Parliamentary Speeches and Public Opinion: Unveiling Economic Sentiment through NLP Techniques

Rudi Narendran, Terinder Bhangal, Zac Payne-Thompson

Working Paper No. 005 August 2025

EMAP Working Paper Series

ECOM189: Big Data and Machine Learning Group Project: Group I

Parliamentary Speeches and Public Opinion: Unveiling Economic Sentiment through NLP Techniques

Rudi Narendran
te23022@qmul.ac.uk

Terinder Bhangal
te221371@qmul.ac.uk

Zac Payne-Thompson
te29003@qmul.ac.uk

ABSTRACT This study investigates the relationship between economic discussions in the UK Parliament and public economic sentiment from 2011 to 2025. Utilising natural language processing and machine learning techniques, we analysed over 897,000 parliamentary speeches to classify economic sentiment as positive, neutral, or negative. These classifications were then compared with weekly public opinion data from YouGov, focusing on the public's prioritisation of economic issues. Our findings reveal a minimal predictive relationship between parliamentary economic discourse and public economic concerns, suggesting that parliamentary debates may operate independently of public sentiment. This raises important questions about the channels through which political discourse influences public opinion and the role of parliamentary debate in modern democracy.

KEYWORDS Natural Language Processing (NLP); Machine Learning; Parliamentary Speech Analysis; Economic Sentiment Analysis; Public Opinion

Introduction

The relationship between political discourse and public opinion is fundamental to democratic governance, yet our understanding of how parliamentary debate shapes - or reflects - public sentiment remains limited. This paper examines the connection between economic discussions in the UK Parliament and public perception of economic issues, leveraging recent advances in natural language processing (NLP) and machine learning to analyse over 897,000 parliamentary speeches between 2011 and 2025.

The UK Parliament serves as both a forum for national debate and a key institution for economic policy-making. Through Hansard, the official record of parliamentary proceedings, we have unprecedented access to the complete text of parliamentary discussions. Simultaneously, regular public opinion polling provides a consistent measure of citizens' economic concerns. This combination of data sources creates a unique opportunity to examine whether parliamentary discourse predicts, follows, or operates independently from public economic sentiment.

Our research employs a novel methodological approach, combining supervised machine learning for sentiment classification with time series analysis to examine the predictive relationship between parliamentary speech and public opinion. We utilize the SetFit framework, a state-of-the-art approach for few-shot text classification, to categorize parliamentary speeches into positive, neutral, and negative economic sentiment. This classification is then analysed alongside YouGov's weekly survey data on the public's most important issues, focusing specifically on economic concerns.

The study makes several contributions to the existing literature. First, we develop and validate a robust framework for classifying economic sentiment in parliamentary speech, achieving 91% accuracy across sentiment categories. Second, we provide empirical evidence on the relationship between political discourse and public opinion, testing both contemporaneous and lagged effects. Finally, we offer methodological insights for analysing political text data at scale, demonstrating the effectiveness of modern NLP techniques in political science research.

Our findings challenge assumptions about the relationship between parliamentary discourse and public sentiment. Despite testing multiple specifications and employing sophisticated modelling approaches, we find minimal predictive power between parliamentary economic sentiment and public economic concerns. This surprising result raises important questions about the channels through which political discourse influences public opinion and the role of parliamentary debate in modern democracy.

This paper proceeds as follows: Section 2 reviews the relevant literature on political discourse analysis and public opinion formation. Section 3 describes our data sources and preliminary analysis. Section 4 details our methodological approach. Section 5 presents our results, and Section 6 discusses their implications and limitations. Section 7 concludes with recommendations for policy and future research.

Literature Review

The recent advent of Artificial Intelligence (AI) and Large Language Models (LLMs) has led to a resurgence in text-based analysis. However, this has been under-utilised in academic literature to analyse political discourse. This literature review examines the use of Hansard transcripts and public sentiment data sources to assess political narratives. The selected studies explore text mining, sentiment analysis and the application of statistical techniques to political speech. The discussion will highlight contributions of previous research and methodological challenges, particularly in alignment between politics and public sentiment. Prior to the ready availability of LLMs, text-based analysis was still widely used across academia and private research to assess political discourse. Studies have demonstrated how topic modelling, sentiment analysis and machine learning techniques can extract meaningful patterns from large text datasets, enabling a deeper understanding of ideological trends, policy priorities and rationale for decision-making (Van Atteveldt, Welbers and Van Der Velden, 2019). This "text as data" approach involves converting textual content into a structured format for quantitative analysis to address research questions, such as partisan divisions, political strategy and policy shifts.

A key enabler in text-based analysis on UK parliamentary discourse is the Hansard Corpus, complied by Alexander, Davies and Dalachy (2015) and accessible here. This dataset has been widely referenced, containing over 1.6 billion words and nearly every speech given in UK Parliament from 1803-2005. Additionally, the official Hansard website offer an application programme interface (API), with other organisations, such as TheyWorkForYou, leveraging and improving this, enabling greater access. Abercrombie and Batista-Navarro (2018) applied sentiment analysis to Hansard transcripts using supervised learning models to classify MPs speeches as positive (supportive) or negative (oppositional) towards legislative motions. This study emphasised the importance of metadata, textual and contextual features, demonstrating that that party alignment is, often, more influential than individual linguistic cues on sentiment polarity.

Whilst Hansard is an incredibly rich corpus for political analysis, it also presents major methodological challenges. Firstly, due to the age of the data source (with the first entries available from 1803), it is subject to changes in process. For instance in 1909, the move from third-person summaries to verbatim reports marked a shift in reporting styles and thus linguistic consistency (Hiltunen, Räikkönen and Tyrkkö, 2020). Their study revealed shifts such as increased use of colloquial conversational expressions e.g. "is going to" instead of "will", greater use of hedging phrases e.g., "I think it", and a decline in passive constructions and formal structures. Moreover, the structure around parliamentary debates naturally imposes constraints on linguistic expression. Rhetorical strategies, such as selective emphasis and political framing complicate sentiment analysis. Abercrombie and Batista-Navarro (2018) noted formalised debate structures, party discipline and intentional vagueness were causes for simple sentiment classification techniques to not be adequate for their Hansard analysis.

Statistics and the use of quantitative measures play a key role in political discourse, often serving as a persuasive device rather than objective facts. Lawson and Lovatt (2020) explored how statistics were used to frame arguments about the NHS winter crisis, identifying rhetorical tropes such as dehistoricisation (separation from history), synecdoche (term for part of something to refer to the whole or vice versa) and enthymeme (fallacy where premise is omitted). This study showed how MPs could shape narratives by preferentially selecting numerical representations, whilst omitting important context. Similarly, Lodge and Norderland (2017) examined the increased use of statistics in parliamentary debate, showing that these are often politicised rather than objective. Their findings align with broader societal concerns about "post-truth democracy" where numerical statements are used to lend credibility to subjective political arguments.

