Using Transformers and Reinforcement Learning as Narrative Filters in Macroeconomics

Vegard H. Larsen   Leif Anders Thorsrud
CAMP, BI Norwegian Business School

Motivation

"It would be nice to point to recognizable events, of the type that is reported by newspapers, as the source of economic fluctuations, rather than to residuals from some equations." Cochrane (1994, p. 298)

A fundamental problem in macroeconomics is identifying the underlying drivers of economic fluctuations and measuring their broader dynamic implications. Speaking to:

- Business cycle measurement - Burns and Mitchell (1946)
- Shocks and propagation
  - SVARs (Sims, 1980)
  - Larger systems - merging SVARs with Burns and Mitchell (1946) - (Stock and Watson, 1989; Bernanke et al., 2005; Bańbura et al., 2010)
- Adding narrative content (Friedman and Schwartz, 1963; Romer and Romer, 1989; Ramey, 2011; Stock and Watson, 2012; Antolín-Díaz and Rubio-Ramírez, 2018)

We take up the same theme - **measurement, structural interpretation, and narrative attribution** - from a different perspective, leveraging Transformer-based neural architectures (Vaswani et al., 2017)

- Allows for join contextualized representations of different data modalities via encoder structures and cross-attention
- Easily captures potentially complex and long-range dependencies and mixed frequency information
- Generative in nature when combined with decoder structures

The idea/contribution

Underlying assumption: Text and time series co-move due to common cause

Thus, we combine: Economic structure/theory + Transformers + BERT = Narrater, which

- **Jointly** processes and integrates **text** (news) + **time series** (macro) via **cross-attention**
- Decodes the multimodal signal into **business-cycle fluctuations** and **textual class + sentiment**
- Allows for **high-frequency** extractive news **summaries** via **Reinforcement Learning** objective
- Provides **structural narrative attribution** - where the structural part is instilled during training based on user preferences (economic structure/theory)

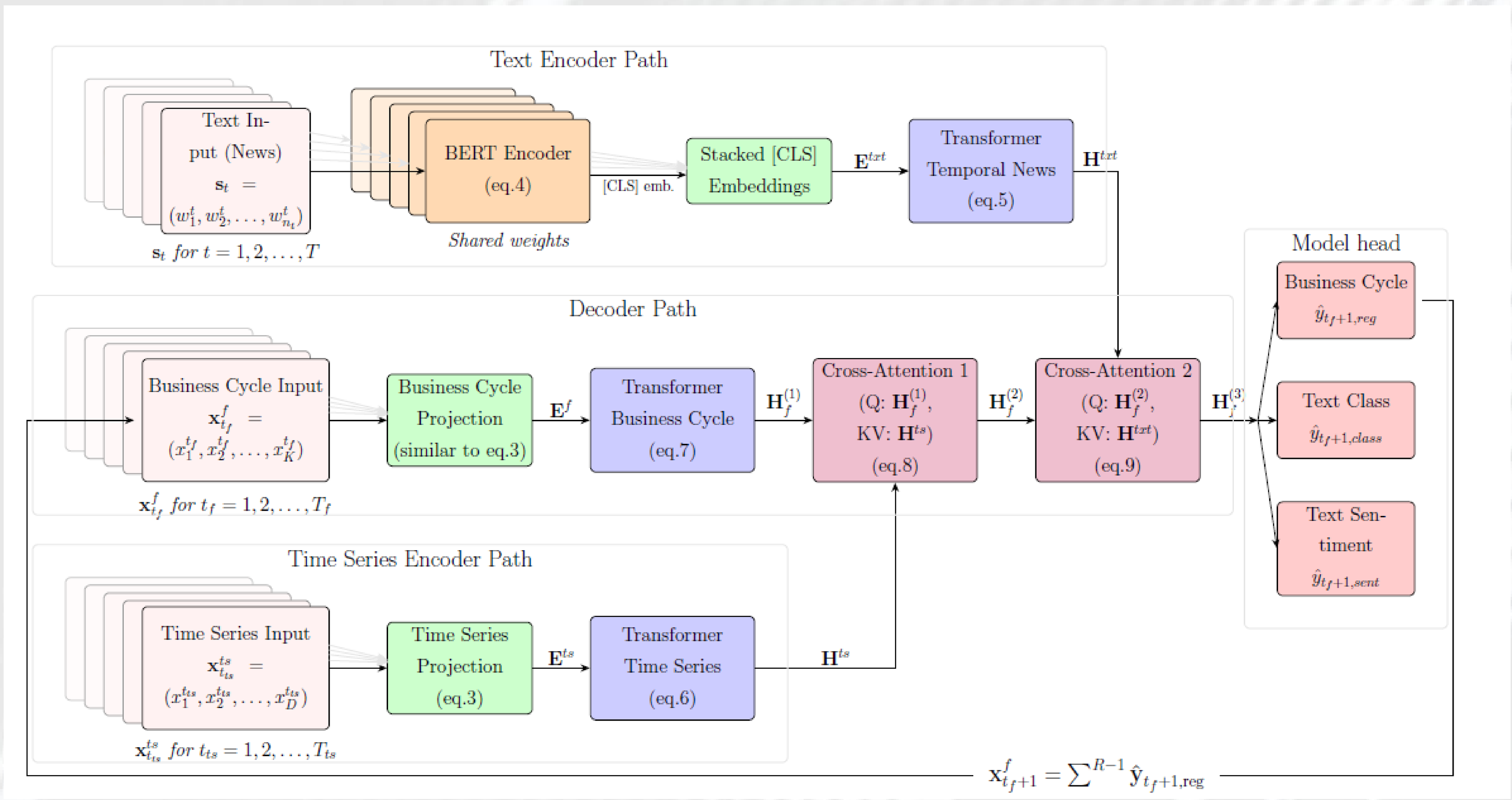
In short: **A multimodal structural narrative filter**

