Emotions in #SaveTheChildren Rallies

Do Won Kim
College of Information Studies
University of Maryland
College Park, MD, United States
dowonkim@umd.edu

ABSTRACT

How can online #SaveTheChildren campaigns mobilize user engagement? This study examines emotional cues in comments and posts on #SaveTheChildren campaigns on the social media platform Parler. Drawing upon theories of emotions in the political process, this study hypothesized that there are positive associations between negative emotions, anger, and fear in the content and levels of engagement that it receives. The findings show that while negative emotions and fear in the content correlate positively with user engagement, anger is associated with lower levels of engagement. Overall, this research contributes to our understanding of how emotions can drive engagement in the context of conspiracy movements.

KEYWORDS

emotion, sentiment analysis, QAnon, #SaveTheChildren, conspiracy theory, social media

1 Introduction

The role of social media in political processes has garnered growing attention in recent scholarly discourse. Studies suggest that information diets on social media can facilitate different forms of political participation including protests (for a review, see [1]). This potential spillover effect of getting information from social media on political behavior is particularly pertinent in the context of conspiracy theories and misinformation, given that their online propagation can incite significant real-world consequences.

The case of QAnon's #SaveTheChildren campaign serves as a compelling context to explore the dynamic between online social movements and political behaviors. QAnon hijacked hashtags such as #SaveTheChildren (henceforth #STC) to spread their conspiracist allegations, and their online campaigns gained significant momentum, potentially leading to a series of offline demonstrations [2].

What is the relationship between conspiracist content on social media platforms and individual engagement with this content? This study aims to address this question in the context of #STC movements. Based on theories on emotions in the political process, this study posits that negative emotions, especially fear and anger, matter in mobilizing individuals' engagements with the #STC-related content.

This study analyzed emotional cues from #STC-related comments and posts on Parler and found that content conveying more negative emotions or fear was associated with higher levels of user engagement, whereas content having more anger was linked with lower user engagement. Overall, this study contributes to our understanding of how emotions in online discourses can catalyze individuals' political involvement in the context of conspiracist social movements.

2 Literature Review

Not all online activities successfully translate into political mobilizations. One important content-level factor is emotion. Several studies have found that emotionally charged and negatively valenced information spreads faster through social networks [3, 4, 5]. That is, how content is emotionally framed and how it resonates with its audience matter.

Broadly, there are three theoretical approaches to studying emotions in the political process [6, 7]. The first is valence theory, in which scholars distinguish emotions based on their valence - whether they are positive or negative. Valence theory suggests that emotions are automatically labeled as positive or negative after exposure to some external stimuli [8]. This automatic like-dislike emotional evaluation is critical in information processing. Furthermore, when it comes to positive versus negative emotions, negativity makes people more likely to engage with the information. For example, a study found cross-national evidence that negative news gets more attention [9].

On the other hand, the appraisal theory views emotions as discrete categories rather than a continuum with positive and negative valence attached [10, 11]. It acknowledges that fear can be different from anger, although both have the same negative valence. However, this theory does not take into account the intensity of discrete emotions.

In this light, the affective intelligence theory becomes useful. Affective intelligence theory posits that different emotional appraisals are important in two distinct judgment processes: an automatic mode characterized by reliance on intuition and habits (similar to System 1 in dual-processing theory), and a departure from this, marked by more deliberation and reflection to deal with higher uncertainties (like System 2 in dual-processing theory) [12, 13].

More specifically, emotional appraisals such as enthusiasm enhance behaviors that help achieve goals, whereas anger causes confrontational reactions in the face of familiar threats [13]. Similarly, Erisen [14] noted that anger tends to make people take risks and participate more in politics. Contrary, in settings with greater uncertainty and novelty, fear tends to be more prominent, and heightened fear levels result in more reflection and information-seeking behaviors [13]. This suggests that, despite both fear and anger having negative emotional valence, people usually feel fear when confronted with new challenges and unfamiliar threats, whereas anger tends to surface in response to recurring and familiar threats.

