

Estimating policy uncertainty within monetary policy debates

Sami Diaf¹, Florian Schütze²

¹ Department of Socioeconomics, University of Hamburg, Germany, ² Department of Socioeconomics, University of Hamburg, Germany.

Abstract

Studying policy uncertainty contained in collections of documents has been a major task for political researchers and economists, who aim at measuring this degree exclusively with wordlists and topic models to feed further econometric inferences or test hypotheses. Such bag-of-word applications constrain the analysis and cannot render a clear picture of uncertainty drivers and their persistence, even if semi-supervised strategies may offer coherent improvements at the topic level. This work proposes a semantic search strategy, using Top2vec, to identify sources of uncertainty, at the debate level, and uncover coherent topics whose representations will be used to get uncertainty prevalence within each debate. Unlike aggregate-level measurements, this strategy is suited to study per speaker contributions at central banks, where uncertainty is regarded as a forward guidance tool and a key strategy when devising monetary policy actions. Applied to FOMC transcripts (1994-2016), the resulting semantic space yields non-overlapping topic vectors indicating a dominance of economic discussions in uncertainty formation within committee meetings, while risks concerns are bounded to financial markets and investments using an investor jargon. Moreover, results demonstrate the importance of experts' contributions in steering the economic debate, hence coloring uncertainty with words not found in traditional uncertainty wordlists and diffusing a significant persistence to uncertainty prevalence during debates that exhibits fractal patterns.

Keywords: *Uncertainty; Semantic search; Topic models; Monetary policy*

1. Introduction

Narrative economics (Shiller, 2017) popularized the already existing interest in extracting information from data using a variety of techniques, often borrowed from machine learning, and use the results as covariates to augment further inferences (Gentzkow, 2019). Social scientists, as opposed to computer scientists, pay attention to results but not the techniques used, resulting in a dominance of bag-of-word methods when it comes to analyzing text data. This strategy, even if it succeeds in uncovering latent patterns, remains suboptimal and constrained (Grimmer and Stewart, 2013). It overlooks the semantic features stemming from documents (Ash et al., 2021), crucial to understanding narrative signals authors try to send throughout their texts.

Advances in natural language processing investigated the importance of context words and the possibility to transform words into vectors encompassing the semantic and syntactic meaning, also known as distributional representation (Mikolov et al., 2013). This concept was later extended at the paragraph or document level (Dieng et al., 2019) and adopted by other architectures (Angelov, 2020) that make use of a dual word-document embedding to efficiently search for meaningful and coherent representations.

Central banking, as an active economic field of interest, witnessed several contributions belonging to the text-as-data fashion. Mostly to gauge sentiments communicated by available corpora or to scale central bankers (Baerg and Lowe, 2020) for an understanding of potential partisanship among monetary policy members or investigate transparency within these committees (Hansen et al., 2018). However, communication-specific characteristics as for ambiguity (Baerg, 2020) and consensus (Meade and Stasavage, 2008) make it difficult to discern intrinsic features as for uncertainty, which is considered as a forward guidance tool in modern central banking (Greenspan, 2004). The existing indices employed to gauge uncertainty based on word counts (Baker et al., 2016) were found to be informative and able to be augmented with further machine learning techniques to improve their accuracy (Tobback et al., 2018). But they fail to determine sources of uncertainty and tie them to specific topics of interest that better explain the context used for such assertions. Moreover, the availability of sentiment dictionary or wordlists for uncertainty (Loughran and McDonald, 2011) cannot guarantee an effective application as context words, being corpus-specific, are often ignored when building such indices, despite being able to confirm hypotheses and meet economic developments.

Particularly, Federal Open Market Committee (FOMC) transcripts have been widely investigated to assess the communicative content as for transparency (Hansen et al., 2018), scaling members' preferences (Baerg and Lowe, 2020) or assessing objectives (Shapiro and Wilson, 2019). Their public release, although with a five-year delay, came as a transparency effort toward a more public-oriented monetary policy, so to end the long-standing secrecy

that prevailed until the end of the 1980s, in what was qualified as *monetary mystique* (Goodfriend, 1986).

