

Mining for Uncertainty

Finding Economic Trends in large Newspaper
Corpora

DoCMA WP #1, #2

2



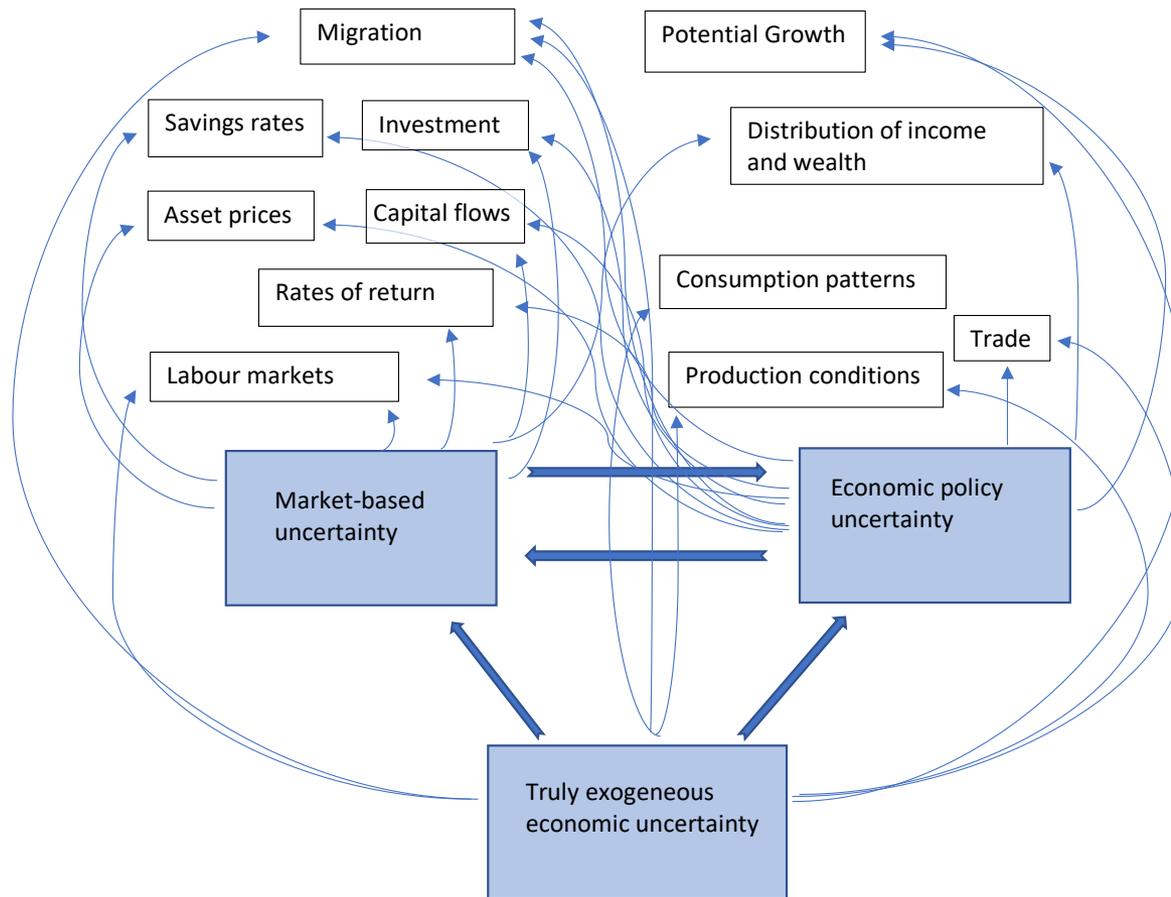
DoCMA Working Paper #2-2020

“I heard the News today, oh Boy”

An updated Version of our Uncertainty Perception Indicator (UPI) – and some general thoughts on news-based economic indicators

Henrik Müller and Nico Hornig*

The Nature of Uncertainty



Source: Müller and Hornig (2020a)

How to Gauge Uncertainty?