Is this useful?

Yes, we think so. A central bank, for example, needs to estimate the business cycle and identify its underlying drivers for appropriate policy actions and justification. However, macro data is lagging and does not provide a narrative per se

Thus, a role for a structural narrative filter than can systematically process high-frequency information in line with a given world view. We focus on macro, but approach potentially useful in other domains where text and time series co-move due to common cause

The Narrater



Applications and inductive bias

- The Narrater is a filter; need to instill a specific inductive bias into the training data - related to the recent literature on prior-fitted networks (Müller et al., 2021; Hollmann et al., 2022; Nagler, 2023)
- Benchmark application; business cycle measurement
  - Simulate training data from Dynamic Factor Model (business cycle & inflation factors)
  - Shocks: demand (prices & output co-move) vs supply (move opposite) + noise
  - Linking text to shocks (historical shock decomposition)
    - Embed news; label via similarity to seed narratives (demand/supply) and polarity
    - Probabilistic draw of class based on shock magnitudes over time
    - Noise = off-topic or weakly matching content
- Underlying data:
  - **Text:** DN (Dagens Næringsliv) news corpus (late 1980s–2023: >500k articles)
  - **Time series (quarterly):** GDP, Investment, Unemployment, CPI, HPI, Term spread, Credit, Oil
- Alternative application; global oil market dynamics as identified in Baumeister and Hamilton (2019)
- **IMPORTANT: Preserve data > 2010 for out-of-sample testing. I.e., training data is generated using information only prior to 2010**

Output generation and estimation protocol

Outputs in benchmark application

- **Cycle regression:** structural components (demand, supply, noise) and sum
  - $\hat{\mathbf{y}}_{t_f+1,\text{reg}} = (\hat{y}^{(demand)}_{t_f+1,\text{reg}}, \hat{y}^{(supply)}_{t_f+1,\text{reg}}, \hat{y}^{(noise)}_{t_f+1,\text{reg}})$, where $\mathbf{x}^{f}_{t_f+1}$ is constructed as a sum over fundamentals (demand and supply)
- **Class probabilities:** demand / supply / noise
- **Sentiment:** continuous score $[-1, 1]$