We propose in this paper a semantic, topic-based approach to assess policy-based uncertainty using *Top2vec* algorithm, to identify relevant topic structures and terms semantically related to uncertainty from the collection of FOMC transcripts. While usual probabilistic bag-of-word methods use word frequencies to learn global topic structures, they do not fit corpora with a debate structure, which requires specifications to learn local topics. Moreover, the use of prior information as for pre-trained embedding models might be useful to get coherent topics but comes embedded with an information bias (Papakyriakopoulos et al., 2020) when applied to domain-specific corpora. We argue that semantic search strategies are better suited to uncover semantic uncertainty as explained by (Szarvas et al., 2012) without further requiring post-hoc inferences or prior information from external sources. Nonetheless, it is possible to detect uncertainty origins and study their persistence within each debate to quantify its memory occurrence using *Rescaled Range* (R/S) Analysis (Mandelbrot and van Ness, 1968).

Applied to the corpus of per-speaker FOMC transcripts in the United States (1994-2016), the methodology uncovered a macroeconomic color of uncertainty, stemming from economic debates and forecast-based discussions at the meetings, while risk concerns seem to target mostly financial markets. Policy discussions confirm the consensus feature of central bankers and remain exclusively steered by FOMC members, who are prone to use a more neutral, ambiguous tone than in economic debates.

Our paper makes two distinct, significant contributions. In topic models, we demonstrate the efficiency of semantic search models and their ability to uncover local topics when analyzing debate-structured corpora, freeing the analysis from potential biases arising when overlooking the debate dimension. In policy analysis, we reinforce the *policy context* by searching topics, documents and words related to a specific domain or task from the source, instead of doing it at the word level. This enables us to get an extended representation of semantic uncertainty (Szarvas et al., 2012). Moreover, we introduce the concept of *subject persistence* within debates by studying the statistical properties of persistence measurements as for Hurst exponent (Mandelbrot and van Ness, 1968), to examine the regular use of uncertainty within monetary policy debates.

2. Methodology

Traditional topic models are based on the Dirichlet distribution, which is not sequential and unable to capture time-based topics and unbiased topic prevalence over time. For the case of debates, the use of a dynamic topic model is problematic, as defining a time frame to learn topics could be biased. This is due to debates might be fierce targeting few topics or with a

broad spectrum, spanning over many topics. The relatively small size of debate contributions makes it very difficult to capture the rhetoric used by the speakers when adopting a bag-of-word strategy. We argue that semantic representations make it possible to quantify/study a given subject, whether be it a topic of a keyword, and track its prevalence throughout the debate by computing a persistence index, as for Hurst exponent (Mandelbrot and van Ness, 1968), that informs us about its relative incidence throughout the debate. A fierce debate dealing with a given subject is likely to demonstrate persistence, denoting a permanent prevalence that needs to be captured via a measure, while a vague debate has mostly anti-persistent, short memory patterns. The Hurst exponent, resulting from the Rescaled Range Analysis (Hurst, 1951), is used to measure the degree of variability associated with a given time series. It is linked to a geometric measure of irregular shapes known as the *fractal dimension* (Mandelbrot and van Ness, 1968).

For the case of monetary policy practices, debates are not known to be fierce, even if dissents are likely to happen but not explicated (Meade and Stasavage, 2008) either in the resulting transcripts of when casting votes. This reinforces the hypothesis that a subject occurrence within a single debate is likely to have a limited variability that could be gauged via a unique fractal dimension. It results, in this paper, the application of the R/S Analysis within each debate on the computed uncertainty scores for each contribution so to give an unbiased measure of how uncertainty was persistent during the FOMC meetings.

3. Data and Results

FOMC transcripts from 1994 to 2016 were gathered from the Federal Reserve website, consisting of 250 documents featuring mostly meetings transcripts as well as transcribed conference calls. Texts were cleaned and decomposed into individual contributions (54,173 entries), so to stress out the debate dimension and an equal weight for each participant at the meetings. Such a scheme maximizes topic detection and will not overlook small contributions that may appear as distinct topics related to the studied task.

Applying Top2vec to the corpus yields 463 local topics related to monetary policy committee inner functioning, where uncertainty-content is correlated to situations where information is vague, ambiguous, or misleading.

Table 1 shows the 10 most correlated topics with the word *uncertainty*. Topics 1, 3, 8 and 9 comprise technical terms used at presentations, usually conducted by economists related to the economic status, while other topics seem to be related to the international environment (topic 2), housing sector (topics 4 and 7), and economic analysis (topics 5 and 9).

