Nowcasting GDP in Real-Time with a Tone-Adjusted, Time-Varying Layered Topic Model

Jasper de Winter Dorinth van Dijk

De Nederlandsche Bank
Economic Policy & Research
Econometrics & Modelling Department

Seminar, Centraal Planbureau, 20 April, 2023

Introduction

Idea

- Can we use newspaper text to track the business cycle and nowcast GDP growth?
- Extract **topics** from newspaper articles using unsupervised machine-learning model
- Extract **sentiment** using lexicon-based method;
- Combine **topics** and **sentiment** in **tone-adjusted time-varying news topics**

Motivation

- Understand what drives business cycle fluctuations;
- Important to have a point of departure: nowcast current pace of growth economy.

Main contributions

- Analyze unique new source of Dutch newspaper texts;
- Extend base topicmodel by including **time-variation** and **layering** in topics;
- Analyze forecasting quality of tone-adjusted time-varying news topics in nowcasting model.

DeNederlandscheBank
EUROSYSTEEM

Outline presentation

1. **Data**
2. **Sentiment**
3. **Model**
4. **Nowcasting**
5. **Concluding remarks**

DeNederlandscheBank
EUROSYSTEEM

1. Data

Data: Financieele Dagblad

- Complete **full-text archive** of Dutch "Financial Times";
- Strong focus on financial-economic news and socio-economic (politics);
- Analyzed period | 36 years | January 1^{st} 1985–January 1^{st} 2021;
- **Clean** and construct **vocabulary** of 2, 153 words over total period 1985 – 2020.


fd. hetfinancieeledagblad

Crisis teistert Wall Street

Koersen kelderen, bank Lehman in surseance, verzekeraar AIG wankelt

De Fed trekt een streep in het zand

Smit wijst bod Boskalis af

Prinsjesdag zonder looneis FNV en CNV

September 16 2008


fd. hetfinancieeledagblad

Grote onrust onder beleggers, beurzen beleven rampweek

AEX-index -11%

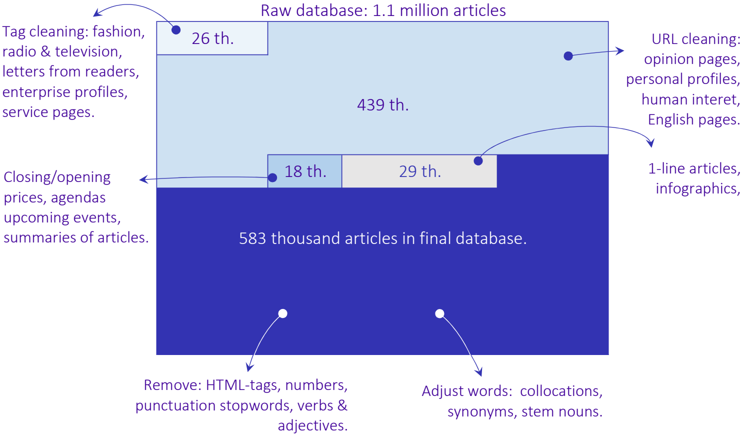
AMX-index -11%

Air France-KLM -23%

Slechtere vooruitzichten voor mondiale economie. Eerste besmetting in Nederland

February 17 2020

Data: Cleanup articles & texts



2. Sentiment

DeNederlandscheBank
EUROSYSTEEM

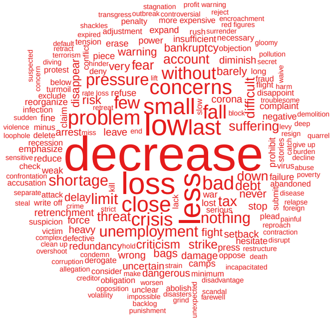
Sentiment: Extraction from newspaper articles

