How Do U.S. Politicians Use Emotional Intensity?

A Text-as-Data Approach to Twitter Engagement

Zhiyang(Chris) Cheng

Introduction

Social media interaction has become an increasingly vital mechanism for expressing opinions, constructing public images, and facilitating engagement between elected officials, constituents, and broad political audiences. Existing research demonstrates that politicians strategically utilize tweets to advance their political goals (Gervais et al., 2020). A critical component of this strategy involves emotion regulation (Wang et al., 2021), and the choice of narrative strategy has been shown to significantly impact public reactions (Rathje et al., 2021).

This report explores the emotional intensity of tweets crafted by elected U.S. politicians, addressing three key questions:

1. What is the effect of using emotionally intense language on public reaction?

2. How does a politician's identity influence their general patterns of language intensity usage?

3. Under what conditions do politicians tend to employ more intense language versus less intense language?

Using tweet data from elected officials collected between May and November 2024, our analysis yields three primary findings.

First, we confirm that emotional intensity is a significant driver of public reaction. This effect reveals a sharp partisan and racial divide: Republicans have higher average public engagement than Democrats and exhibit a higher sensitivity regarding audience responsiveness to the emotional intensity of the language used in the posts. White politicians also have the highest average public engagement as well we audience responsiveness sensitivity.

1

Second, regarding the identity of politicians who use intense language, we found that Republicans use significantly less intense language—averaging roughly 6 points lower than Democrats—despite their audience's high responsiveness. In contrast, Latino politicians use the highest intensity of language ; however, they fail to convert this significantly into active Likes or Comments. White and black politicians has the similar usage of intense language but receive totally different public reactions.

Finally, structural topic modeling reveals that politicians reserve their highest-intensity language for less divisive/consensus topics such as voter registration, assistance information, and safety advocacy. Conversely, for divisive topics such as abortion legislation and crime, intensity scores are lower and centered around the average, suggesting politicians employ a milder tone on controversial subjects to prevent provoking division.

Data and Methods

Data

This report utilizes the Digitally Accountable Public Representation (DAPR) Database, specifically selecting a sample[1] of tweets sent by elected U.S. politicians from May to November 2024. The original data collection was conducted on a weekly basis.

The initial dataset contained approximately 91,000 observations. After attempting to retrieve posts from the URLs provided, only about 43,000 tweets were successfully collected, while roughly 48,000 tweets had been deleted or removed.

Figure 1 illustrates the pattern of these deletions[2].

The deletion of posts likely introduces selection bias[3]. Evidence suggests that posts during the campaign period have lower intensity scores compared to normal weeks, which contradicts the expectation that election-period posts would contain more drastic language and content to attract public attention. Politicians may also strategically publish and then delete posts to capitalize on short-term effects. Consequently, our dataset represents only the content politicians chose to and were able to preserve. If more drastic or offensive posts are disproportionately

---

[1]This can cause sampling bias (Morstatter et al., 2013). Tendencies that exist but are less significant may be omitted; for example, the structural topic model was unable to capture the event of the first assassination attempt on President Donald Trump, even though related posts exist in the corpus.

[2]We found no significant difference for the post deletion between parties. For democrats the number of deletions is 23886 while for republicans it's 22897.

