

From Greetings to Corruption: Politicians, Political Parties, and Tweeting in India

Lia Bozarth, Anmol Panda, Joyojeet Pal

University of Michigan, Microsoft Research

Abstract

We present an in-depth large-scale study of topical discussions on Twitter by 1,711 Indian politicians. We find that politicians collectively more frequently indulge in establishing personal branding through low-substance, personality-focused messaging as opposed to broadcasting policy stances. Further, elites prefer to engage with news media rather than with each other when posting policy-related tweets. Additionally, compared to the party-in-power, opposition politicians collectively post more complex tweets and demonstrate higher negativity, especially regarding corruption. Finally, through contextual examination of the most retweeted messages from two key leaders - the prime minister and the leader of the largest opposition party, we find that there are qualitatively important distinctions between both the styles of key politicians - while the former focuses on positive-themed messaging, the latter employs confrontation and aggressive language with direct attacks on individuals and issues. These suggest that a healthy contrarian, albeit abusive, space for democratic discourse exists online for politicians in India.

From Greetings to Corruption: Politicians, Political Parties, and Tweeting in India

Introduction

Social media have impacted the content and channels of political communications in ways that have dramatically changed both campaign and non-campaign outreach (Baumer, Sueyoshi, & Tomlinson, 2011; Kou, Kow, Gui, & Cheng, 2017; Mascaro & Goggins, 2010). New forms of rhetorical construction, which allow politicians control on what subjects they address, facilitate both the personalization of communication at the cost of policy-relevant discussions, and create a reward mechanism through technology-mediated rapid message dissemination (G. E. Enli, 2013; McGregor, 2018; Meeks, 2016). Political actors can increasingly rely on private resources and networks to provide their own narratives (Casero-Ripoll's, , & P. Franch, 2017; G. Enli, 2017; Fuller, , & C. Fisher, 2018; Kellner, 2018), bypassing the editorial constraints of traditional news media.

Prior research, largely centered on early US and European adopters of social media, has focused on politicians' social media strategies on issues of brand building, and championing for or against salient legislative issues (Chadwick, 2017; Hemphill, Otterbacher, & Shapiro, 2013). In recent years, the growth of access to low-cost mobile devices has brought social media to centerstage for political communication in various parts of the world - with major world leaders such as Joko Widodo, Jair Bolsonaro, Rodrigo Duterte, and Narendra Modi all gaining reputations as masters of campaigning using Twitter (J. Pal, 2015).

Our paper focuses on the run up to the 2019 general election in India. We study Indian political elites' behaviors when tweeting about policy-relevant topics versus non-issues¹. We populated a list of 1711 Indian politicians on Twitter, acquired 4.6 million tweets from their accounts, and annotated each applicable tweet into one of the 6 categories of tweets: **corruption, development, inflation, technology, poverty, and greetings**.

Our paper makes the following contributions. i) To the best of our knowledge, this is the first large scale study that uses both manual and automated approach to identify thousands of

¹ Non-issue tweets include those related to birthdays or festival greetings, condolences etc. While messaging on such 'non-issue' tweets serves an important purpose such as contributing to personal branding, we argue that these are less likely to be immediately policy-relevant.

politicians in an electoral system in a Global South setting. ii) We show that politicians' messages are more frequently on casual outreach subjects such as greetings compared to issue or policy-relevant topics such as technology, development, poverty, or welfare. This suggests that politicians see value in signaling relatability through casual talk rather than lean more heavily towards a political-business centric message. iii) We find that politicians are significantly more likely to retweet each other on substantive topics compared to casual messaging. This behavior was more significant for ruling BJP party politicians which suggests that the party has a more intra-party collaboration, something that has been suggested in the news and academic articles in the past (J. Pal, 2015), but not shown empirically. Further, iv) when politicians talked about substantive issues, they also showed a stronger preference for engaging with news media compared to messaging on casual topics. This suggests strategic employment of the media by politicians when they attempt to influence specific policy areas. Finally, v) our work shows key differences in the ways ruling and opposition parties communicate. We find that the ruling BJP party is relatively less engaged in antagonistic messaging — it has more engagement with sports-related and nationalistic hashtags, while the opposition INC messages relatively more frequently on issue-centric and antagonistic hashtags aimed at the governing party when discuss corruption. Non-governing parties were also found overall to have more negative sentiments in their tweets.