Sentiment Lexicon

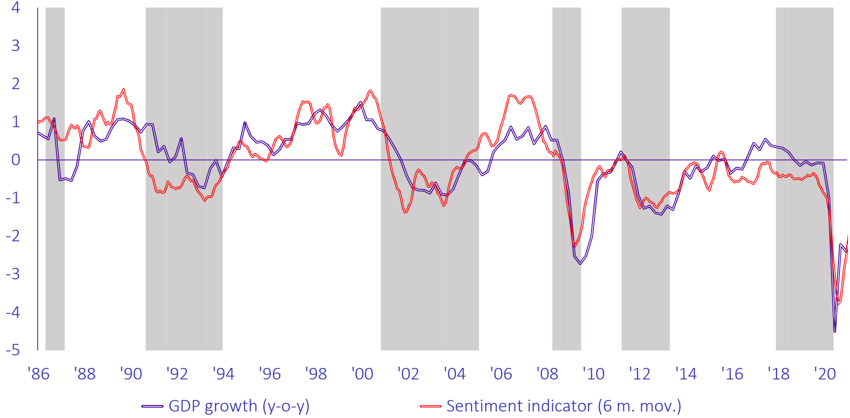
- Customized Dutch version of **Loughran and McDonald (2011)**
- Check for **double negations** i.e.: deficit decreased, unemployment decreased
- Total list: 1,532 words | Positive: 468 | Negative: 1,063
- Sentiment score **per article** (see e.g. Tetlock, 2007 and Shapiro et al., 2020):
  (# positive words - # negative words)/(# words in article)

Positive words

Negative words

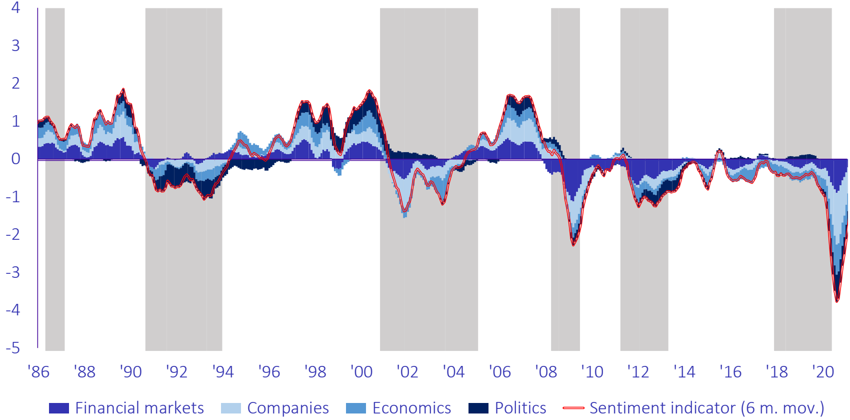


Sentiment: newspaper sentiment and GDP growth



OECD recessions in grey

. . . Next part of presentation: sentiment→ per topic sentiment



OECD recessions in grey

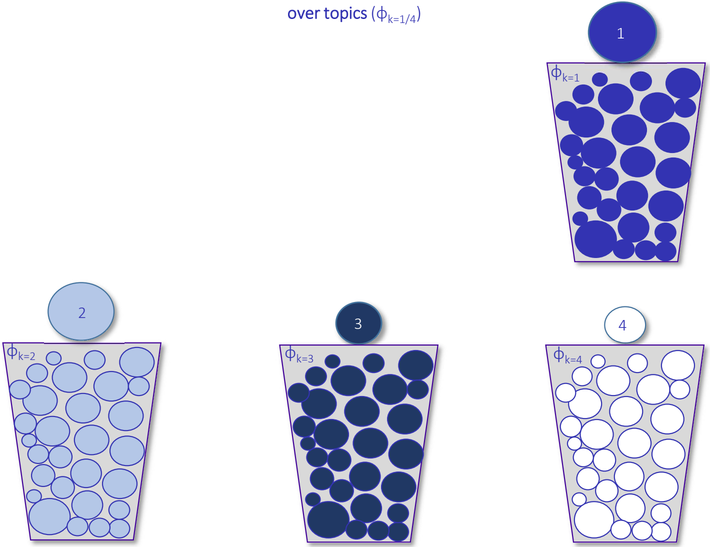
DeNederlandscheBank EUROSYSTEEM

3. Model

DeNederlandscheBank
EUROSYSTEEM

Model: intuition Latent Dirichlet Allocation

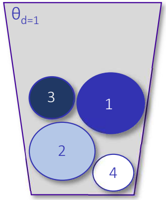
Step 1: Draw prob. distribution for **words over topics** ($\phi_{k=1/4}$)



