Bearish or Bullish? A Predictive Sentiment-Weighted Attention Analysis of the Wall Street Journal

By MATTHEW SENICK AND SAMUEL SCHWARTZBEIN

I. Context and Motivation

In their paper, *Business News and Business Cycles*, (Bybee et al., 2021) Bybee et al. are interested in exploring the intuition behind the predictive relationship between time-series topic attention in the *Wall Street Journal* and key market index indicators. The authors are mainly concerned with finding the amount of explainability that relative topic developments in popular financial literature has on future market performance.

Our extension of Bybee et al.'s investigation dives deeper into exploring this relationship. In the following analysis, further predictive precision regarding general market fluctuations is sought, namely by considering how article sentiment weighting may explain future market performance. Specifically, we seek the effect of article-by-article sentiment weighted topic attention over time and the effect of fluctuation in this metric.

By introducing sentiment as a weight on topic attention, a more descriptive relationship is presented between attention and key market fluctuation indicators. In essence, we believe sentiment to be an important consideration for future directional change of market fluctuations.

II. Contribution/Difference to the Paper

As mentioned, our main contribution is the development of an interpretable sentiment scaling methodology for relative time-series attention. Intuitively, in our adjusted model, sentiment weights give the directional movement of attention, and attention presents the absolute importance of a topic given articles over time.

Another large contribution is that our analysis presents a deeper view of out-of-sample and testing performance when considering topic importance ranking via lasso. Out-of-sample Mean Squared Error (MSE) and R^2 were considered in our ranking of the most informative topics on future behaviour of key market fluctuation indicators. This differs from Bybee et al., who only presented out-of-sample R^2 for a few variables

Smaller variations from the paper also include: ranking sentiment metrics using sentiment transformer output, aggregating daily topic attention by probability of topics rather than by one-hot topic assignment, checking for sentiment/topic bias amongst prolific authors, further exploring the hyperparameter space of the underlying Latent Direchlet Attention (LDA) model, adjusting LDA selection criteria to focus on the coherence/perplexity tradeoff rather than using Bayes Factor, considering more varied article subjects, and leaving out the final Vector Autoregression (VAR) model.

III. Methodology

To begin our analysis, various LDA models were developed using a web-scraped corpus of approximately 80,000 tokenized articles from the *Wall Street Journal*, which were then ranked on both coherence and perplexity scores which is shown in *Figure A1*. Using this figure, we selected the LDA model with 190 topics.[1] This was done so as to choose a locally maximal coherence score while also minimizing perplexity score. We also felt that this was holistically the best model, as it provided many well-segmented, interpretable topics. As a note, different hyperparameters were utilized in the building of this LDA model, specifically, automatically optimized a-priori beliefs on the document-topic and topic-word distributions rather than uniform a-priori beliefs, and 8 passes were done over the corpus rather than 300 passes.

Using the chosen LDA model, relative attention per topic in each month was calculated.[2] For each day, the probability of all articles being in a given topic were aggregated and normalized. These daily values were then aggregated

[1] This differs slightly from Bybee et al.'s 180 topic model.
[2] This was derived using *Equation C1*

by month to produce a time-series of month-by-month attention vectors containing all topics in the model. An example of the output daily aggregation of topic attention is given in *Figure A2* for the topic of "Financial Crisis". As seen, during the 2007 Financial Crisis, there is a very evident increase in relative daily topic weighting.

Adapting this building process for attention, a novel addition is proposed to weight the topic attention of any given article by the predicted sentiment of the article.[3] Given an article, a sentiment transformer is run on the article to predict the sentiment of the text (negative, neutral, positive), which is then mapped to a ternary variable (-1, 0, 1 respectively) which acts as a scaling factor to article attention. So, negative articles will have a directionally negative impact on topic attention and positive articles will have a directionally positive impact on topic attention. Days with conflicting or neutral article sentiment are then considered to be smaller absolute value of attention. This process aids in fully understanding the directional impact of sentiment on attention. An example of this time series sentiment-weighted attention is provided in *Figure A3*, again using the "financial crisis" topic. As shown, the negative nature of this topic appears to be much better preserved.

To predict article-level sentiment, we utilized the RoBERTa sentiment transformer (Heitmann et al., 2020). A transformer is essentially a deep neural network that achieves generally accurate results on a number of NLP tasks. This particular pre-trained sentiment transformer gives a binary positive or negative classification for a given text.

News articles are generally written such that the first paragraph summarizes the article. Thus, sentence-level sentiment of the first five sentences should have a great deal of predictive power for the rest of the article's sentiment. Therefore, we use the transformer to predict the sentiment of each of the first five sentences of each article. We then classified each respective article's overall sentiment as the majority vote of these predictions. Ties and erroneous articles were labelled as neutral. This method was chosen for the rest of this analysis as it achieved better and more representative results than other methods such as predicting sentiment from article headlines, while considering computational limitations.

Finally, having both sentiment-weighted and unweighted topic attention, Lasso regression is employed to rank the importance of topics when considering many different market fluctuation indicators as response variables, such as market volatility, policy uncertainty, taxation uncertainty, etc. which were analyzed by Bybee et al. Our methodology in preparing the Lasso regression remains identical to Bybee et al. except for the fact that we generate an 80/20 train/test split on the aggregated monthly attention data. An 80% split was chosen since we are working in a low-data environment when using month-level attention data.

Building on this data split, relative MSE values were considered alongside R^2 values when ranking the performance of all considered models. Following Bybee et al., data was rescaled such that a 1 standard deviation change in x would cause a 1 standard devation change in y. Additionally, the maximum number of topics was tuned via grid search. In the paper, a maximum of the 5 most informative topics were selected from the Lasso regression. We followed the 5 topic maximum in the replication section. However we considered models which selected 10, 15, and 35 variables in our extension.