While fear may deter people from participating in offline forms of politics like protests, this may not be the case in the context of an online conspiracist movement like #STC. Recently, studies have begun to examine persuasion strategies in conspiracy narratives, such as the use of elements of uncertainty and fear [15]. Chen et al. [16], for example, showed that conspiracy-related videos used visual modalities used in horror films to elicit feelings of fear. Given this, #STC campaigns may seek to instill fear in individuals, potentially to increase concern about the well-being of children to broaden their audience and influence.

In summary, this study expands on previous research by investigating how emotional cues in online #STC campaigns may be associated with increased engagement.

Hypothesis 1. There will be a positive correlation between the level of negative emotions conveyed in #STC-related content and the extent of engagement it receives.

Hypothesis 2. There will be a positive correlation between the degree of anger expressed in #STC-related content and the level of engagement it garners.

Hypothesis 3. There will be a positive correlation between the degree of fear conveyed in #STC-related content and the level of engagement it generates.

3 Methodology

To test the hypotheses, this study uses the Parler dataset of comments and posts, collected from the open repository named "A Large Open Dataset from the Parler Social Network" (for more information, see [17])[1]. Table 1 describes the variables that are relevant to this study.

Name	Description	Note
body	Main text of content (comment or post)	Used in filtering and sentiment analysis
hashtags	A list of hashtags contained within the body	Used in filtering process
createdAtformatted	Timestamp when the content was created	Used when analyzing trends
upvotes	Number of likes received	Engagement metrics (Dependent variables)
impressions	Total number of exposures or views received	
reposts	Number of times the content was reshared.	
followers	Number of followers of the content creator	Control variables
following	Number of accounts that the creator is following	

Table 1: **Variables**

[1] Version 1 of this dataset, which was made available on January 15, 2021, was downloaded from the following site: https://zenodo.org/records/4442460

To focus on content relevant to #STC, I explored the data qualitatively and filtered cases by using keywords including "saveourchildren," "savethechildren," "save the children," "save our children," "savethebabies," and "save the babies." This filtering process resulted in 13,428 rows of content.

The sentiment analysis was then performed on this filtered dataset to extract emotions from each piece of content. Using a DistilRoBERTa-based approach, the proportion of basic emotions—anger, fear, joy, sadness, disgust, and surprise—as well as a neutral class—were extracted from each content [18]. During preprocessing, this process was streamlined for efficiency by removing emojis and limiting the token length to less than 512 characters.

Extracted values represent the intensity of each emotion and sum to one. A text, for example, may have a score of 0.8 for anger and 0.2 for disgust, with negligible scores for other emotions. That is, the given text conveys more anger and less disgust while conveying no other emotional cues (with near zero intensities). These emotions were then operationalized into rate metrics, such as calculating anger as 20% (=0.2*100), and so forth. Additionally, a variable called negative_rate is generated to indicate the cumulative rate of negative emotions (anger, fear, disgust, and sadness). Overall, the emotion-related independent variables were operationalized specifically to align with the study's hypotheses.

Finally, this study tested hypotheses to investigate relationships between emotional cues and engagements like upvotes, impressions, and reposts. In this process, this study controlled the number of followers and followers for each content creator. By taking into account factors like the number of followers, the study tried to identify whether the predictions about user engagement could be attributed to emotional cues, controlling for other external factors.

Due to the highly skewed nature of the dependent variables and the excess cases of zeros (see Appendix), a series of zero-inflated negative binomial regressions are conducted for the hypothesis testing.

4 Results

4.1 Descriptive Analysis

Figure 1 shows the distributions of emotional cues identified by the DistilRoBERTa model. The x-axis corresponds to the percentage of detected emotional cues in each content, whereas the y-axis reflects the count of content exhibiting the respective emotional intensities. As the histograms for each emotion exhibit, most content does not exhibit strong emotional cues and is clustered near zero. Meanwhile, the histogram for 'negative_rate', the composite measure of negative emotions comprising anger, disgust, fear, and sadness rates, shows a contrasting distribution clustered towards the higher end. Overall, such distributions imply that, while any given piece of content may only exhibit one or two emotional cues at most, those that are detected are typically negative. Therefore, the content related to #STC on Parler is likely to convey negative emotions.