Data on public opinions provides a complimentary perspective, enabling insight into the alignment between public sentiment and political discourse. Studies have indicated that policy issues emphasised in Hansard often diverge from public priorities, raising questions around 'elite responsiveness' (Van Atteveldt, Welbers and Van Der Velden, 2019). Additionally, statistical models combining survey responses and textual data have been employed to assess predictive relationships between public option and political debate. However, these models faced challenges in question framing, sample biases and self-selection effects in survey participation (Lodge and Norderland, 2017).

Notable studies have looked at Climate Change and how this has evolved in recent political discourse. Willis (2017) analysed climate change debates in Hansard, finding that MPs over-emphasise economic and technological solutions over environmental and social factors. Similarly, this study also demonstrated the selective use of scientific evidence, omitting important findings to sustain favourable narratives. The Carbon Brief provided additional insights into which UK politicians discuss climate change the most, using Hansard and media appearances (Gabbatiss, 2019). This, simplistic, look at the frequency climate change is discussed and by whom is a valuable aspect to understand political and media influence on public opinion, underlining the value in metadata and contextual features.

The UK government has on-going initiatives to better harness parliamentary debate data, driven by the recent advancements in natural language processing (NLP), AI and LLMs. The Incubator for Artificial Intelligence (I.AI) is a government initiative and includes projects like Parlex and Lex, which leverage machine learning models to extract insights from legislative records (AI.gov.uk, 2025). These innovations may contribute to improved text classification and entity recognition and showcase the value in enhancing the efficient in political discourse analysis.

Our project contributes and adds value to the literature by integrating corpus analysis of Hansard with public opinion, using machine learning techniques and automated data processing.

Data

For this analysis we have selected two main data sources. Our primary data source is 'Hansard', the official source of data for all Parliamentary debates, publicly available here (Hansard, 2025). For public opinion data, we have used data made available through YouGov, a survey and polling company which conducts online surveys and publishes the outputs of some of these online.

Hansard – TheyWorkForYou API

Hansard is an incredibly rich source of text-data; however, the magnitude of the data presents challenges to its use. A Hansard corpus from 1803-2005 was created by Alexander, Davies and Dalachy (2015) containing nearly every speech from 1803 -2005, totalling over 1.6 billion words, as part of the Semantic Annotation and Mark-up for Enhancing Lexical Searches (SAMUELS) project funded by the UK Arts and Humanities Research Council (AHRC), demonstrating the value in this data asset.

Hansard provide a Representational State Transfer (REST) API for access, but development and maintenance of this connection has paused. TheyWorkForYou is an organisation run by the charity mySociety with the aim to make Parliament more accessible. They provide their own API to access Hansard, providing data from 1912 to today (with a day's lag).

Data compiled via the Hansard corpus has been "cleaned" and tagged with more extensive metadata than is available via the API (e.g., person speaking, type of debate etc), likely due to maintenance on the underlying Hansard REST API having been paused for some time (2018). Nonetheless, the underlying Hansard data can be used on its own to extract interesting insights. Figure 1 shows the number of mentions of the economy per million words combining the Hansard Corpus and TheyWorkforYou API, showing how this has varied over time with increased emphasis in the modern era.

The Hansard Corpus is a powerful tool which allows for simple string searches, viewing strings in context, comparing strings, collocation of strings and key words in context, producing both tabular and visual results. However, as this is limited to 2005 and an old interface that we are not able to connect to, our work uses the TheyWorkForYou API. For data quality and assurance processes, we have run comparison between the data to ensure consistency, Figure 1 is the product of one of these checks.

YouGov Survey – The most important issues facing the public

From Feb-2011 to Feb-2025, YouGov have conducted a near-weekly online survey of the most important issues facing the public, where respondents are asked to select up to 3 of their top issues from a set list (over time this list has ranged from 17-19 including "None of these" and "Don't Know" options).

The dataset for this survey is publicly available at this link, (YouGov, 2025b). The number of respondents each week varies, ranging from 971 – 226 over the entire dataset. However, as Figure 2 shows, the number of respondents is clustered around the mean, 1803, and median, 1749. YouGov outline their method of surveying here (YouGov, 2025a), and use "active sampling" which gives them control on the type of respondent to ensure a representative panel is being used to construct datasets.

For the most important issues survey, the following demographic data is collected, the granularity of this category available in the publicly available dataset is given in […] brackets.

- Age ([18-24], [25-49], [50-49], [65+])
- Gender ([Men], [Women])
- Politics ([Conservative], [Labour], [Liberal Democrat], [Remain], [Leave])
- Region ([London], [Rest of South], [Midlands], [North], [Scotland], [Wales])

- Social Grade ([A,B,C1], [C2,D,E] – note this aligns with Market Research Societies published social grade descriptors, used by the ONS for the census, here.

Each demographic corresponds to a 'tab' in the excel containing the publicly available YouGov survey data (in addition to an "All_Adults" tab), with a percentage against each issue. As such, it is not possible to accurately apply multiple filters to the data, a measure likely put in place to address GDPR and identification issues.

As part of the data exploration, we first assessed correlation between issues given that respondents can pick "up to three" issues. Figure 3 shows the correlation matrix for the Top 10 issues selected by respondents by mean average across the entire dataset. This shows there is relatively poor correlation between any issues, with most values close to 0. This may indicate that public opinion on issues is distinct from each other, such that if a certain opinion became the "most important issue" it may not necessarily increase another issue. Notably, there do seem to be some notable negative correlations. Inspecting values with an absolute value greater than 0.7, we can observe a negative correlation between "Britain leaving the EU" and "The economy" and between "The environment" and "Immigration & Asylum", this may indicate strong divergent opinions between groups.

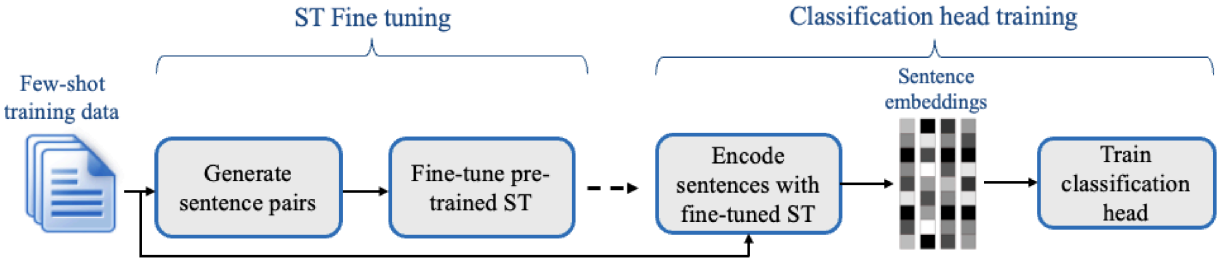
Exploring this further, we can assess how opinions diverge between groups on a specific topic. We have chosen to look at "the economy" as a consistently high-scoring top-issue and a focus for this report. Figure 4 shows that, between demographics, opinions on whether the economy is a top issue varies widely. Here, we look at the maximum average percentage point difference between a demographic and "All Adults" across the surveying period. This shows some expected results based on social tropes: Men, older age groups and higher social grade status, generally, view the economy as a more important issue than their counterparts in that demographic. But there are some unexpected results, all respondents with a party political affiliation view the economy as more of an issue than those without, respondents 65+ tend to view the economy as less of an issue (likely substituting this for Pensions) and regional differences are difficult to explain.

