GEORGETOWN UNIVERSITY | McCourt School of Public Policy

Playing to the Gallery: Emotive Rhetoric in Parliaments

MORITZ OSNABRÜGGE, SARA B. HOBOLT, TONI RODON

PPOL-6801 Text As Data

Amber Ni
Tian Tong

Introduction

Background:

- **Many prominent politicians rely heavily on emotional appeals in electoral compaigns and public speeches under recent research**
- **We know less about whether and how emotions are used strategically inside the legislature, particularly in formal parliamentary debates**

Research questions:

- **When and why do legislators use emotive rhetoric in parliament?**

Hypothesis:

- **Legislators make greater use of emotive language in high profile parliment debates than in routine, low-profile debates (In high-profile context, parliamentarians have an incentive to *'play to the gallery'* by using emotionally charge language to attract attention and signal alignment with public sentiment.)**
- **This challenges the idea that political language is always technocratic or policy-focused — it's also theatrical, performative, and audience-aware.**

Methods

Dataset:

- A dataset of nearly one million parliamentary debates between 2001 and 2019 in the lower house of the UK Parliament, the House of Commons. The dataset includes:
  - Debate types: Prime Minister's Questions （PMQs), Ministerial Question Times, Queen's Speech debates, and Urgent Questions
  - Speaker's identity/Party affiliation/Role in government/Date

Methods:

- The key outcome variable is emotive rhetoric, which the authors quantify using dictionary-based methods:
- Supervised machine learning: training a machine learning algorithm on the hand-annotated data to predict the labels of unseen speeches. [costly]
- Affective Norms for English Words dictionary [not developed for political texts]
- Word-embedding [add content specific words to the dictionary]

To create a domain specific dictionary

Methods

1. ANEW Dictionary

- Identify seed words in the ANEW(Affective Norms for English Words) dictionary
- ANEW provides human-rated scores for words along emtional dimensions

2. Word Embedding

- Train word vectors on parliamentary speeches using *skip-gram* (window = 10, dim = 250).
- Score each word by comparing cosine similarity to emotive vs. neutral seed words.
- Build vocabulary of 2,015 emotive and 2,095 neutral words.
- Calculate emotive rhetoric score as [% emotive words − % neutral] words (range: –100% to +100%).