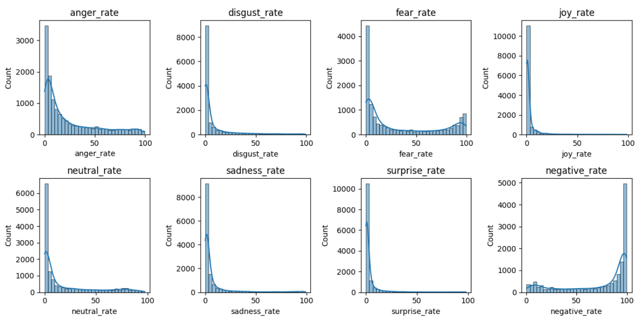

Figure 1: **Histograms of Emotional cues**

Figures 2–4 show time series trends aggregated at the month level to further explore how variables of interest have fluctuated. Figure 2 illustrates the number of #STC-related contents over time. The sharp rise in content volume starting in 2020 indicates that the online #STC campaign flourished at that time, and there was increased discussion about the topic on the Parler. Figure 3 depicts the average rate of different emotional cues over time. The fluctuations before 2020 seem to be because there was not a lot of content available before then. However, there was a clear upward trend in fear over time.

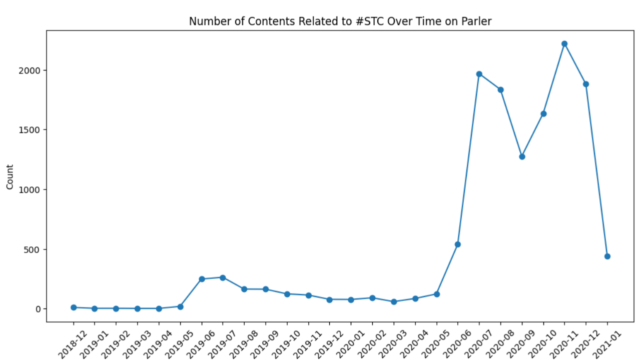

Figure 2: **Number of contents shared on Parler over time**

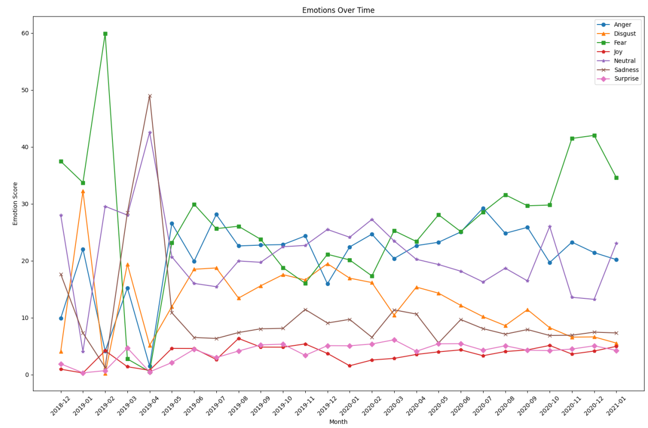

Figure 3: **Average emotional cues of contents shared on Parler over time**

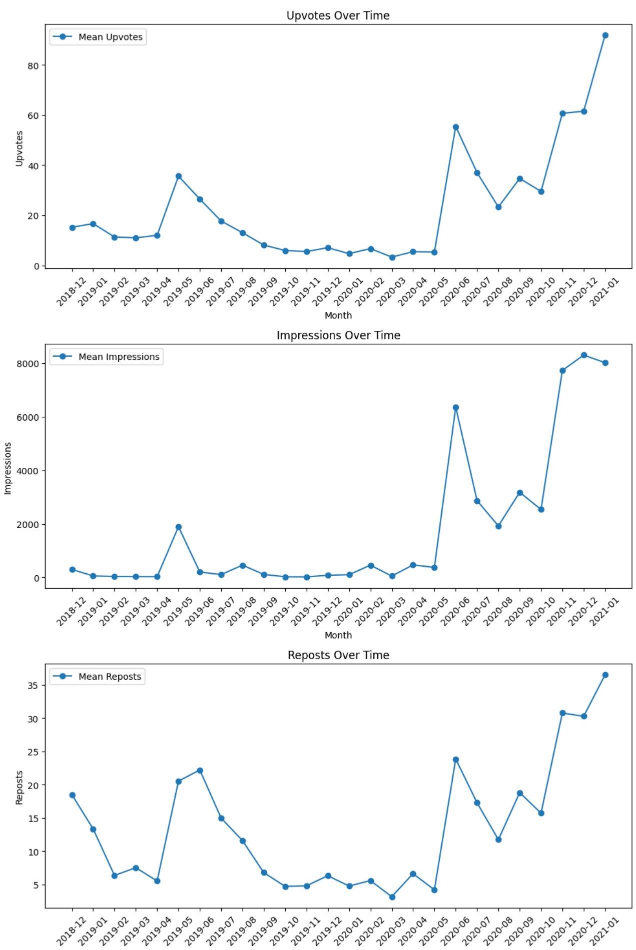

Figure 4: **Average engagements of contents shared on Parler over time**

Figure 4 presents the trends in mean upvotes, impressions, and reposts per content. The spikes in these engagement metrics may be due to the increases in content volume shown in Figure 2, implying that as more content is produced, it naturally attracts more attention and interaction from users. Yet, these spikes may also be correlated to the emotional trends seen in Figure 3. In this case, content eliciting specific emotions, particularly fear, may foster user engagement. Overall, the presence of specific emotional cues that may resonate with users could result in increased engagement with the content.

