Occasional Paper

ISSUE NO. 350 **MARCH 2022**

Social Media and Political Leaders: An Exploratory Analysis

Shamika Ravi and Mudit Kapoor

Abstract

Politicians have always used the media to communicate with the electorate. What makes social media unique is the scale, speed, and minimal cost at which leaders can do this task. This paper analyses the relationship between political leaders and the microblogging site, Twitter. It provides a comprehensive and systematic analysis of the use of Twitter by two contemporary Indian political leaders—Narendra Modi and Rahul Gandhi. It describes trends on how often these leaders use Twitter and how their audiences respond; explores the relationship between the sentiment of a tweet and its amplification; and studies the impact of a change in Twitter global policy in late 2020 that added friction before users could engage with tweets. This framework of analysis can easily be used for political parties and other officials who use the platform for outreach.

Introduction

Amidst the ongoing digital revolution, social media platforms such as Twitter and Facebook are permanently altering the nature of political discourse in many parts of the world. This is particularly true in democracies, where freedoms of speech and expression are guaranteed, albeit in varying degrees. Across the world, political leaders, like corporations, use these platforms regularly to reach out to large populations. Politicians in the past have exploited traditional media platforms such as newspapers, radio, and television to serve their political interests. Social media channels are unique because of the scale, speed, and minimal cost at which political leaders can directly—i.e., independent of their office—reach out to people to disseminate their message.

Twitter, a microblogging website, is a key actor in this phenomenon. Its stated purpose is to provide a "fast, free, and fun" platform for safe and free expression.[1] However, political thinkers like Ronald Deibert believe that these platforms, if unchecked, could eventually lead society to an era of "digital unfreedom".[2] Deibert outlines what he calls the "three painful truths" surrounding political authority and social media. First, social media platforms are private enterprises whose primary objective is to maximise shareholders' value and profitability. Their business model is centred around pushing advertisements to individuals by continuous and clever surveillance of personal data. Second, even though individuals have "consented" to social media "unwittingly," these platforms are designed as "addiction machines" and tend to manipulate by appealing to emotions at the expense of reason. Manipulation masquerades as persuasion. Third, the algorithms that induce attention are the same instruments that facilitate authoritarian and anarchist tendencies to seduce ordinary people by sowing seeds of "confusion, ignorance, prejudice, and chaos" to "manipulate and undermine authority."

More recent empirical research[3] using data provided by Twitter itself, validates some of these concerns and sheds light on the dark side of human nature. Vosoughi and co-authors[3] have shown that false news, typically more novel than the truth, perhaps to attract attention, "diffused significantly faster, farther, deeper, and more broadly" than the truth in all categories of information. This effect was more pronounced for political news. For example, truth takes six times longer to reach 1,500 people than falsehoods. Interestingly, they show that falsehoods were more likely to be spread by humans than by robots.

In light of this, it becomes imperative to understand the relationship between political leaders and social media platforms. This paper provides a comprehensive and systematic analysis of the use of Twitter by two contemporary Indian political leaders: Narendra Modi (NM), the prime minister of India—the largest democracy in the world; and Rahul Gandhi (RG), the de facto leader of the country's largest opposition party, the Indian National Congress. NM

has more than 75 million followers on Twitter, while RG, over 20 million. This analysis is a pioneer in that it offers a sentiment analysis of the tweets posted by the two leaders. It also seeks to understand whether tweets with different sentiments (positive, negative, and/or neutral) are amplified differently in terms of the number of retweets and likes (counts of 'favorites').

The paper has three primary objectives. First, it will outline comprehensive and systematic trends on (a) how often the two political leaders use Twitter to communicate with a larger audience, and (b) how many retweets, likes, and quotes are generated by their tweets. Second, it presents a sentiment analysis of both political leaders' tweets as mentioned earlier. The analysis uses *VADER* (Valence Aware Dictionary for sEntiment Reasoning), a parsimonious rule-based model for sentiment analysis of social media text.[4] Each tweet's sentiment score was computed to determine whether it reflects a positive, negative, or neutral sentiment.

The authors then use regression analysis to determine the relationship between the sentiment of a tweet and its amplification based on the number of retweets and likes. In particular, the analysis investigates whether tweets with negative sentiments are more (or less) likely to be amplified. The third objective is in the context of the US presidential election in 2020, ahead of which Twitter altered its global policy on retweets and likes[5] to encourage more careful consideration before a tweet is amplified. Twitter added some friction before people could retweet, and prevented "liked by" and "followed by" recommendations from people they did not follow. This paper studies the impact of these policy changes on the number of retweets and likes generated from a tweet by the two political leaders. Furthermore, it explores the differential effect of this change in policy on the amplification of tweets based on their sentiment. While this paper focuses on these two specific political leaders, the analysis can be easily replicated for political parties and other government officials who use this platform for political and social outreach.