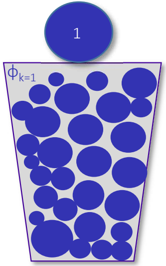
DeNederlandscheBank
EUROSYSTEEM

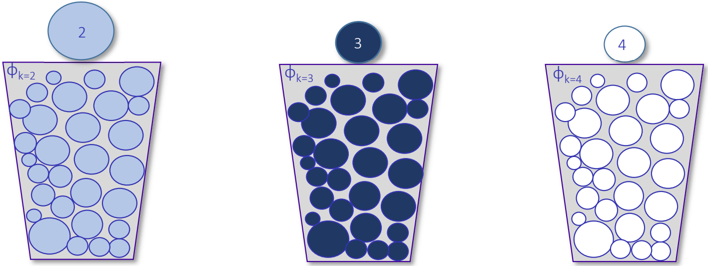
Model: intuition Latent Dirichlet Allocation

Step 2: Draw prob. distribution for **topics over document** for document d_1 ($\theta_{d=1}$)



Step 1: Draw prob. distribution for **words over topics** ($\phi_{k=1/4}$)



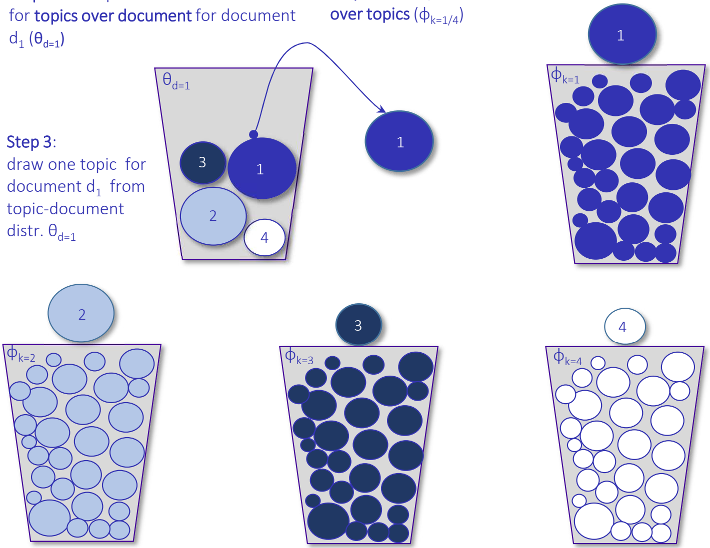


DeNederlandscheBank
EUROSYSTEEM

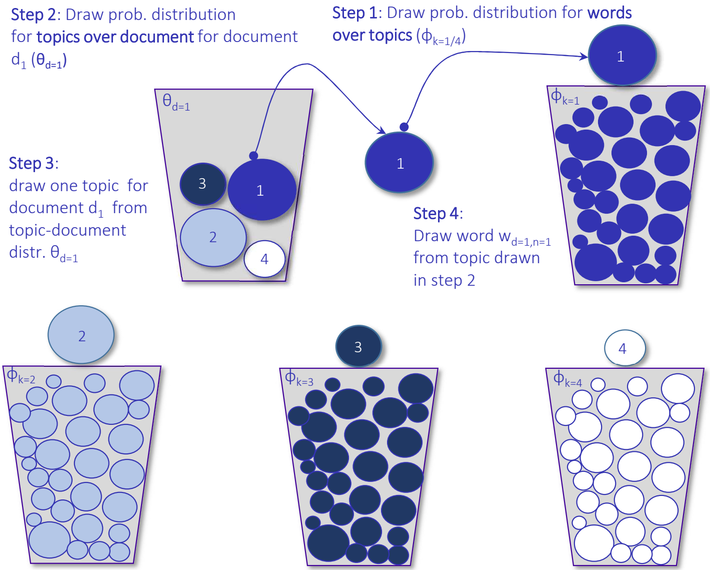
Step 2: Draw prob. distribution for **topics over document** for document d_1 ($\theta_{d=1}$)

Step 1: Draw prob. distribution for **words over topics** ($\phi_{k=1/4}$)

Step 3: draw one topic for document d_1 from topic-document distr. $\theta_{d=1}$