Related Work

The enthusiastic embrace of social media that followed spontaneous and affect-driven organized by citizens during events such as the 'Arab Spring' (Lotan, Graeff, Ananny, , & I. Pearce, 2011; Morozov, 2009; Tusa, 2013) has since been tempered by the possibility that social media actually enhanced established political actors' ability to systematically use social media to monitor citizens (David, 2018; Gayo-Avello, 2017; Schäfer, 2016; Zittrain et al., 2017), or to feed them a preferred narrative in an organized and sustained way (Curato, 2017; Ohm, 2015). Here, we first i) survey literature on individual politician's communication strategies. Next, ii) we discuss studies that address how political parties advance party-based agendas. Finally, iii) we examine relevant research on online journalism, and the interactions

between political elites and traditional news media.

Social Media and Personalized Communication by Political Elites. Prior studies have observed an increased effort from established politicians and candidates in utilizing social media for personal branding. While most of the studies have focused on western societies, there is a growing body of work that examines the role of social media in the communications of politicians in the Global South (S. Ahmed, , & J. Cho, 2016; S. M. Ahmed, 2014; Correa, 2017; Lee, 2017,?; Lufkens, 2016; J. A. Pal, 2016). Within the context of India, Jaffrelot (2015), for instance, note that "the second largest independent PR firm in America, charged over US\$25,000 a month to manage Modi's [media] account", underscoring the importance of digital branding to Prime Minister Narendra Modi. Additionally, Pal et al. (2016) observe politicians' selective engagement with topics (policy-relevant or policy-irrelevant) in tweeting, depending on the stage of the electoral cycle. Furthermore, political candidates are not only using data mining techniques to classify potential supporters into granular groups and then personalize their campaign messages targeted at these groups accordingly, these messages are also becoming increasingly focused on "lifestyle" messages instead of being policy and issue-based (Bennett, 2012; Bimber, 2014; Howard, 2006; Kreiss, 2016).

Researchers including Bennett and Pfetsch (Bennett & Pfetsch, 2018) stressed the importance of examining and understanding the changes in political discourse within the new media environment, and how these changes affect democratic institutions at large. Luckily, the digital prints, such as policy tweets on Twitter, left by these political entities (Wilkerson & Casas, 2017), provide interested parties opportunities to study individual politicians at a unprecedented scale and depth.

Party based Collaborative Strategies on Social Media. In the west, Libby Hemphill et al (2013) show that Republican and Democratic officials collectively use distinct hashtags on Twitter to frame or counter-frame policy-relevant discussions. Active collaboration using hashtags can reinforce individual politicians' stances, or to garner momentum for discussions on these issues. Studies also show that members of a party can perform collective antagonistic action, and that this depended on their position of power. For instance, Evans et al. (2014)

studying the 2012 presidential election demonstrate that the Republican party was significantly more likely to attack President Obama than the Democratic party to Romney, suggesting a difference between the party in power and the opposing party. Additionally, political elites are more likely to interact with others from the same party (S. M. Ahmed, 2014; DâĂŹheer & Verdegem, 2014) than those from opposing parties, perhaps in an attempt to increase each others' visibility (DâĂŹheer & Verdegem, 2014).

Moreover, Boggild and Pedersen (2018) contend that while many recent studies indicate that politicians are using a more individualized approach in campaigns and elections, politicians from systems in which parties have direct control over the nomination process are less likely to engage in personalized campaigning. Thus, in India where candidate selection is typically managed by parties, social media usages by members of the same party are potentially more uniform, and the interactions or collaborations between same-party politicians more significant.