Introduction

" Empirical research has shown that on social media, truth can take six times longer to reach 1,500 people than falsehoods. "

Description of Data

The timeline data of tweets for the analysis comes from using *snscrape*—a python-based scraper for social networking services.[6] The data contains the date and time when the tweet was created, the tweet's text, the tweet's language, whether the tweet was a retweet or not, the number of retweets, likes, and quotes that a tweet generates. It also contains data on the number of followers. Each tweet has a unique identification number which is called *status id*.

This paper limits itself to tweets posted between 1 January 2019 and 31 December 2021. It includes only tweets they post, and excludes retweets (or tweets generated by others). During this period, 11,312 tweets were posted by the two leaders—1,835 tweets by RG (16 percent of all tweets analysed), and 9,477 (84 percent) tweets by NM. The study considered four outcomes: (i) total engagement, or the sum of retweets, likes, and quotes; (ii) retweets; (iii) likes; and (iv) quotes.

> " This paper studies the tweets posted by Modi and Gandhi between 1 January 2019 and 31 December 2021. It covers only tweets they post, and not retweets. "

Sentiment Analysis

There is hardly any doubt that words play an important role in politics. Political leaders with many years of experience become more adept at choosing their words carefully to express opinions, propose reforms, or oppose their peers, especially the incumbent. Thus, they are of academic interest to political scientists.[7] One possible way to analyse political text is by focusing on the sentiment of words (positive or negative) by using sentiment lexicons such as Linguistic Inquiry and Word Count (LINW),[8] General Inquirer,[9] or that developed by Zhang, Wang, and Liu,[10] where the emphasis is on the semantic orientation of the word, independent of the context. Other lexiconic approaches, such as Affective Norms for English Words (ANEW)[11] and SenticNet[12] focus on words and the associated intensity of sentiments.

However, these approaches are not suitable for lexical features of sentiments in microblog-like contexts such as that of Twitter, where opinions are expressed in fewer characters. In the modern digital era, political leaders use microblogging sites such as Twitter to express themselves frequently in a very limited number of words. Therefore, this paper performs a sentiment analysis of tweets based on *VADER*.[4] *VADER* is also useful because it considers grammatical and syntactical conventions that are typically used for expressing the intensity of the sentiment.

Furthermore, research using *VADER* suggests that it outperforms individual human raters in classifying sentiments of tweets into 'positive', 'negative', and 'neutral' classes. Additionally, it is easy to implement, is quick, and computationally economical "without sacrificing accuracy." For further details on *VADER*, see the paper by Hutto and Gilbert.[13]

Much of the sentiment analysis is performed in the English language. In the context of India, this poses a challenge because many tweets posted by the two leaders—28 percent of NM's and 49 percent of RG's—were not in English. Overall, the analysis used 7,778 tweets in the English language, divided as 933 tweets by RG, and 6,845 by NM.

> **"** Political leaders choose their words carefully to express opinions, propose reforms, or oppose their peers. **"**
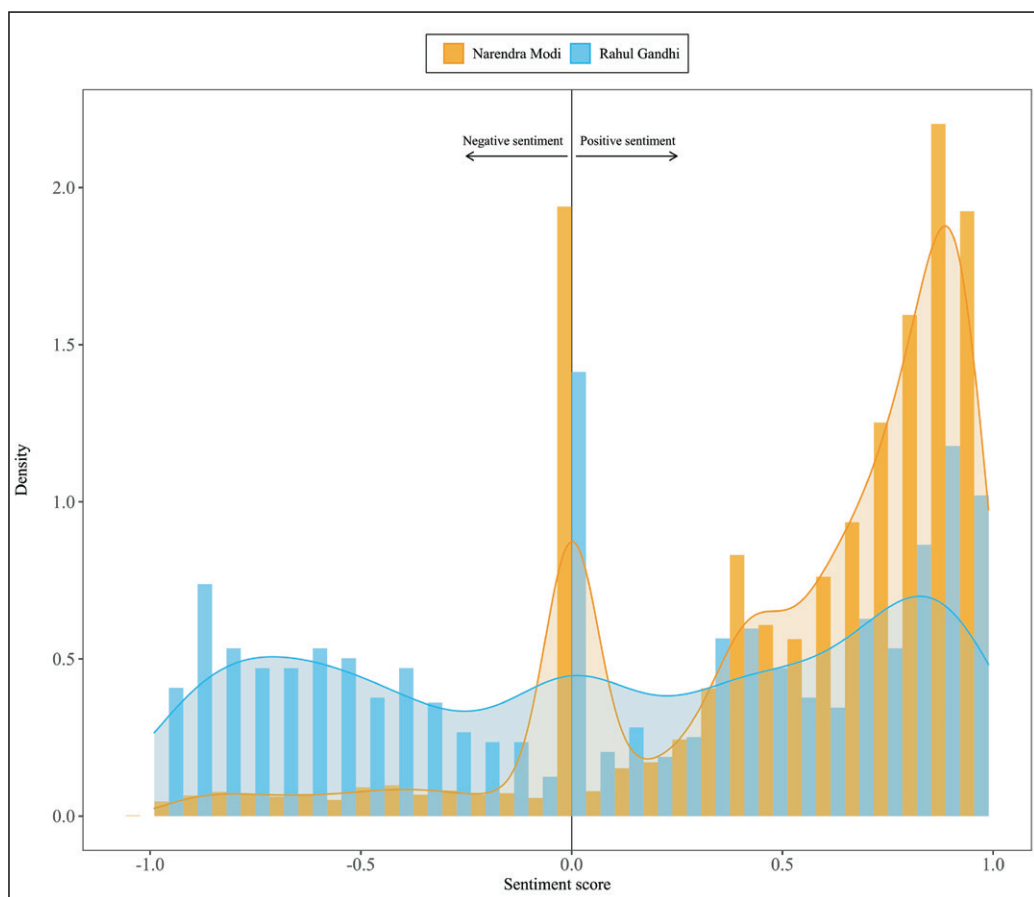
Statistical Analysis and Results