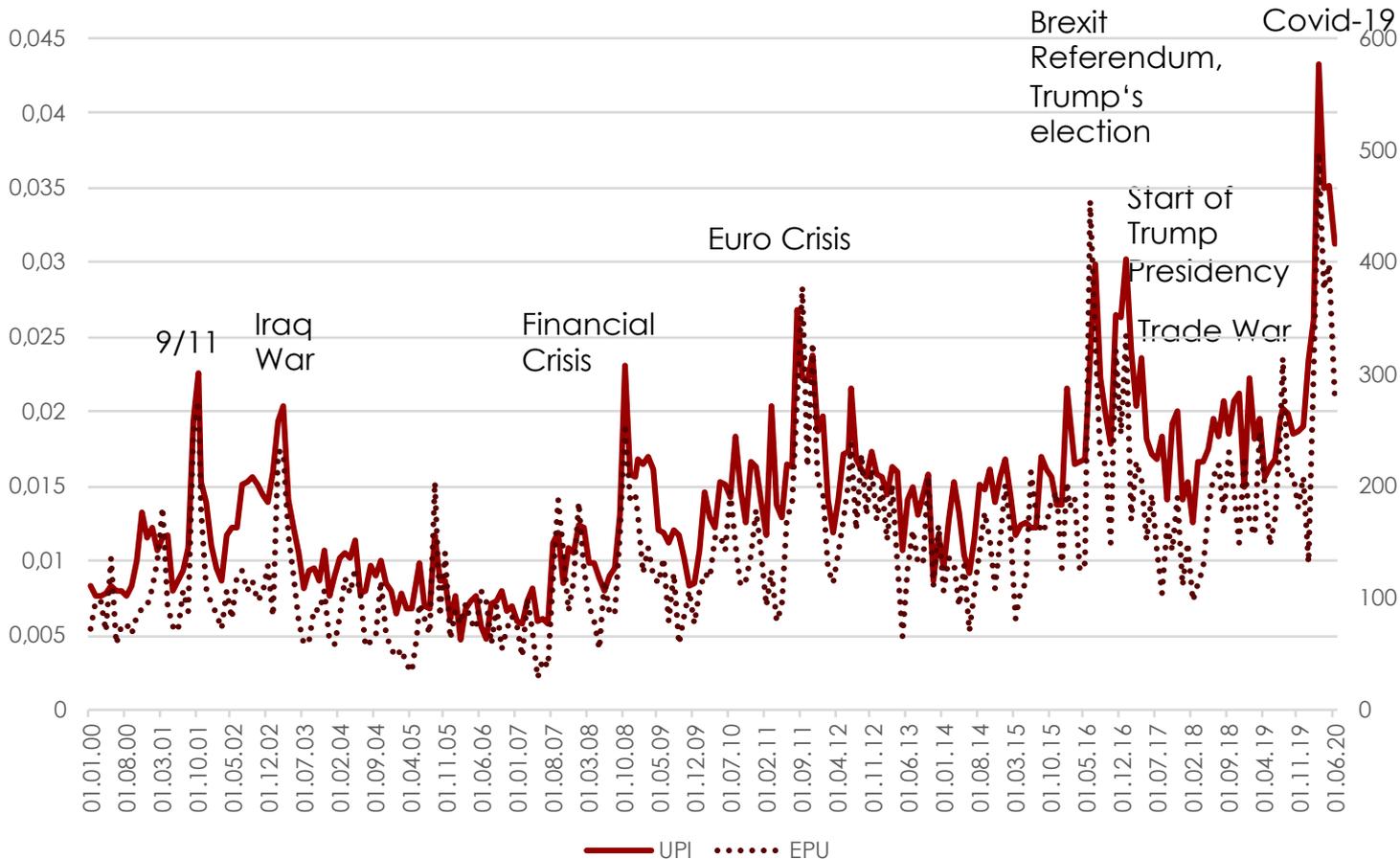
- Since uncertainty is, by definition, an **exogenous** phenomenon economic variables, financial market data, or survey data will not do the job.
- **News-based indicators** have been applied, with the Economic Policy Uncertainty Index (EPU) being the most popular (and widely cited) example (Baker et al. 2016).
 - Rationale: Newspapers report all the news early on, particularly about politics.
- The EPU is basically a count of articles that contain a query of search terms related to uncertainty, the economy and certain areas of economic policy (fiscal, monetary, regulation, taxation).
- Researchers make **ex ante assumptions** about the sources of uncertainty, i.e. economic policy in a narrow sense.
 - *good* because it enables Baker et al. to build a large set of national and regional indicators over long periods of time; *bad* because it filters for *known unknowns*, sources of uncertainty that we have experienced in the past, but not the new stuff that's happening
- Therefore, the EPU only responds when these policy areas are effected by an uncertainty shock, that is, **rather late**.

