

# Central bank communication on economic activity

## 3rd IFC Workshop on Data Science in Central Banking

**Eniko Gabor-Toth<sup>1</sup>** Sercan Eraslan <sup>1</sup>

<sup>1</sup>Deutsche 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.

# Eraslan & Gabor-Toth (2023)

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

## Finanzen der sozialen Pflegeversicherung\*



## Finanzierungsquellen, Wertpapierkäufe



\* Die soziale Pflegeversicherung ist ein Zweig der Versicherungsarten des Sozialrechts. Sie ist eine der drei Versicherungszweige des Sozialrechts. Die Beiträge werden von den Versicherten und den Arbeitgebern getragen. Die Ausgaben werden von den Versicherten und den Arbeitgebern getragen.

schrittweise Satz sank zum Jahresbeginn aber nur auf 1,0%. Selbst mit den neuen Belastungen aus dem Terminservice- und Versorgungsgesetz könnten sonst wieder Überschüsse entstehen.

Diskutiert wird derzeit, die Beiträge auf Beitragsebenen und andere Versorgungsleistungen zu setzen. Dies würde die Krankenversicherung finanziell stark belasten. Würden die genannten Beiträge halbiert (wie vor 2004 der Fall)<sup>1)</sup> würden Entnahmen von etwa 3 Mrd € pro Jahr entfallen.<sup>2)</sup> Dies entspricht Mindestanteilen in Höhe von etwa 0,2 Beitragsatzprozentspunkten.

Definierte Beiträge entziehen von Beitragsnehmern ein erhebliches Einkommenspotenzial und werden

Begründet werden die Forderungen nach Beitragsermäßigungen mit dem Argument, dass es derzeit zu einer „Doppelbelastung“ komme. Beiträge würden sowohl in der Anlage<sup>3)</sup> als auch in der Rentenschufe (Bsp. Wenzel kommt es jedoch lediglich in einigen Ausnahmefällen<sup>4)</sup> in der Regel unterliegen. Zuführungen zu Versorgungseinrichtungen oder Aufstockungen von Pensionsrückstellungen kalten-GKV-Beiträgen<sup>5)</sup> würden ebenfalls ein fälliger, selbständiger und andere Versorgungsbezüge wie hinausgehendere Arbeitsentgelte zu verbieten und keine Beitragsermäßigung vorzuziehen.

Dies derzeit gute Finanzlage mit Überschüssen und hohen Reserven verdeckt absehbaren demographischen Druck. Leistungsanforderungen oder Begründungen mögen zwar kurzfristig finanziell einsehen. Ungeplante Leistungsanforderungen oder steigende Beitragsätze, die die gesellschaftliche Entwicklung belasten, werden aber damit künftig in noch stärkeren Maße erforderlich. Dies sollte bei den aktuellen Diskussionen nicht aus den Augen verloren werden.

## Soziale Pflegeversicherung

Das Defizit der sozialen Pflegeversicherung stieg im Jahr 2018 um 1 Mrd € auf 316 Mrd €.<sup>6)</sup> Die Entnahmen wuchsen um 4%, darunter die Beiträge der Beschäftigten um gut 41% und die

<sup>1)</sup> Mit dem GKV-Mindestrentenentgelt wird seit 2004 der volle GKV-Beitrag erhoben.  
<sup>2)</sup> Vgl. Bundesgesetzblatt, 704041 vom 21. September 2010, S. 4.

<sup>3)</sup> Ein etwa ein-fünftel Teil der Beiträge an die Versorgungseinrichtungen, die Beiträge an die Versorgungseinrichtungen oder Pensionen, die über 4% der Beitragsbemessungsgrenze in der gesetzlichen Rentenversicherung (BRG) hinausgehen (§ 1 Abs. 1 Nr. 1 Sozialversicherungs-Gesetz (Sonderausgaben)).  
<sup>4)</sup> Vgl. Deutscher Bundestag, Wissenschaftliche Dienste, Verfügung der 12. „Doppelbelastung“ von Beitragsnehmern mit Doppelbeiträgen und anderen Versorgungsleistungen in der gesetzlichen Kranken- und Pflegeversicherung – 20. Arbeitshilfskommission und Fortschreibung, 2016, S. 11.  
<sup>5)</sup> In der Regel Beitragsentgelte gegenüber niedrigeren Lohnentgelten, wie wird die Erhöhung über den Pflegeversicherungssatz beschlossen, der zum Vermögensnachlass 2015 jährlich Überschüsse entzieht. Zum Jahresende 2014 betragen die kumulierten Rücklagen 3,1 Mrd €.