For the purposes of this analysis, we are going to focus on the data from "All Adults" as we are most interested in overall public opinion over time. Figure 5 shows the top 5 most important issues (highest mean averages across the dataset) plotted over time with a moving average trend line, in addition to a stacked area chart to show the proportions over time (note that the y-axis scale for each issue is not the same). This shows that there is great variance over time, which is expected as public sentiment is driven by current socio-economic climates and news of the day. Notably, we can see some "trade-off" between the top issues, notably drops in the economy and immigration & asylum as top issues balanced by the rise in "Britain leaving the EU". Additionally, the environment has increased sharply since 2018 but has seen a recent drop. Finally, immigration & asylum has seen a steady increase since 2020, expected given immigration statistics' coverage in the media. Other trends may be linked to major socio-political events, e.g., elections (drop in the economy post-2014 election), the COVID pandemic (sharp rise in health and the economy) or major events (Britain negotiating with the EU). In-depth analysis of this is out of scope for this report but are considered when interpreting results.

Methodology

This section outlines our empirical strategy for analysing the relationship between parliamentary economic discourse and public economic sentiment. We first present our text classification framework, followed by our econometric specification for time series analysis.

Text Classification Framework



We implement a two-stage classification approach following Tunstall et al. (2022). The first stage employs contrastive learning to fine-tune sentence embeddings, while the second stage trains a classification head on these embeddings.

Contrastive Learning Stage

Let $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{K}$ denote our labeled dataset, where x_i represents parliamentary speech text and $y_i \in \mathcal{Y}$ represents the economic sentiment label, with $\mathcal{Y} = \{positive, neutral, negative\}$. To train our model effectively, we generate structured sentence pairs:

1. **Positive pairs**: Sentences with the same sentiment category.
2. **Negative pairs**: Sentences with different sentiment categories.

By balancing these pairs across sentiment categories, we create a dataset that helps the model distinguish nuanced economic sentiment in speech.

Embedding Generation

To convert textual data into a numerical format, we use a sentence transformer that generates dense vector representations (embeddings) for each speech. These embeddings allow comparisons between speeches based on semantic similarity. The training process adjusts these embeddings to make similar sentiment pairs closer in space while pushing apart sentences with opposing sentiments.

The model optimises this process through a contrastive loss function, which penalises pairs that should be close but are too far apart and rewards pairs that maintain appropriate distances.

Classification Head

Once embeddings are generated, we apply a classification model to assign sentiment labels. This classifier estimates the probability that a given speech belongs to a particular sentiment category using learned representations. The final classification layer allows us to systematically categorise parliamentary economic discourse into positive, neutral, or negative sentiment classes.

Econometric Specification

Base Model

Our initial specification examines the contemporaneous relationship between parliamentary sentiment and public economic sentiment:

$$y_t = \alpha + \beta_p s_t^p + \beta_n s_t^n + \epsilon_t$$

where:

- y_t is the YouGov Economic Index at time t
- s_t^p is the proportion of positive economic statements
- s_t^n is the proportion of negative economic statements
- ϵ_t is the error term

Dynamic Specification

To account for potential temporal dynamics, we extend our base model to include lagged effects:

$$y_t = \alpha + \sum_{l=0}^{L} \beta_p^l s_{t-l}^p + \sum_{l=0}^{L} \beta_n^l s_{t-l}^n + \epsilon_t$$

where L represents the maximum lag length, empirically set to $\{7,\ 30,\ 90\}$ days to capture weekly, monthly, and quarterly effects.

Incorporating lagged effects allows us to capture the influence of past parliamentary economic statements on current public sentiment. This approach recognises that public reactions may not be immediate and can evolve over time. By empirically setting the maximum lag length to 7, 30, and 90 days, we aim to identify whether the impact of parliamentary discourse manifests on a weekly, monthly, or quarterly basis.

Feature Engineering

For our enhanced specifications, we construct the following feature sets:

Time-based Features

Let $\mathbf{x}_t$ denote our feature vector at time t:

$$\mathbf{x}_t = [m_t, d_t, y_t, s_t^p, s_t^n, s_t^{p*} m_t, s_t^{n*} m_t]^\top$$

where m_t, d_t, and y_t represent month, day-of-week, and year indicators respectively, and $s_t^{p*} m_t$ and $s_t^{n*} m_t$ represents sentiment-time interactions.

Rolling Statistics

For each sentiment measure, we compute:

$$\bar{s}_t^k(w) = \frac{1}{w} \sum_{j=0}^{w-1} s_{t-j}^k$$

$$\sigma_t^k(w) = \sqrt{\frac{1}{w-1} \sum_{j=0}^{w-1} (s_{t-j}^k - \bar{s}_t^k(w))^2}$$

where $k \in \{p, n\}$ and $w \in \{7, 30, 90\}$ days.

Model Selection and Validation

We employ R^2 scores and out-of-sample validation for model selection. Models are compared using both R^2 and RMSE (Root Mean Square Error). All models are estimated using a temporal train-test split, with the final 20% of chronologically ordered observations reserved for testing.

Results

Our analysis explored whether parliamentary speech sentiment could predict public opinion on economic matters. Key findings include:

- Parliamentary speech shows little predictive power for public economic sentiment (R^2 consistently near 0)
- More sophisticated modelling approaches, including time-lagged features and advanced feature engineering, failed to improve predictive power
- The relationship remains poor even when testing reverse causality

- The vast majority (88.35%) of parliamentary economic speech is neutral in tone
- These findings suggest parliamentary economic discourse operates independently of public sentiment

Data Quality and Initial Approach

Our analysis faced several methodological challenges stemming from the structure and accessibility of the parliamentary data. The Hansard API's limitations prevented filtering to specific parliamentary sessions such as Prime Minister's Questions (PMQs), requiring us to work with a broader corpus of parliamentary speech. This constraint potentially introduced noise into our analysis, as economic sentiment expressed during PMQs might differ systematically from that expressed in other parliamentary contexts.