Table 1. Top 10 topics correlated with "uncertainty" and their 10 most related words

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
error	war	tendencies	weak	moderation
bands	terrorist	projections	ui	premia
intervals	invasion	italics	housing	empire
errors	iraq	narratives	household	reassuring
fan	kuwait	assessments	deleveraging	downside
frb	knightian	tendency	labor	trajectory
interval	geopolitical	panels	availability	risk
stochastic	military	your	drags	outlook
width	gulf	column	households	portend
model	scandals	clustered	consumer	benign
Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
contacts	burn	fan	okun	unhinged
sixth	resets	charts	law	smidgen
atlanta	tapped	bands	nairu	mishkin
directors	subprime	nondissenting	gap	blips
anecdotal	mortgages	obligated	relationship	bad
retailer	conduits	dissenters	model	knightian
reports	delinquencies	errors	intercept	nimble
optimism	jumbo	histogram	statistical	loop
regards	sheets	width	regularity	pray
reported	uninsured	nonvoters	arguable	very

Source: Authors' own calculations.

The top 10 topics in table 1 are weakly correlated with the word "uncertainty" and the correlation coefficient ranging from 17% to 25%. Table 2 confirms the dominance of the

forecasting jargon, where the top 20 words correlated with the term *uncertainty* are given, along the cosine similarity of their respective vectors, taken from the *Top2vec* semantic space.

Table 2. Top 20 words correlated with the word "uncertainty" and their correlations given by the cosine similarity.

Word	Correlation	Word	Correlation
uncertainties	0.694	regarding	0.550
uncertain	0.659	risks	0.547
surrounding	0.637	face	0.546
confidence	0.624	potential	0.545
many	0.593	there	0.545
around	0.576	outlook	0.545
see	0.565	sense	0.544
about	0.561	of	0.543
given	0.557	still	0.541
considerable	0.554	more	0.540

Source: Authors' own calculations.

The cluster of documents formed by the reduced semantic space will be used to get the correlation of each individual level contribution with the word “uncertainty”. In other terms, each speaker in each meeting gets a score that translates the degree of uncertainty used in his/her contribution during the FOMC meetings.

It results in the estimation of uncertainty persistence at each meeting using rescaled range analysis, namely, the Hurst exponent, which provides a good measurement of persistency for time series data. For values of Hurst exponent ranging from 0 to 0.5, the series is said to be an anti-persistent, mean-convergent process, while values from 0.5 to 1 indicate a persistent process that digresses from the mean. A 0.5 value indicates a memoryless process, known as the Brownian motion.

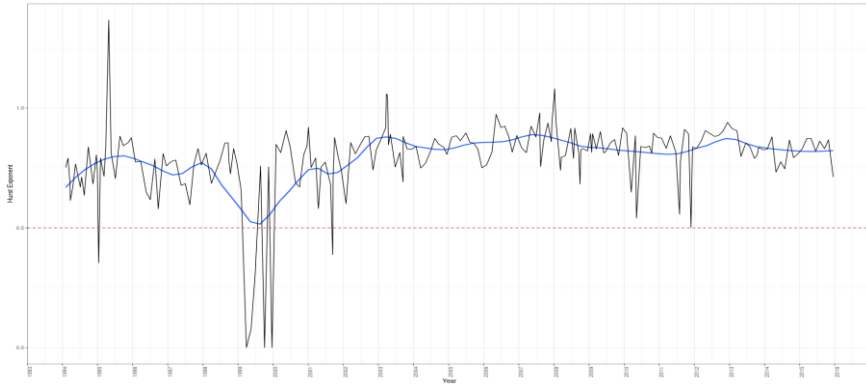


Figure 1. Estimated uncertainty persistence using Hurst exponent within each debate. Blue line is the Loess smoothed curve. Source: Authors' own calculations.

Figure 1 shows an overall erratic, highly volatile persistence, with an anti-persistent episode during the period 1999–2001 that could indicate a frequent use of unsure wording/rhetoric in the debates. Overall, Hurst exponent demonstrates a significant occurrence of uncertainty-related contributions, which could be interpreted as a specific consensus shared by some members, or a *herding*, at the contrast of a general *consensus* (Baerg, 2020; Meade and Stasavage, 2008) that was found prevailing at the FOMC meetings. In other terms, a member hinting uncertainty is likely to be followed by others who do the same, therefore leading to a Hurst exponent exceeding 0.5. Rarely, the Hurst exponent records values at 0.3 or below, which could be interpreted as a polarization of the debate around uncertainty.