The statistical analysis for total engagement, retweets, quotes, and likes was based on negative binomial regression, given that the dependent variable was a non-negative count.[14] The authors used STATA 16 for the analysis.[15] The average values were computed using the *margins* command in STATA, and the standard errors were estimated using the delta method.

As of 31 December 2021, Narendra Modi (NM) had more than 75 million followers on Twitter compared to Rahul Gandhi's (RG) over-20 million. From 1 January 2019 to 31 December 2021, NM posted 9,477 tweets, while RG posted 1,835 tweets. On average, NM posted 8.6 tweets daily (95% Confidence Interval [CI]: 8.3 to 9.0), while RG posted 1.7 tweets (95% CI: 1.6 to 1.8). In terms of the total engagement (sum of retweets, likes, and quotes), a tweet by RG on average generated 57,618 engagements (95% CI: 55,694 to 59,543), while a tweet from NM on average generated 33,801 engagements (95% CI: 33,305 to 34,298).

The difference between NM and RG was statistically significant at a P-value <0.001. Similarly, looking at the average number of retweets, the authors found that a tweet by RG had 10,034 retweets (95% CI: 9,722 to 10,345), while for NM, it was 4,554 retweets (95% CI: 4,492 to 4,616). Meanwhile, in terms of likes, a tweet by RG on average generated 43,455 likes (95% CI: 41,984 to 44,927), while for NM, it was 28,095 likes (95% CI: 27,677 to 28,514). The difference for RG and NM in terms of retweets and likes was statistically significant at a P-value <0.001.

In terms of the sentiments of the tweets, the *VADER* compound score—which ranges from -1 (the most negative sentiment) to +1 (the most positive sentiment)—shows that the mean sentiment of tweets posted by NM was 0.54 (95% CI: 0.53 to 0.55), while for RG it was 0.09 (95% CI: 0.07 to 0.12). The difference was statistically significant, where P value <0.001.

Figure 1

Distribution of sentiment score of English-language tweets by Narendra Modi and Rahul Gandhi

Sentiment Scores: NM vs. RG

Table 1, and Figures 2a and 2b plot the average number of daily tweets and the mean sentiment score over time for both the leaders, NM and RG.

Table 1
Summary

	Narendra Modi	Rahul Gandhi	P-value (Difference)
Total followers	75+ million	20+ million	
Tweets	9,477	1,835	
Daily tweets			
Average	8.6 (8.3 to 9.0)	1.7 (1.6 to 1.8)	<0.001
Tweet amplification (Average)			
Total engagements	33,801 (33,305 to 34,298)	57,618 (55,694 to 59,543)	<0.001
Retweet	4,544 (4,492 to 4,616)	10,034 (9,722 to 10,345)	<0.001
Likes	28,095 (27,677 to 28,514)	43,455 (41,984 to 44,927)	<0.001
Quotes	217 (212 to 222)	820 (779 to 861)	<0.001
Sentiment score of tweets*			
Mean	0.54 (0.53 to 0.55)	0.09 (0.07 to 0.12)	<0.001

*Note: *Sentiment score is based on VADER (Valence Aware Dictionary for sEntiment Reasoning).[4] It ranges from -1 to +1, where negative scores reflect negative sentiments and a positive score a positive sentiment. 95% confidence intervals are in parentheses.*