Model: intuition Latent Dirichlet Allocation



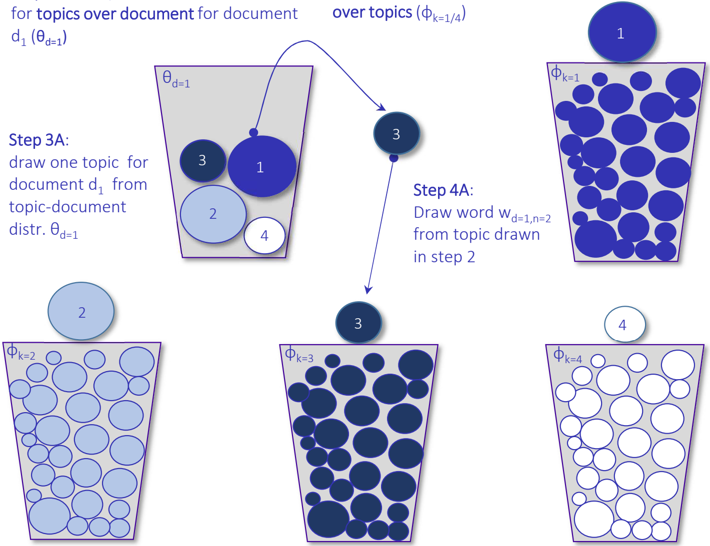
DeNederlandscheBank
EUROSYSTEEM

Step 2: Draw prob. distribution for **topics over document** for document d_1 ($\theta_{d=1}$)

Step 1: Draw prob. distribution for **words over topics** ($\phi_{k=1/4}$)

Step 3A: draw one topic for document d_1 from topic-document distr. $\theta_{d=1}$

Step 4A: Draw word $w_{d=1,n=2}$ from topic drawn in step 2

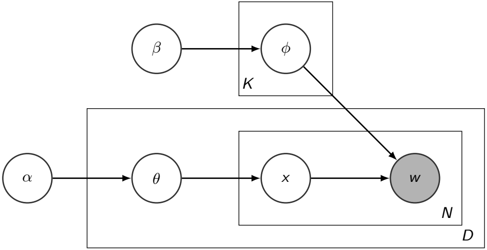


Repeat steps 3A and 4A until document d_1 is filled with N_1 words.

DeNederlandscheBank EUROSYSTEEM

Model: formal description topic model

- Determine the topic-word distribution, i.e.: sample from $\phi_k \sim$ **Dirichlet**(β), where $k \in \{1, \ldots, K\}$.
- Determine the document-topic distribution, i.e.: sample from $\theta_d \sim$ **Dirichlet**(α), where $d \in \{1, \ldots, D\}$.
- For each word $w_{d,n}$, where $n \in \{1, \ldots, N_d\}$, and $d \in \{1, \ldots, D\}$
  - From the document-topic distribution, θ_d, sample one topic, i.e.: choose a topic $x_{d,n}$ for each word $w_{d,n} \sim$ **Multinomial**(θ_d);
  - Sample a word from the topic-word distribution,ϕ_k, conditional on the topic, i.e.: choose a word $w_{d,n} \sim$ **Multinomial**(ϕ_k),



Latent Dirichlet Allocation using plate notation

Model: Inference of model parameters

Bayesian inference of model parameters

- Derive joint distribution of the document-topic distributions θ_d the topic-word distribution ϕ_k and the allocation of words $w_{d,n}$ to topics k in all documents;
- Several methods: Expectation-Maximization (e.g. Hoffman, 2001), variational inference (e.g. Blei et al., 2003) and Gibbs sampling (Griffiths and Steyvers, 2004);
- Bayesian inference via Gibbs sampling feasible but quite costly computationally, i.e.: $Pr(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{x} | \boldsymbol{w}, \alpha, \beta)$;
- Collapsed Gibbs sampling reduces computations: $Pr(\boldsymbol{x} | \boldsymbol{w}, \alpha, \beta)$. Steyvers and Griffiths (2004) show that collapsed Gibbs sampler simplifies to parsimonious **counting rule**. Key to this result is the fact that Dirichlet priors are **conjugate priors** to the multinomial distribution of the likelihood function;
- Based on topic-assignments easy to derive predictive document-topic distribution ($\hat{\theta}_d$) and topic-word distribution ($\hat{\phi}_k$);