In the end, we estimate our 190 topic fitted model using 5 and 10 variables when only considering sentiment-weighted topic attention, and fitted using 15 and 35 topic variables for the combination of attention and sentiment-weighted attention variables. These considerations were made to balance interpretability with performance. Lastly, our top fitted models were tested against the out-of-sample dataset, which was split chronologically.

IV. Data

Article data for all analysis contained within this paper was exclusively sourced from the *Wall Street Journal*. To obtain this textual data, a multi-staged web scraper was built. In the first stage, a day-by-day article directory was identified, from which day-level article links were scraped, obtaining nearly 150,000 article links in total. In the second stage, text from $min(\text{total articles for the day}, 40)$ articles were scraped per day. However, due to lim-

[3]Given by *Equation C2*

itations such as a paywall/login screen, anti-crawler CAPTCHAs and text availability, just under 100,000 article texts were able to be successfully scraped.[4] In addition to scraping article text, the title, author, and date of each article was also collected.

Before following through with the rest of the LDA and time-series attention analysis, authorship bias was considered. A Naive Bayesian classification model was built such that it would search for a specific number of the most prolific authors in the database, after which a supervised learning problem is defined to attempt classifying each of these presented authors. As shown in *Figure A4*, the model is extremely effective at predicting the authorship of a given article. This may raise concerns with the bias of topics and sentiments in the suggested analyses. Most informative words are also identified, usually leading to suggest that prolific authors have a large amount of classifiability within their own topic.

We then cleaned the data according to the steps laid out in Internet Appendix A of Bybee et al.. We adhere precisely except for the following alterations: We keep all 2017 dates rather than the removing articles after June of that year as our corpus is already far smaller. Although best effort was put forth to filter out topics irrelevant to this research, some extraneous articles made it into our final corpus. Additionally, the original paper only used articles in the WSJ's three main sections, whereas we include articles outside of these sections. This can be attributed to the original authors having access to the WSJ's article metadata. The authors concatenated articles with the same accession number as these are chained articles. Since we do not have these numbers, this was not done in our analysis. Also, the authors performed custom lemmatization whereas we used SpaCy's lemmatizer. This choice was made due to the high degree of similarity between the two lemmatization rules and computational constraints. A threshold of 200 for making bigrams was selected as Bybee et al. does not provide guidance on this. Notably, we follow Bybee et al. in removing all words that appear in less than 0.1% of articles as well as words that are less than three characters.

All things considered, 79,910 articles fit the criteria to make it past all cleaning steps and be passed into the LDA models. Of course, as we are only considering a sample of at most 40 articles per day, the relevance of chosen articles cannot be guaranteed. However, a small amount of filtering based on article titles was performed. Unfortunately, due to data and processing limitation, articles from 2009, the first third of 2014, and 2015 were unable to be scraped. However, a representative sample for all other days from January 1st, 2006 until December 31st, 2017 was collected.

During all stages of the analysis, data was ordered and aggregated chronologically. Additionally, in the Lasso regression, a training/test split was decided upon chronologically. All of these considerations were made in order to preserve the time-series and predictive nature of the collected data.

Finally, we estimate our models on a selection of response variables included in Bybee et al.. First, CBOE Volatility Index was sourced as a measure of market volatility. Second, total employment and industrial production was gathered from FRED. We then converted total employment and industrial production into growth rates, following the original paper. Also, market volatility was aggregated to its mean monthly value to follow the rest of our data. Third, we follow the original paper in estimating our model on Baker, Bloom and Davis (2016)'s economic policy uncertainty indices (BBD). These specific variables were chosen as they are easily accessible.

V. Results and Comparison to the Discussed Paper

Replication

Our topic attention model achieved similar in-sample R^2 to Bybee et al. for many variables utilized by the authors. By using R^2 as the primary performance measure of the model. Market Volatility, Industrial Production Growth from FRED and Financial Regulation, Taxes, Health Care and Government Spending from BBD all achieved in-sample R^2 within 5 percentage points of the original paper.[5] However, the original model achieved higher

[4] This scraped dataset is available upon request

[5] The regression results are shown in *Table B1* in the appendix

in-sample R^2 on average. An example of the in-sample replication for market volatility is given in *Figure A5*.

Only FRED variables were presented with out-of-sample estimates in the original paper. Our out-of-sample estimate of Market Volatility was 0.36, which is 0.08 below the original estimate. Lasso regressions for Industrial Production Growth and Employment Growth were not predictive out-of-sample. This points to the advantage of the original model's larger corpus. For example, their most predictive variables for employment growth included "Iraq", and "Clinton". Both Bill Clinton and the Iraq war preceded the articles in our data.

There are mixed results for our most predictive variables in the replication. As seen in *Tables B2-6*, topic attention to "government budget" had the most predictive influence as measured by the size of its coefficient in four of our five best models. Also, "financial crisis" was highly predictive of market volatility. Thus, our most predictive variable for each variable mirrored Bybee et al. These most predictive variables in our replication models are possibly correlated with the dependent variable. For instance, attention to war was associated with tax uncertainty and elections with taxes and health care. However, our model included more extraneous variables than the original paper. Some variables such as sports entertainment in the market volatility regression would have been omitted with a slightly harsher regularization penalty. Thus, our model captures the main themes of topics, but does not perform as well as the original.

<u>Extension</u>

Our methodology for sentiment weighted topic attention achieved strong results. *Figure A6* shows the mean sentiment weighted topic attention for each topic. Polarized topics in this graph have both consistent sentiment and a high degree of attention. We see that the metric does an excellent job ordering topics by their sentiment attention. The most negative topics include criminal investigations, police shootings, financial crisis, and product recalls. The most positive topics include positive business culture, new technology, and economic growth. We can also get a sense of WSJ's bias, as it has favourable sentiment towards real estate development, mergers and acquisitions and corporate executives.