[3]Deletions peaked three weeks before voting, two weeks after voting, and two weeks after the results were released. The cause of the deletions during the 35th, 38th, and 39th weeks is unclear.
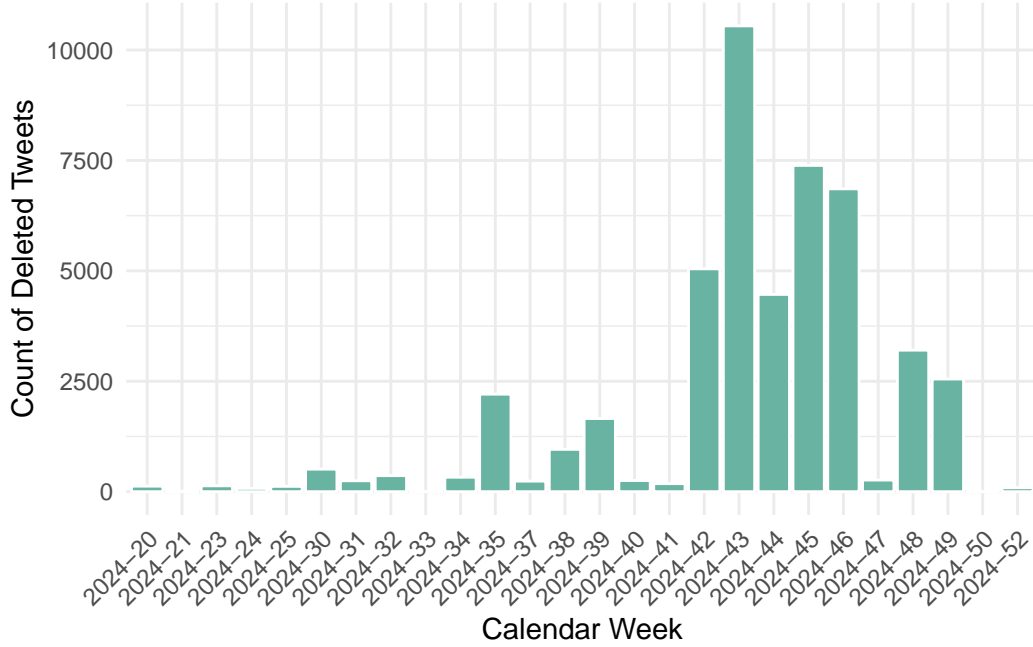
Figure 1: Pattern of tweet deletions over time

removed, our observation that politicians use mild language on certain topics may be biased; it is possible that only the mild posts survived.

Best-Worst Scaling for Intensity Measurement

To measure tweet intensity, this research adopted the Best-Worst Scaling (BWS) method. We first created sets containing a small number of texts, ensuring each text appeared a specific number of times alongside different texts in each appearance. We then labeled which text in each set was the most emotionally intense and which was the least intense. Based on the labels for these sets, we calculated a score for each text.

$$\text{BWS Score}_j = \frac{(\text{Number of times chosen as Best}) - (\text{Number of times chosen as Worst})}{\text{Total number of times the item appeared}}$$

For this research, we used a sample dataset of 3,000 texts, created sets of 4 texts, ensuring each text appeared for 14-16 times, which generated approximately 12,000 sets. We then **labeled** these sets using the OpenAI API. After generating the training data, we took a pre-

trained RoBERTa model[4], removed its three sentiment analysis layers, and replaced them with a single-layer regression head. After training this model on the labeled sample, we applied it to the full dataset and rescaled the resulting scores from -1 to 1 to a range of 0 to 100. Figure 2 shows the final distribution of Best-Worst Scaling scores.[5]
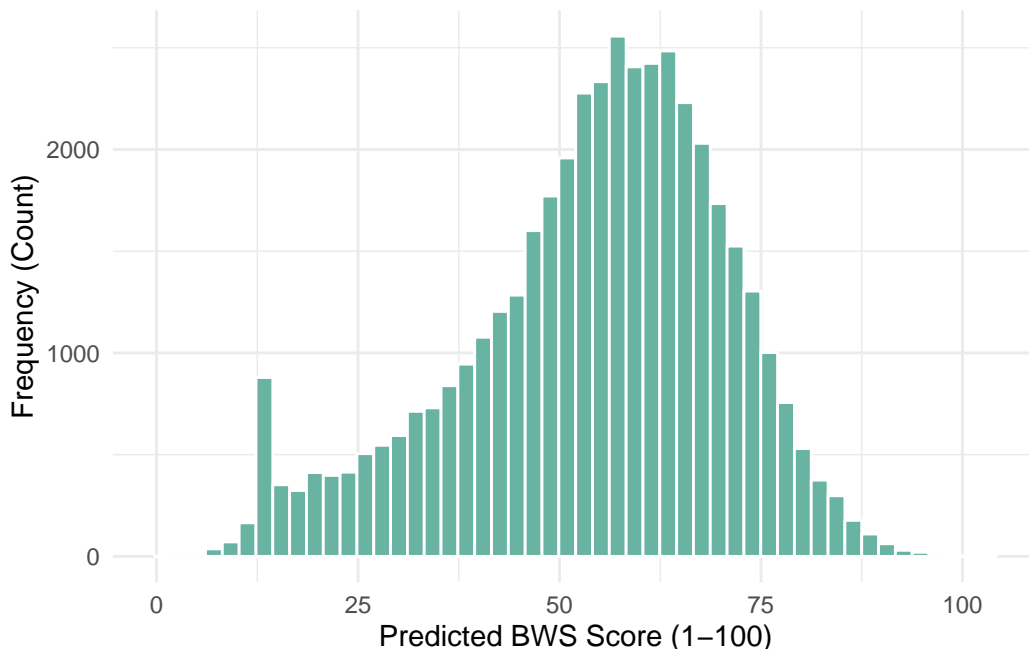