Our initial approach employed topic modelling using a modified TF-IDF methodology, supplemented by large language models to assist in topic identification. However, this unsupervised learning approach failed to produce interpretable results that could be meaningfully correlated with public economic sentiment. This led us to pivot toward a supervised learning approach.

Supervised Learning Implementation

We implemented a SetFit classifier, a supervised learning approach that can achieve high accuracy with limited training data. This classifier was trained to categorise parliamentary speeches into three sentiment categories: positive, neutral, and negative economic sentiment. We used `BAAI/bge-small-en-v1.5` as our base sentence transformer, which was chosen for its efficient architecture while maintaining strong performance.

The model was trained on a curated set of examples that reflected typical parliamentary language patterns when discussing economic matters. One of the key advantages of SetFit in this context is its contrastive learning approach, which generates sentence pairs from limited training data to create a richer training signal. As demonstrated by Tunstall et al. (2022), this contrastive fine-tuning step allows SetFit to maximize the utility of small datasets by learning to distinguish between semantically similar and dissimilar examples within the same class.

Annex A.1 shows our classification evaluation metrics, which demonstrate strong performance across all three sentiment categories. The model achieved an overall accuracy of 91%, with particularly strong precision for neutral economic sentiment (1.00) and high recall for both positive and negative economic sentiments (0.93). This balanced performance across classes is notable given the inherent complexity of economic sentiment analysis in parliamentary speech, where nuanced language and technical terminology are common.

The confusion matrix in Annex A.2 reveals that misclassifications were minimal and primarily occurred between adjacent sentiment categories (neutral-negative and positive-negative), which is intuitive given the potential subtlety in economic sentiment expressions. This pattern aligns with the findings of Tunstall et al., who demonstrated SetFit's ability to make fine-grained distinctions in sentiment analysis tasks.

One of the key benefits of the SetFit methodology is its efficiency in both training and inference. Unlike larger language models that require extensive computational resources, our implementation achieves high accuracy while maintaining practical deployment requirements. This efficiency, combined with the strong performance metrics, makes the model particularly suitable for real-world applications in parliamentary speech analysis where both accuracy and computational efficiency are important considerations.

This allows us to achieve robust classification performance despite the limited size of our labelled dataset, demonstrating the effectiveness of SetFit's two-step approach: first fine-tuning the sentence transformer in a contrastive manner, followed by training a classification head on the resulting embeddings. The high F1-scores across all categories (ranging from 0.88 to 0.93) suggest that the model has successfully learned to capture the nuanced differences in economic sentiment expression in parliamentary language.

Table 1 shows the distribution of sentiment classifications across the parliamentary speech corpus and Figure 6 shows these series over time.

Table 1: Sentiment proportions across the full series of parliament text data

Sentiment Category	Count	Percentage
Neutral Economy	792,771	88.35%
Negative Economy	70,275	7.83%
Positive Economy	34,239	3.82%

This distribution aligns with expectations about parliamentary speech patterns, where formal and procedural language often dominates.

Time Series Analysis and Model Development

Our analysis progressed through increasingly sophisticated models to capture the relationship between parliamentary discourse and public economic sentiment. Beginning with our base model that examined contemporaneous relationships ($y_t = \alpha + \beta_p s_t^p + \beta_n s_t^n + \epsilon_t$), results showed minimal explanatory power, suggesting that immediate parliamentary sentiment had little direct impact on public economic perception.

The dynamic specification, which incorporated lagged effects up to 90 days, revealed some temporal patterns but with modest coefficient magnitudes. The most influential lagged coefficients were observed in the three-month positive sentiment ($\beta_p^{90} = -0.011$), one-week positive sentiment ($\beta_p^{7} = -0.0109$), and contemporaneous positive sentiment ($\beta_p^{0} = -0.0109$). However, these effects were not economically significant despite their statistical significance at conventional levels.

Our enhanced specification incorporating time-based features ($\mathbf{x}_t$) and rolling statistics ($\bar{s}_t^k(w)$ and $\sigma_t^k(w)$) actually led to deteriorating model performance. The inclusion of month-sentiment interactions ($s_t^{p*} m_t$) and various window lengths for rolling averages failed to improve the model's predictive power. R^2 scores remained consistently low across all specifications, and out-of-sample validation on the 20% test set confirmed poor generalisation performance.

Given these results, we conducted a reverse causality analysis to examine whether public economic sentiment might instead predict parliamentary discourse. This alternative specification yielded similarly poor results, suggesting the absence of a strong directional relationship in either temporal direction. All model performances are detailed in Table 2, consistently showing low explanatory power across our various specifications.

Table 2: Sentiment proportions across the full series of parliament text data

Approach	Features Used	Best Model	Best R^2	Best RMSE	Key Issues
Basic Sentiment	Daily proportions	Linear	0.0002	0.1746	No clear signal
Time-Lagged	7/30/90 day lags	Linear	-1.14	0.1098	No delayed effects
Time Interactions	Sentiment × time	Linear	-27.70	0.4021	Severe over-fitting
Feature Selection	Top 10 features	Lasso	-1.13	0.1096	No predictive value
Short-term Patterns	3/5/7-day rolling	Ridge	-1.34	0.1148	High volatility
Combined Approach	Selected + regularized	Lasso	-1.13	0.1096	No improvement
Reverse Causality	Sentiment categories	Random Forest	0.0002	0.1746	No predictive relationship

Our progression through increasingly sophisticated modelling approaches reveals several important insights. The lack of predictive power in basic sentiment analysis led us to explore time-lagged relationships, considering that public opinion might respond to parliamentary discourse with a delay. When time-lagged features proved insufficient, we developed more complex feature engineering approaches, hypothesising that

the relationship might be captured through interaction effects and rolling patterns. Even with careful feature selection and regularization to prevent over-fitting, no meaningful predictive relationship emerged. Our final attempt to understand the relationship by testing reverse causality—whether public opinion predicts parliamentary sentiment—also showed no significant connection.

These results consistently demonstrate that parliamentary economic sentiment, regardless of how it is measured or modelled, has minimal predictive power for public economic sentiment as measured by the YouGov index. The persistence of poor predictive performance across multiple modelling approaches and feature engineering strategies suggests this is a robust finding rather than an artefact of any particular analytical choice. This disconnect between parliamentary discourse and public sentiment raises important questions about the effectiveness of parliamentary debate in reflecting or influencing public economic concerns.

Discussion and Limitations

Limitations

Our analysis faces several methodological and data limitations that warrant consideration. First, the Hansard API's constraints prevented filtering to specific parliamentary sessions such as Prime Minister's Questions (PMQs), requiring us to work with a broader corpus of parliamentary speech. This potentially introduced noise into our analysis, as economic sentiment expressed during PMQs might differ systematically from that expressed in other parliamentary contexts.