Statistical Analysis
and Results

Figure 2a

**Average number of daily tweets by
Narendra Modi and Rahul Gandhi
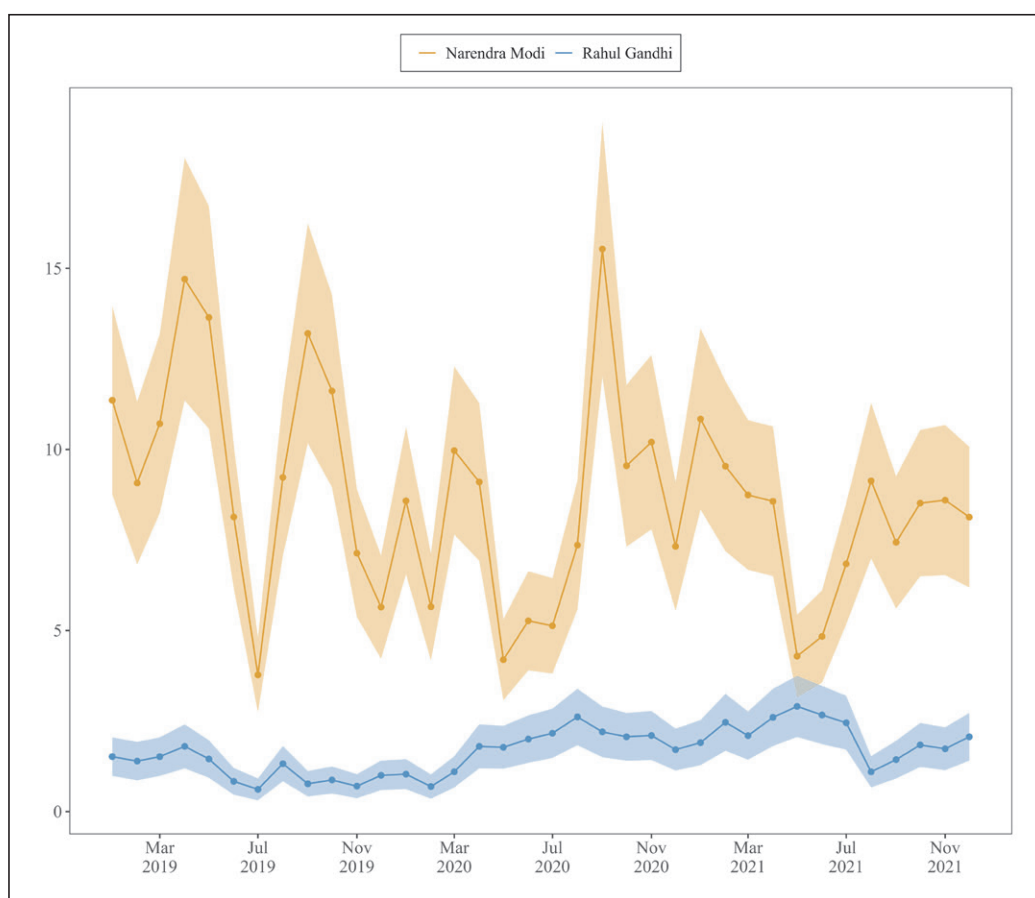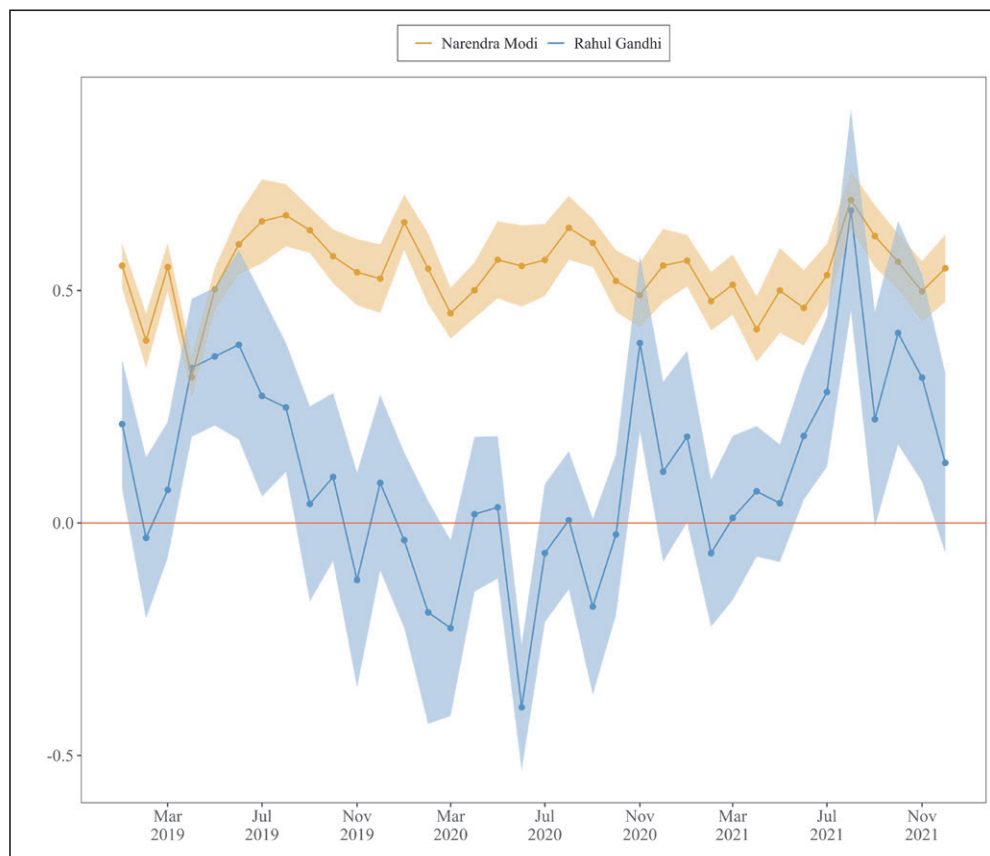(March 2019 – November 2021)**