The figure also provides evidence that topic attention is effectively incorporated. "Terrorism" is the 10th most negative sentiment weighted attention topic. This topic has a less negative ranking than other topics such as "criminal investigations" because terrorism occurs relatively infrequently. Therefore, even though terrorism covers more negative events than the criminal investigation, it is ranked lower because terrorism receives far less attention on average.[6][7]

The Lasso regression fit with our sentiment weighted topic attention data on market volatility with 9 features performed better out-of-sample than any of Bybee et al.'s regressions, achieving an R^2 of 0.572 on the test sample. *Table B8* in the appendix shows the regression table for this model. *Figure A7* and *Figure A8* visualize the in-sample and out-of-sample fits respectively. The selected variables in this regression show a number of interesting relationships. For instance, an association that markets become less volatile after quarterly earnings reports are released. However, it is difficult to interpret coefficients in general with this metric because any topic can be positive or negative even if it receives more attention. Since quarterly earnings reports are one of the top 3 most positive topics, it is likely that this coefficient suggests that volatility decreases when major earnings reports are released.

Aside from market volatility, the training R^2 for this model are on average higher than then our paper's attention model and similar to the original paper. However, there are more variables included in this model, which may lead to overfitting the training set. We also see that the out-of-sample R^2 is generally worse than in our replication model outside of the market volatility result.

Our final model combines topic attention estimates from our replication datasets and the

[6] The criminal investigations topic received the 4th highest average topic attention, whereas terrorism ranked 73rd

[7] It is possible that articles were misclassified due to key sentimental words. However, having reviewed its classifications manually, it appears to classify article sentiment using the majority vote decision rule with a high degree of accuracy. *Figure A11* provides an example of the transformer's predictions of an article.

sentiment weighted topic attention. We estimate models with 15 and 35 features each respectively. In doing so, verify whether we could achieve better out-of-sample predictions by combining pure topic attention alongside sentiment-weighted attention predictions. *Table B10* shows R^2 and MSE for each of the response variables estimated by the Lasso selection of the top 15 variables. An example of in-sample and out-of-sample fits for market volatility are shown in *Figure A9* and *Figure A10* respectively.

We see that with the 15 variable specification in *Table B10*, training R^2 and MSE are generally better than the other two estimations. Meanwhile, none of the test R^2 or MSE values are significantly better than the other models. It is notable that both sentiment and attention variables are chosen by Lasso. Sentiment weighted variables continue to have predictive power even when selected alongside non-sentiment weighted variables of the same topic. This shows that for some regressions it is better to weight by sentiment whereas for some pure topic attention is sufficient. As we increase the number of variables, the test scores increase for Taxes, Government Spending and Entitlement Programs. The test R^2 is optimized at a 35-feature specification. In the end, it is found that out-of-sample R^2 is maximized and MSE is minimized for all dependent variables with the combination of pure attention and sentiment weighted attention variables, except for market volatility, which is best predicted with only sentiment weighted variables.

VI. Conclusion

By using a simple aggregated sentiment-weighted time-series attention model to predict key market fluctuation indicators, we have found an interpretable yet effective way to combine topic attention and directional sentiment in the context of future-looking market prediction. As shown in the analysis, using sentiment-weighted attention is a viable candidate for further enhancing big-data models in the context of market prediction using text analysis.

Importantly, it is shown that while unweighted attention (as used by Bybee et al.), is a good proxy for market movement, it is possible that directional information of these market movements is being lost. The authors explain this by saying topics should be moving in one direction. But consider the case where attention during a day is devoted to positive news about financial crises. The original unweighted model would have lost this information and incorrectly predicted downward trends with increased attention. By introducing sentiment-weighted attention, this directional movement can be recaptured and considered in a more realistic and accurate manner.

While the comparison of Bybee et al.'s model to our own is somewhat unfair due to data and computational limitations, we believe that we have found an original way to adapt the attention representation presented in their paper. There are many facets with which our model is performing just as well in capturing market fluctuation indicators as Bybee et al.'s model, with far less data. Even more interestingly, there are some respects with which our model is performing better out-of-sample than the authors in-sample estimation, for example with our sentiment-weighted attention model for Market Volatility, which produced an R^2 over 0.57.

As a result, further inspection into this route of research is encouraged. Augmenting attention by using sentiment in perhaps more interesting ways, or even using sentiment on larger samples are two advisable routes for further research. It appears that the omission of formally and explicitly including topic sentiment in Bybee et al.'s model is potentially limiting the significance of their results.

We conclude that sentiment is an essential consideration when assessing forward-looking market index fluctuations based on the discovered predictive capacity of sentiment augmented attention. Particularly, topical financial news publications are concluded to reveal predictive direction and magnitude information on market index fluctuations to a significant extent.

REFERENCES

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. "Measuring economic policy uncertainty." *The quarterly journal of economics*, 131(4): 1593–1636.

Bybee, Leland, Bryan T Kelly, Asaf Manela, and Dacheng Xiu. 2021. "Business News and Business Cycles." National Bureau of Economic Research Working Paper 29344.

Heitmann, Mark, Christian Siebert, Jochen Hartmann, and Christina Schamp. 2020. "More than a feeling: Benchmarks for sentiment analysis accuracy." *Available at SSRN 3489963*.