Social Media, Mainstream News Media and Journalism in India. The news media - whether print, radio, television, or indeed social media - has broadly been seen as playing the role of checks and balances against the political establishment. However, this role is predicated upon both the state of free press and the traditions of public commentary by political actors. The growth of social media in the Global South has offered a new set of continuities to work on journalism in developing countries, which has long argued that media studies needs to be de-westernized and seen as posing distinct challenges in such settings (Park, 2000). In part, the ability of social media to play an arbiter of accountability is problematised by issues such as colonial legacies of mainstream media as a propaganda tool (Agrawal, 2006) or histories of political control of public discourse through state media ownership (Rahman, 2014). For Indian, work (Rao, 2010) has shown that political journalists still treated public officials with a colonial legacy of deference.

Commentators have also noted the intersection of social media with the emergence of a sensationalism in the news in India, where investigative or probing journalism has been pushed towards dramaturgic events such as sting operations (Sundaram, 2015). Issues of public importance such as corruption and development are mediated through the performance

of outrage by celebrities at these events on social media instead (U. Rodrigues, 2014). Further, with politicians setting up their direct channels of communication on social media, their ability to be heard can be decoupled from both traditional journalism and the substantive matters of the issues these politicians are willing to address online. Other work has also shown that politicians' social media messaging reverberate on both television (Kumar, 2015) and print (Chakraborty, Pal, , & D. M. Romero, 2018). That is, politicians can dramatically reduce their non-social media output and still find voice on traditional media. Indeed, several politicians have turned social media into the primary driver of their signature campaigns (U. M. Rodrigues, 2017).

The growth of social media as a source of news has exacerbated the challenges with quality news access, creating a heteronomic effect on journalistic practices, essentially enabling political hegemons to dominate the media discourse (Maheshwari, 2018). We find elements of this trend in the domination of social media political conversation by a few key parties in this research. Studies have also argued that the weakened professional press corps has diluted the quality of discussion on issues of core policy importance — research on the Indian 2014 elections showed that the notion of development is increasingly discussed in the abstract, as a brand, rather than in specifics of issues such as poverty or inequality (Mudgal, 2015). This work aligns closely with what we are doing in this paper in terms of how issues such as corruption, development and poverty get talked about on Twitter, and how these in turn, have resonance in the public.

Summary. This rich body of work shows how a range of disciplines has approached the questions of what politicians do and say online as both individuals and collaborators in a party system. In this work, we dived deep into the what and how of politicians' speech on Twitter². Unlike prior studies that focused on a small set of key politicians or political actors

² While there are several popular social media platforms all of which garnered significant interest from the academia, in the paper, we focused on Twitter due to the following reasons. Twitter has had early success in political use by key politicians, and has been used widely by various political parties since the 2014 general elections (S. Ahmed, Jaidka, & Cho, 2016). Additionally, Twitter, as pointed out by prior work (Kwak, Lee, Park, & Moon, 2010), differs from other social media platforms due to its short effective diameter and non-power-law follower distribution. In other words, Twitter is ideal for information broadcasting. Consequently,

without an official role (e.g. bloggers or activists), we leveraged Twitter networks to classify a substantially large number of influential Indian politicians both at national and regional levels. Our study has a dataset spanning 4 years and we use both qualitative and quantitative methods to examine social media strategies of these political entities.

Data

Our dataset consisted of 4.6 million tweets in both English (2.94M) and Hindi (1.66M) from Jan, 2014 to October, 2018, contributed by 1711 Indian politicians from both state and national levels. We note that 779 accounts were Bharatiya Janata Party (BJP) politicians who supplied 2.34M tweets (or, 51.0% of total tweets); 417 accounts were from the principle national opposition the Indian National Congress (INC) politicians with aggregated sum of 1.25M tweets (or, 26.9% of total tweets); and the remaining 515 politicians from other parties posted 1.02M tweets (or, 22.1% of total tweets). The data is thus significantly skewed towards BJP politicians. This underlines the party's dominance of social media, partly led by its leadership's diktat that only politicians with over a certain threshold of followers on social media could be offered party seats for elections Bureau (2018).