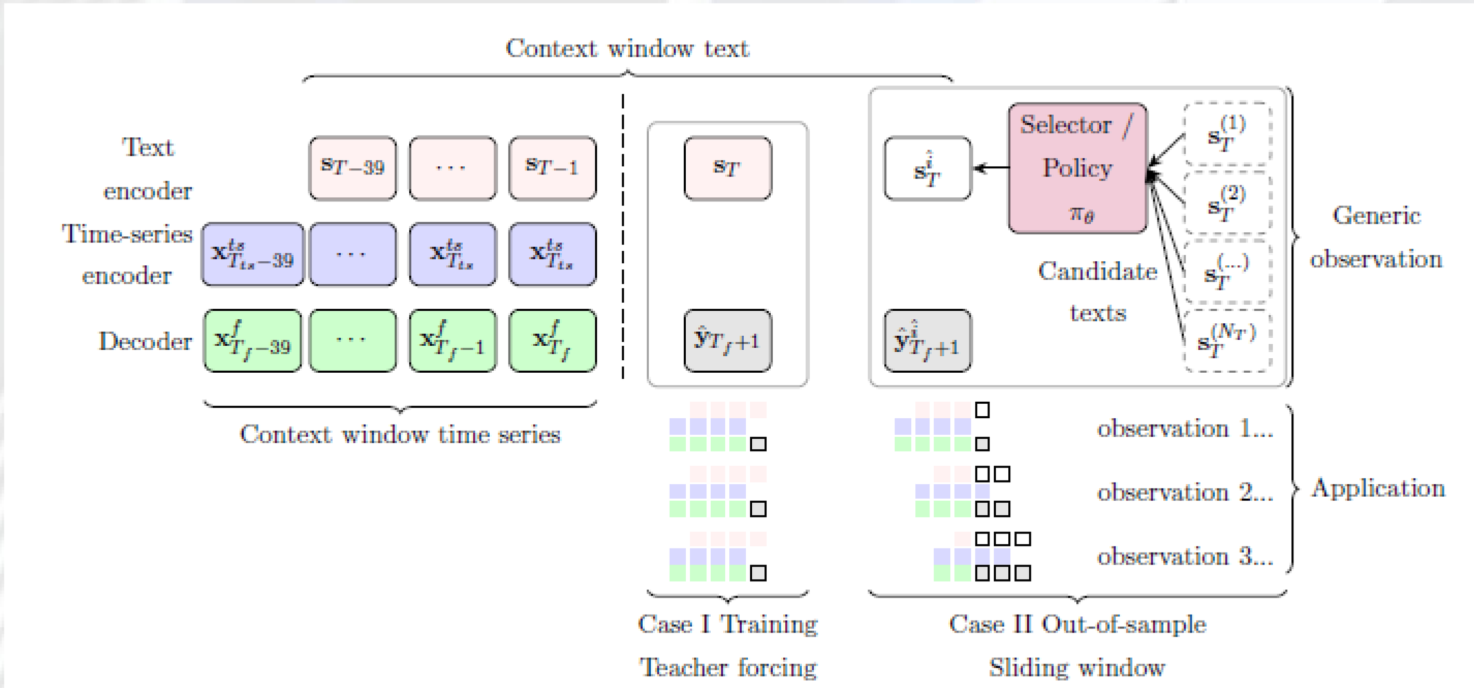
Estimate end-to-end

- Adam, batch 16; lr 5e-4 (BERT at 1/5 rate); early stopping
- Loss: weighted sum (reg 0.5, class 0.25, sent 0.25)
- Dropout + L2; 90/10 train/val; fine-tune BERT is important

Multi-task training objective

$$\mathcal{L}_{\text{total}} = \omega_r \|\hat{y}^{\text{reg}} - y^{\text{reg}}\|_2^2 + \omega_s \|\hat{y}^{\text{sent}} - y^{\text{sent}}\|_2^2 + \omega_c \Big(- \sum_{c=1}^{C} y^{(c)}_{\text{class}} \log \hat{y}^{(c)}_{\text{class}} \Big)$$