Figure 2b
Average sentiment score of tweets by Narendra Modi and Rahul Gandhi (March 2019 – November 2021)

Impact of Twitter Policy Change: NM vs. RG

The next set of results was based on the impact of a policy change implemented by Twitter in October 2020 on total engagement, retweets, likes, and quotes for both leaders. The analysis found that the policy change had a negative impact on the amplification of the tweet in terms of total engagements, retweets, likes, and quotes. However, the impact was not uniform for both the leaders. For Rahul Gandhi, the total engagements on average declined from 65,123 (95% CI: 53,721 to 76,524) to 44,880 (95% CI: 37,104 to 52,656), a drop of approximately -31% (95% CI: -51% to -3.0%). However, for Narendra Modi, the decline in total engagements on average was from 36,354 (95% CI: 33,872 to 38,836) to 31,533 (95% CI: 28,702 to 34,364), a decline of approximately -13% (95% CI: -25% to 1.3%), which was statistically insignificant at the conventional 95% confidence interval.

Statistical Analysis and Results

Statistical Analysis and Results

Table 2

Impact of policy change on total engagements, retweets, likes, and quotes for Narendra Modi and Rahul Gandhi

	Prior to policy change	Post-policy change	Percent change
Total Engagements			
Narendra Modi	36,354 (33,872 to 38,836)	31,533 (28,702 to 34,364)	-13% (-26% to 1%)
Rahul Gandhi	65,123 (53,721 to 76,524)	44,880 (37,104 to 52,656)	-31% (-51% to -3%)
Retweets			
Narendra Modi	5,113 (4,768 to 5,458)	3,991 (3,675 to 4,308)	-22% (-32% to -10%)
Rahul Gandhi	12,518 (10,265 to 14,771)	7,060 (6,044 to 8,077)	-44% (-59% to -22%)
Likes			
Narendra Modi	30,074 (28,007 to 32,140)	26,422 (24,015 to 28,828)	-12% (-25% to 3%)
Rahul Gandhi	47,685 (39,338 to 56,031)	34,976 (28,782 to 41,171)	-27% (-48% to 4%)
Quotes			
Narendra Modi	193 (179 to 206)	289 (239 to 338)	50% (18% to 89%)
Rahul Gandhi	962 (738 to 1,186)	594 (420 to 768)	-38% (-63% to 4%)