The aforementioned 1.7K politician Twitter accounts are identified using an mixed-methods approach. Furthermore, we also used an cosine similarity-based unsupervised process to assign tweets into the following 7 non-exclusive categories: 1) information and communication technologies (Technology) related tweets, 2) poverty & welfare (Poverty), 3) economics and development (Development), 4) corruption, scams, and bribery (Corruption), 5) inflation and fuel price surge (Inflation), 6) greetings and holiday celebrations (Greetings), and 7) other (Other). Both of these classification procedures are described below.

Twitter Account Classification. We used both manual and supervised learning methods to generate a comprehensive list of Indian politicians on Twitter. Our classification covered a considerable number of politicians from 20 distinct parties (both national and

much research has been done to examine the adoption of Twitter by political elites. See (Jungherr, 2016) for a review. Yet, in-depth studies that focus on the Global South are still lacking.

regional) including BJP, INC, Shiv Sena, CPI(M), AAP, etc.

Manual Approach: We used Twitter's native search function to query for politician Twitter accounts using the names of elected parliamentary representatives, key party officials (such as party presidents, general secretaries), chief ministers of states, and state legislators. We manually looked through the top 20 results returned for each query and identified all accounts matched to actual politicians. This procedure gave us a total of 1002 manually collected political handles.

Supervised Learning Method: We used a 2-step process to identify accounts of additional Indian politicians. First, we built a well labeled dataset of 1002 known politicians (positive class) and the top 800 retweeted non-political accounts (negative class). We then trained a LogisticRegression binary classifier adopting the common Bag-of-Words Ngram-based approach (Joulin, Grave, Bojanowski, & Mikolov, 2016) using user profile description text³. We achieved a precision score of 0.967, recall score of 0.659 and f1 score of 0.762. In the second step, we aggregated all Twitter accounts that are followed by at least one politician in the labeled data. We then used our classifier to label these additional accounts. For accounts that are classified as politicians, we filtered out all that had fewer than 1000 tweets or 1000 followers. Finally, we manually examined all remaining accounts and discarded all false positives.

Tweet Classification. Keywords and/or Cosine similarity measurement based document clustering has been used extensively by many prior work Adamic and Glance (2005-08); Conover et al. (2011); Stieglitz and Dang-Xuan (2013) focused on political communication on social media. In this paper, we used a similar approach to cluster tweets into different issue-based and personal-appeal-based categories (details in the Appendix).

Briefly, we generate a numeric vector for each unique word. For instance, the word "govt" can have the corresponding vector of [-1.9078784 1.654422 1.8524699...]. Vectors of words that share more similar contexts are closer together in space (i.e. having higher cosine similarity score). Here, we first manually selected a handful keywords related to Technology,

³ For instance, words including "mp", "minist", and "parliamentari" in user profile have the largest positive weights, while the words "super", "andhra", and "columnist" have the largest negative weights.

such as "technology", "digital", et cetera, and then used cosine similarity scoring to discover additional keywords that share comparable semantic meanings to the selected words. Next, for each tweet, we categorized it as **Technology** if it contains 1 or more **Technology** keywords. We repeated the process for other categories. Using this approach, we generated 157 keywords for **Technology**, 91 for **Poverty**, 131 for **Development**, 151 for **Corruption**, 37 for **Inflation**, and 275 for **Greetings**. Due to the overlapping nature of political discourse on Twitter, some tweets were classified under multiple themes. Though, if a tweet has no matching keywords, it's assigned to **Other**. A table containing the complete lists of keywords will be made available online after paper publication.

Using this approach, we labeled 187.7K or 3.4% tweets as **Technology**, 188.1K or 3.4%, as **Poverty**, 568.1K or 10.2%, as **Greetings**, 327.3K or 5.9% as **Development**, 171.0K or 3.1% as **Corruption**, 63.2K or 1.1% as **Inflation**, and the remaining **Other** tweets. This distribution of tweets suggested that politicians generally tweet more about holiday celebrations and greetings in comparison to actually policies related to technology and development. In order to assess the performance of our classification, for each category of tweets, we randomly select a representative sample (e.g. we calculate sample size using 95% confidence level and $\pm 3\%$ confidence interval) and manually assess whether each tweet indeed focuses on technology, poverty, and/or development related topics. We see an 80.3% precision for **Technology**, 82.0% for **Poverty**, 90.0% for **Development**, and comparable results for others. Finally, we aggregate all categories of tweets together and observe a combined precision and recall of 85.4%, and 82.0% respectively.