DeNederlandscheBank EUROSYSTEEM

Essence collapsed Gibbs sampling algorithm

- **Random initialization** of algorithm: topic assignments of **all words in the corpus** from random draws of Multinomial document-topic distribution (θ_d) drawn from Dir(β) and Multinomial topic-word distribution (ϕ_k) drawn from Dir(α)
- **First pass Gibbs sampler**: Re-assigning each word to a topic on the counting rule. The rule calculates how **likely** the word is for a topic and the **dominance** of a topic in a document based on the assignment of all other words to topics, i.e.:

$$Pr(x_{i,n} = K | \mathbf{x}_{-i}, w_i, d_i, .) \propto \text{“likeliness”} \times \text{“dominance”}$$

▸ Algorithm

DeNederlandscheBank
EUROSYSTEEM

Model: Extension time-variation

Time variation enables trends in topics

- 15 year time-slices of corpus of articles for $t = (1, \ldots, T)$ | time slice 1 spans the period Jan. $1^s t$ 1985 until Dec. $1^s t$ 2000 | time slice T spans the period Jan. $1^s t$ 2006 until Dec. $1^s t$ 2021;
- **Intuition**: use predictive document-topic distribution ($\hat{\theta}_d$) and topic-word distribution ($\hat{\phi}_k$) in time slice t to initialize collapsed Gibbs sampler in time slice $t+1$, repeat for time slice $[t+2, \ldots, T]$;
- **Implementation**: time-invariant vocabulary based on total sample, so words can slowly gain importance in topics by appearing in the time slice ... remember $Pr(x_i = k|\mathbf{x}_{-i}, w_i, d_i, .) \propto$ **"likeliness"** $\times$ **"dominance"**;
- Standard LDA models will underestimate trends because the model is estimated over the whole time period $1985 - 2020$. E.g. word "Corona" has very low "likeliness" when measured over total sample but very high "likeliness" when measured in last time-slice;
- We do **not** impose word-dynamics as in dynamic topic models in e.g. Blei and Lafferty (2006) and Bitterman and Rieger (2022);

DeNederlandscheBank EUROSYSTEEM

Model: Extension layering

Layering enables deeper understanding topics

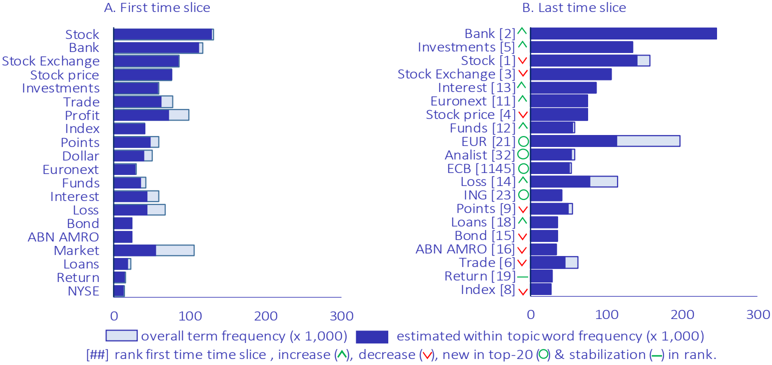
- **Intuition**: estimate new topic model in second layer for words assigned to each topic in first layer. E.g. 4 topics in first layer. Estimate four new topic models for words assigned to each of these 4 topics;
- **Implementation**: We construct three layers with four topics for ex-positional purposes;
- Use random initialization in first time slice and posterior distribution from previous time slice in subsequent periods;
- We impose hierarchy. Different from hierarchical topic models where hierarchy is based on correlations (Griffiths et al., 2003).