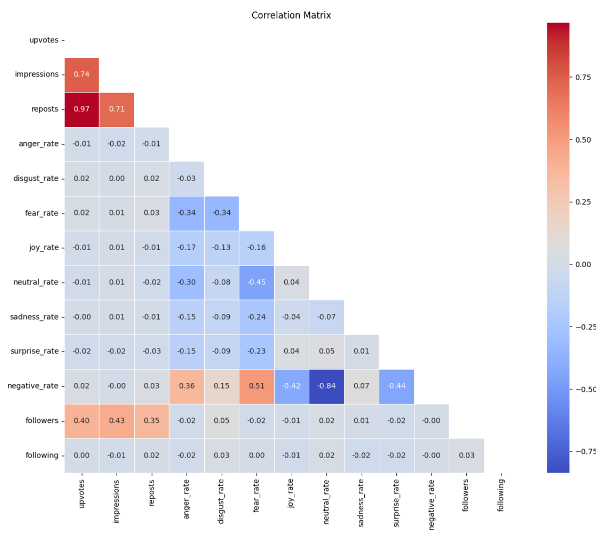

Figure 5: **Correlation plot**

While aggregated time trends can be useful, the main unit of analysis is at the content level. Thus, it is worthwhile to investigate the correlations between variables, as shown in Figure 5. Overall, there is a positive relationship between the number of followers and engagement metrics, implying that users with more followers receive more engagements. Aside from that, emotional cues have little to no correlation with engagements.

4.2 Main Analysis

This study conducted zero-inflated negative binomial regression analyses to test the hypotheses. The results are shown in Figure 6. Tables are presented in the Appendix.

Figure 6 shows the marginal effects with 95% confidence intervals. Green bars indicate positive associations, while red bars indicate negative associations. Generally, the number of followers had the largest marginal effects on all engagement metrics. Consistent with hypotheses 1 and 3, *negative_rate* and *fear_rate* had positive and statistically significant marginal effects on the engagement metrics, indicating that content laden with negative emotions or fear was more likely to engage users on the platform. Contrary to hypothesis 2, *anger_rate* had negative marginal effects on engagement metrics, showing that higher levels of anger did not necessarily increase user engagement.