Analyses

A look at the overall descriptive statistics from the sample gives a sense of the landscape of political twitter in India. Figure 1 shows that Prime Minister Narendra Modi was clearly the dominant figure both in terms of his following, and the extent to which his tweets got retweeted. However, there are a number of other political leaders with significant presence and influence online. For instance, the main leader of the rival INC party, Rahul Gandhi, gets more average retweets than Modi despite the later currently has several times as many followers.

We also see a trend that all of the accounts with above-500-median-retweets were leaders of specific parties (e.g. Akhilesh Yadav of the Samajwadi Party, Arvind Kejriwal of the Aam Admi Party), even if they weren't necessarily the most followed leaders from within the overall sample. This suggests a centralizing tendency for social media communications to coalesce around party leaders. Additionally, given that recent work observed high bot activities on Twitter (Varol, Ferrara, Davis, Menczer, & Flammini, 2017), we also randomly sample a subset of (sample size corresponds to 99% confidence-level and +/-3 confidence internal) Narendra Modi and Rahul Gandhi followers and then use Botometer (Davis, Varol, Ferrara, Flammini, & Menczer, 2016) to determine their bot scores (technical details are in the Appendix). We observe that 63% and 54.7% of randomly sampled Modi and Gandhi followers are bots. These numbers are much higher than the reported 9% to 15% by another study (Varol et al., 2017), suggesting high follower number inflation which occurs across parties⁴.

The Analyses section is structured as follows. We first assess the distribution of original tweets (i.e. not including retweets) for each category of tweets (e.g. **Greetings**, **Inflation**, etc.). We then examine how politician accounts interact with each other and news media on Twitter via Retweeting (RT @) and Mentioning (@). Finally, through linguistic analysis, we assess how politicians differ when talking about policy-related and non-substantive issues.

Frequency Distribution of Direct Contributions for Each Category of Tweets

Results are summarized in Figure 2. We observe that BJP contributed 348.9K **Greetings** tweets, accounting for 12.8% of its total tweets. This is much higher than the INC which only contributed 124.3K or 8.1% of its total tweets to **Greetings** (similar observation for politicians in 3rd parties). These numbers suggest that the governing party members are more frequently engaged in non-issue-based greetings (over 50% more than the main opposition party). Moreover, BJP also demonstrate a higher preference to discuss **Development** related issues such as economic development and GDP. INC, on the other hand, shows a higher affinity in discussing **Corruption** and **Inflation**: 83.5K or 5.4% of its total

⁴ Another possible explanation is that Botometer has a higher false positive rate for professionally managed organizational accounts (Davis et al., 2016).

tweets are focused on **Corruption** while the number is only 2.2% for both BJP and 3rd-party politicians. Similarly, the median number of tweets by BJP accounts on **Technology** is 83 and **Development** 167, significantly higher than those of INC (48 and 107). In comparison the median number of tweets posted by INC on **Poverty** is 101, **Inflation** 46, and **Corruption** 118, whereas the numbers are 76, 14 and 48 for BJP. Results suggest that opposition parties use social media more extensively for critical messaging aimed at the governing party.