Model: outcome Time Varying Layered (TVL) LDA model, layer 1 - layer 3

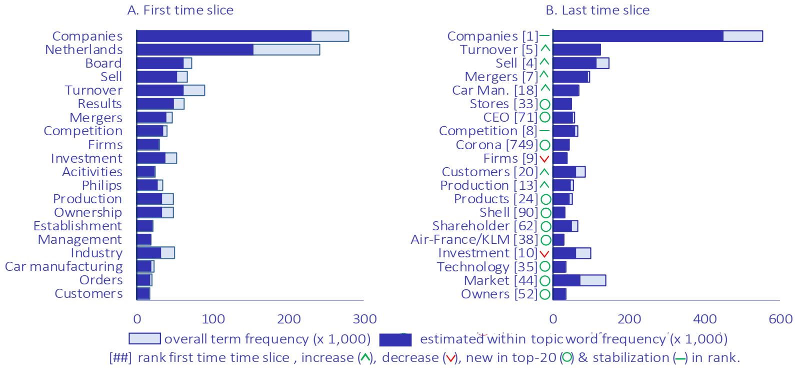
Layer 1	Financial Markets	Firms	Economics	Politics
Layer 2	Markets	Infrastructure	Elections	Parliament
Layer3	1. Raw materials	17. Chemical & Pharma	33. Elections	49. Politics
	2. Exchanges	18. Indices	34. Easten Europe	50. Budgetary policy
	3. International	19. Mobility	35. Africa & Asia	51. Cabinets
	4. Monetary Policy	20. Company Results	36. United States	52. Ministries
Layer 2	Financials	Multinationals	Indicators	National
Layer3	5. Corporate Finance	21. Telecom	37. International	53. Justice
	6. Financials	22. Customers	38. Europe	54. Pensions & Healthcare
	7. Banks	23. Big Tech	39. Trading Partners	55. Supervision
	8. Insurance	24. Media	40. Fiscal Policy	56. Education & Research
Layer 2	News	Construction & Energy	Raw Materials	Lower Government
Layer3	9. Emissions	25. Construction	41. Asia	57. Housing
	10. Takeovers	26. Logistics	42. Oil & gas	58. Public-private
	11. Trade	27. Energy	43. Conflicts	59. Agriculture & Fishery
	12. Insurers	28. Industry	44. Emerging Markets	60. Transport
Layer 2	Fin. Indices	Demography	European Union	Social Partners
Layer3	13. Stock Markets	29. Retail	45. Germany	61. Wage Negotiations
	14. Euronext	30. Bankruptcies	46. European Union	62. Labor Market
	15. Analists	31. Listed	47. Italy & Spain	63. Entrepeneurs
	16. Results	32. International	48. France	64. Social Security & Pensions

DeNederlandscheBank
EUROSYSTEEM

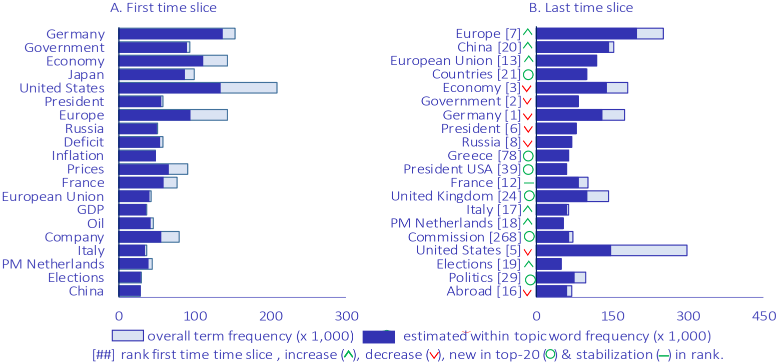
Model: outcome layer 1: “Financial Markets”



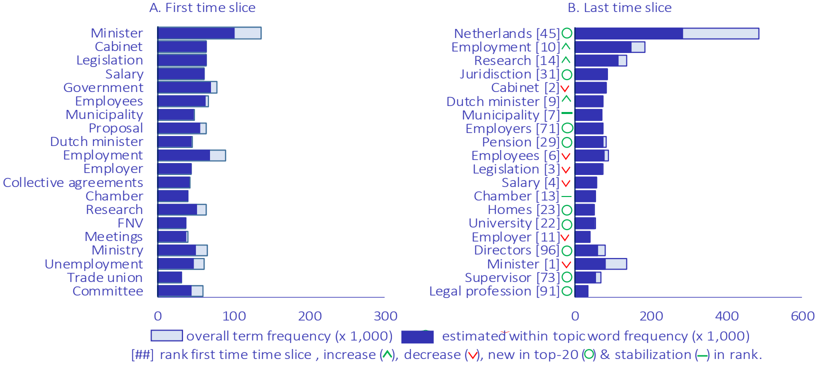
Model: outcome layer 1: "Firms"



Model: outcome layer 1: "Economics"