Figure 2: Distribution of Emotional Intensity Scores

While this score provides a comparative measure of intensity within the corpus, its interpretation requires nuance. The lowest scores correspond to short, content-light observations (e.g., "This.", "no", "Yep"), which are objectively low in intensity.[6] However, the interpretation of the highest-scoring texts is less straightforward. Even observations with the highest BWS scores are relatively mild compared to "intense language" in a broader sense; for example, the highest-scoring texts in our corpus often involve sharing disaster relief information.

The loss of absolute intensity information means we cannot quantify the precise degree of change in intensity represented by each score increment. If we assume that absolute intensity is evenly distributed[7], the rapid drop in area between gaps in Figure 2 suggests that politicians

---

[4]Twitter-roBERTa-base for Sentiment Analysis

[5]The spike at the score between 12 - 14 is due to the post with empty crawled content, which means the post only contains videos or/and images.

[6]As verified manually, most of these posts actually come with videos and images that are not captured. This is actually a source of bias.

[7]It is unclear if this presumption holds for the Best-Worst Scaling method in this context, as the absolute intensity

are generally less likely to use intense language and are more willing to use milder language compared to common discourse.

To measure the effect of intense language usage, we examined four public engagement metrics: comments, retweets, likes, and views. We hypothesize that higher counts in these metrics indicate a stronger public reaction to emotionally intense content.

Analysis

This research applies regression analysis and Structural Topic Modeling (STM) to explore the data.

Regression Analysis

We utilized multiple equations, treating likes, comments, retweets, and views as dependent variables. We applied a logarithmic transformation to rescale these public reaction variables due to their right-skewed distribution. The intensity score was centered around the mean to facilitate interpretation; consequently, the differences in intercepts between different identity groups represent the estimated gaps in engagement for a tweet of average intensity, holding other covariates constant. A dummy variable, 'campaign_period', was created to indicate posts sent between week 42 and week 49 of 2024. The formula is as follows:

$$\log(\text{Metric}_i + 1) = \beta_0 + \beta_1 \text{Score}_c + \beta_2 \mathbf{D}_i + \beta_3 (\text{Score}_c \times \mathbf{D}_i) + \beta_4 \log(\text{Followers}_i + 1) + \epsilon_i$$

where Di is a vector defined as politicians' identity factors:

$$Di = [Party\_i, Race\_i, Gender\_i]$$

---

intervals between scores can vary. For example, imagine a resentful complaint within a corpus of children's literature: it would receive the highest score, but the gap in absolute intensity between it and the second-most intense text may be significantly larger than the distance between the second and the lowest. Consequently, while we know politicians are more likely to use less intense language than the general baseline, we cannot determine the precise magnitude of this difference. It is possible that posts with scores near the majority are also closer in absolute levels; in other words, absolute intensity might be highly dense around the center while sparse at the tails. While we can draw inferences, they must be premised on the assumption that the distance between scores is equidistant (or at least similar). We are unable to provide evidence for this here without using human coders to label the absolute intensity for each knot score.

Next, we treated the intensity score of the post content as the dependent variable to identify the factors influencing its usage. The formula is as follows:

$$\text{Intensity}_i = \beta_0 + \beta_1\text{Gender} + \beta_2\text{Party} + \beta_3\text{Race} + \beta_4(\text{Party} \times \text{Race}) + \beta_5\log(\text{Followers}) + \beta_6\text{Campaign} + \epsilon_i$$

Structural Topic Model (STM)

The corpus was pooled by the weak and politicians. To handle the interaction between race and party, we created a new variable combining both factors, filtering out empty categories. Testing revealed that a topic number of 16 best captured the characteristics of the corpus. The formula for STM is as follows:

$$\text{Prevalence}_{d,k} = \alpha_k + f_1(\text{Intensity}_d) + f_2(\text{Date}_d) + \gamma_{\text{Race} \times \text{Party}}$$

Results

Figure 3 displays the results when treating public reactions as dependent variables. We found a consistent positive association: the more intense the language usage, the greater the public reaction. The magnitude of this sensitivity varies by engagement type, ranking from highest to lowest: Views, Retweets, Likes, Comments.