The top 10 contributors for each category of tweets determined by the absolute number of tweets are shown on Table 1. Unsurprisingly, politicians associated with the tech industry such as the ministers of Electronics (RS Prasad), Science and Technology (Harsh Vardhan), and Commerce (Suresh Prabhu) contributed significantly to the Twitter conversation on **Technology**, as well as Andhra Pradesh Chief Minister Chandrababu Nadu, and INC member, tech billionaire Nandan Nilekani. On **Development** discussions we intersections with the most vocal leaders on issues of technology and poverty. Dharmendra Pradhan, Piyush Goyal and Suresh Prabhu who featured in the **Technology** list also contributed significantly in discussing **Poverty**, suggesting that these are co-occurring themes, an idea that has purchase in a body of 'ICTD' work that proposes technology as fundamentally tied to the idea of development in the Global South. In the **Greeting** related tweets, Prime Minister Narendra Modi emerges as the leader, and all the highest tweeting politicians in this set are from the ruling party. This aligns well with previous research which suggests that Narendra Modi's online presence aims to build an image of national harmony (Kaur, 2015), but also shows that parties in power are more likely to highlight a positive discourse. The top contributors of **Corruption** are all key party spokespersons, rather than the top ranking leaders themselves, suggesting that the less savory topics are left to be addressed by party mouthpieces. The discussion on **Poverty** is fairly diverse in terms of the actors involved—a significant presence of opposition figures (Arun Yadav, Sitaram Yerchuri, Ashok Gehlot), and high contributions by the central government ministers who deal with issues of poverty and rural development as part of their office requirements.

The discussion on **Inflation** is also dominated by non-BJP politicians. However, the list of biggest contributors to the subject is led by Dharmendra Pradhan, who is a false positive

due to his position as petroleum minister (he dominates the list by repeatedly tweeting about petroleum which is otherwise typically an inflation-related subject), and Sanju Verma, a BJP economist whose tweets are largely about defending the party against inflation-related charges. The lukewarm tweeting of BJP members on inflation and corruption reaffirms that the party in power is less inclined to engaging with negative topic.

This is comparable to work on topical spread of politicians' tweets by Hemphill et al., who find that about 9% of US Congresspeoples' tweets are on soft topics (2% thanks or regards and 7% on telling a story about their day), while the two major parties in India - the BJP scores in greetings at 12.8% while the opposition INC has about 8.1%.

Politicians' Engagement with Each Other and News Media

Retweeting (RT @) and Mentioning (@) are two of the most prevalent user interaction types on Twitter. Prior studies (S. Ahmed, , & J. Cho, 2016; Jungherr, 2016) focused on Western politics demonstrate that elites on Twitter interact more frequently with others of the same party and news media compared to interacting with their constituents or politicians from competing parties. Yet, these studies often do not address the contextual information surrounding these interactions such as what topics (e.g. Greetings, Corruption, etc.) are under discussion. In this section, we focus on comparing and contrasting politicians' engagement with each other to their engagement with news media.

Here, We first identify a list of 605 influential news media Twitter accounts (see the Appendix for details). We then utilize patterns of engagement to assess the significance of an account. We use 2 distinct metrics: indegree centrality and pagerank. Indegree measures the number of engagements, or edges, directed at an account and captures the account's popularity or reputation (Bonacich, 1987). Pagerank considers both the number of engagements as well as the quality of these accounts (Kwak, Lee, Park, & Moon, 2010-04)⁵.

First, we use the following procedure to build *directed* retweeting and mentioning

⁵ To elaborate, an account that is retweeted by x number of insignificant politicians (e.g. lesser known state officials) has a lower pagerank than another who may be retweeted by fewer than x politicians, but the ones who did retweet him or her are high profile (e.g. being retweeted by both Narendra Modi and Sushma Swaraj).

networks of politicians and news media on Twitter for each category of tweets. Let graph $G_{retweet,technology} = \{V, E\}$ where V is the list of politician and news media accounts. A directed edge $e_{u,v}$ exists in E if politician u retweeted one of v 's Technology tweets on Twitter⁶. Likewise, we build $G_{mention,technology}$, $G_{retweet,poverty}$, et cetera. This give us a total of 14 distinct networks. Next, we apply *networkx*, a python library, to generate the indegree and pagerank centrality scores of each node for all 14 networks⁷.