FIGURES

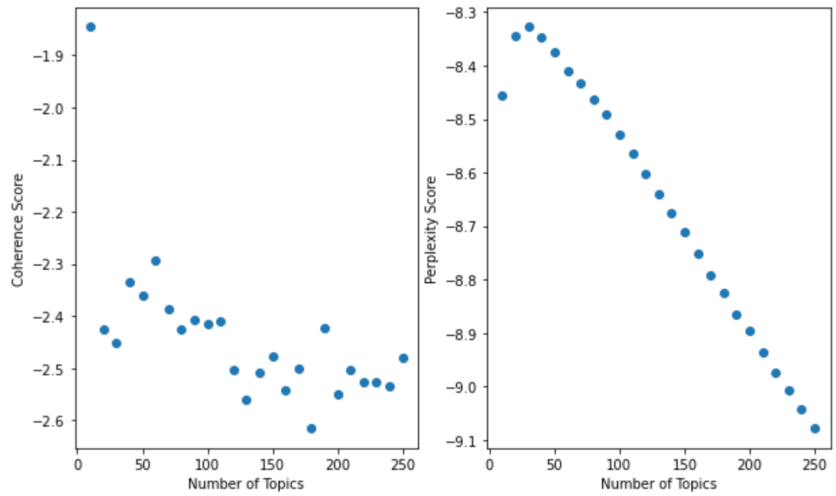

FIGURE A1. LDA MODEL RANKING

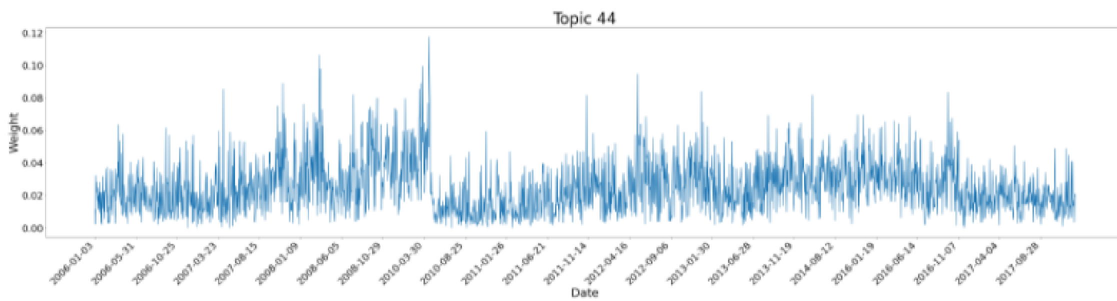

FIGURE A2. DAILY-AGGREGATED ATTENTION FOR 'FINANCIAL CRISIS'

Note: There are various cases of missing daily data in this time-series

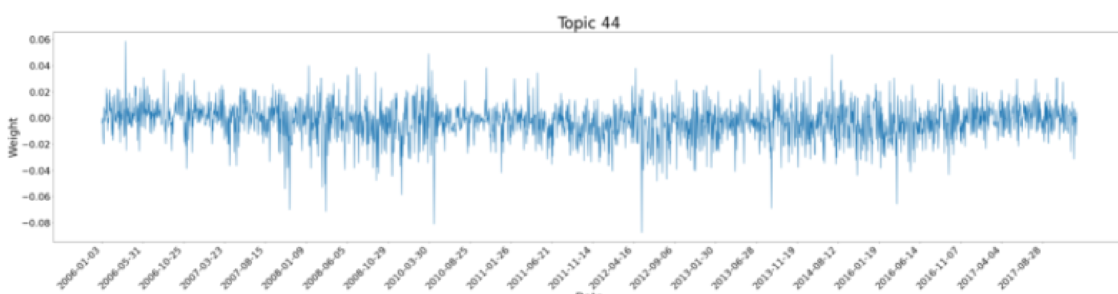

FIGURE A3. DAILY-AGGREGATED SENTIMENT-WEIGHTED ATTENTION FOR 'FINANCIAL CRISIS'