Information structure



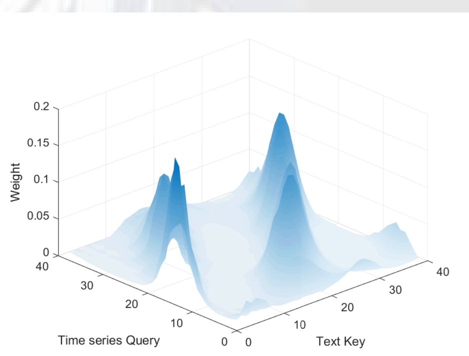
Validation and ablation experiments

Table. Panel A reports classification and sentiment performance of the *Narrater* relative to a Naive Bayes (NB) classifier and the Lasso sentiment regression, as well as relative performance measures on all tasks compared to three different LLMs. A score below 1 indicates the *Narrater* performs better. Panel B reports the performance of the benchmark model relative to several alternative model specification; without a text encoder (*NoTxt*); without a time series encoder (*NoTS*); for two different embedding dimensions (*TxtComp* and *TsExp*); estimating the model without simultaneously fine-tuning the language encoder (*NoFt*); removing the sentiment and classification model heads (*NoMTL*).

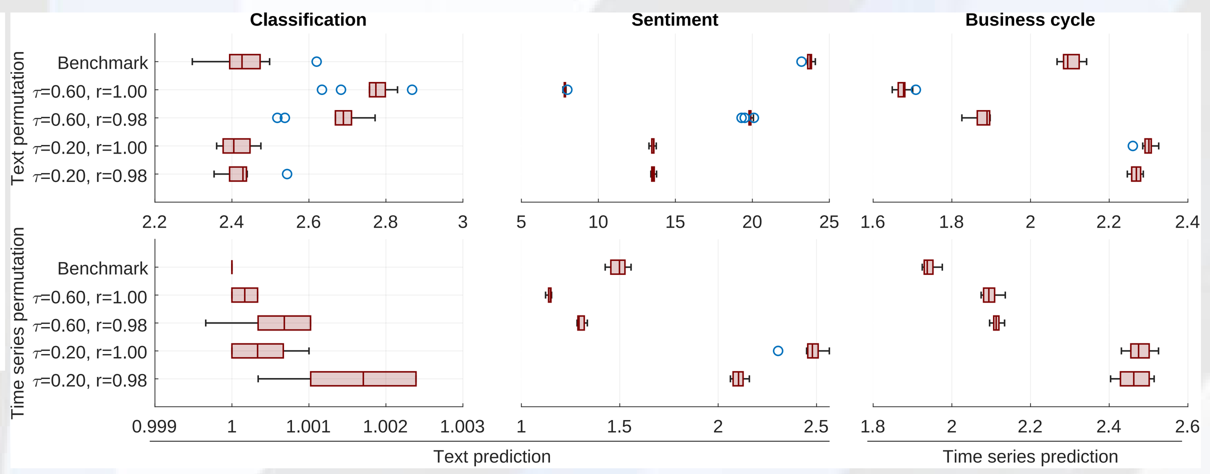
Panel A	NB / Lasso		LLM (zero-shot)			Panel B	Ablation		Hyperparameter		Fine tuning	No MTL
			Gemini	o3	R1		NoTxt	NoTS	TxtComp	TsExp	NoFt	NoMTL
Class	0.95		0.42	0.34	0.32		0.36	1.00	1.00	1.00	0.94	
Sent.		0.29	0.05	0.04	0.07		0.08	0.70	0.63	0.63	0.16	
BusC.			0.55	0.68	0.00		0.83	0.96	1.00	0.97	0.92	0.94

Attention example and permutation tests

(a) Text to time series (Bidirectional)



(b) Permutation tests



Reinforcement Learning for out-of-sample inference

One article per quarter enters state ⇒ affects all future predictions = Reinforcement Learning (RL) problem for article selection and extractive news summaries

- Challenge: Thousands of articles per quarter ⇒ adopt a contextual bandit approximation with short rollouts of horizon H
- Rewards balance accuracy, narrative priors, decisiveness, non-redundancy.
- KL-regularized policy optimization ensures stability
- Relates to (Bertsekas and Tsitsiklis, 1996; Ziebart, 2010; Todorov, 2006; Haarnoja et al., 2018; Peters et al., 2010; Schulman et al., 2015; Abdolmaleki et al., 2018)