Result

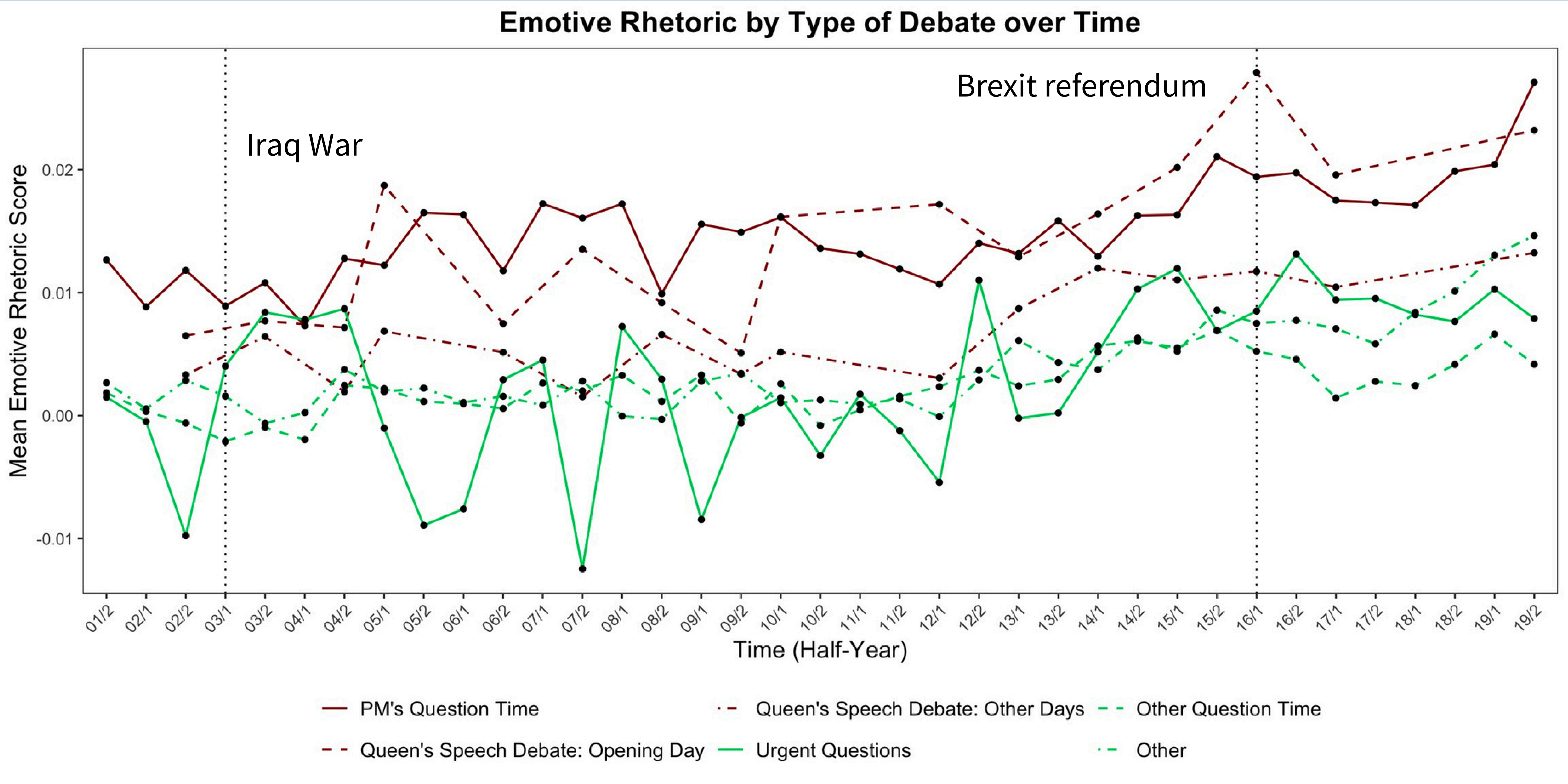
Example: Emotive and Neutral Speeches - a speech with high and a speech with low levels of emotive rhetoric.

TABLE 2. Examples: Emotive and Neutral Speeches

Score	Text	Speaker
43	Evil happens when good people stand by and do nothing. There is evil running through and infiltrating the Labour party, but it is full of good people and they are trying to do something about it. I commend them, appreciate them and have nothing but respect for them.	Alec Shelbrooke, MP
−25	When used with old-fashioned copper wires, 10 megabits can become a lot less than that. We need a superfast fibre infrastructure instead of copper wires.	Geoffrey Clifton-Brown, MP

Results

The development of emotive rhetioric in the period 2001 until 2019



- **The level of emotive rhetoric is highest for PMQs and the opening day of the Queen's Speech debate**

Result

Five regression models with different specifications [Depedent variable: emotive rhetoric score]