Note: There are various cases of missing daily data in this time-series

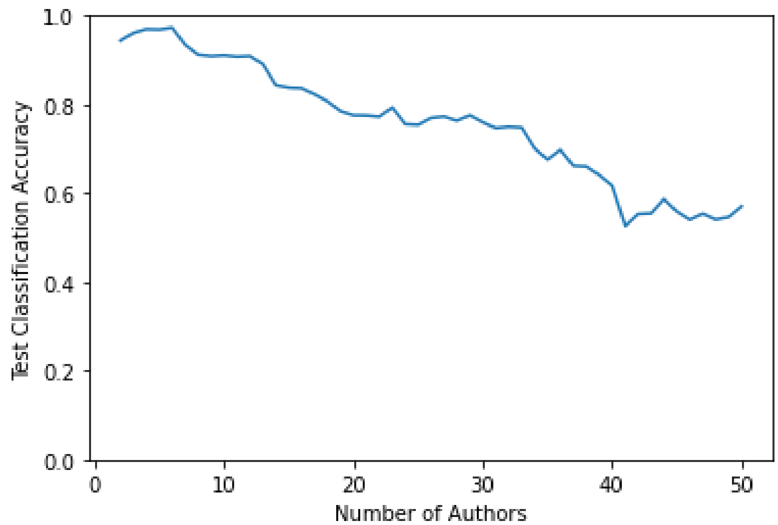

FIGURE A4. PROLIFIC AUTHORSHIP BIAS

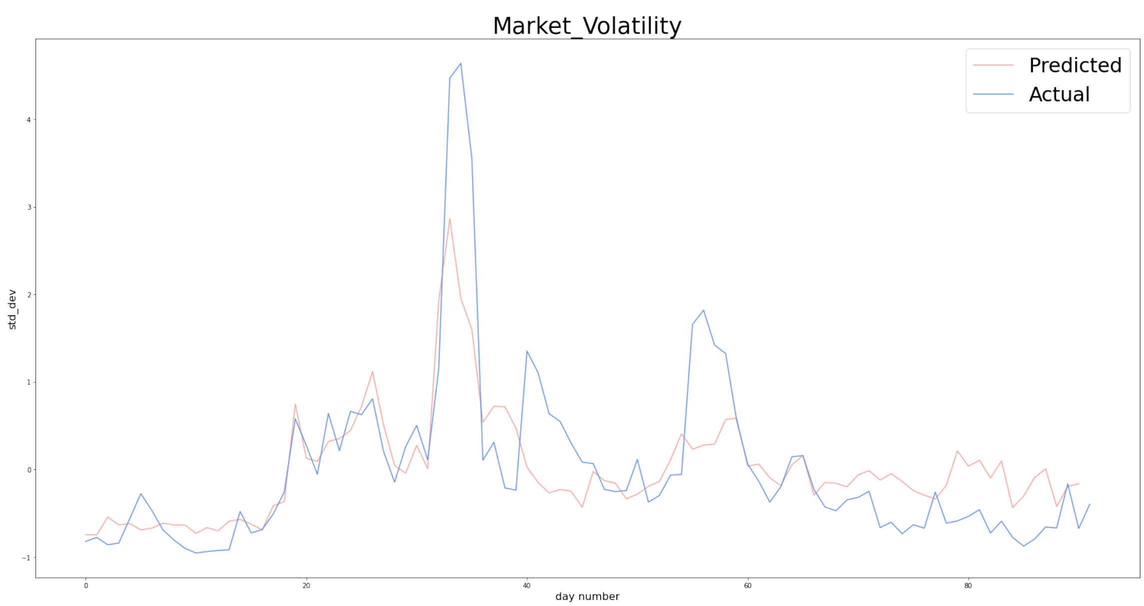

FIGURE A5. IN-SAMPLE MARKET VOLATILITY REPLICATION

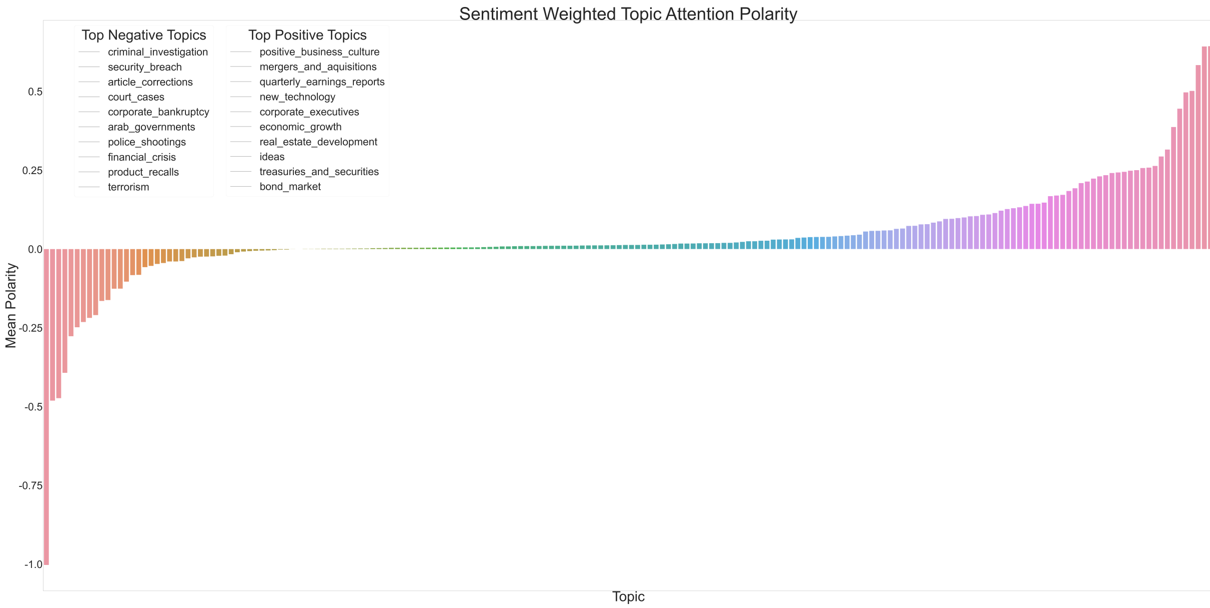

FIGURE A6. POLARIZED SENTIMENT-WEIGHTED ATTENTION BY TOPIC

Note: Topics on the extreme right represent highly positive sentiment weighted attention topics, while topics on the extreme left represent highly negative sentiment weighted attention topics.