EPU vs. UPI

- The Uncertainty Perception Indicator (UPI) uses a **more open query** (no specific policy areas, just uncertainty and the economy).
- Different newspapers: EPU for Germany is based on FAZ and Handelsblatt, the UPI on Süddeutsche, Handelsblatt, Welt: 2,9 million articles, applying query yields 37.000
- A **topic-modelling approach** (LDA, Blei et al. 2003) is used to differentiate between various sources of uncertainty.
 - We ran models for K -values of 6, 8, 10, 12, 14. With 14 = best results.
 - Closely related topics may be merged.
- The results of the UPI exercise highlight not only the **size** of shocks but also their **origins** as well as the interactions between **different kinds of uncertainty**.

Shocking Times

German EPU (Index, rhs) vs. UPI (share in corpus), 2000-2020 Q2
(source: policyuncertainty.org, Müller and Hornig, 2020b)



Results

Topic Analysis

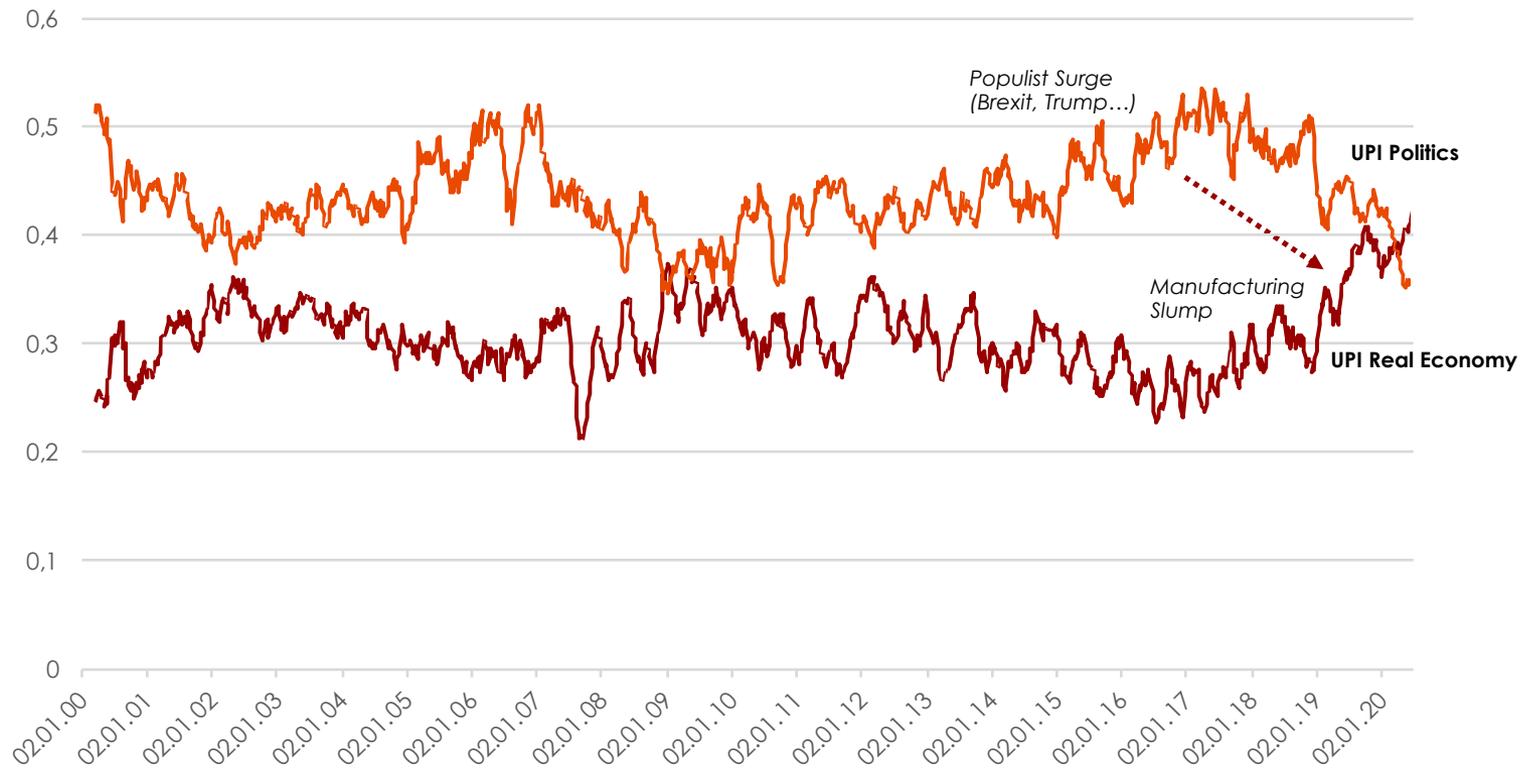
Topics and Labels

Topic No.	Label (equivalent to topic no. in old LDA ^a)	Share in analysis corpus (per cent)	Content	Type of uncertainty	Part of Uncertainty Factor...
1	German Politics (equ. T 4)	6	Political developments in Germany (national level)	Economic Policy	UPI Politics
2	Financial Markets I (no direct equ.)	7	Private Investing, financial market developments	Market-based	UPI Financial Markets
3	Human Resources (no direct equ.)	7	Education and Knowledge, Career issues	Market-based, Economic Policy	UPI Real Economy