TABLE 3. Regression Analysis of Emotive Rhetoric

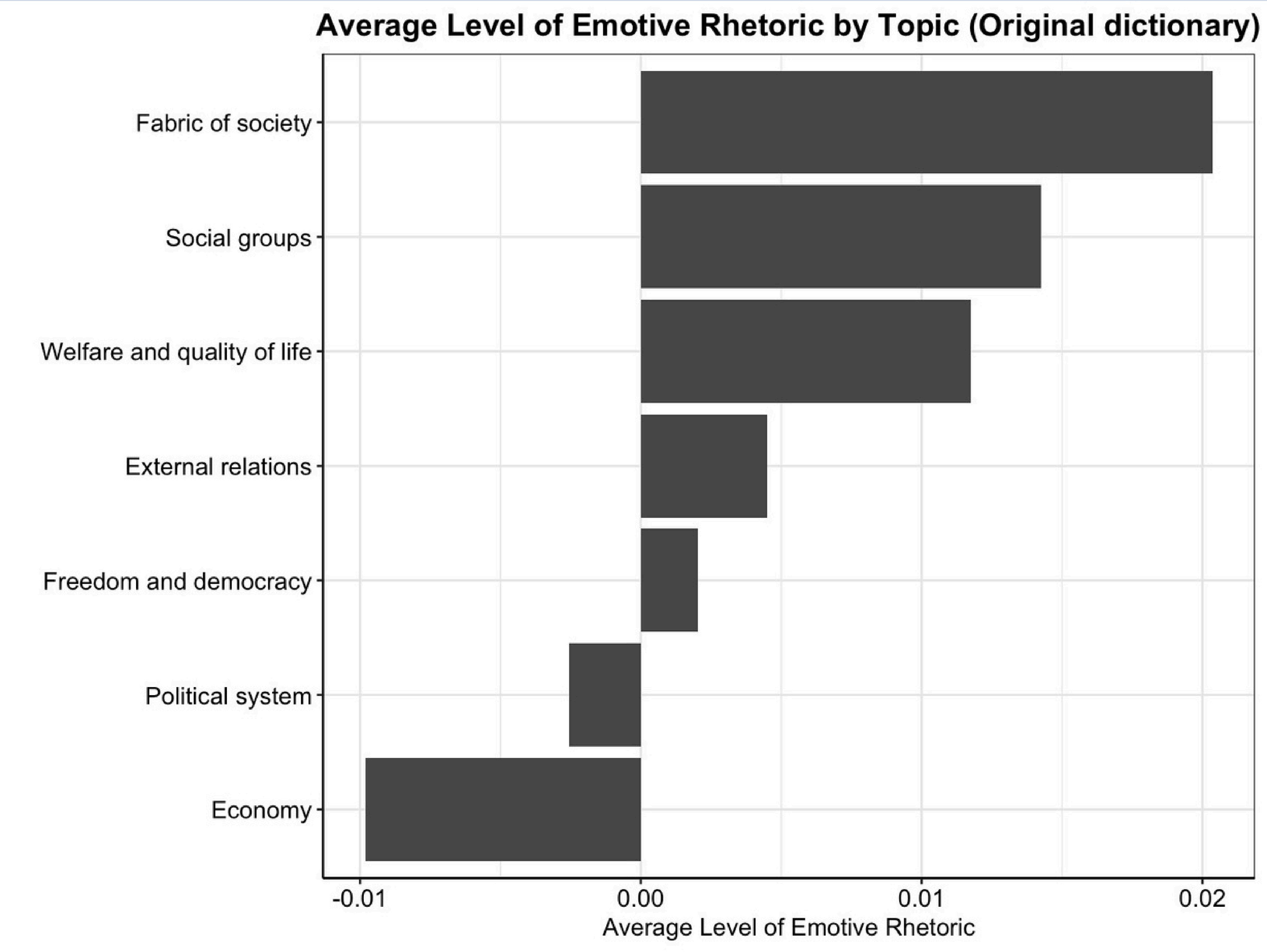
	(1)	(2)	(3)	(4)	(5)
PMQs	2.226**	2.198**	1.562**	1.618**	1.663**
	(0.128)	(0.124)	(0.137)	(0.167)	(0.214)
Queen's Speech Debate: Opening Day	2.172**	2.166**	1.943**	2.061**	2.064**
	(0.214)	(0.208)	(0.180)	(0.194)	(0.166)
Queen's Speech Debate: Other Days	0.581**	0.623**	0.515**	0.643**	1.090**
	(0.106)	(0.104)	(0.083)	(0.095)	(0.078)
Ministerial Question Time	−0.293**	−0.354**	−0.433**	−0.378**	−0.207**
	(0.074)	(0.072)	(0.053)	(0.069)	(0.057)
Urgent Questions	0.610**	0.148	0.009	0.162	0.044
	(0.116)	(0.105)	(0.089)	(0.099)	(0.096)
Constant	0.850**	−0.075	0.033	−1.664**	0.188*
	(0.072)	(0.101)	(0.082)	(0.361)	(0.080)
Linear time trend		X	X	X	X
MP fixed effects			X		X
Controls				X	
Weighting by speech length					X
N	958,925	958,925	958,925	958,925	958,925
R^2	0.003	0.006	0.041	0.014	0.094

Note: Standard errors, clustered by speaker, in parentheses; $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

- **The effect across the type of debates, we observe that the effect size of the PMQs and the opening day of the Queen's Speech debates is the largest**

Result

Issue-Specific Effect: The average level of emotive rhetoric by topic



Some topics are simply more emotive than others?