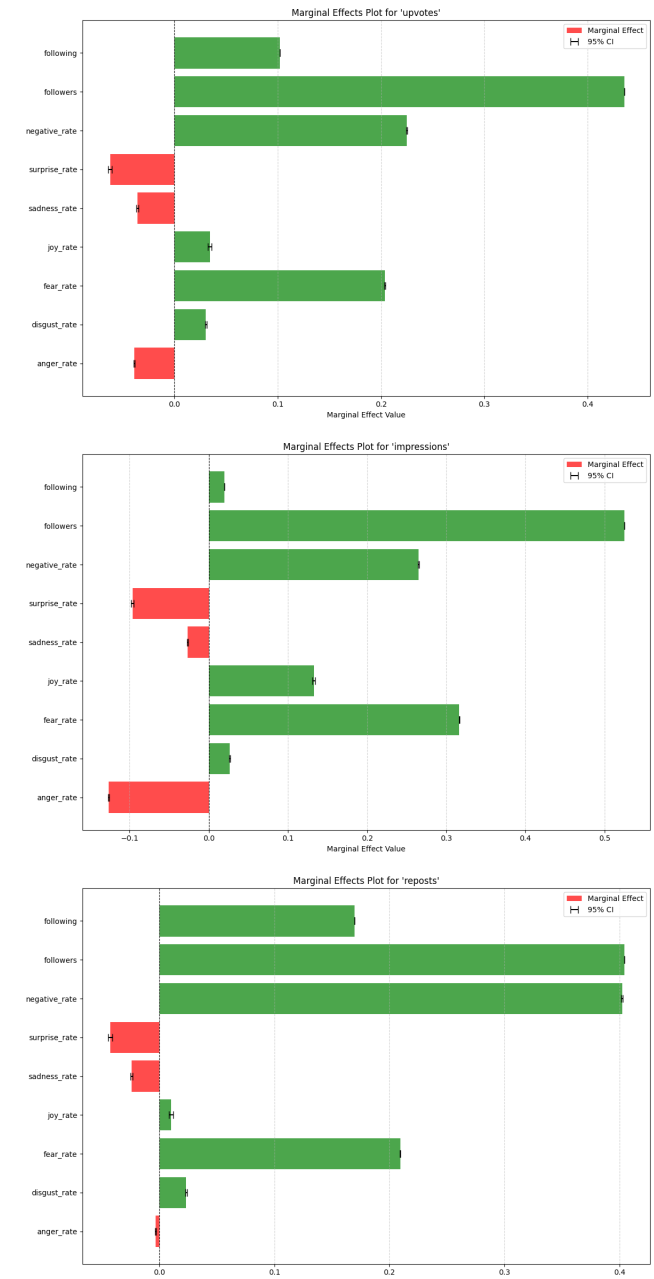

Figure 6: **Marginal effects plot**

5 Limitations and Future Work

This study has some limitations. First, the findings are at best correlational, which is due to the observational nature of the data and methods. Future research can employ advanced causal inference techniques like propensity score matching, or conduct experiments to address this issue.

Second, the Parler dataset does not represent public opinion on #STC, but rather a subset of more extreme opinions. This selection bias limits the generalization of this study's findings. Future studies could incorporate data from a variety of social media platforms to mitigate the bias of using data from a platform that may not represent general public opinion.

Finally, this study did not distinguish between original posts and user comments. Posts frequently set the tone for discussion and can serve as a barometer for the framing strategy of the campaign, whereas comments typically reflect people's responses and engagement. Future research could benefit from a more granular investigation of these two types of content separately to uncover the nuanced dynamics in the collective mobilization process.

REFERENCES

[1] Ekaterina Zhuravskaya, Maria Petrova, & Ruben Enikolopov. (2020). Political Effects of the Internet and Social Media. Annual Review of Economics, 12(1), 415-438.

[2] Cody Buntain, Michael Deal Barlow, Matthew Bloom, & Michael A. Johns. (2022). Paved with Bad Intentions: QAnon’s Save the Children Campaign. Journal of Online Trust and Safety, 1(2).

[3] William J. Brady, Julian A. Wills, John T. Jost, Joshua A. Tucker, & Jay J. Van Bavel. (2018). Emotion shapes the diffusion of moralized content in social networks. Proceedings of the National Academy of Science, 114(28), 7313–7318.

[4] William J. Brady, Jackson C. Jackson, Brindusa Lindström, & Molly Crockett. (2023). Algorithm-Mediated Social Learning in Online Social Networks [Preprint]. Open Science Framework.

[5] William J. Brady, Kevin L. McLoughlin, Marilie P. Torres, Kevin F. Luo, Maria Gendron, & Molly J. Crockett. (2023). Overperception of moral outrage in online social networks inflates beliefs about intergroup hostility. Nature Human Behaviour.