Figure 3a: Total engagements, retweets, likes, and quotes prior to, and post-Twitter policy change
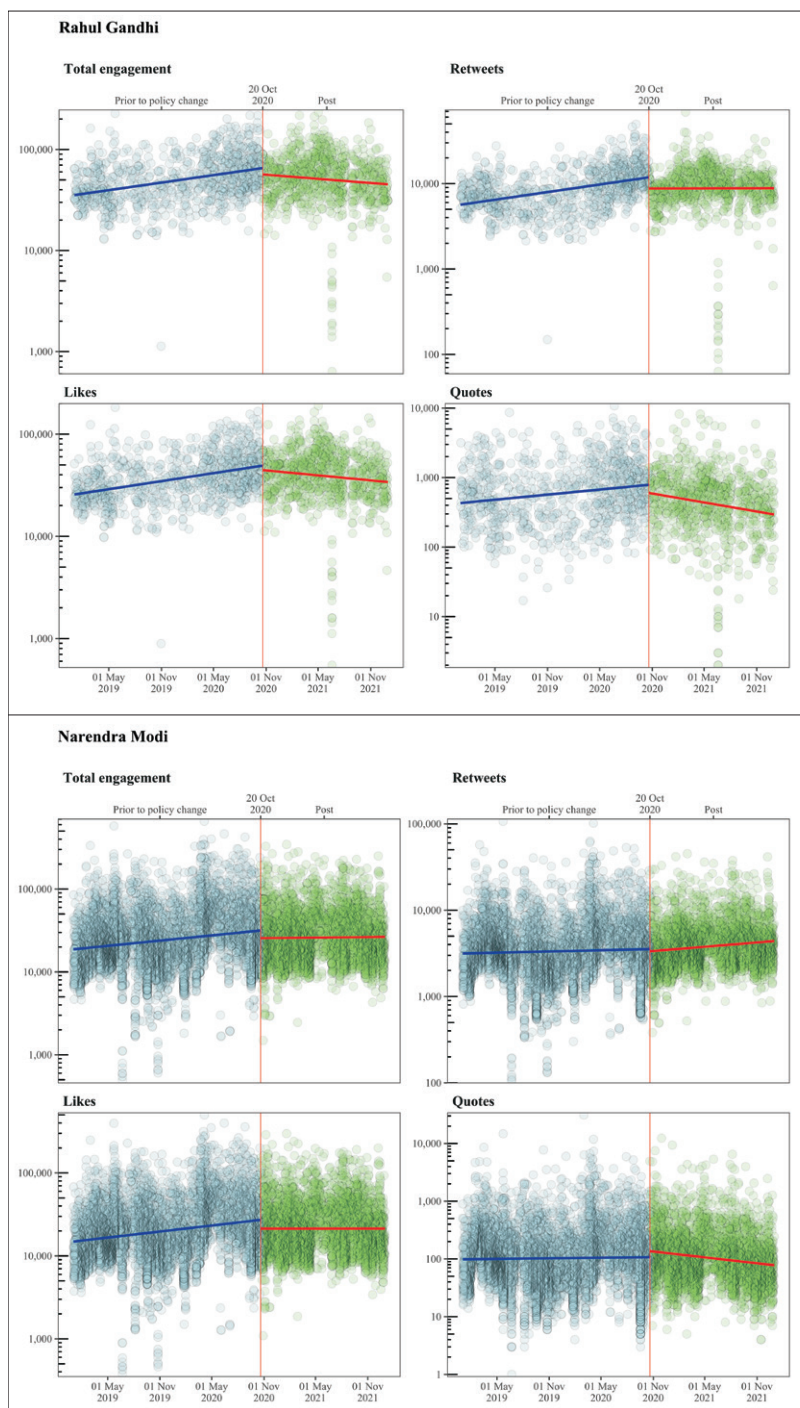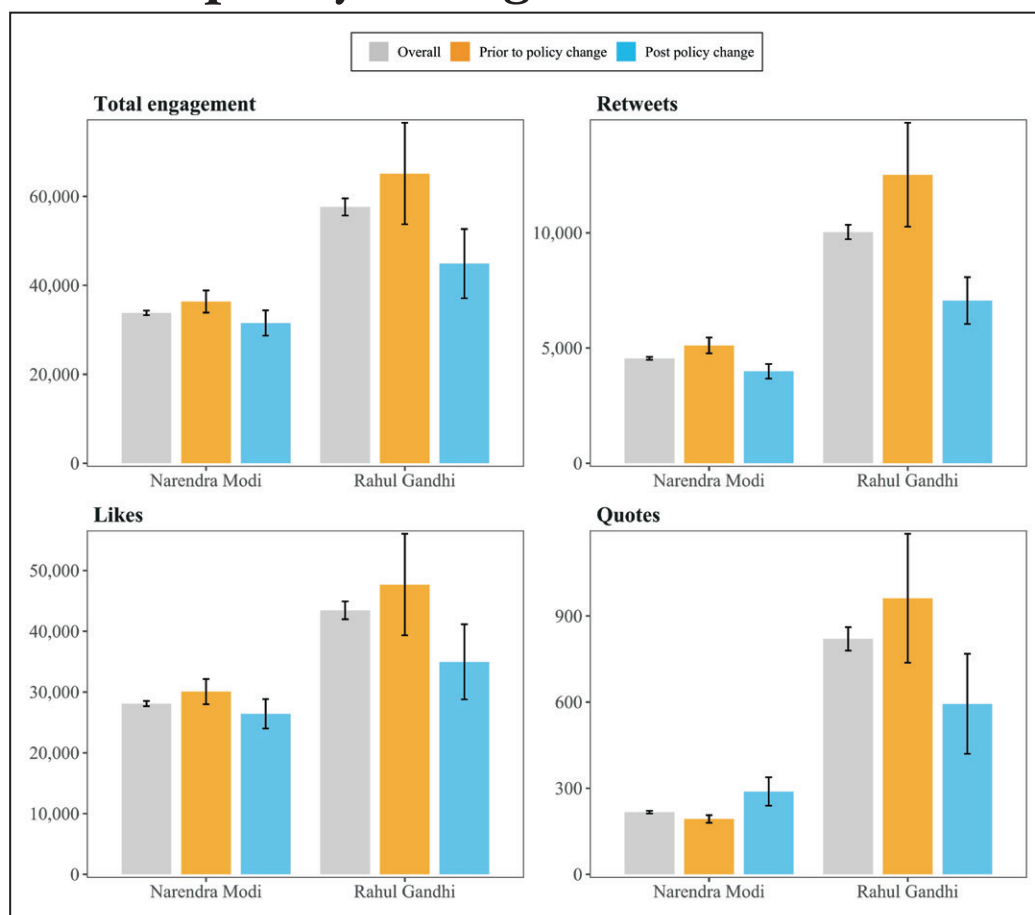
Statistical Analysis
and Results