This ranking is intuitive as it mostly reflects the difficulty of each action—it is easier to view a post than to interact, and clicking is easier than commenting. Therefore, for the same level of intensity, the more effort an action requires, the lower its sensitivity. However, it is confusing that retweets showed a stronger response than likes; we presume that making audience advocate for the post is more demanding compared to having them sharing the post.

This tendency is significantly differentiated by political party (Figure 4 and Figure 5) and racial identity (Figure 6), while the rank for each action remains the same.

Results show that Republican politicians not only have significantly higher public interaction on average (Figure 4) but also have a higher sensitivity regarding audience responsiveness to the language used in their tweets (Figure 5). Audience engagement for Republicans is nearly 90% higher than for Democrats. Across interaction types, while comments still exhibit the smallest disparity, the magnitudes of the other three categories are much closer to one another compared to the results from Figure 3.
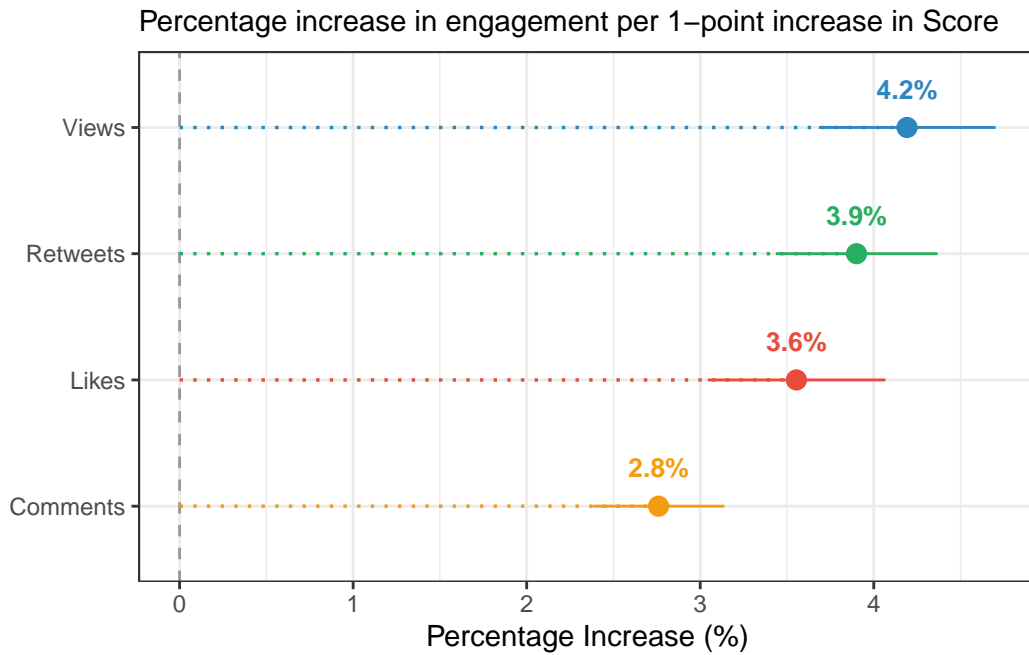
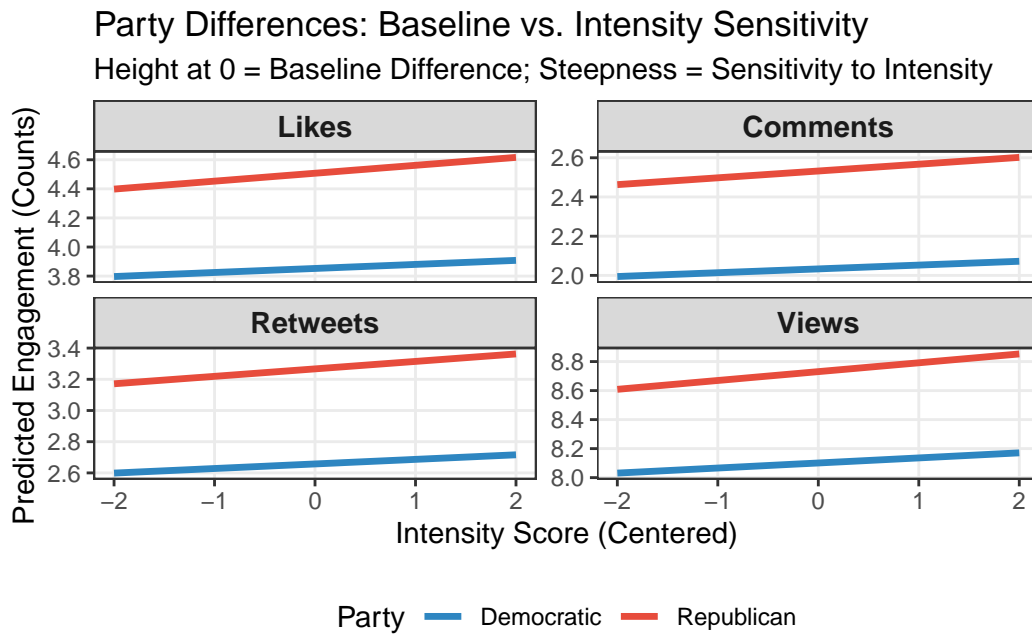Figure 3: Sensitivity of Public Interactions to BWS Score
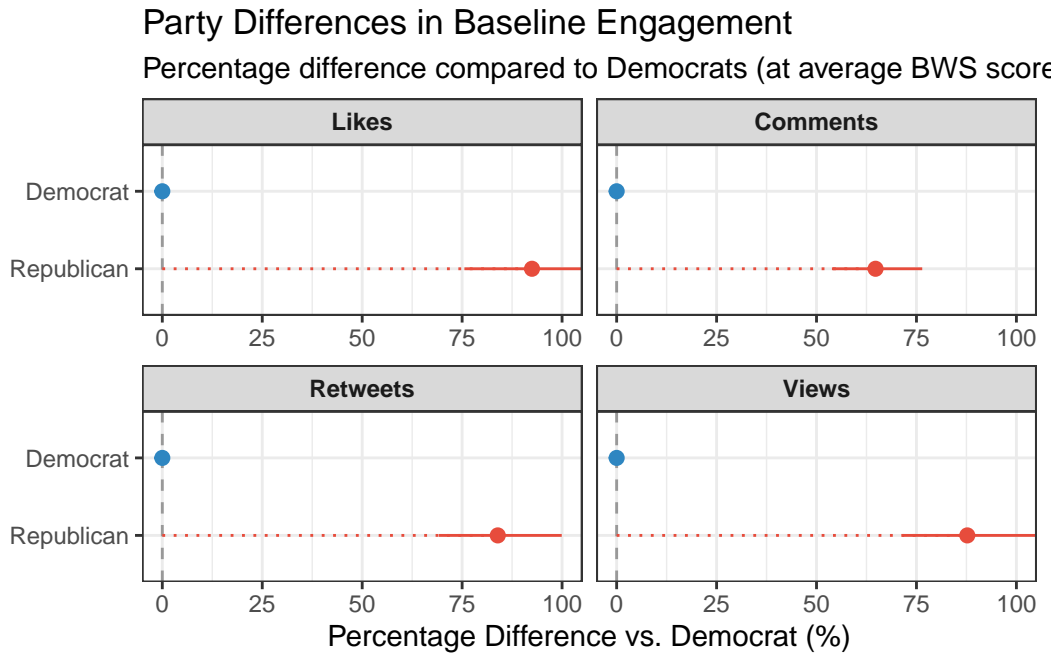