4	Central Banks (equ. T 2)	6.8	ECB, Fed etc. actions against crises	Economic policy	UPI Politics
5	EU Conflicts (equ. T 11)	5.6	Brexit, Greece debt etc.	Economic policy	UPI Politics
6	Miscellaneous	10.9	Diverse	–	–
7	Companies & Markets (equ. T 1)	3.5	German Corporates in trouble	Market-based	UPI Real Economy
8	Society (partly equ. T 12)	11	Debates on Capitalism, Globalization, Democracy, Populism, Immigration, national identity	Economic policy	UPI Politics
9	Energy & Climate Change Mitigation (equ. T 7)	6.14	Energy market developments, transition to sustainables etc.	Economic policy/market-based/truly exogenous	UPI Real Economy
10	German Economy (equ. T 5)	7.7	Business cycle developments, forecasts, surveys	Market-based	UPI Real Economy
11	Geopolitics (equ. T 9)	7.1	Conflicts involving US, China, Russia, Turkey, Middle East...	Economic policy	UPI Politics
12	Financial Markets II (equ. T 10)	7.6	Up and down at the bourses	Market-based	UPI Financial Markets
13	Big Business (equ. T 1)	6.55	Developments at quoted international corporates	Market-based	UPI Real Economy
14	Legal Risks (equ. T 6)	7.18	Regulations and court rulings affecting businesses	Economic policy	UPI Politics

see Müller and Hornig (2020b, p. 13). We combine 1 and 14, 2 and 12, 7 and 13 due to their proximity. For topics highlighted in gray equivalents in Müller and Hornig (2020) can be traced.