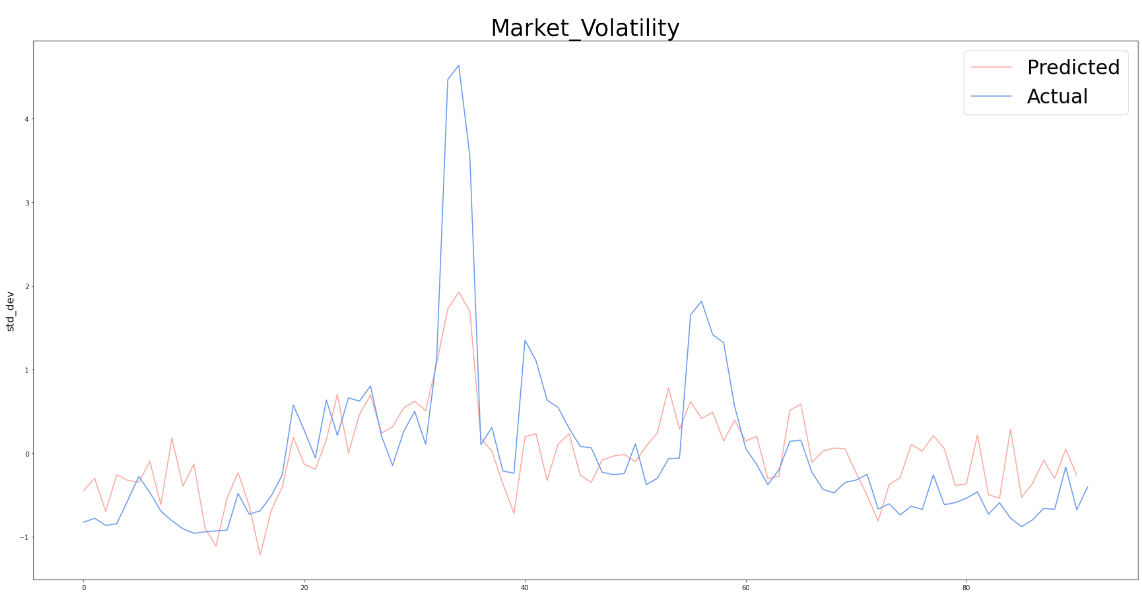

FIGURE A7. SENTIMENT ONLY MODEL IN-SAMPLE MARKET VOLATILITY PREDICTION

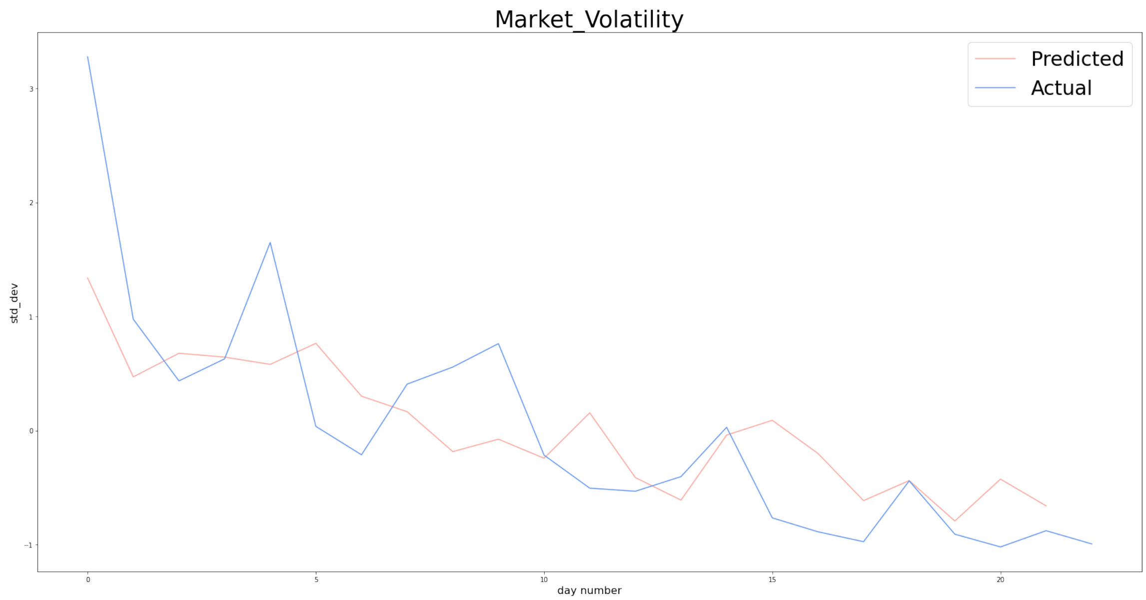

FIGURE A8. SENTIMENT ONLY MODEL OUT-OF-SAMPLE MARKET VOLATILITY PREDICTION

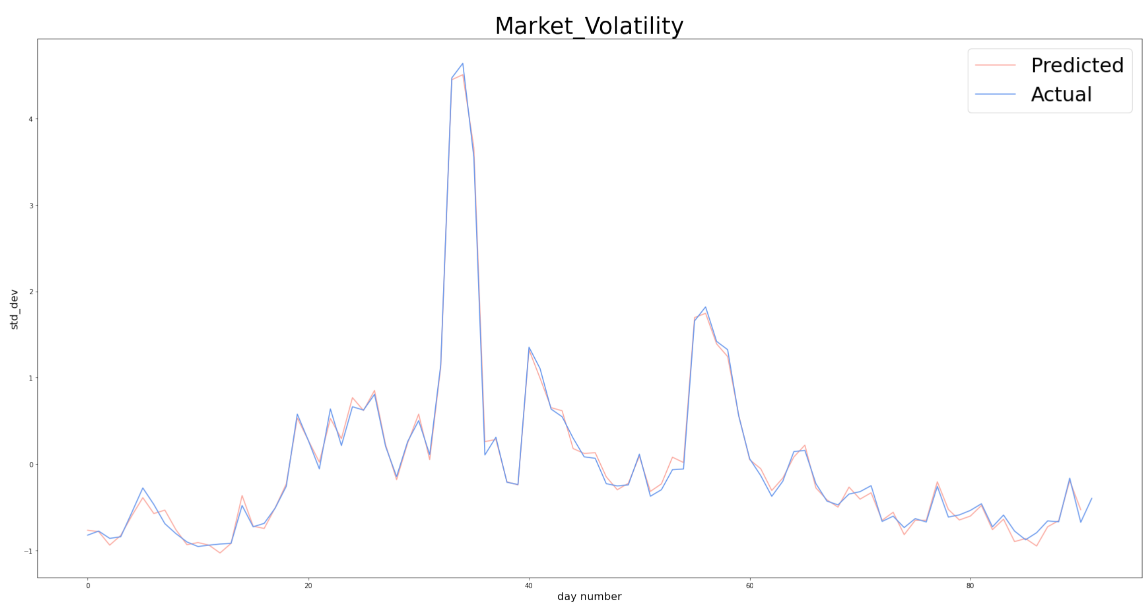

FIGURE A9. COMBINED MODEL IN-SAMPLE MARKET VOLATILITY PREDICTION

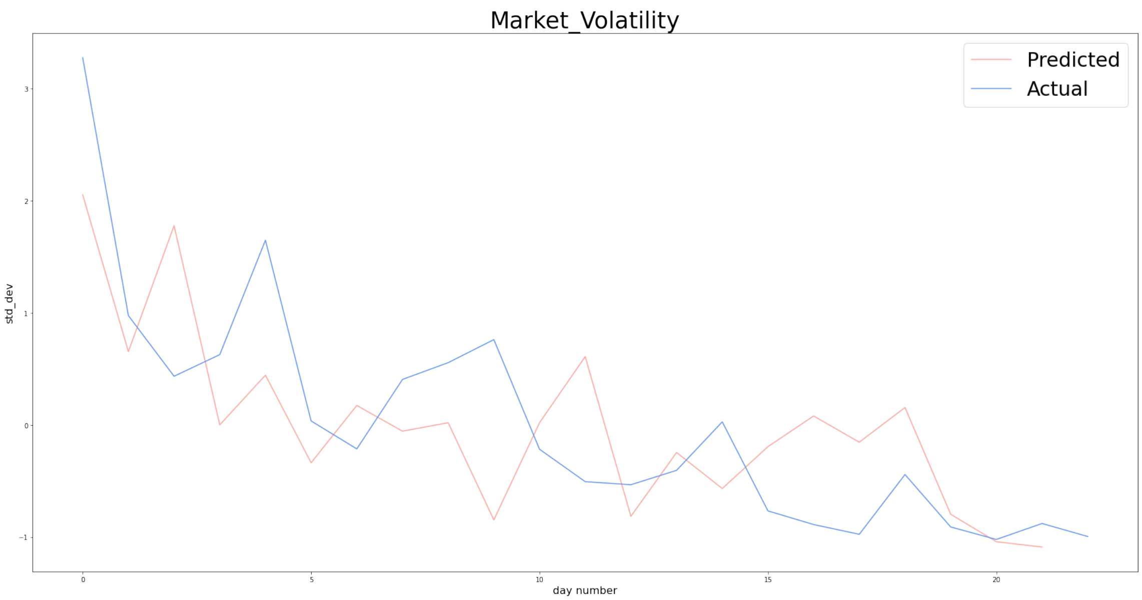

FIGURE A10. COMBINED MODEL OUT-OF-SAMPLE MARKET VOLATILITY PREDICTION

```
['NEGATIVE', 'NEGATIVE', 'POSITIVE', 'NEGATIVE', 'NEGATIVE']
[" The New York region's three major commuter railroads are struggling to meet a federally mandated deadline to install a sys
tem that reduces the risk of derailments and crashes.", 'The issue was underscored this week with the fatal derailment of a s
peeding Amtrak passenger train in Washington state.', 'Long Island Rail Road, Metro-North Railroad and NJ Transit carry a tot
al of about 450,000 rail commuters on an average weekday.', 'If they fail to meet the December 2018 deadline, they risk huge
fines and exposure to liability should a preventable accident occur.', 'The railroads have had almost a decade to get the sys
tem, known as positive train control, up and running.']
```

FIGURE A11. EXAMPLE OF SENTIMENT TRANSFORMER CLASSIFICATION

Note: Sentences 1, 2, 4, and 5 were classified as negative and sentence 3 was classified as positive by the transformer. Thus, the majority vote decision would classify this article as negative.

TABLES

TABLE B1—REPLICATION RESULTS

Dependent Variable	R2_train	R2_test	MSE_train	MSE_test
Economic_Policy_Uncertainty	0.444823	0.015660	0.555177	0.984340