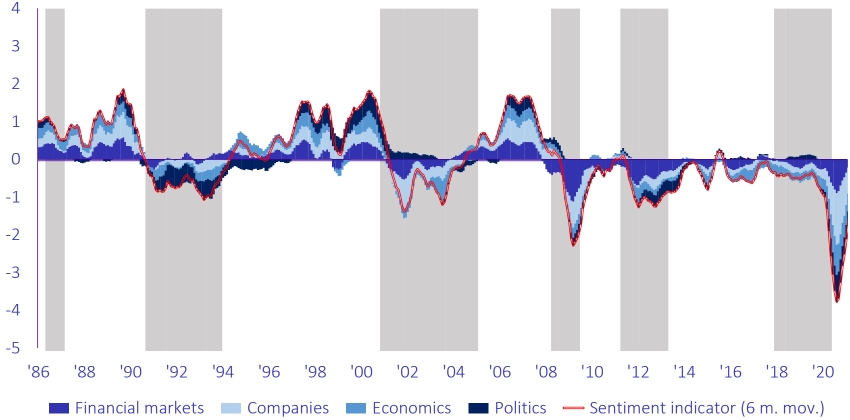


Model: outcome layer 1: “Politics”



DeNederlandscheBank EUROSYSTEEM

Model: newspaper sentiment per topic



OECD recessions in grey

DeNederlandscheBank
EUROSYSTEEM

4. Nowcasting

Nocasting horse-race: benchmark model

- Pseudo real-time exercise: using short-term forecasts of dynamic factor model **with** and without textual data to see if newspaper sentiment adds to forecast accuracy;
- **Intuition** dynamic factor model: summarize information in large set of indicators ("curse of dimensionality") in few factors with dynamics in factors (VAR). Use factors to forecast GDP;
- Estimate model-coefficients dynamic factor model with **quasi maximum-likelihood** taking into account **differences in frequencies** (GDP: quarterly, indicators: monthly) and **publication delays** (see e.g. Hindrayanto, de Winter and Koopman, 2016);
- **Estimation**: start 1996M1 | Evaluation: 2003Q1–2020Q3 | 70 monthly indicators | (max) 64 newspaper sentiment indicators.

Nocasting horse-race: design

- GDP released with 45-day lag | first release Dutch GDP growth first quarter 2023 will be released on May 15^{th};
- We analyze forecast quality for 8 different forecasting horizons;

Example: Naming GDP forecasts first quarter 2023

Nr.	Forecast type	Month	Forecast made on the 1^{st} of
1	Forecast	M1	October
2		M2	November
3		M3	December
4	Nowcast	M1	January
5		M2	February
6		M3	March
7	Backcast	M1	April
8		M2	May

DeNederlandscheBank EUROSYSTEEM

Nowcasting GDP growth with(out) newspaper sentiment

Total sample

- **Economic** significance (relative RMSFE ≤ 0.95), **Statistical** significance (DM-test);
- TVL-LDA increases the forecasting accuracy of the dynamic factor model;
- Particularly when it pertains to current or previous quarter;
- Parsimonious configuration topic model more accurate (4×4).

		DFM	DFM **with** tone-adjusted topics		
topicmodel type		-	TVL-LDA	TVL-LDA	TV-LDA
# topics per layer (X)		-	4 X 4 X 4	4 X 4	4
		(absolute RMSFE)	(relative RMSFE against DFM)		
Backcast	M2	*0.78*	0.98	0.86 *	0.94
	M1	*0.95*	1.11	0.95 *	0.99
Nowcast	M3	*1.31*	1.03	0.93 *	0.98 *
	M2	*1.47*	1.05	1.00	0.99 *
	M1	*1.71*	1.04	1.01	1.00 **
Forecast	M3	*1.73*	1.01	1.00 *	1.00 *
	M2	*1.65*	1.00	1.00	1.00
	M1	*1.58*	1.01	1.00	1.00

RMSFE: Root Mean Squared Forecast Error, 2003Q3-2020Q3.