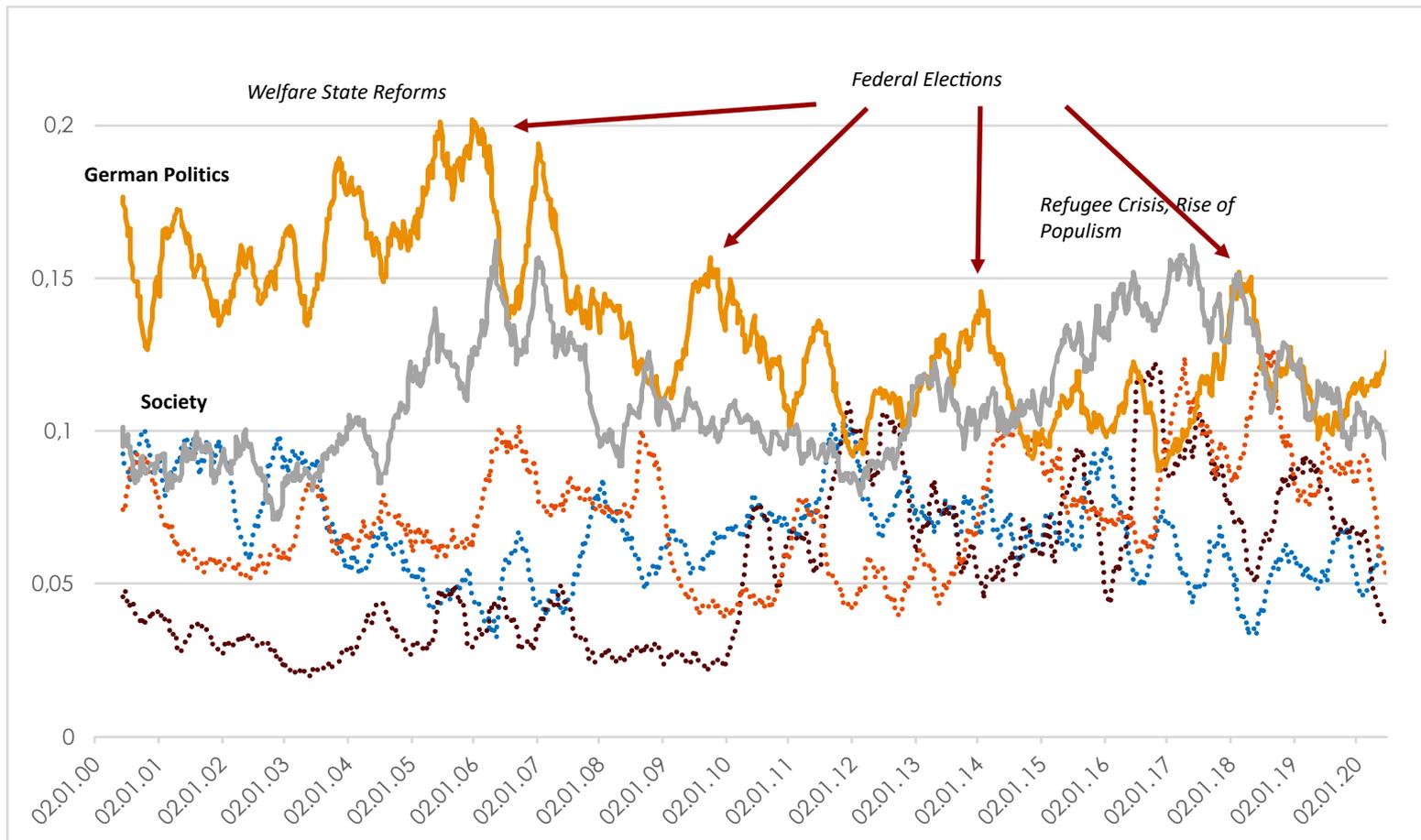
Causes and Consequences?