4. Conclusion

Uncertainty, as an adopted strategy by central bankers, remains difficult to pin down with automated text analysis. Given the necessity to perform such granular tasks, learning local structures via topic vectors helped to identify sources of uncertainty at FOMC meetings, consisting mainly of the use of forecasters' jargon by economists, rather than policymakers. The idea of topic persistence was introduced to gauge the variability of uncertainty-related contributions within each meeting. The Hurst exponent, as a persistency measurement, indicated a clear persistent trend, if we except periods of recession/crises that showed an anti-persistent behavior. These findings confirm consensus being a key feature of modern central banking communication. Latent differences among speakers regarding uncertainty were found to be linked to the economic outlook, usually debated at the opening of the meetings, with the use of proper words related to economic forecasting.

References

- Airoldi, E. M., Blei, D., Erosheva, E. A., & Fienberg, S. E. (Eds.). (2014). *Handbook of Mixed Membership Models and Their Applications*. Hoboken: Taylor and Francis.
- Angelov, D. (2020). Top2Vec: Distributed Representations of Topics. ArXiv. Retrieved from <https://arxiv.org/pdf/2008.09470>
- Ash, E., Gauthier, G., & Widmer, P. (2021). RELATIO: Text Semantics Capture Political and Economic Narratives. ArXiv. Retrieved from <https://arxiv.org/pdf/2108.01720>
- Baerg, N. (Ed.). (2020). *Crafting consensus: Why central bankers change their speech and how speech changes the economy*. New York, NY, United States of America: Oxford University Press.
- Baerg, N., & Lowe, W. (2020). A textual Taylor rule: estimating central bank preferences combining topic and scaling methods. *Political Science Research and Methods*, 8(1), 106–122. doi:10.1017/psrm.2018.31
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty*. *The Quarterly Journal of Economics*, 131(4), 1593–1636. doi:10.1093/qje/qjw024
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *J. Mach. Learn. Res.*, 3(null), 993–1022.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, June). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. doi:10.18653/v1/N19-1423
- Dieng, A. B., Ruiz, F. J. R., & Blei, D. M. (2019). Topic Modeling in Embedding Spaces. Retrieved from <https://arxiv.org/pdf/1907.04907>
- Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, 57(3), 535–574. doi:10.1257/jel.20181020
- Goodfriend, M. (1986). Monetary mystique: Secrecy and central banking. *Journal of Monetary Economics*, 17(1), 63–92. doi:10.1016/0304-3932(86)90006-1
- Greenspan, A. (2004). Risk and Uncertainty in Monetary Policy. *American Economic Review*, 94(2), 33–40. doi:10.1257/0002828041301551
- Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, 21(3), 267–297. doi:10.1093/pan/mps028
- Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach*. *The Quarterly Journal of Economics*, 133(2), 801–870. doi:10.1093/qje/qjx045
- Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1), 770–799.
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65. doi:10.1111/j.1540-6261.2010.01625.x

- Mandelbrot, B. B., & van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10, 422–437.
- Meade, E. E., & Stasavage, D. (2008). Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve. *The Economic Journal*, 118(528), 695–717. doi:10.1111/j.1468-0297.2008.02138.x
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. ArXiv. Retrieved from <https://arxiv.org/pdf/1301.3781>
- Papakyriakopoulos, O., Hegelich, S., Serrano, J. C. M., & Marco, F. (2020). Bias in word embeddings. In M. Hildebrandt (Ed.), *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 446–457). doi:10.1145/3351095.3372843
- Shiller, R. J. (2017). Narrative Economics. *The American Economic Review*, 107(4), 967–1004. Retrieved from <http://www.jstor.org/stable/44251584>
- Szarvas, G., Vincze, V., Farkas, R., Móra, G., & Gurevych, I. (2012). Cross-Genre and Cross-Domain Detection of Semantic Uncertainty. *Computational Linguistics*, 38(2), 335–367. doi:10.1162/COLI_a_00098
- Tobback, E., Naudts, H., Daelemans, W., Junqué de Fortuny, E., & Martens, D. (2018). Belgian economic policy uncertainty index: Improvement through text mining. *International Journal of Forecasting*, 34(2), 355–365. doi:10.1016/j.ijforecast.2016.08.006