- **Method: Supervised learning to implement cross-domain policy classification**
- **Result: The highest emotiveness is found in speeches on the topic fabric of society, which includes discussions about the national way of life, traditional morality, and law and orde.**

Result

Issue-Specific Effect: The average level of emotive rhetoric by topic

TABLE 4. Regression Analysis of Emotive Rhetoric with Topic Fixed Effects

	(1)	(2)	(3)	(4)	(5)
PMQs	2.020**	1.995**	1.426**	1.383**	1.525**
	(0.156)	(0.117)	(0.143)	(0.153)	(0.208)
Queen's Speech Debate: Opening Day	2.298**	2.287**	2.012**	2.132**	2.303**
	(0.221)	(0.216)	(0.171)	(0.201)	(0.154)
Queen's Speech Debate: Other Days	0.554**	0.605**	0.491**	0.628**	1.033**
	(0.096)	(0.094)	(0.080)	(0.087)	(0.073)
Ministerial Question Time	−0.436**	−0.489**	−0.539**	−0.492**	−0.319**
	(0.063)	(0.061)	(0.045)	(0.060)	(0.049)
Urgent Questions	0.390**	−0.094	−0.168*	−0.068	−0.111
	(0.108)	(0.097)	(0.084)	(0.091)	(0.086)
Constant	0.908**	−0.069	−2.571**	0.899^{+}	−2.513**
	(0.064)	(0.089)	(0.085)	(0.517)	(0.079)
Linear time trend		X	X	X	X
MP fixed effects			X		X
Controls				X	
Weighting by speech length					X
Topic fixed effects	X	X	X	X	X
N	958,925	958,925	958,925	958,925	958,925
R^2	0.048	0.052	0.078	0.057	0.158

Note: Standard errors, clustered by speaker, in parentheses; $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$.

- **Effect size is slightly smaller**

Extension - Trained Dictionary

Why the Trained Dictionary is a Meaningful Extension

- **Model Hyperparameters Differ**

```
train_word2vec(
  train_file = "corpus.txt",
  output_file = "vectors.bin",
  vectors = 100,       # embedding dimension
  threads = 4,
  window = 5,          # context window
  iter = 3,            # number of passes
  min_count = 8,       # ignore rare words
  force = TRUE         # overwrite if file exists
)
```

- **Dictionary Construction Differ**

The NRC (National Research Council Canada Emotion Lexicon) is a list of English words and their associations with eight basic emotions *(anger, fear, anticipation, trust, surprise, sadness, joy, and disgust)* and two sentiments (negative and positive).

Size: ~14,000 English words(~1,000 to 2,000 for ANEW)

Extension - Trained Dictionary

Training Process:

Extract emotive and neutral seed words

```r
# Get emotive words (only keep those with value = 1 in emotion categories)
emotive_emotions <- c("anger", "fear", "joy", "disgust", "sadness", "trust", "surprise", "anticipation")

emotive_seeds <- nrc %>%
  filter(emotion %in% emotive_emotions, value == 1) %>%
  distinct(word) %>%
  pull(word)

# Get sentiment labels (positive/negative only)
sentiment_wide <- nrc %>%
  filter(emotion %in% c("positive", "negative")) %>%
  pivot_wider(names_from = emotion, values_from = value, values_fill = 0)

# Neutral words = words with both pos & neg == 0
neutral_seeds <- sentiment_wide %>%
  filter(positive == 0 & negative == 0) %>%
  pull(word)
```

Keep only words that exist in model

```r
emotive_seeds <- emotive_seeds[emotive_seeds %in% rownames(model)]
neutral_seeds <- neutral_seeds[neutral_seeds %in% rownames(model)]
```

Build vocabulary of 2,044 emotive and 2,044 neutral words with 4% threshold

Extension - Trained Dictionary

Why word embedding is an important approach for sentiment analysis

- **Capture Parliamentary-specific semantics**

Political language contains domain-specific usage (e.g., “war,” “freedom,” “justice”) that generic lexicons like ANEW or NRC may misrepresent. Word embeddings trained on the corpus help to capture how these words are actually used in parliamentary discourse.