UPI (shares in analysis corpus, three-month mov. av.)
(source: Müller and Hornig 2020b)



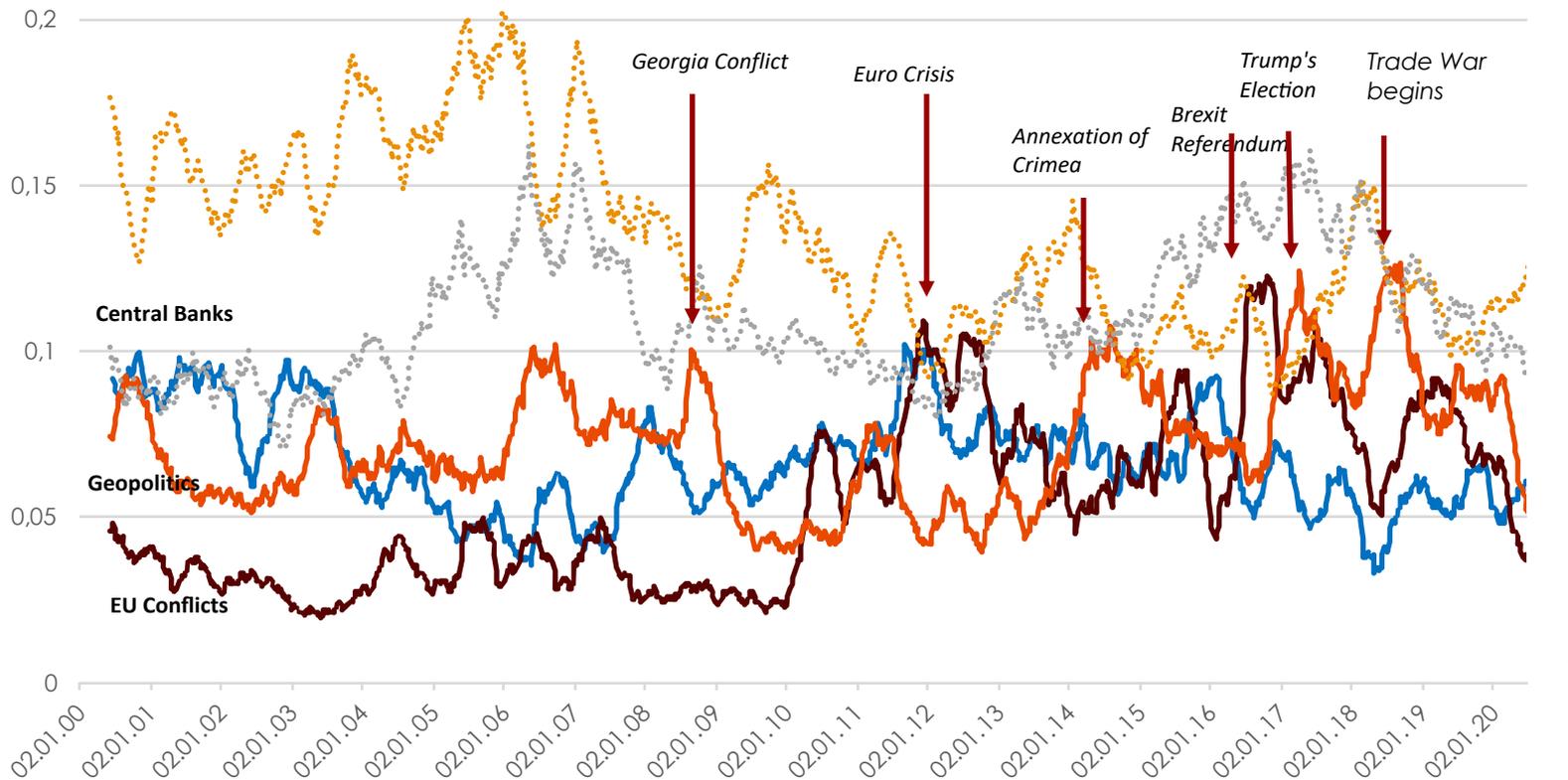
Political Uncertainties, national topics

UPI Politics (shares in analysis corpus, three-month mov. av., source: Müller and Hornig 2020b)