The temporal resolution of our data sources also presents a limitation. While Hansard provides daily data, the YouGov survey is conducted weekly, requiring temporal aggregation that may mask short-term relationships between parliamentary discourse and public sentiment. Additionally, the survey's structure allowing respondents to select up to three issues may introduce measurement challenges, as the relative importance between selected issues cannot be determined.

Our sentiment classification approach, while achieving high accuracy metrics, relies on a simplified three-category framework (positive, neutral, negative) that may not capture the full complexity of economic discourse. Parliamentary discussions often contain mixed or nuanced sentiments that resist clear categorisation. The high proportion of neutral classifications (88.35%) might partly reflect this limitation rather than truly neutral content.

Finally, our analysis focuses solely on linguistic content without incorporating important contextual factors such as the speaker's role, party affiliation, or the broader political context. The impact of parliamentary discourse on public opinion likely depends on these factors, as well as on mediating influences such as media coverage and social media amplification, which our current analytical framework does not capture.

These limitations suggest opportunities for future research, particularly in developing more sophisticated classification schemes and incorporating additional contextual data sources. However, they do not fundamentally undermine our core finding regarding the weak predictive relationship between parliamentary economic discourse and public sentiment.

Policy Implications

This disconnect has several implications for MPs, policy makers and political commentators. Despite our analysis only looking at the issue of the economy in detail, the robustness of the finding means we have generalised this to all issues for the policy implications set out below.

McCombs and Shaw (1972) set out theories of "agenda-setting", wherein political institutions play a role in shaping public priorities. However, our findings indicate that parliamentary discussions may not significantly influence public perception. This may indicate that the topics MPs are choosing to raise are not aligned with the emerging concerns of their constituents and gives rationale to reassessing engagement with constituents. It should be noted, however, that the make-up of a constituency will not be the same as sample used in our analysis.

Additionally, this raises questions about the dominant channels through which political discourse may influence public opinion. For instance, alternative mechanisms, such as media framing of political issues may be more dominant and exert a stronger influence than parliamentary debates (Gibson, 1989). This view is supported by Krosnick and Kinder (1990), who introduced the concept of "issue-priming", suggesting that media coverage of issues may be more influential than direct parliamentary debate. Policy makers should be cautious that parliamentary discourse alone may not be sufficient to help address public concerns on topics.

Our results also suggest a potential gap between political discussion and public priorities. This may indicate systemic inefficiencies in the process, namely that legislative policy discussions are not addressing key public concerns. Research suggests that governments tend to be more responsive to persistent issues over time (Wlezien, 1995). Therefore, policy makers may need to ensure use of a more diverse range of public opinion data to capture emerging issues and ensure alignment between government actions and public concerns.

The data processing and pipeline developed through our analysis could be an asset to policy makers to address inefficiencies. Officials contacted for this research indicated that understanding public opinion and debates in parliament are largely manual processes. There are resource savings that could be made by leveraging data processing pipeline as we have developed here. This value can be justified through the Parlex and Lex project being completed by I.AI (AI.gov.uk, 2025).

Further research is needed to better understand the links between political discourse and public opinion. Our analysis could be extended to explore the alternative channels mentioned, leveraging more advanced processing techniques. This could include media coverage and social media trends (through transcription and web-scraping) to see if these provide stronger predictive signals to public priorities. Additionally, network or sentiment analysis of Hansard may yield deeper insights into how issues are framed and how narratives change around issues.

Finally, this analysis has focused on looking at overall public opinion but could be extended to understand how close or far topics discussed in parliament are to different demographics. For example, the YouGov survey shows the economy tends to be a stronger issue for men rather than women, so the results indicate that the frequency this is discussed is more aligned to men than women.

Conclusions

Whilst our analysis did not indicate that text data from Hansard was a good predictor of public priorities, this does not necessarily mean political discourse is irrelevant to public opinion. Instead, it may indicate that it is not the dominant channel through which the public form their views and the complexity of this process. It does highlight the need for policy makers to utilise multiple channels to attempt to predict emerging priorities and provides rationale to review public engagement within government. Ensuring a more representative and responsive democratic system may require greater integration of real-time data from various sources that can help to feed into the policy making process.

References

- Abercrombie, G. and Batista-Navarro, R. (2018) "Aye' or 'No'? Speech-level Sentiment Analysis of Hansard UK Parliamentary Debate Transcripts', in N. Calzolari et al. (eds) *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. LREC 2018, Miyazaki, Japan: European Language Resources Association (ELRA). Available at: https://aclanthology.org/L18-1659/ (Accessed: 22 February 2025).
- AI.gov.uk (2025) *Parlex and Lex, Incubator for Artificial Intelligence - GOV.UK*. Available at: https://ai.gov.uk/projects/parlex-and-lex (Accessed: 22 February 2025).
- Alexander, M., Davies, M. and Dalachy, F. (2015) *English-Corpora: Hansard*. Available at: https://www.english-corpora.org/hansard/ (Accessed: 22 February 2025).
- Gabbatiss, J. (2019) 'Analysis: The UK politicians who talk the most about climate change', *Carbon Brief*, 11 September. Available at: https://www.carbonbrief.org/analysis-the-uk-politicians-who-talk-the-most-about-climate-change/ (Accessed: 22 February 2025).
- Gibson, W. (1989) 'Review of *News That Matters: Television and American Opinion.*', *American Journal of Sociology*, 94(4), pp. 882–884.
- Hansard, U. gov (2025) *Hansard - UK Parliament*. Available at: https://hansard.parliament.uk/ (Accessed: 22 February 2025).
- Hiltunen, T., Räikkönen, J. and Tyrkkö, J. (2020) 'Investigating colloquialization in the British parliamentary record in the late 19th and early 20th century', *Language Sciences*, 79, p. 101270. Available at: https://doi.org/10.1016/j.langsci.2020.101270.
- Krosnick, J.A. and Kinder, D.R. (1990) 'Altering the Foundations of Support for the President Through Priming', *The American Political Science Review*, 84(2), pp. 497–512. Available at: https://doi.org/10.2307/1963531.
- Lawson, B. and Lovatt, M. (2020) 'Towards a rhetorical understanding of statistics in politics: Quantifying the National Health Service "Winter Crisis" ', *European Journal of Communication*, 36(2), pp. 110–124. Available at: https://doi.org/10.1177/0267323120966842.
- Lodge, M. and Norderland, M. (2017) *The use of statistics and facts in contemporary political discourse*. Produced for the UK Statistics Authority. *The Centre for Analysis of Risk and Regulation (carr) [LSE]*. Available at: https://www.lse.ac.uk/accounting/assets/Documents/news/carr-report-for-UKSA-final.pdf.
- McCombs, M.E. and Shaw, D.L. (1972) 'The Agenda-Setting Function of Mass Media', *The Public Opinion Quarterly*, 36(2), pp. 176–187.
- Tunstall, L. et al. (2022) 'Efficient Few-Shot Learning Without Prompts'. *arXiv*. Available at: https://doi.org/10.48550/arXiv.2209.11055.
- Van Atteveldt, W., Welbers, K. and Van Der Velden, M. (2019) 'Studying Political Decision Making With Automatic Text Analysis', in *Oxford Research Encyclopedia of Politics*. Oxford University Press, pp. 1–11. Available at: https://doi.org/10.1093/acrefore/9780190228637.013.957.
- Willis, R. (2017) 'Taming the Climate? Corpus analysis of politicians' speech on climate change', *Environmental Politics*, 26(2), pp. 212–231. Available at: https://doi.org/10.1080/09644016.2016.1274504.
- Wlezien, C. (1995) 'The Public as Thermostat: Dynamics of Preferences for Spending', *American Journal of Political Science*, 39(4), pp. 981–1000. Available at: https://doi.org/10.2307/2111666.