- **Support Scalable and Data-Driven Scoring**

Word embeddings let us compute the emotive scores for all words in the corpus - even unseen ones and scale scoring across hundreds of thousands of speeches without manual labeling.

- **Enable Robust Extension and Comparison**

Demonstrate how small changes in embedding-based dictionary construction can affect the measurement of emotive rhetoric, potentially revealing new patterns missed by the original setup.

Result

Example Test: Emotive and Neutral Speeches

Table 1: Examples: Emotive and Neutral Speeches (Original vs. Extended Dictionary)

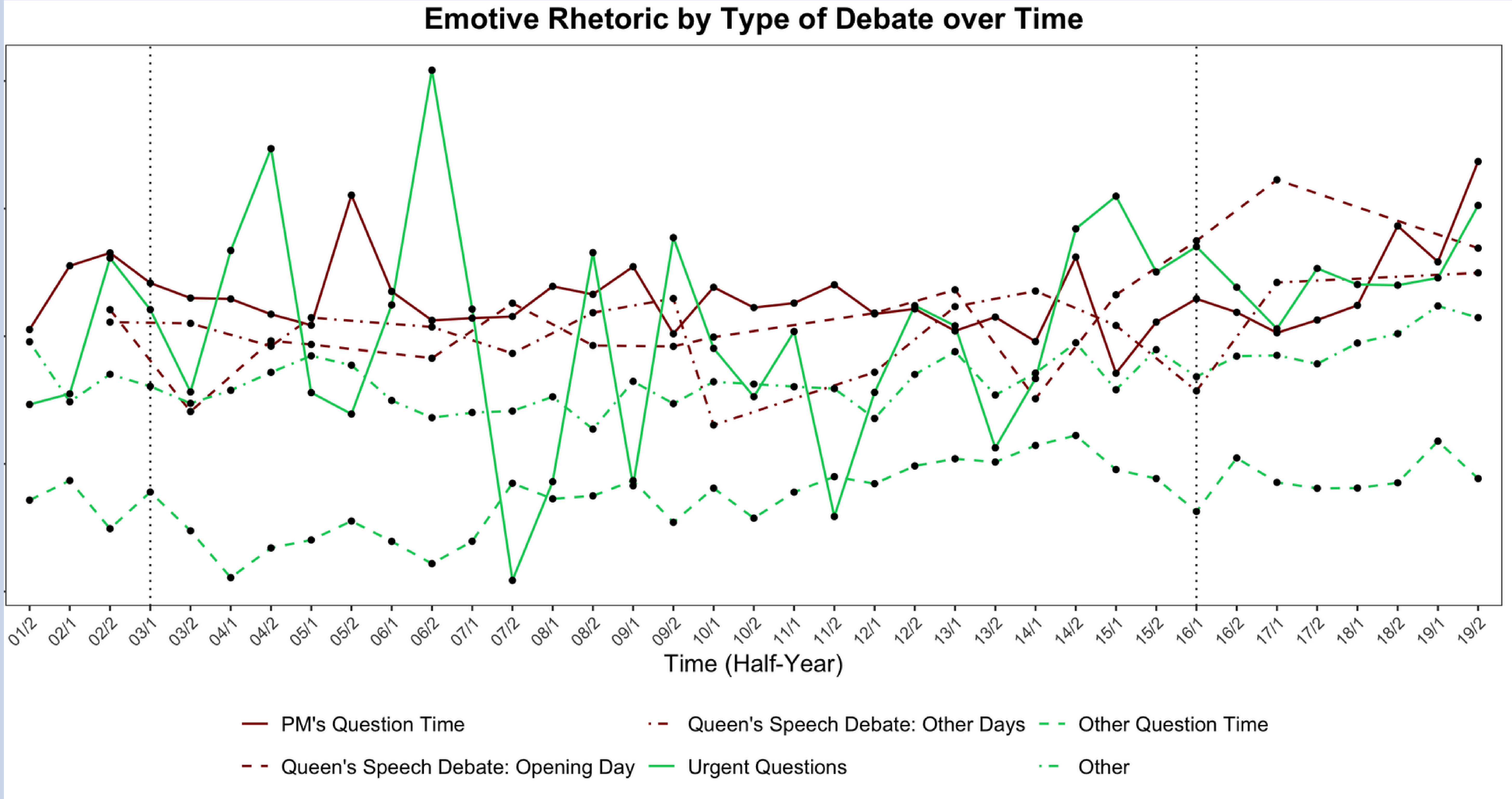
Original	Extended	Text	Speaker
43	20	Evil happens when good people stand by and do nothing. There is evil running through and infiltrating the Labour party, but it is full of good people and they are trying to do something about it. I commend them, appreciate them and have nothing but respect for them.	Alec Shelbrooke, MP
–25	–16.67	When used with old-fashioned copper wires, 10 megabits can become a lot less than that. We need a superfast fibre infrastructure instead of copper wires.	Geoffrey Clifton-Brown, MP