Centrality scores are summarized in Figure 3. As shown, the median indegree centrality of news media accounts in graph $G_{mention,technology}$ (i.e. the network where politicians mention other accounts when tweeting about technology) is 2 times more than the median indegree centralities for politicians. Further, news media's indegree scores are higher in **Corruption** and **Inflation**, are comparable to politicians' for **Development** and **Poverty**, but are much lower for **Greetings**. That is, politicians mentioned news media more or just as much as they did other politicians when discussing policies and issues related to technology, development, social welfare, etc. But, politicians seldom mentioned news media when they spoke about celebrations, and birthday wishes. Similarly, the median pagerank for news media is actually slightly higher than or comparable to politicians' for substantive topics including **Inflation**, **Corruption**, **Technology**, **Development**, and **Poverty**, which suggest that politicians preferentially retweeted or mentioned news media rather than each other when discussing important policy-related issues. Both indegree centrality measurements and pagerank scores for new media in the **Greetings** tweet categories were, however, noticeably smaller.

Related work observe increased use of social media by politicians for personal branding and political narratives of their own (G. E. Enli, 2013; McGregor, 2018; Meeks, 2016). Here we expanded on these studies by demonstrating that when actual policies are in discussion, politicians were strategically choosing to rely on news media either via retweets or mentions instead of talking about the topic themselves. Indeed, this may also be interpreted as a proxy for relying on reports on these subjects coming from professional journalists, rather than

⁶ Note that while u has to be a politician, v can either be a politician or news media.

⁷ Here, we also used a *weighted* version of these metrics to account for politicians' overall activity levels (i.e. how often each individual politician retweet or mention others). Our observations are robust.

talking about them directly. In fact, when politician tweet about soft, "lifestyle" topics, they rarely engaged with the media.

Text Attributes Comparison Across Parties

Here, i) we first analyze the most commonly used hashtags by each political party given that prior research observe parties strategically use hashtags to frame or counter-frame specific narratives on controversial issues (Hemphill et al., 2013). ii) We then determine the difference in language complexity between tweets of substantive topics and **Greetings** tweets. Finally, sentimental appeal in political communication has important consequences (e.g. altering voter behavior) (Sharma & Moh, 2016-12; Soroka, Young, & Balmas, 2015). Thus, iii) we also conduct sentiment analysis on each category of tweets to evaluate politicians' strategic use of emotional appeals and assess whether party based differences exist.

Hashtag Usage. In this section, we focus on BJP and INC politicians' use of hashtags in **Greetings** and **Corruption** tweets: the former category is a non-issue while the later is a very contentious topic in India (Bussell, 2012).

First, we hand-coded each of the top 50 hashtags from either BJP or INC under the thematic heading of **Greetings** into 5 subtypes: i) party-based greetings (such as birthday greetings to party members), ii) nationalistic greetings (e.g. greetings for soldiers on independence day), iii) sports-related, iv) Hinduism-related, and v) generic. We then apply the following regression model to show the usage likelihood of these hashtags as a comparison between BJP and INC. First, we write H as the entire set of hand-coded hashtags, and $P_{greetings}$ as accounts who posted at least 1 **Greetings** tweet. For a given hashtag $h \in H$ and politician $p \in P_{greetings}$, we assign dependent variable $y_{p,h} = 1$ if p used h at least 1 time, and $y_{p,h} = 0$ otherwise. We denote $type_h$ as h 's subtype, and $party_p$ as p 's party. Further, we write $status_count_p$, $follower_count_p$, and $friend_count_p$ to control for p 's overall Twitter activity level. Finally, we apply logistic regression equation:

$$y_{h,p} = \beta_0 + \beta_1 * type_h * party_p + \beta_2 * status_count_p + \beta_3 * follower_count_p + \beta_4 * friend_count_p \quad (1)$$

As shown in Table 2, INC politicians are far less likely to make nationalistic greetings or greetings to sports-related events. Additionally, INC is also slightly less likely to make greetings for Hindu festivals. To elaborate, the likelihood of a nationalistic hashtag being used by a politician is reduced by a log odds of 1.34 if the politician is from INC. We can conclude that the use of nationalistic greetings as well as shout-outs to sports are a consistent part of the brand image of the BJP in the period studied. This is consistent with our prior observations: the BJP party as a whole, lead by Narendra Modi, is more invested in branding itself as optimistic and personable.