Figure 4: Party Differences

Party Differences in Baseline Engagement
Percentage difference compared to Democrats (at average BWS score

Figure 5: Party Differences in sensetivity

Racial Differences in Likes: Baseline vs. Intensity Sens
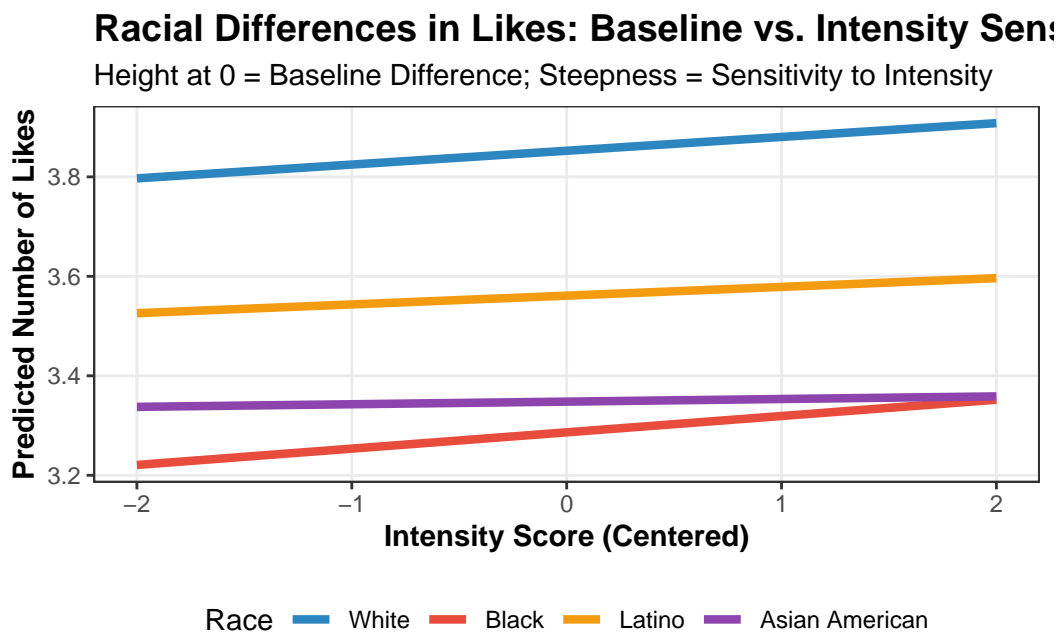Height at 0 = Baseline Difference; Steepness = Sensitivity to Intensity

Figure 6: Racial Differences in Engagement: Likes

The difference based on the racial identity of the politicians is also substantially significant. White politicians generally elicit both the highest public interaction on average and the highest audience responsiveness sensitivity. Black politicians have the lowest public interaction on average but have a similar level of public response sensitivity as White politicians. Meanwhile, Asian politicians are lower in both measurements.

Next, we explored the determinants of the intensity score itself (Figure 6)—how a politician's identity influences their likelihood of using intense language.
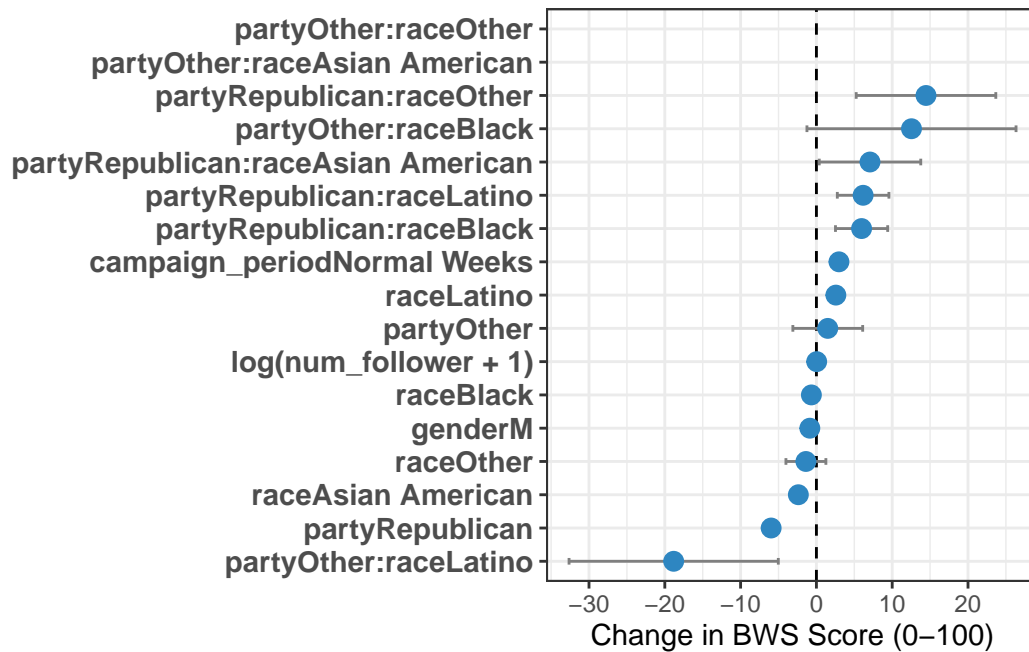

Figure 7: Predictors of Emotional Intensity (BWS Score)

The result shows that Although Republican politicians enjoy the highest rewards for using intense language, they are actually less likely to use it. On average, Republican tweets score roughly 6 points lower on the intensity scale than Democratic tweets. This suggests that while the Republican electorate is highly reactive to intensity, the politicians themselves maintain a lower average intensity baseline.