Figure 3b
Change in Total engagements, retweets, likes, and quotes prior to, and post-Twitter policy change

Amplification Post-Twitter Policy Change

The next set of results related to the sentiment score of the tweet, which is the VADER compound score, and its amplification in terms of retweets, likes, and quotes, for both the leaders. For Narendra Modi, before the policy change, a tweet with negative sentiment (Vader compound score of -0.7) on average generated 6,010 retweets (95% CI: 5,352 to 6,668), while a tweet with positive sentiment (Vader compound score of 0.7) generated 4,981 retweets (95% CI: 4,584 to 5,378). Meanwhile, for Rahul Gandhi, before the policy change, a tweet with negative sentiment (Vader compound score of -0.7) on average generated

11,829 retweets (95% CI: 9,232 to 14,427), while a tweet with positive sentiment (Vader compound score of 0.7) generated 9,492 retweets (95% CI: 7,418 to 11,566).

The analysis also found that due to Twitter's policy amendment, for Rahul Gandhi, there was a decline in the average number of retweets for a tweet with negative sentiment. In contrast, there was a decline for tweets with positive sentiment, but it was statistically insignificant at the conventional 95% confidence interval. However, the decline was higher for tweets with negative sentiment than tweets with positive sentiment: -42% (95% CI: -64% to -6%) versus -31% (95% CI: -57% to 12%). While for Narendra Modi, the decline in retweets was much lower for tweets with negative sentiment than for those with positive sentiment; -3% (95% CI: -23% to 21%) versus -24% (95% CI: -36% to -9%).

Table 3
Total engagement, retweets, likes, and quotes according to Vader score, prior to policy change, post-policy change, and percentage difference

Narendra Modi

	Vader score (compound)	Prior to policy change	Post-policy change	Percentage change
Total engagement	-0.7	37,880 (33,617 to 42,142)	46,558 (38,453 to 54,664)	23% (-4% to 57%)
	0	30,916 (28,302 to 33,530)	27,722 (24,381 to 31,064)	-10% (-26% to 9%)
	0.7	34,863 (32,086 to 37,639)	30,088 (26,568 to 33,607)	-14% (-29% to 5%)

Statistical Analysis and Results

Retweets	-0.7	6,010 (5,352 to 6,668)	5,808 (4,890 to 6,725)	-3% (-23% to 21%)
	0	4,809 (4,405 to 5,212)	3,822 (3,408 to 4,235)	-21% (-34% to -4%)
	0.7	4,981 (4,584 to 5,378)	3,798 (3,400 to 4,196)	-24% (-36% to -9%)
Likes	-0.7	31,581 (27,987 to 35,176)	40,213 (33,070 to 47,356)	27% (-1% to 63%)
	0	25,959 (23,743 to 28,174)	23,638 (20,729 to 26,547)	-9% (-25% to 11%)
	0.7	29,732 (27,345 to 32,119)	26,056 (22,944 to 29,169)	-12% (-28% to 7%)
Quotes	-0.7	270 (227 to 312)	630 (446 to 814)	134% (59% to 242%)
	0	194 (173 to 215)	255 (199 to 311)	32% (-4% to 80%)
	0.7	192 (173 to 210)	221 (174 to 269)	15% (-15% to 57%)

Statistical Analysis and Results

Rahul Gandhi

	Vader score (compound)	Prior to policy change	Post-policy change	Percentage change
Total engagement	-0.7	53,649 (42,424 to 64,873)	38,685 (25,933 to 51,437)	-28% (-57% to 21%)
	0	50,562 (39,984 to 61,139)	45,377 (30,462 to 60,292)	-10% (-46% to 50%)
	0.7	52,709 (41,730 to 63,688)	45,424 (30,770 to 60,078)	-14% (-49% to 44%)
Retweets	-0.7	11,829 (9,232 to 14,427)	6,823 (4,866 to 8,780)	-42% (-64% to -6%)
	0	10,500 (8,204 to 12,796)	7,605 (5,438 to 9,771)	-28% (-55% to 18%)
	0.7	9,492 (7,418 to 11,566)	6,540 (4,722 to 8,358)	-31% (-57% to 12%)
Likes	-0.7	40,882 (32,416 to 49,348)	31,508 (20,786 to 42,230)	-23% (-54% to 30%)
	0	39,130 (31,020 to 47,240)	37,305 (24,639 to 49,970)	-5% (-44% to 61%)
	0.7	42,384 (33,646 to 51,122)	38,675 (25,771 to 51,580)	-9% (-46% to 54%)
Quotes	-0.7	1,102 (791 to 1,413)	447 (194 to 701)	-59% (-82% to -9%)
	0	969 (693 to 1,245)	665 (286 to 1,044)	-31% (-70% to 55%)
	0.7	694 (499 to 889)	523 (231 to 815)	-25% (-67% to 70%)

Discussion and Conclusion

Political leaders are known to carefully choose their words when expressing opinions, proposing policies, or offering contrary views from other officials. Political scientists have always been interested in analysing these words. In the contemporary digital era, political leaders are shifting from traditional platforms such as the radio, television, or newspapers, to newer social media platforms such as Twitter and Facebook, which can often provide more direct access to the public. Unlike traditional media—which largely operate according to implicit and explicit codes of conduct, as well as checks and balances on freedom of expression—the digital social media platforms, at least today, have no such measures in place. Indeed, one may argue that they are moving in a direction where all is permitted.