For candidate article $a \in \mathcal{A}_T$: reward = accuracy + prior alignment + fundamental + persistence

$$r_T(s_T, a) = -\mathrm{RMSE}\Big(\hat{GDP}^{(i)}_{1:T_f}, GDP_{1:T_f}\Big) + \lambda_N \,\mathrm{NPRIOR}(a, \tilde{a}_T) + \lambda_F \,\mathrm{FUND}(a) + \lambda_P \,\mathrm{PER}(a, \mathcal{H}_{T-1}),$$

Truncated expected returns define a target distribution:

$$q_T(a) \;\propto\; \mathbb{E}\Big[\sum_{h=0}^{H-1} \gamma^h \, r_{T+h}(s_{T+h}, a_{T+h}) \;\Big|\; a_T = a\Big]$$

Policy (softmax over features)

$$\pi_\theta(a \mid s_T) = \frac{\exp\{\beta\, \phi(s_T, a)^\top \theta\}}{\sum_{a' \in \mathcal{A}_T} \exp\{\beta\, \phi(s_T, a')^\top \theta\}}$$

KL-regularized projection to q_T (trust-region style):

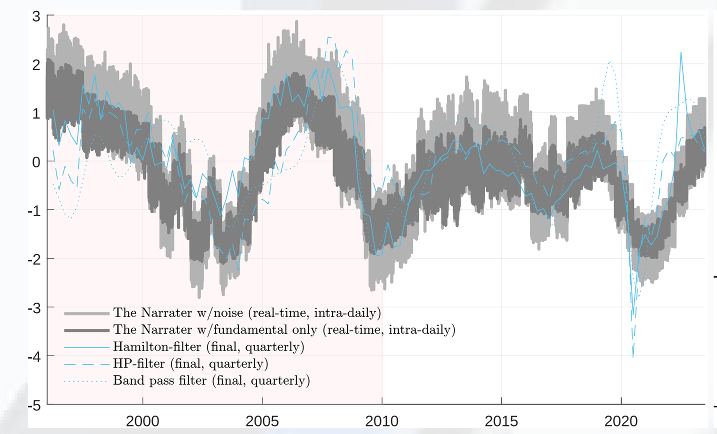
$$\min_\theta \sum_{t_s} \Big[\mathrm{KL}(q_{t_s} \,\|\, \pi_\theta(\cdot \mid s_{t_s})) + \lambda_{\mathrm{KL}} \,\mathrm{KL}(\pi_\theta \,\|\, \pi_0) - \lambda_H H(\pi_\theta)\Big]$$

Greedy bandit baseline - fast, transparent baseline driven purely by predictive error

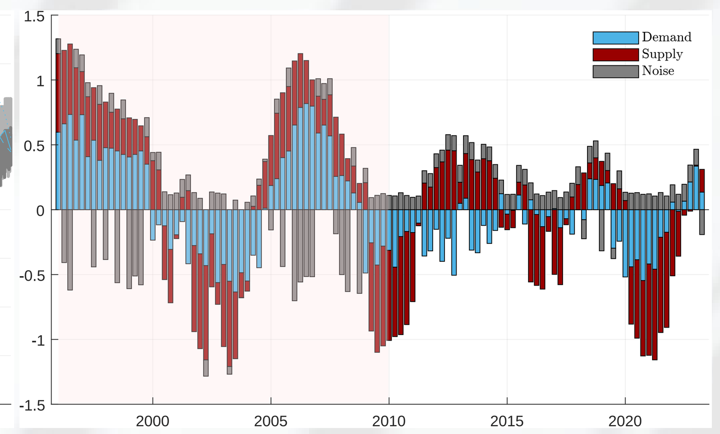
$$\hat{a}_T \;:=\; \arg\min_{a \in \mathcal{A}_T} \; \mathrm{RMSE}(\widehat{\mathrm{GDP}}^{(i)}_{1:Tf}, \mathrm{GDP}_{1:Tf})$$