[6] Ted Brader & George E. Marcus. (2013). Emotion and political psychology. In Leonie Huddy, David O. Sears, & Jack S. Levy (Eds.), The Oxford Handbook of Political Psychology (pp. 165–204). New York, NY: Oxford University Press.

[7] David Redlawsk & Daniel Pierce. (2017). Emotions and Voting. In Kai Arzheimer, Jocelyn Evans, & Michael S. Lewis-Beck (Eds.), Sage Handbook of Electoral Behaviour (pp. 406–432). Thousand Oaks, CA: SAGE.

[8] Milton Lodge & Charles Taber. (2005). The automaticity of affect for political candidates, groups, and issues: An experimental test of the hot cognition hypothesis. Political Psychology, 26(3), 455–482.

[9] Stuart Soroka, Patrick Fournier, & Lilach Nir. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. Proceedings of the National Academy of Sciences, 116(38),

[10] Richard S. Lazarus. (1991). Progress on a cognitive-motivational-relational theory of emotion. American Psychologist, 46(8), 819–834.

[11] Ted Brader. (2006). Campaigns for Hearts and Minds: How Emotional Appeals in Political Ads Work. Chicago, IL: University of Chicago Press.

[12] George E. Marcus, W. Russell Neuman, & Michael MacKuen. (2000). Affective Intelligence and Political Judgment. Chicago, IL: University of Chicago Press.

[13] George E. Marcus, Nicholas A. Valentino, Pavlos Vasilopoulos, & Michel Foucault. (2019). Applying the Theory of Affective Intelligence to Support for Authoritarian Policies and Parties. Political Psychology, 40, 109-139.

[14] Cengiz Erisen. (2020). Anger in Political Decision Making. Oxford Research Encyclopedia of Politics. Oxford University Press.

[15] Jan-Willem van Prooijen & Karen M. Douglas. (2017). Conspiracy theories as part of history: The role of societal crisis situations. Memory Studies, 10(3), 323–333.

[16] Kathryn Chen, Sun Joo Kim, Qian Gao, & Sebastian Raschka. (2022). Visual Framing of Science Conspiracy Videos: Integrating Machine Learning with Communication Theories to Study the Use of Color and Brightness. Computational Communication Research, 4(1).

[17] Max Aliapoulios, Emmi Bevensee, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, & Savvas Zannettou. (2021). An Early Look at the Parler Online Social Network. arXiv. eprint 2101.03820.

[18] Jochen Hartmann. (2022). Emotion English DistilRoBERTa-base. Hugging Face. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base

