CS 224N, Winter 2023
Stanford University

# Dynamic Fed Attention

## Custom Final Project

Amar Venugopal
amarvenu@stanford.edu

## Abstract

The Federal Reserve (“the Fed”) is the principal monetary policy body of the United States, and its policies have ramifications throughout the economy. As a result, significant attention is paid to the statements and speeches of its principal officers and committees. Past literature has sought to apply static topic models to these texts in an attempt to gauge the attention of the Federal Reserve (Jegadeesh and Wu, 2017).

This project seeks to extend this literature to the use of dynamic embedded topic models (D-ETM), formulated by Dieng et al. (2019), comparing the results to dynamic latent Dirichlet allocation (D-LDA) established in Blei and Lafferty (2006). I show that D-ETM can significantly outperform D-LDA in terms of quantitative measures such as topic coherence and diversity and investigate the variable effect of using different classes of embedding models for D-ETM.

## Data

The primary dataset for this project is the corpus of FOMC meeting statements and speeches made by Federal Reserve Governors. Speeches are obtained from an existing repository and span from 1996-2020, while the meeting statements are pulled from the FedTools API and span from 1994-2023.

The preprocessing steps taken on this dataset include:

- removing stopwords
- employing a part-of-speech tagger to remove words corresponding to “uninformative” text that do not correspond to underlying economic discussion (e.g. proper nouns, etc.)
- lemmatizing words to remove the effect of specific conjugations
- removing words that occur in fewer than 1% or greater than 90% of the documents. The resulting cleaned dataset consists of 4938 unique words across 1650 documents.

## Models

**D-LDA:** The benchmark model

**SBERT:** Static word embeddings derived from the provided BERT embedding weight matrix.

**CBERT:** Static-analog word embeddings derived from averaging last hidden layer contextual word embeddings from occurrences of a given word across a subset of documents in the corpus.

**GloVE:** Static, pretrained word embeddings.

**T5P:** Static-analog embeddings obtained from averaging all hidden layers of the model when a given word is passed through.

**T5L:** Static-analog embeddings obtained from the last hidden layer of the model when a given word is passed through.

## Quantitative Evaluation Methods

**Topic Coherence:** I adopt the $C_V$ coherence metric, which combines an indirect cosine measure of word similarity with a normalized log-ratio measure of word co-occurrence probabilities. This measures the semantic similarity between a topic’s top words.

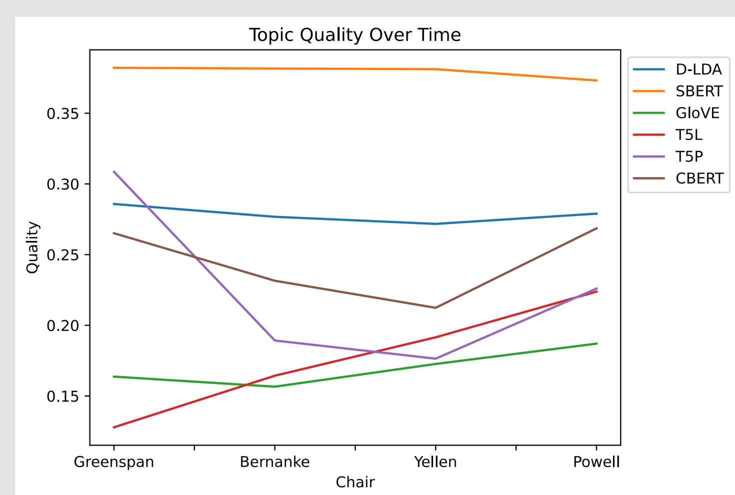
**Topic Diversity:** I implement a metric known as ”proportion unique words”, which measures the percentage of unique words in the top words of all topics.