- YouGov (2025a) *Methodology | YouGov*. Available at: https://yougov.co.uk/about/panel-methodology (Accessed: 22 February 2025).
- YouGov (2025b) *The most important issues facing the country*. Available at: https://yougov.co.uk/topics/society/trackers/the-most-important-issues-facing-the-country (Accessed: 22 February 2025).

Figures

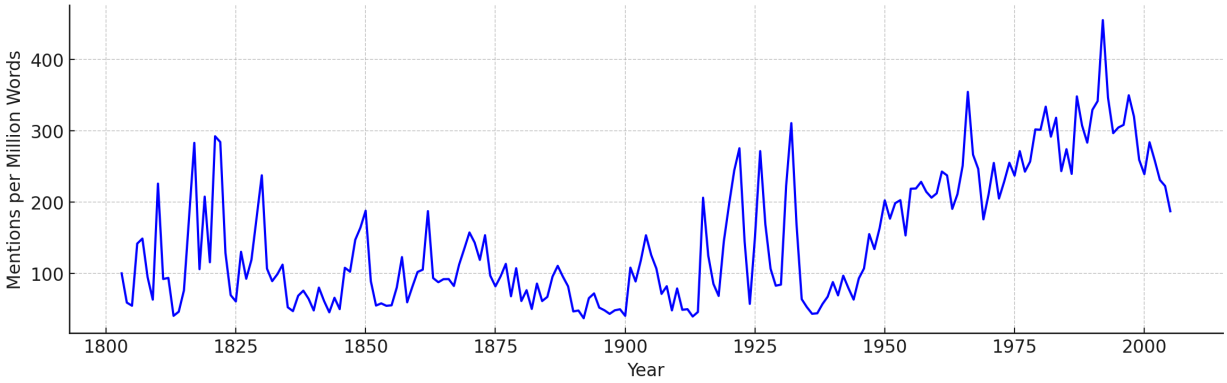


Figure 1: Mentions of "Economy" per million words

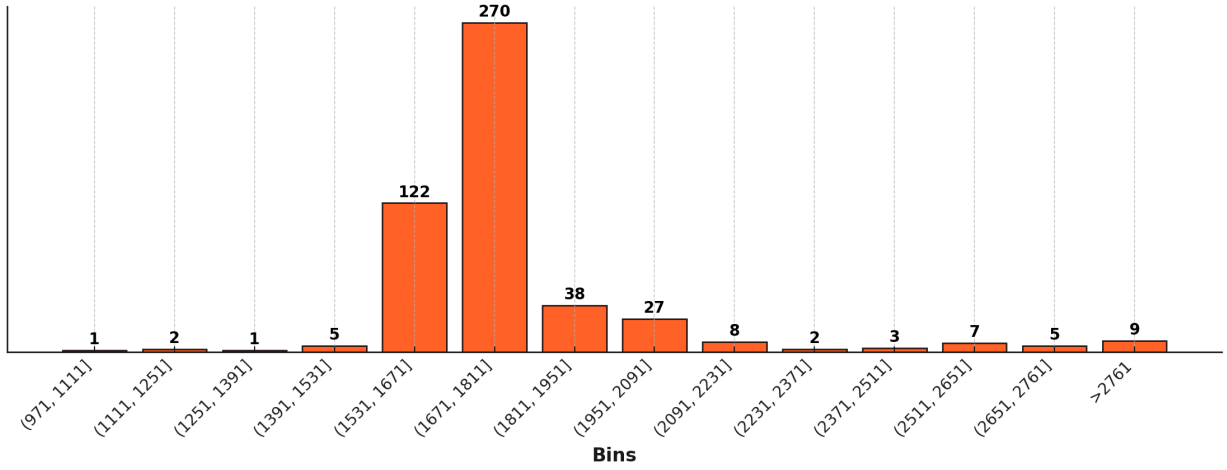


Figure 2: YouGov - The Most Important Issues Weekly Survey Respondent Number Histogram