Note: Scores represent the emotive rhetoric score computed using either the original dictionary (as in Osnabruegge et al.) or the extended dictionary built with Word2Vec-trained embeddings on parliamentary speech.

Descriptive & technical?

Result

The development of emotive rhetioric in the period 2001 until 2019 on trained dictionary



- **Urgent Questions: Event-driven, unscheduled and topically diverse nature**

Result

Five regression models with different specifications [Depedent variable: emotive rhetoric score:

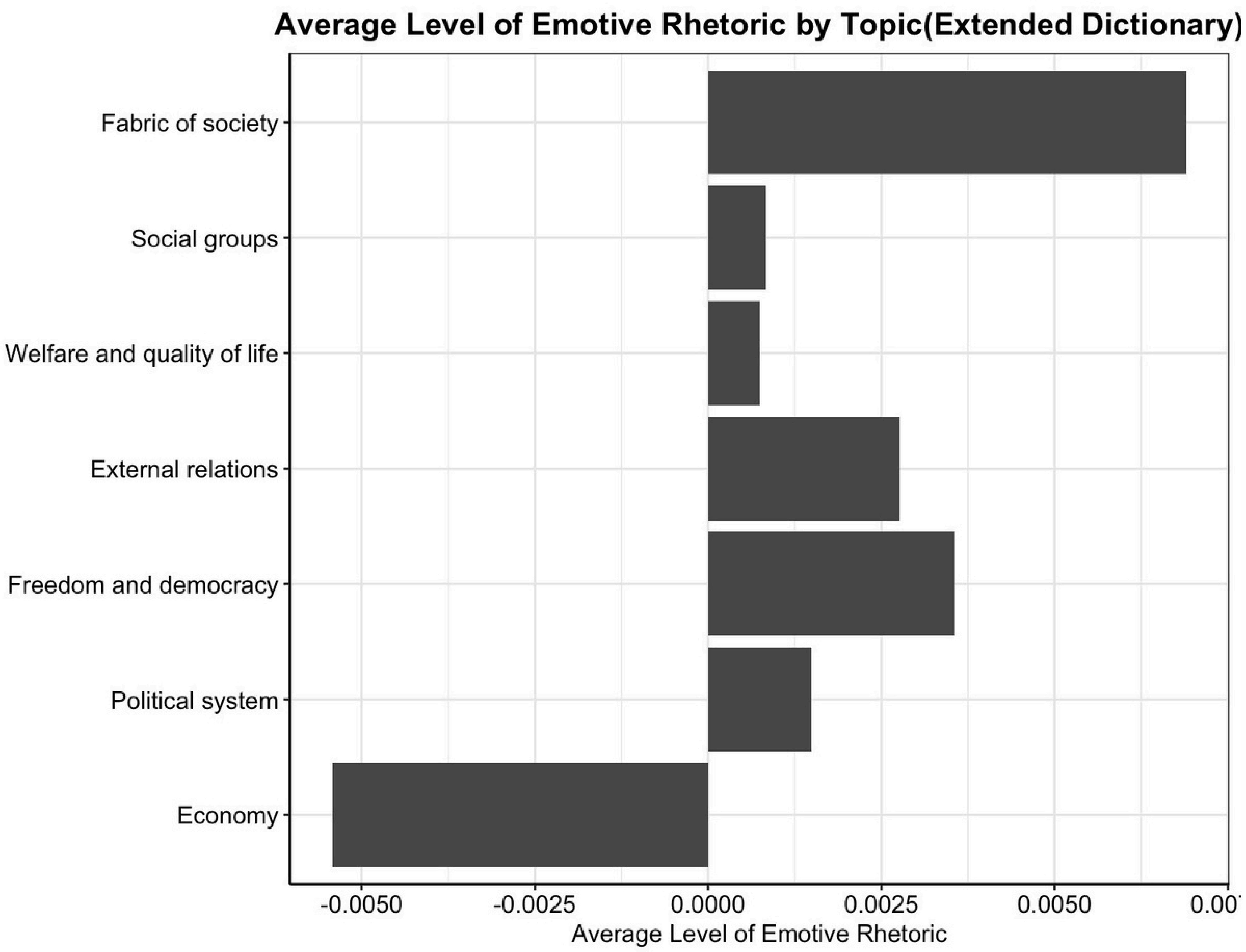
The larger the coefficient, the more emotive language MPs tend to use during that type of parliamentary event.