Elected officials are aware of this drift and are increasingly adopting these new platforms and adapting to their cultures. However, these platforms, often referred to as micro-blogging sites, are not designed for meaningful discussions on complex and often paradoxical issues that can have an impact on people's lives. Rather, they typically cater to affairs that will draw the mercurial attention of users who seek out these spaces primarily for entertainment.[16]

It therefore becomes important to analyse the behaviour of politicians on these digital platforms. This paper attempted this by systematically and comprehensively analysing two Indian politicians on the microblogging site, Twitter: Narendra Modi, prime minister of India and leader of the Bharatiya Janata Party, and Rahul Gandhi, Member of Parliament and de facto leader of the Indian National Congress—the leading opposition party and the oldest political party in India. The paper used *VADER* (Valence Aware Dictionary for sEntiment Reasoning) to conduct a sentiment analysis,[17] and found that tweets with a negative sentiment are more likely to be amplified primarily in terms of retweets when compared to tweets with a positive sentiment.

> "Micro-blogging sites are not designed for meaningful discussions on complex and often paradoxical issues that can impact people's lives."

Discussion and Conclusion

If a politician's objective is to attract attention, then these results indicate that tweets with negative sentiment attract more attention than those with positive sentiment. The second important finding is that platforms such as Twitter, being private enterprises that are at liberty to alter their policies, exert an outright large influence on how messages are propagated. For example, as discovered in this analysis, the change in policy to add friction before a tweet is amplified, caused a significant drop in the number of retweets. This effect was even more substantial for tweets with a negative sentiment.

At the same time, however, this raises an important question for free societies: Who should control the propagation of political expressions—should it be left to the whims of private enterprises with minimal public oversight, or should they be brought into the ambit of regulation where an explicit code of conduct is established? As social media continue to play a more significant role in public and political life, these issues will become central to debates surrounding these platforms and political discourse. ORF

Dr Shamika Ravi *is Vice President at ORF.* **Mudit Kapoor** *is Associate Professor at Indian Statistical Institute, Delhi.*

Endnotes

1   "About Twitter | Our company purpose, principles, leadership," https://about.twitter.com/en/who-we-are/our-company.html.

2   Ronald Diebert, "The Road to Digital Unfreedom: Three Painful Truths About Social Media," *J. Democr* 30, no. 1 (2019): 25–39, https://www.journalofdemocracy.org/articles/the-road-to-digital-unfreedom-three-painful-truths-about-social-media/.

3   S. Vosoughi, D. Roy and S. Aral, "The spread of true and false news online," *Science* 359, no. 6380 (2018): 1146–151, https://www.science.org/doi/10.1126/science.aap9559.

4   C.J. Hutto, E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media* 10, no. 1 (2014), https://ojs.aaai.org/index.php/ICWSM/article/view/14550.

5   "Additional steps we're taking ahead of the 2020 US Election," https://blog.twitter.com/en_us/topics/company/2020/2020-election-changes.html.

6   JustAnotherArchivist, snscrape, 2022.

7   J. Wilkerson and A. Casas, "Large-Scale Computerized Text Analysis in Political Science: Opportunities and Challenges," *Annu. Rev. Polit. Sci.* 20, no. 1 (2017): 529–44, http://dx.doi.org/10.1146/annurev-polisci-052615-025542.

8   Linguistic Inquiry and Word Count, http://liwc.wpengine.com/.

9   Inquirer Home Page, http://www.wjh.harvard.edu/~inquirer/.

10   L. Zhang, S. Wang and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 8, no. 4 (2018), https://doi.org/10.1002/widm.1253.

11   Center for the Study of Emotion and Attention, https://csea.phhp.ufl.edu/media/anewmessage.html.

12   SenticNet, https://sentic.net/.

13   C.J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text."

14   A. Cameron and P. Trivedi, "Microeconometrics Using Stata," *undefined*, 2009.

15   Stata | StataCorp LLC, https://www.stata.com/company/.

16   R. Dolan, J. Conduit, J. Fahy, and S. Goodman, "Social media engagement behaviour: A uses and gratifications perspective," *Journal of Strategic Marketing*

24, no. 3 (2015): 261–77, https://www.tandfonline.com/doi/full/10.1080/096525
4X.2015.1095222.

17   C.J. Hutto, E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment
Analysis of Social Media Text."

Endnotes

ORF OBSERVER RESEARCH FOUNDATION

Ideas . Forums . Leadership . Impact