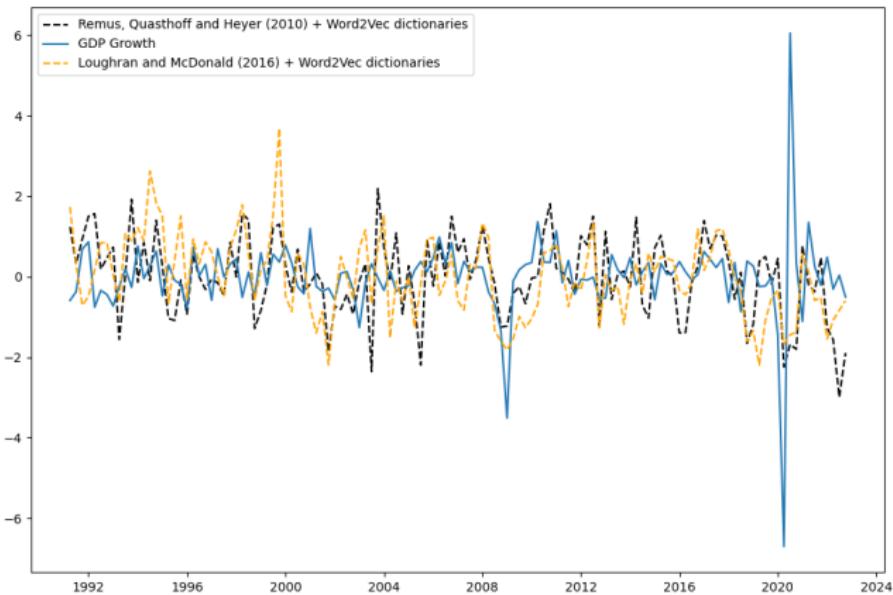
$$s_{n,t} \equiv \frac{W_{n,t}^+ - W_{n,t}^-}{W_{n,t}}$$

where $W_{n,t}^+$, $W_{n,t}^-$ and $W_{n,t}$ are the numbers of positive, negative and all words in topic n in month t .

Result: sentiment, uncertainty and readability indices...

- for topics as they are written in monthly reports, such as industry, labour market, financial/securities markets, prices etc.
- for the overall economic conditions and the whole monthly report

SI economic conditions vs. GDP growth



Mixed-frequency dynamic factor model

We use a mixed-frequency dynamic factor model (MF-DFM) for the forecasting exercise.

$$y_t = Hs_t + u_t, \quad u_t \sim N(0, \Sigma_{u_t})$$
$$s_t = Fs_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma_{\varepsilon_t})$$

... and follow Mariano & Murasawa (2003) to deal with the quarterly indicator. Assuming $\ln Y_t^q = \frac{1}{3}(\ln Y_t^m + \ln Y_{t-1}^m + \ln Y_{t-2}^m)$ and taking three period differences:

$$y_t^q = \frac{1}{3}y_t^m + \frac{2}{3}y_{t-1}^m + y_{t-2}^m + \frac{2}{3}y_{t-3}^m + \frac{1}{3}y_{t-4}^m$$

with $y_t^q := \Delta_3 \ln Y_t^q$ (observed every third period) and $y_t^m := \Delta_3 \ln Y_t^m$ (unobserved).

Estimation and forecast setup

Dataset

- Quarterly GDP and 13 monthly macroeconomic and financial indicators (gdp, ip, new orders, construction, retail sales, exports, unemployment, cpi, hwwi, dax30, corp. spread, ifo c, ifo pp, ifo ee)
- monthly sentiment indices (balance of payments and foreign trade, construction, economic conditions, financial/securities markets, industry, labour market, monetary policy, prices, public finance)
- from 1991M1 to 2021M6
- real-time vintages from 2007M7 - 2021M6

Model specifications: MF-DFMs without sentiment indices vs. MF-DFMs with sentiment indices

Estimation: State space form by means of Kalman filter
($n_f = 3, p = 3, q = 2$), expanding window

Forecast

- up to 2 quarters ahead
- real-time forecast evaluation period 2007Q3–2021Q2
- Evaluation via MAD (calculated with first releases)

Forecast evaluation results

$h =$	2Q-ahead			1Q-ahead			Nowcast			Back
	8	7	6	5	4	3	2	1	0	-1
GDP	0.95	0.96	0.95	0.98	0.96	1.01	0.95	0.88	1.16	1.20
IP	0.91	0.95	0.99	1.01	0.90	0.96	0.89	0.85	1.00	1.05
NewOr	0.92	0.98	0.96	1.01	0.98	1.03	0.95	0.88	0.93	0.98
Const	1.00	1.04	1.08	1.08	1.01	1.02	0.93	0.99	1.07	0.99
Sales	1.00	1.00	1.02	1.00	1.02	0.99	0.87	1.04	0.96	1.04
Expo	0.94	0.97	0.98	1.05	0.99	1.03	0.93	0.93	1.02	1.05
UnEmp	0.96	0.95	1.02	1.00	0.93	0.90	1.00	1.00	1.00	-
CPI	1.00	1.00	1.00	0.98	1.01	0.99	0.95	1.02	0.97	-
HWWI	0.97	0.97	1.00	0.95	0.93	0.97	0.96	0.93	0.93	-
DAX30	0.97	1.04	1.05	1.02	1.06	0.99	1.01	1.07	0.98	-
CorSp	1.00	1.03	1.04	0.98	1.02	0.93	1.00	0.97	0.99	-

Notes: This table plots relative MADs, such that values less (greater) than 1 hint to a better (worse) forecast performance for the MF-DFM with sentiment indices.

Summary and relevance of the results

Key take-aways:

- Sentiment indices appear to lead to (slight) increases in forecast accuracy of some key macroeconomic indicators
- Monthly reports seem to have an additional information content on economic activity on top of those contained in macroeconomic and survey-based indicators
- This helps the private sector better forecast economic developments and thus can enhance the transmission of monetary policy