Monetary_policy	0.227946	-0.107392	0.772054	1.107392
Fiscal_Policy	0.413173	0.213897	0.586827	0.786103
Taxes	0.442393	0.263392	0.557607	0.736608
Government_spending	0.502365	0.225916	0.497635	0.774084
Health_care	0.480570	0.369256	0.519430	0.630744
National_security	0.385382	-0.171831	0.614618	1.171831
Entitlement_programs	0.388513	0.208571	0.611487	0.791429
Regulation	0.123955	0.031914	0.876045	0.968086
Financial_Regulation	0.399585	-0.277346	0.600415	1.277346
Trade_policy	0.115496	0.071070	0.884504	0.928930
Sovereign_debt_currency_crises	0.270470	0.020505	0.729530	0.979495
Employment_growth	0.074409	-0.018862	0.925591	1.018862
Industrial_Production_growth	0.179363	-0.018760	0.820637	1.018760
Market_Volatility	0.615923	0.359021	0.384077	0.640979

TABLE B2—FISCAL POLICY

Topics	coefs
wars	0.001
time	0.029
confidential sources	0.047
government budget	0.333
electoral process	0.023
R^2_{train}	0.413
R^2_{test}	0.214

TABLE B3—TAXES

Topics	coefs
wars	0.054
economic growth	0.052
confidential sources	0.033
government budget	0.318
electoral process	0.09
R^2_{train}	0.442
R^2_{test}	0.263

TABLE B4—HEALTH CARE

Topics	coefs
wars	0.009
time	0.061
random	0.12
government budget	0.295
electoral process	0.082
R^2_{train}	0.481
R^2_{test}	0.369

TABLE B5—ENTITLEMENT PROGRAMS

Topics	coefs
random	0.002
film industry	0.015
confidential sources	0.053
government budget	0.325
market pricing	-0.014
R^2_{train}	0.389
R^2_{test}	0.209

TABLE B6—MARKET VOLATILITY

Topics	coefs
financial crisis	0.598
stock market decline	0.006
sports and entertainment	0.01
geopolitics natural resources	0.006
china	-0.01
R^2_{train}	0.616
R^2_{test}	0.359