RMSFE ≥ 1.05 | 0.95 < RMSFE < 1.05 | RMSFE ≤ 0.95 | * DM-test significance

DeNederlandscheBank EUROSYSTEEM

Nowcasting GDP growth with(out) newspaper sentiment

Excluding crises quarters (2009Q1, 2020Q2 and 2020Q3)

- Some advantage but either **not** statistically or economically significant;
- Value added LDA mainly attributable to crisis, when its critical;

		DFM	DFM **with** tone-adjusted topics		
topicmodel type		-	TVL-LDA	TVL-LDA	TV-LDA
# topics per layer (X)		-	4 X 4 X 4	4 X 4	4
		(absolute RMSFE)	(relative RMSFE against DFM)		
Backcast	M2	*0.53*	0.93	1.01 *	0.98
	M1	*0.54*	0.94	1.00 *	0.99
Nowcast	M3	*0.55*	0.96	1.00 *	0.99 *
	M2	*0.56*	0.97	1.00	0.99 *
	M1	*0.57*	1.00	0.99	0.99 **
Forecast	M3	*0.57*	1.01	0.97 *	1.00 *
	M2	*0.59*	1.04	0.99	1.00
	M1	*0.59*	1.08	0.97	1.00

RMSFE: Root Mean Squared Forecast Error, 2003Q3-2020Q3, excluding 2009Q1, 2020Q2 and 2020Q3.

RMSFE ≥ 1.05 (red) 0.95 < RMSFE < 1.05 (light green) RMSFE ≤ 0.95 (green) * DM-test significance

DeNederlandscheBank EUROSYSTEEM

Nowcasting GDP growth with(out) newspaper sentiment

Excluding crises (2009Q1, 2020Q2 and 2020Q3)

- LDA with 16 topics in different configurations;
- No clear advantage besides nowcasts of TVL-LDA, but not statistically significant;

DFM **with** tone-adjusted topics

topicmodel type		LDA	TVL-LDA	TV-LDA	L-LDA
# topics per layer (X)		16	4 X 4	16	4 X 4
		(absolute RMSFE)	(relative RMSFE against LDA)		
Backcast	M2	*0.66*	1.02	1.00	1.03
	M1	*0.91*	0.98	0.99	1.02
Nowcast	M3	*1.31*	0.93	0.97	0.96
	M2	*1.55*	0.94	0.98	0.97
	M1	*1.80*	0.96	0.98	0.97
Forecast	M3	*1.72*	1.00	1.00	1.00
	M2	*1.65*	1.00	1.00	1.00
	M1	*1.59*	1.00	1.00	1.00

RMSFE: Root Mean Squared Forecast Error, 2003Q3-2020Q3.

RMSFE ≥ 1.05 0.95 < RMSFE < 1.05 RMSFE ≤ 0.95 * DM-test significance

DeNederlandscheBank EUROSYSTEEM

Concluding remarks

Five main takeaways from our research

1. Newspaper sentiment is a conincident indicator of the business cycle;
2. Tone-adjusted time varying layered topics add “story-telling” layer to newspaper sentiment;
3. Tone-adjusted time varying layered topics embody information not captured in other monthly indicators, especially during periods of high volatility;
4. Evidence (at best) mixed on increase in nowcasting accuracy for layering and time-variation in topic tone-adjusted topic model;
5. For nowcasting purposes use a parsimonious topic model;

DeNederlandscheBank
EUROSYSTEEM

Thank you for your attention!

J.M.de.Winter@dnb.nl

Algorithm LDA collapsed Gibss sampler

Draws from the posterior distribution $Pr(x|w)$ are obtained by sampling from

$$Pr(x_i = K|\mathbf{x}_{-i}, w_i, d_i, .) \propto \underbrace{\frac{n^{(j)}_{-i,K} + \beta}{n^{(.)}_{-i,K} + V\beta}}_{\text{"likeliness"}} \times \underbrace{\frac{n^{(d_i)}_{-i,K} + \alpha}{n^{(d_i)}_{-i,.} + k\alpha}}_{\text{"dominance"}}$$

V	number of words in the vocabulary;
(j)	indicates w_i is equal to the jth term in the vocabulary, $j = [1 \dots, V]$;
$n^{(j)}_{-i,K}$	freq. of the jth term assignment to topic K without the ith word;
d_i	document in the corpus to which word w_i belongs;
x_{-i}	vector of current topic membership of all words without the ith word w_i;
$(.)$	summation over index;
α	prior of the Dirichlet distribution of the topic-word distribution (ϕ_k);
β	prior of the Dirichlet distribution of the document-topic distribution (θ_d).

One draw for all words in the corpus equals one iteration of the Gibbs sampler. Based on topic-assignments you can calculate estimated predictive document-topic distribution ($\hat{\theta}_d$) and topic-word distribution ($\hat{\phi}_k$).

Go Back