Benchmark application output

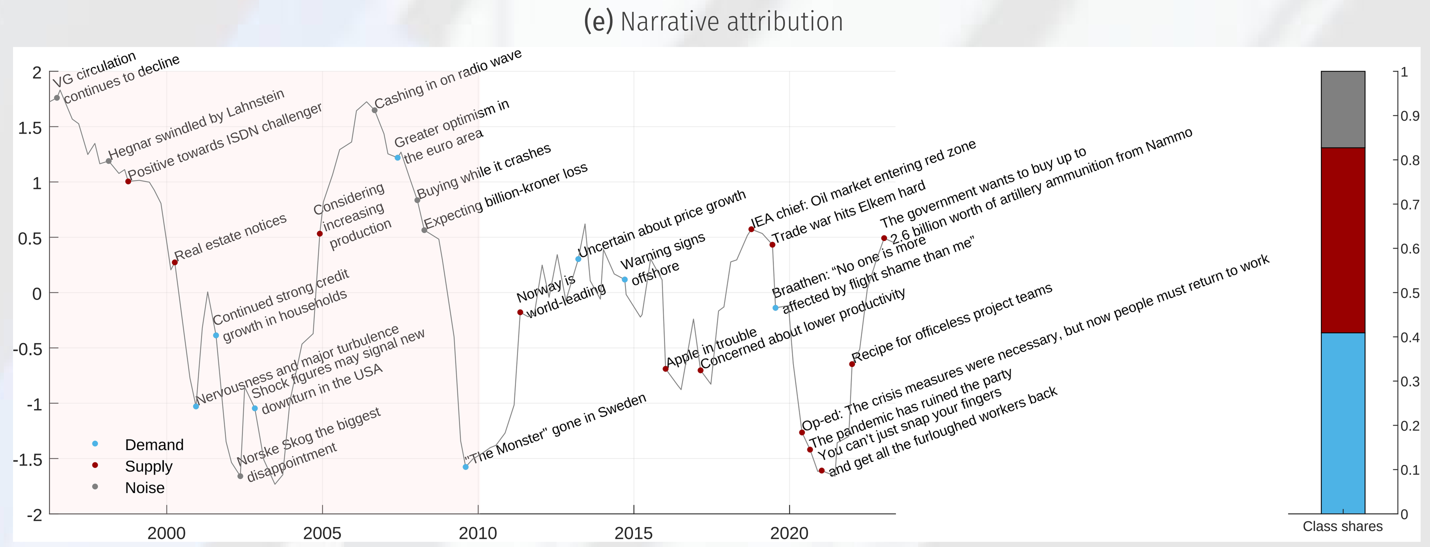
(c) Business cycle



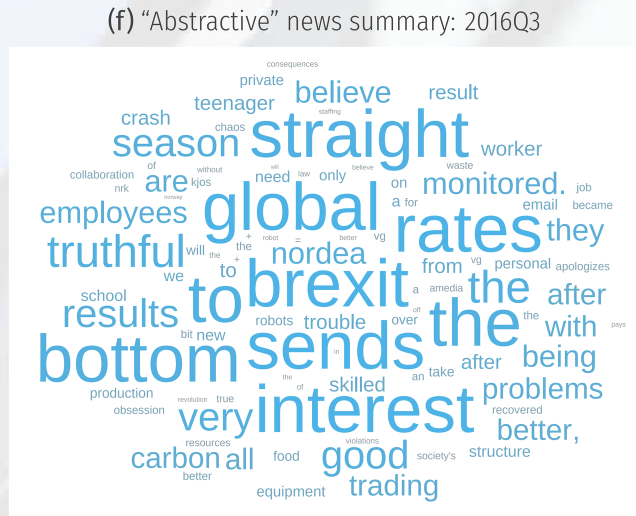
(d) Business cycle decomposition



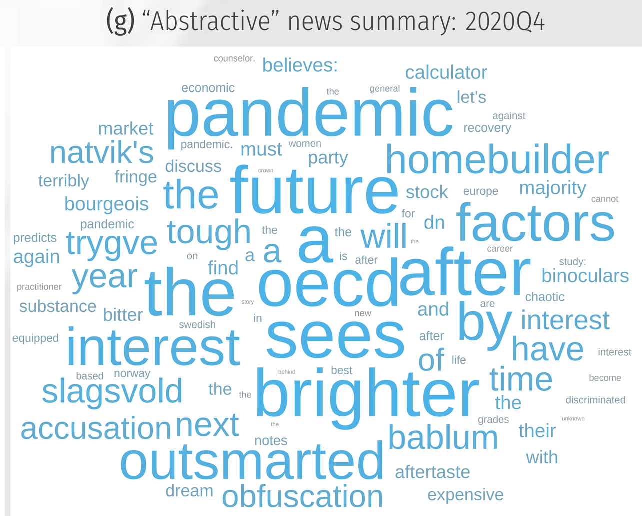
(e) Narrative attribution



(f) "Abstractive" news summary: 2016Q3



(g) "Abstractive" news summary: 2020Q4



References

Abdolmaleki, A., J. T. Springenberg, Y. Tassa, R. Munos, N. Heess, and M. Riedmiller (2018). Maximum a posteriori policy optimisation. *arXiv preprint arXiv:1806.06920*.
Antolín-Díaz, J. and J. F. Rubio-Ramírez (2018). Narrative sign restrictions for SVARs. *American Economic Review 108*(10), 2802–29.
Bańbura, M., D. Giannone, and L. Reichlin (2010). Large Bayesian Vector Auto Regressions. *Journal of applied Econometrics 25*(1), 71–92.
Baumeister, C. and J. D. Hamilton (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review 109*(5), 1873–1910.
Bernanke, B., J. Boivin, and P. S. Eliasz (2005). Measuring the effects of monetary policy: A Factor-augmented Vector Autoregressive (FAVAR) approach. *The Quarterly Journal of Economics 120*(1), 387–422.
Bertsekas, D. P. and J. N. Tsitsiklis (1996). *Neuro-Dynamic Programming*. Belmont, MA: Athena Scientific.
Burns, A. F. and W. C. Mitchell (1946). *Measuring Business Cycles*. Number 46-1 in NBER Books. National Bureau of Economic Research, Inc.
Cochrane, J. H. (1994). Shocks. *Carnegie-Rochester Conference Series on Public Policy 41*, 295–364.
Friedman, M. and A. J. Schwartz (1963). *A Monetary History of the United States, 1867-1960*. Princeton University Press.
Haarnoja, T., A. Zhou, P. Abbeel, and S. Levine (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, pp. 1861–1870.
Hollmann, N., S. Müller, K. Eggensperger, and F. Hutter (2022). TabPFN: A transformer that solves small tabular classification problems in a second. *arXiv preprint arXiv:2207.01848*.