Racial identity also drives usage in ways that do not align with engagement rewards. Black politicians show no statistically significant difference in intensity usage compared to White politicians yet receive significantly lower engagement on average(Figure 4). Asian American politicians are the least likely to use intense language, correlating with the lower engagement and sensitivity of their audience. Latino politicians have the highest usage of intense language but fail to convert this significantly into active Likes or Comments.
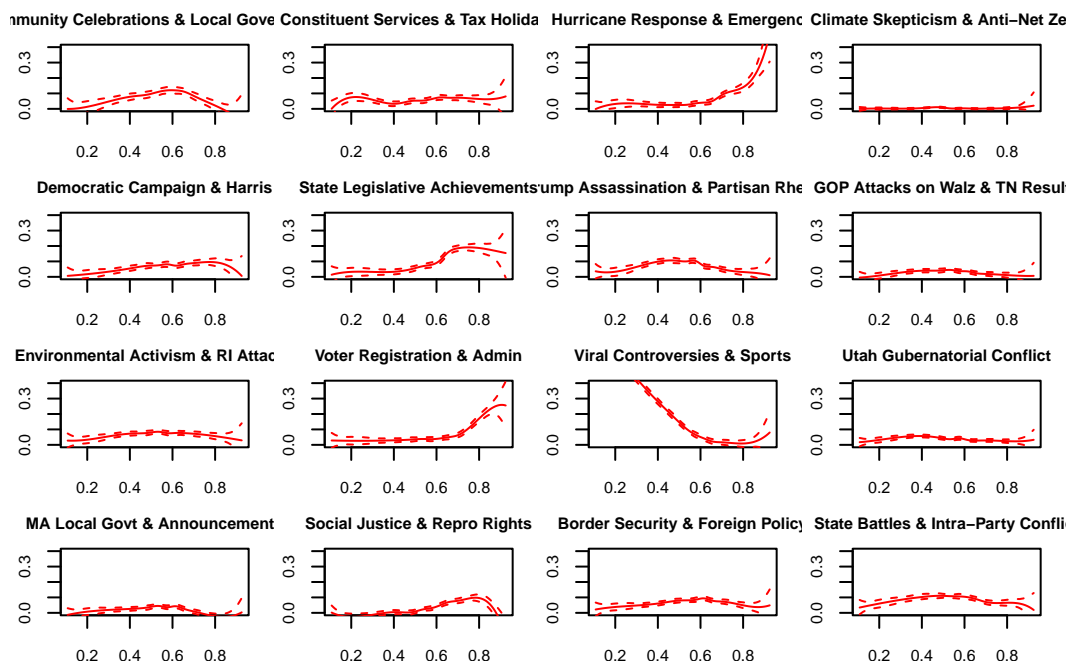
9

Figure 8: Topic Proportions varying by Emotional Intensity (BWS Score)

Figure 7 shows the results from the Structural Topic Model regarding topic proportions differentiated by intensity score. We found that for less divisive topics—such as calls for voter registration, sharing assistance information, expressing sympathy for hurricanes, and safety advocacy—intense language occupies the posts with the highest intensity scores. Conversely, for divisive topics—such as the detention of Paul Watson, attacks on Tim Walz and Katie Hobbs, abortion legislation, and criticism of California crime—intensity scores are lower and centered around the average. This suggests politicians expand their public image by adopting intense language generally but employ a milder tone for controversial topics to prevent provoking division.

Discussion

Summary

Using tweet data from elected officials from May to November 2024, our findings indicate that increased language intensity correlates with a consistent positive increase in public reaction. The magnitude of this sensitivity varies by engagement type, ranking from highest to

lowest: Views, Retweets, Likes, and Comments. These tendencies are significantly differentiated by political party and racial identity. Republicans start with nearly 90% higher baseline engagement and exhibit a higher sensitivity to emotional intensity compared to Democrats. Among racial groups, White politicians elicit the highest public interaction on average and the highest audience responsiveness sensitivity, followed by Black, Latino, and Asian American politicians.