TABLE B7—SENTIMENT WEIGHTED TOPIC ATTENTION LASSO RESULTS

Dependent Variable	R2_train	R2_test	MSE_train	MSE_test
Economic_Policy_Uncertainty	0.384743	0.083543	0.615257	0.916457
Monetary_policy	0.288240	0.035117	0.711760	0.964883
Fiscal_Policy	0.390686	0.136459	0.609314	0.863541
Taxes	0.357330	0.164948	0.642670	0.835052
Government_spending	0.431728	0.118194	0.568272	0.881806
Health_care	0.417006	0.075195	0.582994	0.924805
National_security	0.287030	-0.035553	0.712970	1.035553
Entitlement_programs	0.494734	0.267758	0.505266	0.732242
Regulation	0.334446	0.005711	0.665554	0.994289
Financial_Regulation	0.203811	0.013387	0.796189	0.986613
Trade_policy	0.246006	0.029056	0.753994	0.970944
Sovereign_debt_currency_crises	0.420221	-0.069273	0.579779	1.069273
Employment_growth	0.213536	-0.050978	0.786464	1.050978
Industrial_Production_growth	0.358408	0.111593	0.641592	0.888407
Market_Volatility	0.562948	0.571919	0.437052	0.428081

TABLE B8—MARKET VOLATILITY

Topics	coefs
real estate market	-0.067
market capitalization	-0.167
company market interactions	-0.036
economic growth	-0.243
global markets	-0.016
quarterly earnings reports	-0.042
hedgefunds	-0.09
investments portfolio management	-0.149
president usa	0.059
R^2_{train}	0.563
R^2_{test}	0.572

TABLE B9—ENTITLEMENT PROGRAMS

Topics	coefs
international trade	0.041
corporate executives	-0.072
housing market	0.116
corporations	-0.163
time	-0.1
europe	-0.144
film industry	0.136
government budget	-0.047
credit cards	-0.02
R^2_{train}	0.495
R^2_{test}	0.268

TABLE B10—COMBINED MODEL RESULTS: 15 FEATURES

Dependent Variable	R2_train	R2_test	MSE_train	MSE_test
Economic_Policy_Uncertainty	0.645287	0.007984	0.354713	0.992016
Monetary_policy	0.474864	-0.162172	0.525136	1.162172
Fiscal_Policy	0.690242	0.400367	0.309758	0.599633
Taxes	0.642487	0.428526	0.357513	0.571474
Government_spending	0.619047	0.230384	0.380953	0.769616
Health_care	0.655733	0.468972	0.344267	0.531028
National_security	0.549376	-0.022370	0.450624	1.022370
Entitlement_programs	0.652001	0.372540	0.347999	0.627460
Regulation	0.508588	-0.023625	0.491412	1.023625
Financial_Regulation	0.566271	-0.377065	0.433729	1.377065
Trade_policy	0.305587	0.064984	0.694413	0.935016
Sovereign_debt_currency_crises	0.665261	-0.298827	0.334739	1.298827
Employment_growth	0.412661	-0.229633	0.587339	1.229633
Industrial_Production_growth	0.450482	0.074603	0.549518	0.925397
Market_Volatility	0.814982	0.428998	0.185018	0.571002

TABLE B11—COMBINED MODEL RESULTS

	R2_train	R2_test	MSE_train	MSE_test
Economic_Policy_Uncertainty	0.796291	-0.018862	0.203709	1.018862
Monetary_policy	0.759725	-0.155711	0.240275	1.155711
Fiscal_Policy	0.843092	0.410820	0.156908	0.589180
Taxes	0.852025	0.502880	0.147975	0.497120
Government_spending	0.788814	0.302991	0.211186	0.697009
Health_care	0.772131	0.450492	0.227869	0.549508
National_security	0.707125	0.030155	0.292875	0.969845
Entitlement_programs	0.843526	0.414414	0.156474	0.585586
Regulation	0.831664	0.067140	0.168336	0.932860
Financial_Regulation	0.762118	-0.239551	0.237882	1.239551
Trade_policy	0.661387	0.103252	0.338613	0.896748
Sovereign_debt_currency_crises	0.855900	-0.179209	0.144100	1.179209
Employment_growth	0.757062	-0.690079	0.242938	1.690079
Industrial_Production_growth	0.709218	0.121779	0.290782	0.878221
Market_Volatility	0.922965	0.430225	0.077035	0.569775

EQUATIONS

$$Attention = \begin{bmatrix} \sum_{d \in M_1} \frac{\sum_{a \in d} p_T(a)}{|d|} \\ \vdots \\ \sum_{d \in M_N} \frac{\sum_{a \in d} p_T(a)}{|d|} \end{bmatrix} \tag{C1}$$

where M_n is the collection of all articles in month n, N is the number of topics, d is the collection of all articles in a given day, T is the amount of all topics in an LDA model, $p_T(a)$ is a vector of predicted probabilities in order from 0 to T of an article a belonging to that topic, and $|\cdot|$ is the cardinality operator.

$$Sentiment\text{-}Weighted\ Attention = \begin{bmatrix} \sum_{d \in M_1} \frac{\sum_{a \in d} p_T(a)s(a)}{|d|} \\ \vdots \\ \sum_{d \in M_N} \frac{\sum_{a \in d} p_T(a)s(a)}{|d|} \end{bmatrix} \tag{C2}$$

where M_n is the collection of all articles in month n, N is the number of topcis, d is the collection of all articles in a given day, T is the amount of all topics in an LDA model, $p_T(a)$ is a vector of predicted probabilities in order from 0 to T of an article a belonging to that topic, $|\cdot|$ is the cardinality operator, and

$$s(a) = \begin{cases} -1 & \text{a is predicted negative sentiment} \\ 0 & \text{a is predicted neutral sentiment} \\ 1 & \text{a is predicted positive sentiment} \end{cases}$$