Table 1: Regression Analysis of Emotive Rhetoric (Replication)

	(1) Base Model	(2) + Time Trend	(3) + MP Fixed Effects	(4) + Controls	(5) + Weighted by Speech Length
PMQs	1.664***	1.647***	1.347***	1.634***	2.081***
	(0.00008)	(0.00008)	(0.00033)	(0.00009)	(0.00008)
Queen's Speech: Opening Day	1.224***	1.227***	0.915***	1.056***	1.404***
	(0.00026)	(0.00026)	(0.00023)	(0.00026)	(0.00014)
Queen's Speech: Other Days	0.947***	0.982***	0.800***	0.912***	1.226***
	(0.00012)	(0.00012)	(0.00013)	(0.00012)	(0.00007)
Ministerial Questions	-2.180***	-2.217***	-1.753***	-1.919***	-1.693***
	(0.00003)	(0.00003)	(0.00010)	(0.00003)	(0.00003)
Urgent Questions	1.651***	1.365***	1.541***	1.694***	2.155***
	(0.00007)	(0.00007)	(0.00016)	(0.00007)	(0.00007)
Constant	0.00172***	0.00111***	0.00003	0.00321***	0.00161***
	(0.00002)	(0.00003)	(0.00008)	(0.00005)	(0.00001)
Linear Time Trend		X			
MP Fixed Effects			X		
Controls				X	
Weighted by Speech Length					X
Observations	956,801	956,801	956,684	956,806	956,801
R^2	0.0074	0.0079	0.0404	0.0111	0.0053

Extension

The average level of emotive rhetoric by topic using extended dictionary



The differences suggest the dictionary might be more conservative in identifying emotion in certain social/policy domains — this could be due to how the seed words influence vector space expansion with bias.

Limitations

1.Dependence on Dictionary Quality

Even with a trained dictionary, the emotive/neutral classification still depends on the quality of seed words and assumptions about emotion categories and the full complexity is hard to be captured with word-level embeddings alone.

2. Context-Agnostic Scoring

Our approach scores speeches based on individual words, without considering sentence structure or context. This can misrepresent the speaker's intent or emotional tone.

"War is a terrible thing, but we must be calm and deliberate in our response."

Words like "war" and "terrible" may score as emotive, but the overall tone is actually measured, analytical and even restrained.

3. Static Embedding Limitation

Even though Word2Vec captures some semantic relationships, it uses static embeddings, meaning a word has the same vector regardless of context. For instance, *"charge"* in "he was charged emotionally" vs. "electric charge" means very different things, but the vector representation is identical.

Future Directions

1.Incorporate contextual language models

Use models like BERT to capture the context and tone of entire sentences, not just individual words.

2. Apply to other political settings

Test whether emotional rhetoric patterns are generalized or context-specific in environments like US Congress speeches or different time periods.

3. Link to outcomes

Investigate whether emotive rhetoric affects media coverage, public reactions or legislative success. Understanding impact moves beyond description to political consequences.

The End: Q&A