Second, we explored how identity affects intensity usage. We found that Republican politicians use significantly less intense language than Democrats (averaging roughly 6 points lower), despite their audience's high responsiveness. Racial identity also drives usage in ways that do not align with engagement rewards. Black politicians show no statistically significant difference in intensity usage compared to White politicians but receive significantly lower engagement on average. Asian American politicians are the least likely to use intense language, correlating with lower engagement and sensitivity. In contrast, Latino politicians use the highest intensity of language but fail to convert this usage significantly into active Likes or Comments.

Finally, structural topic modeling reveals a strategic use of emotion based on content. Politicians reserve their highest-intensity language for less divisive consensus topics, such as voter registration, assistance information, and safety advocacy. Conversely, for divisive topics such as abortion legislation and crime, intensity scores are lower and centered around the average, suggesting politicians employ a milder tone on controversial subjects to prevent provoking division.

Next Steps

The first priority for future research is to expand the time coverage of the corpus as well as collecting more completed data to address the sampling bias and selection bias. Including a longer time scale would introduce more observations and topics, allowing for further verification of the general trends identified.

Measurements could also be improved. For the intensity score, human labellers are needed to assess absolute intensity based on specific criteria at various score levels to capture more granular information.

Public reaction also requires a more robust theoretical framework. For instance, an increase in "likes" does not necessarily equate to advocacy, as controversial content may draw both likes and "unlikes" (expressed via comments). The content of comments is therefore crucial. A significant challenge is endogeneity resulting from the omission of text divisiveness. Our STM analysis shows divisiveness is associated with intensity scores (likely negatively), but since divisiveness presumably increases public reaction, the current model specification may be biased, so the "controversy level" of topics should be controlled.

Finally, the omission of videos and images presents a significant challenge, as these media contain rich information currently treated as noise. Bridging the gap between text and visual media is difficult; for a computer, the text "a national flag" is identical to a description, but for a human, seeing the image is a distinct experience. A basic solution would be to include a variable indicating the presence of multimedia. A more advanced solution would be transcribing visual content and inserting special tokens (e.g., `<IMG_START>`) into the text analysis.

Further exploration should focus on narrative strategies. For instance, future work could examine under which specific topics politicians invoke national versus partisan identity, or when they choose to attack opponents versus signaling bipartisan cooperation. We also expect narrative strategies to evolve over time or under changing circumstances.

Reference

Ali, Rao Hamza, Gabriela Pinto, Evelyn Lawrie, and Erik J. Linstead. "A Large-Scale Sentiment Analysis of Tweets Pertaining to the 2020 US Presidential Election." *Journal of Big Data* 9, no. 1 (2022): 79. https://doi.org/10.1186/s40537-022-00633-z.

Gervais, Bryan T., Heather K. Evans, and Annelise Russell. "Tweeting for Hearts and Minds? Measuring Candidates' Use of Anxiety in Tweets During the 2018 Midterm Elections." *PS: Political Science & Politics* 53, no. 4 (2020): 652–56. https://doi.org/10.1017/S1049096520000852.

Morstatter, Fred, Jürgen Pfeffer, Huan Liu, and Kathleen Carley. "Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose." *Proceedings of the International AAAI Conference on Web and Social Media* 7, no. 1 (2013): 400–408. https://doi.org/10.1609/icwsm.v7i1.14401.

Rathje, Steve, Jay J. Van Bavel, and Sander Van Der Linden. "Out-Group Animosity Drives Engagement on Social Media." *Proceedings of the National Academy of Sciences* 118, no. 26 (2021): e2024292118. https://doi.org/10.1073/pnas.2024292118.

Tai, Yuehong Cassandra, Nitheesha Nakka, Khushi Navin Patni, et al. "The Digitally Accountable Public Representation Database: Online Communication by U.S. Officials." *Scientific Data* 12, no. 1 (2025): 1556. https://doi.org/10.1038/s41597-025-05857-1.

Wang, Meng-Jie, Kumar Yogeeswaran, Sivanand Sivaram, and Kyle Nash. "Examining Spread of Emotional Political Content among Democratic and Republican Candidates during the 2018 US Mid-Term Elections." *Humanities and Social Sciences Communications* 8, no. 1 (2021): 300. https://doi.org/10.1057/s41599-021-00987-4.

Data availability

All codes and data can be accessed through the github repository for this [project](#)