APPENDIX

Figure A: **Histogram of log-transformed dependent variables**

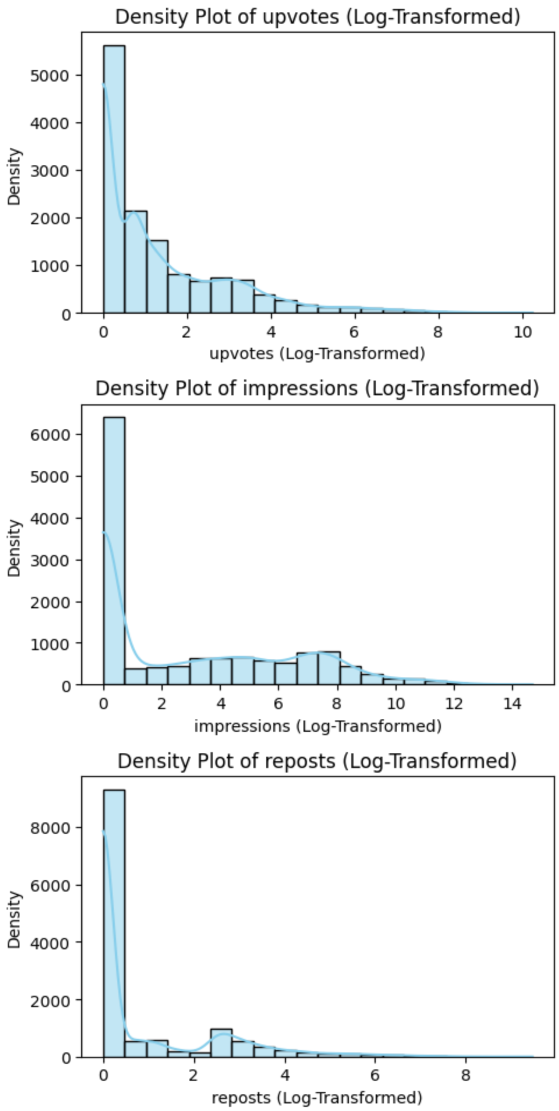

Table A: **Regression Tables**

Dep. Variable:	upvotes	No. Observations:	13428
Model:	GLM	Df Residuals:	13419
Model Family:	NegativeBinomial	Df Model:	8
Link Function:	Log	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-48454.
Date:	Mon, 11 Dec 2023	Deviance:	3.8257e+06
Time:	18:50:12	Pearson chi2:	5.41e+05
No. Iterations:	100	Pseudo R-squ. (CS):	0.8922
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
const	1.6666	0.032	52.785	0.000	1.605	1.729
anger_rate	-0.0016	0.000	-5.512	0.000	-0.002	-0.001
disgust_rate	0.0032	0.000	8.158	0.000	0.002	0.004
fear_rate	0.0062	0.000	28.781	0.000	0.006	0.007
joy_rate	0.0084	0.001	10.010	0.000	0.007	0.010
sadness_rate	-0.0047	0.000	-10.556	0.000	-0.006	-0.004
surprise_rate	-0.0135	0.001	-15.032	0.000	-0.015	-0.012
negative_rate	0.0031	0.000	9.924	0.000	0.002	0.004
followers	2.175e-05	5.29e-08	411.041	0.000	2.16e-05	2.19e-05
following	1.576e-05	4.1e-07	38.455	0.000	1.5e-05	1.66e-05

Dep. Variable:	reposts	No. Observations:	13428
Model:	GLM	Df Residuals:	13419
Model Family:	NegativeBinomial	Df Model:	8
Link Function:	Log	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-41381.
Date:	Mon, 11 Dec 2023	Deviance:	1.2159e+06
Time:	18:50:14	Pearson chi2:	6.65e+05
No. Iterations:	100	Pseudo R-squ. (CS):	0.8553
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.8522	0.034	25.390	0.000	0.786	0.918
anger_rate	-0.0001	0.000	-0.472	0.637	-0.001	0.000
disgust_rate	0.0025	0.000	6.005	0.000	0.002	0.003
fear_rate	0.0064	0.000	28.677	0.000	0.006	0.007
joy_rate	0.0024	0.001	2.689	0.007	0.001	0.004
sadness_rate	-0.0032	0.000	-6.892	0.000	-0.004	-0.002
surprise_rate	-0.0093	0.001	-9.610	0.000	-0.011	-0.007
negative_rate	0.0055	0.000	16.797	0.000	0.005	0.006
followers	2.019e-05	5.3e-08	381.226	0.000	2.01e-05	2.03e-05
following	2.623e-05	4.11e-07	63.852	0.000	2.54e-05	2.7e-05

Dep. Variable:	impressions	No. Observations:	13428
Model:	GLM	Df Residuals:	13419
Model Family:	NegativeBinomial	Df Model:	8
Link Function:	Log	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-1.0892e+05
Date:	Mon, 11 Dec 2023	Deviance:	1.0778e+09
Time:	18:50:13	Pearson chi2:	3.05e+06
No. Iterations:	100	Pseudo R-squ. (CS):	0.9180
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
const	6.0757	0.030	205.911	0.000	6.018	6.134
anger_rate	-0.0053	0.000	-18.976	0.000	-0.006	-0.005
disgust_rate	0.0028	0.000	7.362	0.000	0.002	0.004
fear_rate	0.0097	0.000	46.915	0.000	0.009	0.010
joy_rate	0.0324	0.001	40.684	0.000	0.031	0.034
sadness_rate	-0.0035	0.000	-8.483	0.000	-0.004	-0.003
surprise_rate	-0.0210	0.001	-27.020	0.000	-0.022	-0.019
negative_rate	0.0036	0.000	12.470	0.000	0.003	0.004
followers	2.621e-05	5.29e-08	495.693	0.000	2.61e-05	2.63e-05
following	2.996e-06	4.07e-07	7.360	0.000	2.2e-06	3.79e-06