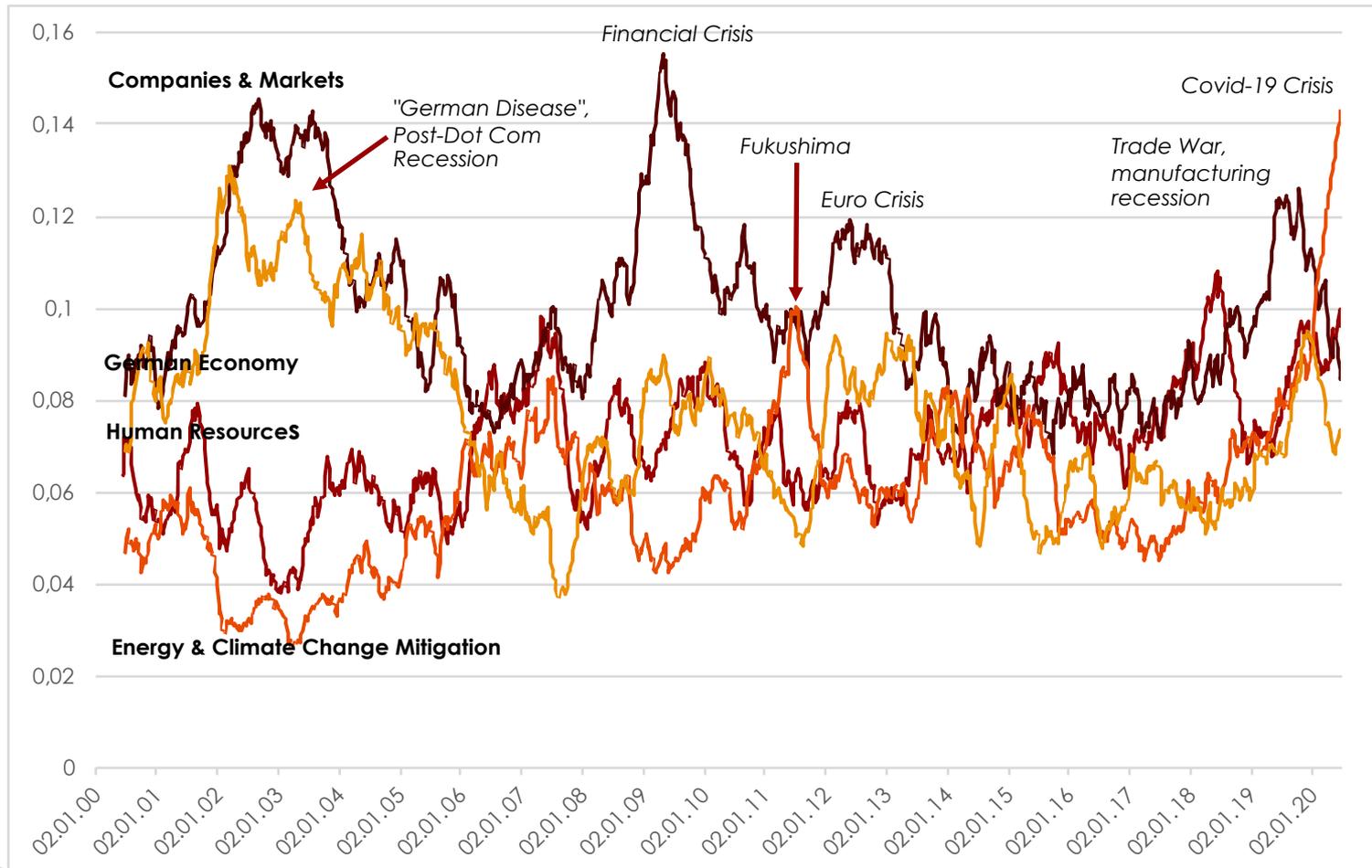
Political Uncertainties, international and European topics

UPI Politics (shares in analysis corpus, three-month mov. av., source: Müller and Hornig 2020b)



Where is Covid?