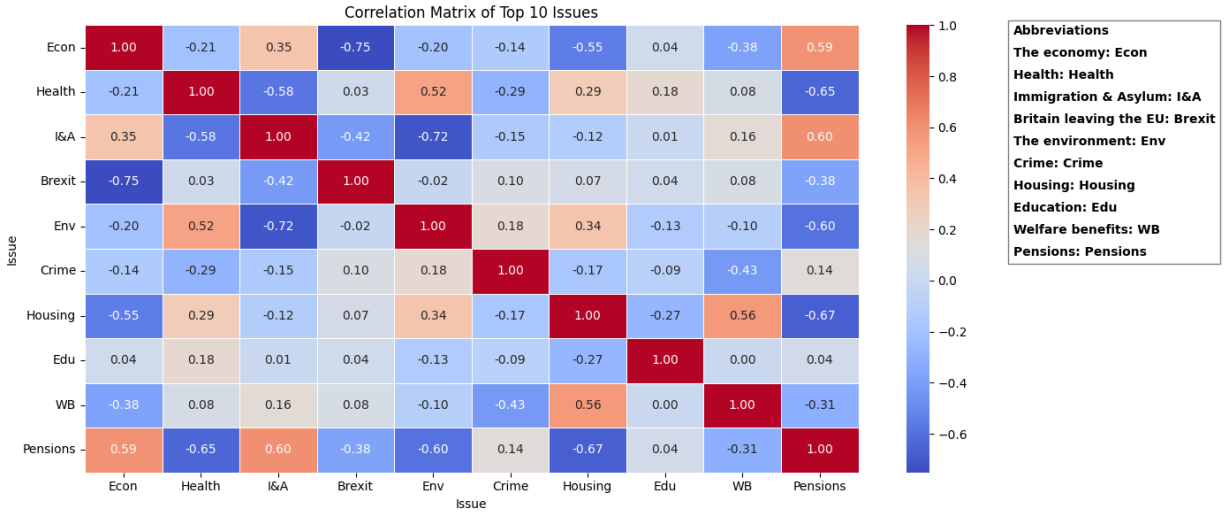


Figure 3: Correlation Matrix for Top 10 Issues for All Adults

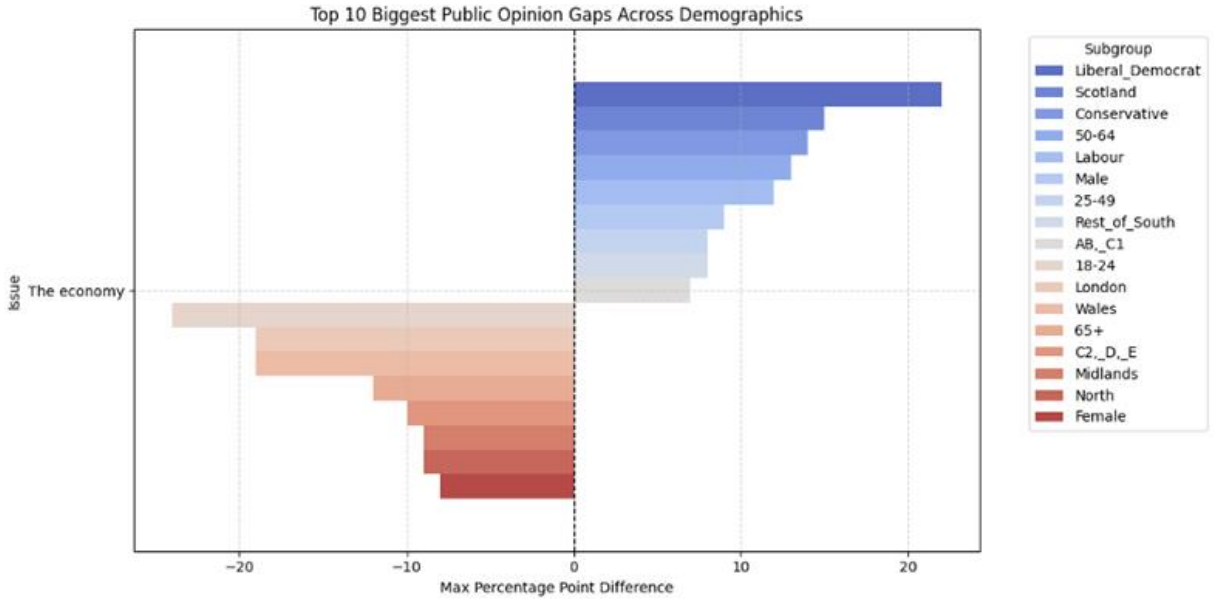


Figure 4: YouGov - Top 5 Most Important Issues Over time

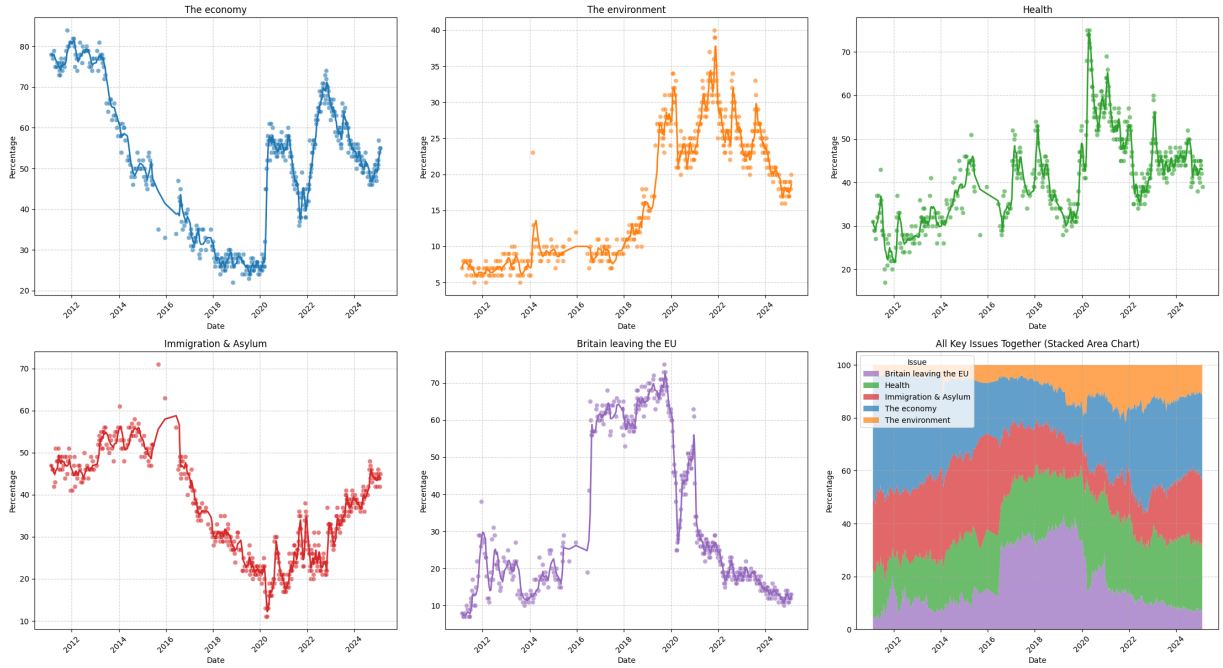


Figure 5: Maximum difference between demographic and All Adults Baseline for "The economy"