Angleichung der höchsten Beiträge mit dem „Zusatzbeitrag“

Angleichung der höchsten Beiträge mit dem „Zusatzbeitrag“

Gehaltene Ostler hat Aufgaben auszuweisen

- 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

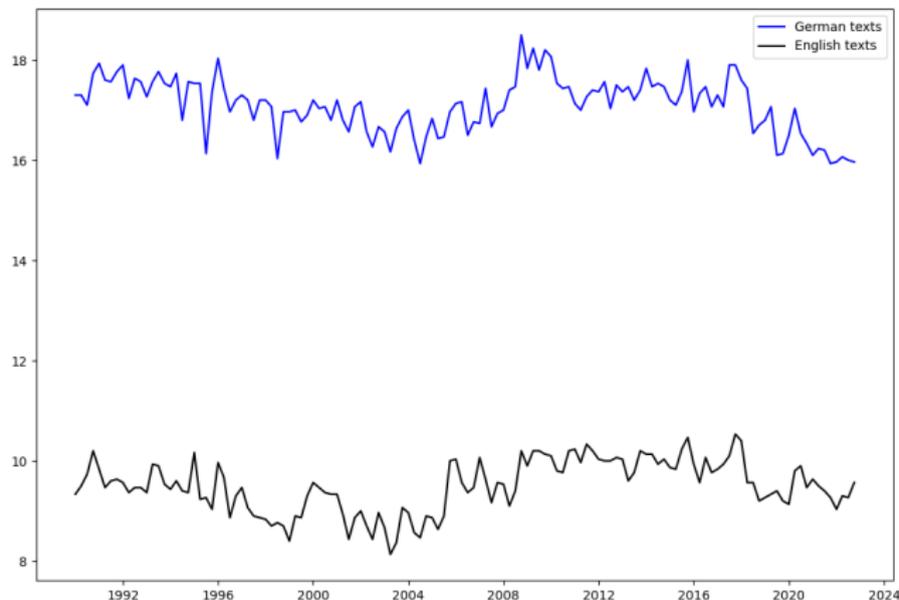


# Main themes over time: TF-IDF

- 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:

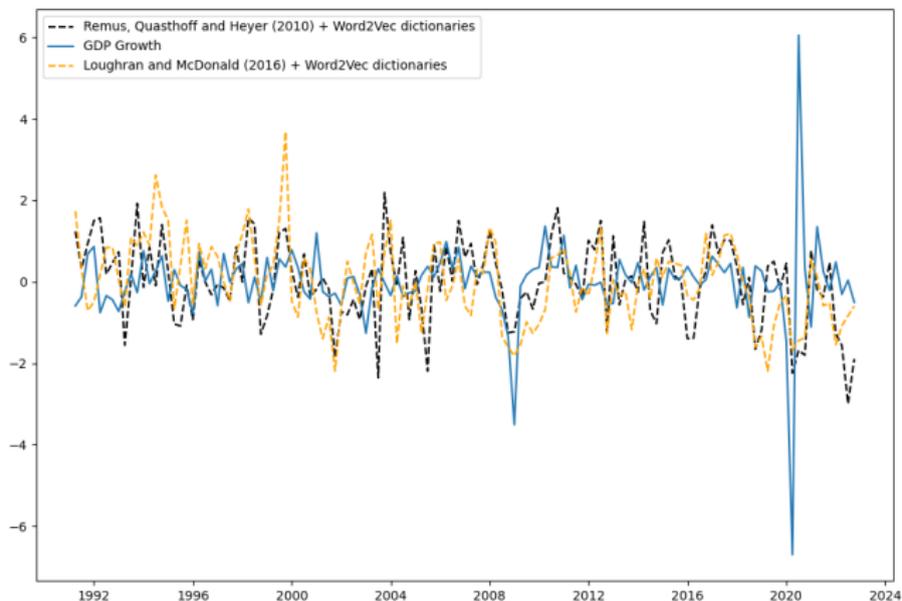
$$s_{n,t} \equiv \frac{W_{n,t}^+ - W_{n,t}^-}{W_{n,t}}$$

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.

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

... 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
<b>GDP</b>	0.95	0.96	0.95	0.98	0.96	1.01	0.95	0.88	1.16	1.20
<b>IP</b>	0.91	0.95	0.99	1.01	0.90	0.96	0.89	0.85	1.00	1.05
<b>NewOr</b>	0.92	0.98	0.96	1.01	0.98	1.03	0.95	0.88	0.93	0.98
<b>Const</b>	1.00	1.04	1.08	1.08	1.01	1.02	0.93	0.99	1.07	0.99
<b>Sales</b>	1.00	1.00	1.02	1.00	1.02	0.99	0.87	1.04	0.96	1.04
<b>Expo</b>	0.94	0.97	0.98	1.05	0.99	1.03	0.93	0.93	1.02	1.05
<b>UnEmp</b>	0.96	0.95	1.02	1.00	0.93	0.90	1.00	1.00	1.00	-
<b>CPI</b>	1.00	1.00	1.00	0.98	1.01	0.99	0.95	1.02	0.97	-
<b>HWWI</b>	0.97	0.97	1.00	0.95	0.93	0.97	0.96	0.93	0.93	-
<b>DAX30</b>	0.97	1.04	1.05	1.02	1.06	0.99	1.01	1.07	0.98	-
<b>CorSp</b>	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