Müller, S., N. Hollmann, S. Pineda Arango, J. Grabocka, and F. Hutter (2021). Transformers can do bayesian inference. *arXiv preprint arXiv:2112.10510*.
Nagler, T. (2023). Statistical foundations of prior-data fitted networks. In *Proceedings of the 40th International Conference on Machine Learning*, Volume 202, pp. 25660–25676.
Peters, J., K. Mülling, and Y. Altün (2010). Relative entropy policy search. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI 10)*, pp. 1607–1612.
Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics 126*(1), 1–50.
Romer, C. D. and D. H. Romer (1989). Does monetary policy matter? A new test in the spirit of Friedman and Schwartz. *NBER macroeconomics annual 4*, 121–170.
Schulman, J., S. Levine, P. Abbeel, M. Jordan, and P. Moritz (2015). Trust region policy optimization. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pp. 1889–1897.
Sims, C. A. (1980). Macroeconomics and reality. *Econometrica 48*(1), 1–48.
Stock, J. and M. Watson (2012). Disentangling the channels of the 2007-2009 recession. *Brookings Papers on Economic Activity Spring 2012*, 81–135.
Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. In O. J. Blanchard and F. Stanley (Eds.), *NBER Macroeconomics Annual*, NBER Chapters, pp. 351–394. Cambridge, MA: The MIT Press.
Todorov, E. (2006). Linearly-solvable markov decision problems. *Advances in Neural Information Processing Systems*, 1369–1376.
Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, Ł. Kaiser, and I. Polosukhin (2017). Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000–6010.
Ziebart, B. D. (2010). *Modeling Purposeful Adaptive Behavior with the Principle of Maximum Causal Entropy*. Ph. D. thesis, Carnegie Mellon University.