UPI Real Economy (shares in analysis corpus, three-month mov. av.; source: Müller and Hornig 2020b)



Our Query Seems to Work

- Compared to the EPU our more open search terms seem to be **more sensitive** to new stuff that's happening.
- In Feb. 2020 the UPI with the broader query shows hikes in uncertainty **earlier**.
- Still, as we have seen, the UPI is strongly driven by **political developments**, too.
 - Could be due to reporting patterns by newspapers that tend to focus on politics and the quarrels among its proponents
 - But at the same time, a lot has been happening in the political sphere over the period (international landscape, populism, Brexit, trade wars) – rise in UPI politics is plausible.
- Furthermore, additional issues become apparent, like environment, energy and climate.

Methodology

- LDA turns **content into data**.
 - Elegant way to **measure narratives** (Shiller 2017)
- Topic modelling works for uncertainty analysis (as already shown in Larsen 2017, Müller et al. 2018, Müller 2020).
 - Need to enhance our corpus by adding further papers and media.
- LDA sorts corpora into thematic clusters in an **unsupervised** fashion. This allows for a more open search.
 - Assumptions at the outset of the analysis can be relaxed (broader query viable).
- In our analysis we calculated four models for different values of K . Results for $K=10$ and 12 are broadly similar, but $K=12$ better suited for our purposes (Müller and Hornig 2020).
- LDA yields topics that are **clear-cut** and **interpretable**, driven by **specific events**.

Open Questions

- Will we be able to use LDA to build **stable time-series**?
- Will apparent causalities pass **causality tests**?
- Will our *Uncertainty Factors* and/or individual topics work as **predictors** of economic variables?
- Will our results be replicable for **other countries**?
- So, a lot of work in progress...

References

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636. <https://doi.org/10.1093/qje/qjw024>

Baker, S. R., Bloom, N., Davis, S. J., Terry, S. (2020). Davis, Terry (2020): COVID-induced economic uncertainty and its consequences. (13 April 2020) <https://voxeu.org/article/covid-induced-economic-uncertainty-and-its-consequences>

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.

Bloom, N., Bond, S., & Van Reenen, J. (2007). Uncertainty and Investment Dynamics. *The Review of Economic Studies*, 74(2), 391–415. <https://doi.org/10.1111/j.1467-937X.2007.00426.x>

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica*, Vol. 77, No. 3 (May, 2009), 623–685

Bolton, P., Despres, M., Pereira da Silva, L. A., Svartzman, R., & Samama, F. (2020). *The green swan: Central banking and financial stability in the age of climate change* (Bank for International Settlements, Ed.).

da Silva, L. A. P. (2020, May 14). *Green Swan 2 - Climate change and Covid-19: Reflections on efficiency versus resilience*. <https://www.bis.org/speeches/sp200514.htm>

Knight, F. H. (1921). Risk, Uncertainty, and Profit (Boston MA: Hart, Schaffner and Marx; Houghton Mifflin, 1921). 23.6.2020. <https://oll.libertyfund.org/titles/306>

References

Larsen, V. H. (2017). *Components of Uncertainty* (SSRN Scholarly Paper ID 2955655). Social Science Research Network. <https://doi.org/10.2139/ssrn.2955655>

Müller, H. (2020). Donald Trump als wirtschaftspolitischer Unsicherheitsfaktor. *Ifo Schnelldienst*, 73(1/2020), 26–29.

Müller, H., & Hornig, N. (2020a). Expecting the Unexpected. *DoCMA Working Paper;1-2020*. <https://doi.org/10.17877/DE290R-21089>

Müller, H., & Hornig, N. (2020b). "I heard the News today, oh Boy"*DoCMA Working Paper;2-2020*. <http://dx.doi.org/10.17877/DE290R-21669>

Müller, H., von Nordheim, G., Boczek, K., Koppers, L., & Rahnenführer, J. (2018). Der Wert der Worte – Wie digitale Methoden helfen, Kommunikations- und Wirtschaftswissenschaft zu verknüpfen. *Publizistik*, 63(4), 557–582. <https://doi.org/10.1007/s11616-018-0461-x>

Shiller, R. J. (2017). Narrative Economics. *American Economic Review*, 107(4), 967–1004. <https://doi.org/10.1257/aer.107.4.967>

Vielen Dank für
Ihr Interesse!