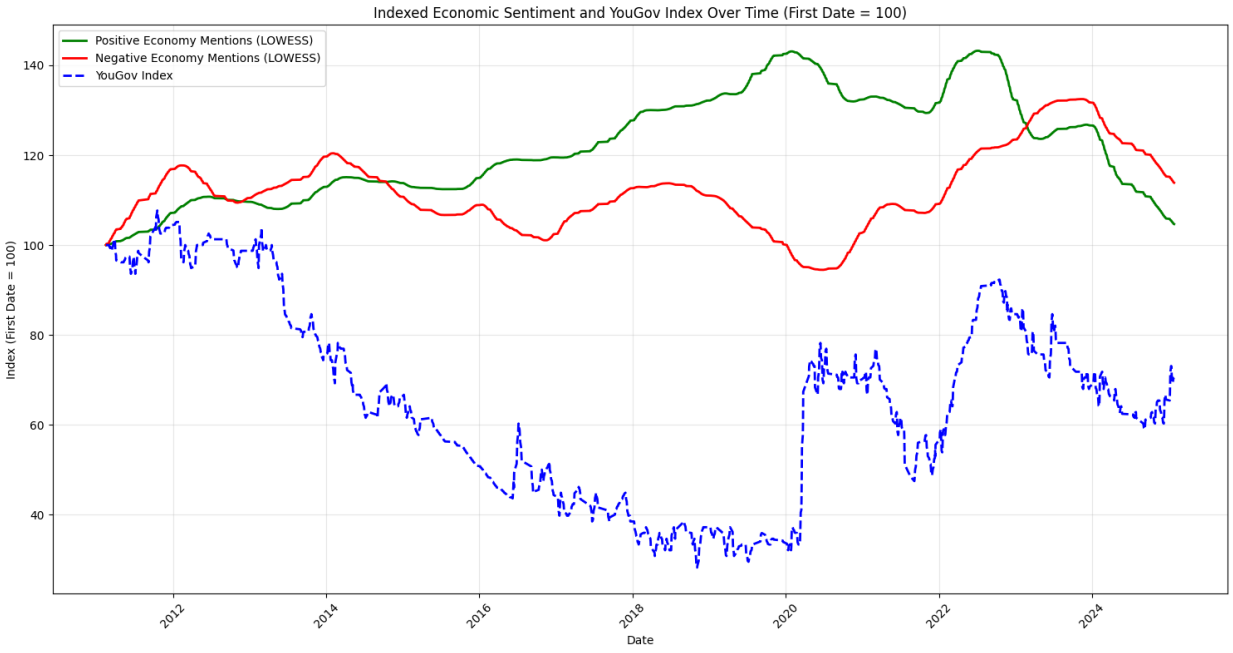


Figure 6: Time series of Parliament vs YouGov sentiment

Annexes

Annex A

A.1

Table 3: Classification Metrics

Class	Precision	Recall	F1-Score	Support
positive_economy	0.93	0.93	0.93	15
neutral_economy	1.00	0.87	0.93	15
negative_economy	0.82	0.93	0.88	15
Accuracy			**0.91**	45

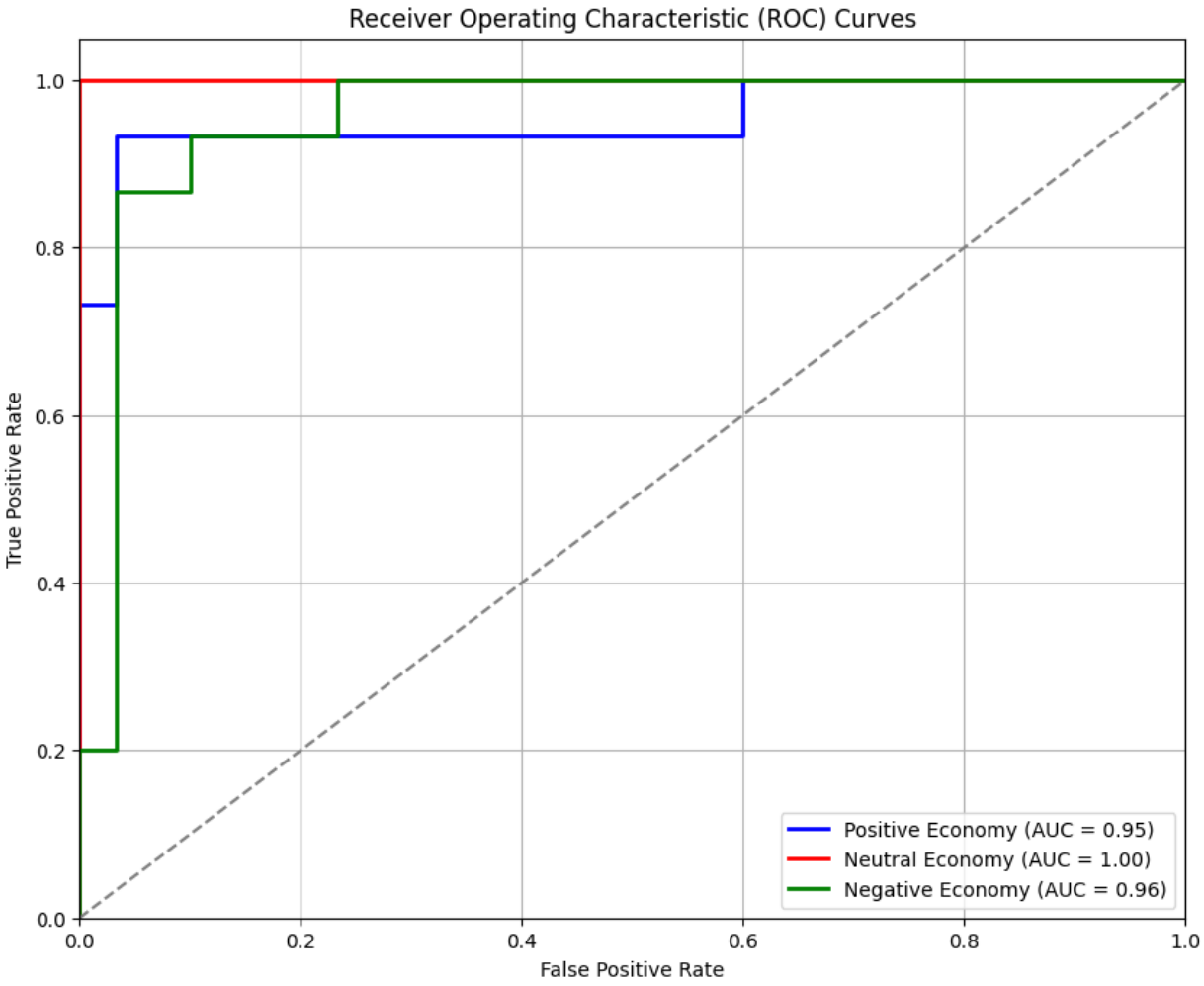
A.2 Confusion Matrix

The table below represents the confusion matrix of the trained classifier, displaying the number of correctly and incorrectly classified instances.

Actual / Predicted	**Positive_Economy**	**Neutral_Economy**	**Negative_Economy**
Positive_Economy	**14**	0	1
Neutral_Economy	0	**13**	2
Negative_Economy	1	0	**14**

- The diagonal elements (**bold values**) represent the correctly classified instances for each class.
- The off-diagonal elements indicate misclassifications.
- The model performed well in classifying "positive_economy" (93.3% accuracy in this class) and "negative_economy" (93.3%), with some confusion in distinguishing "neutral_economy."

A.3 ROC Curves



School of Economics and Finance

Queen Mary University of London

This working paper is based on project work undertaken by EMAP apprentices

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/