**Topic Quality**: The product of **coherence** and **diversity**

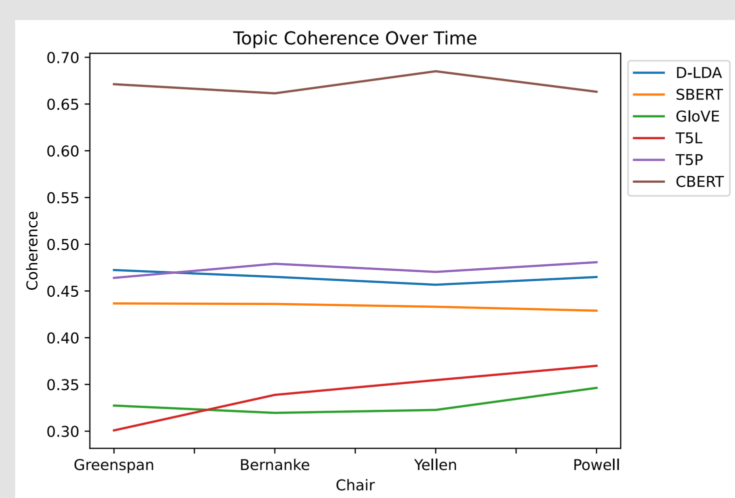
For all quantitative metrics, a higher score is more desirable.

## Quantitative Results

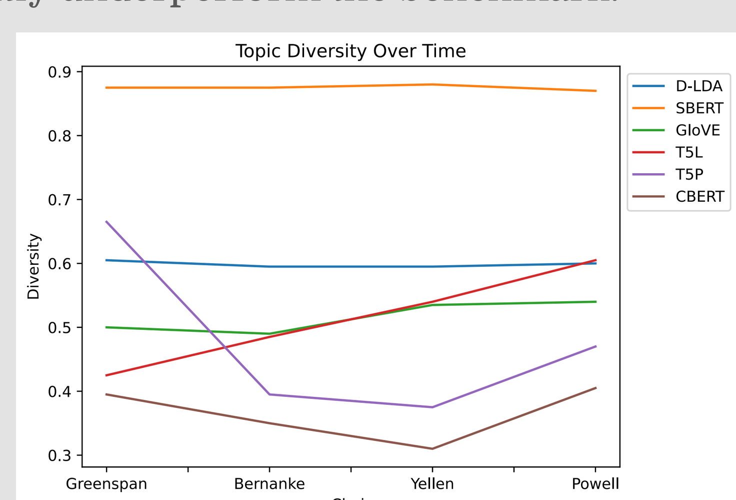
SBERT clearly dominates in terms of quality, while GloVE, T5 and CBERT all lag behind the benchmark D-LDA:




CBERT demonstrates strong coherence, while GloVE lags again:



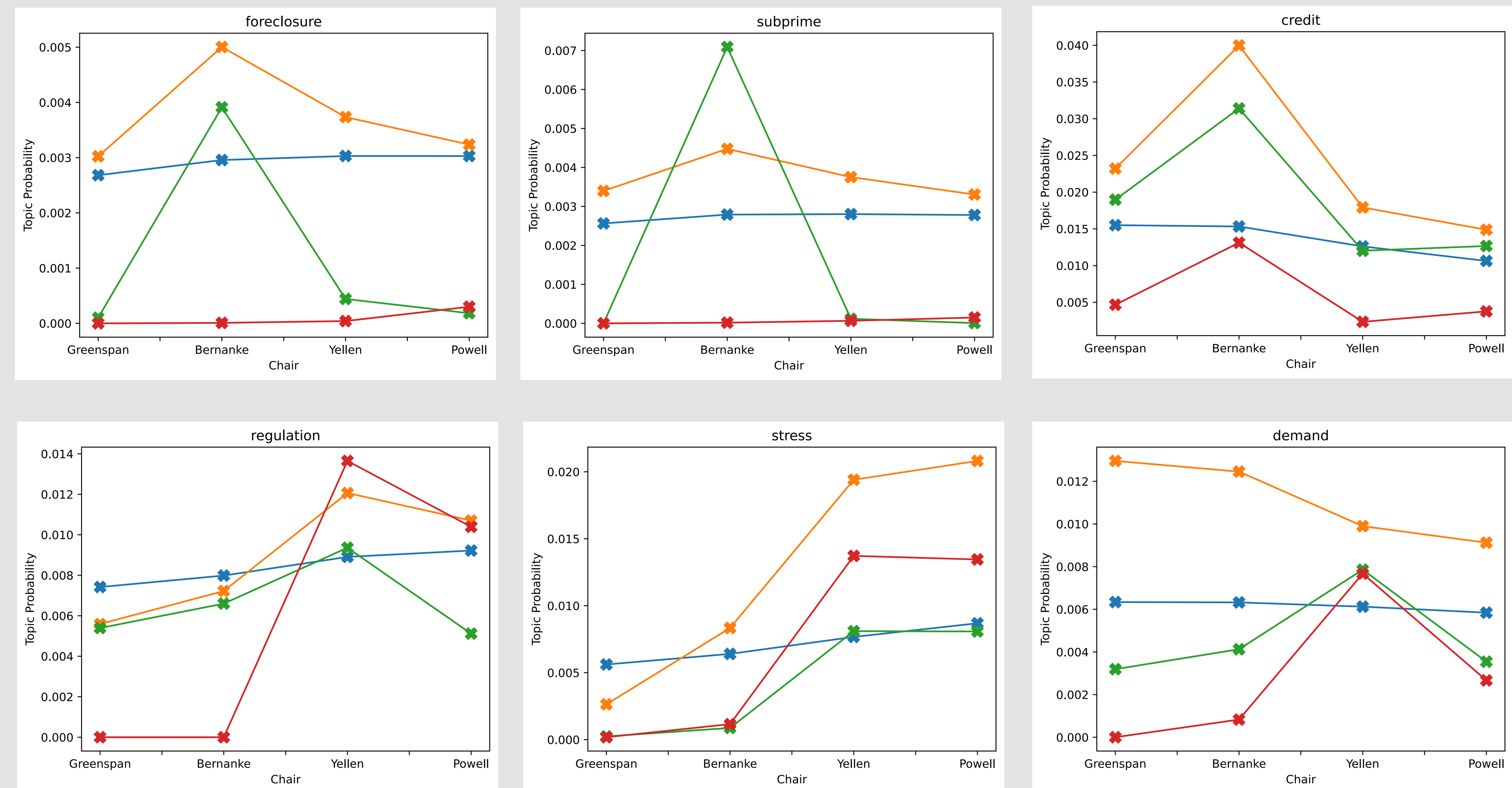

SBERT is the clear winner in diversity, while CBERT and T5 significantly underperform the benchmark:




## Qualitative Results

SBERT, D-LDA, T5L, GloVE

Plot the evolution of word probabilities with their most closely associated topic over time to view how the varying models capture key trends in the economic environment. SBERT appears to perform the best consistently.




## Conclusion + Future Work

- D-ETM is capable of significantly performing D-LDA, even when applied to a small dataset such as Fed speeches and statements,
- However, D-ETM is highly sensitive to the choice of embedding used, appearing to favor static implementations over those that seek to leverage in-sample context.
- Even so, static embeddings produced by BERT, a transformer model, significantly outperform those produced by GloVE, despite the latter being specifically designed to produce static word vectors.
- This suggests that the increased sophistication of transformer models translates into higher quality static word representations, and that the demonstrated limitations of those models in this project is reflective more of the process through which static representations were created, rather than the power of transformer architecture for this use case.
- Given these findings, a key direction for future research would therefore be to utilize the static embeddings produced by Gupta and Jaggi (2022), who put forth a methodology to generate high-quality static word embeddings from contextual transformer models, including BERT. The results of this project suggest that this approach has the potential to provide significant performance improvements and yield more intuitive and useful topics.

## References

Narasimhan Jegadeesh and Di Wu. 2017. Deciphering Fedspeak: The information content of FOMC meetings. *Monetary Economics: Central Banks–Policies & Impacts eJournal.*

Adji B. Dieng, Francisco J. R. Ruiz, and David M. Blei. 2019a. The Dynamic Embedded Topic Model. ArXiv:1907.05545 [cs, stat].

David M. Blei and John D. Lafferty. 2006. Dynamic topic models. In *Proceedings of the 23rd International conference on Machine learning*, ICML ’06, pages 113–120, New York, NY, USA. Association for Computing Machinery.

Prakhar Gupta and Martin Jaggi. 2021. Obtaining Better Static Word Embeddings Using Contextual Embedding Models. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5241–5253, Online.