Similarly, we manually labeled each of the top 50 hashtags from either BJP or INC in **Corruption** into 3 subtypes: i) personal (targeting an individual or a group), ii) issue-based (targeting a particular incident or issue), and iii) generic (neither personal nor issue-specific). We observe that BJP is significantly more likely to talk about corruption using generic terms #SaafNiyatSahiVikas (Clean Intent Good Development), #AntiBlackMoneyDay, #IndiaFightsCorruption, and #BlackMoneyCrackdown. These hashtags address corruption but do not have an explicit target or villain. In contrast, the INC is significantly more likely to tweet about corruption using hashtags that are issue-centric such as #GreatRafaleCoverUp or #ModiRafaleLiesExposed, both of which refer to the Rafale defence deal which has been a center of a corruption controversy involving the prime minister's office. INC politicians also use hashtags that are personality-centric such as #ModiRobbsIndia, or #ChowkidarChorHai (The Watchman is the thief), both of which are direct personal attacks on the prime minister, Narendra Modi, implying that he himself is corrupt. Indeed, as shown in Table 2 (regression results generated using a comparable model), INC politicians are significantly more likely to use a personal corruption hashtag by a log odds ratio of 1.5.

To summarize, our results on hashtags extend prior work focused on the West: similar to U.S. political actors, Indian politicians are also strategically using hashtags on Twitter for brand and narrative building purposes (Hemphill et al., 2013; J. Pal, 2015). Here, the ruling BJP party is using optimistic and nationalistic hashtags to market itself as a party that can move India forward. In comparison, INC is more focused on advancing the narrative that BJP is corrupt and unfit to govern.

Language Complexity. Going beyond text similarity measures at high level, we also assessed the complexity of language for each category of tweets using 3 measurements of readability: Flesch-Kincaid Reading grade, Smog Index, and Gunning Fog Index (Si & Callan, 2001)⁸. These metrics were used by many prior studies focused on political communications (Flaounas et al., 2013; Ott, 2017). Here, for each original English tweet with more than 2 word tokens i , we use the ReadCal python library to compute all three scores and denote as the value as $readability_{i,flesch}$, $readability_{i,smog}$, $readability_{i,gunning}$. Then, we run OLS treating $readability$ as the dependent variable, and $category_i$, and $party_i$ as the independent variables.

As shown in Table 3, compared to **Greetings**, all other categories of tweets except for **Inflation** are correlated with higher language complexity (e.g. **Development** tweets have roughly 2-grades higher reading-level than **Greetings**), suggesting that tweets about substantive issues such as corruption, development, poverty and welfare are associated with higher language complexity. Further, we see that INC and 3rd party politicians contribute tweets with higher reading-levels. While a clear conclusion cannot be drawn from this alone, the suggestion is that the BJP exhibits a simpler messaging approach using more straightforward language.

Sentiment Appeal. Here, we first used VADER (Valence Aware Dictionary and sEntiment Reasoner), a sentiment analysis library (Hutto & Gilbert, 2014-06) specifically attuned to social media, to generate positive and negative sentiment scores of each tweet⁹. We also obtained the compound affect score which indicates a text's overall valence. Looking at the visual representation of the differences between each category of tweets in Figure 4, we can see that median affect scores are more positive for both **Greetings** and **Technology**, indicating that politicians, as a whole, tend to utilize more positive emotional appeals when discussing celebrations or information technology. Furthermore, we also saw that INC

⁸ For all 3 metrics, a higher score indicates that a reading material is more complex. For instance, a Flesch-Kincaid reading grade of 10.0 suggests that the text corresponds to 10th grade reading-level.

⁹ To assess VADER's performance on our dataset, here we randomly sample 50 tweets with positive sentiment scores higher than negative scores, we observe 41 tweets are positive; similarly, we randomly sample 50 tweets with higher negative scores, 40 out of which are indeed negative.

demonstrates noticeable negativity in their tweets about corruption, inflation, and poverty comparing to BJP and 3rd party politicians. Next, for each tweet i , we apply OLS with pos_i and neg_i as the dependent variables (pos_i and neg_i are the positive, negative sentiment scores of i), and $type_i$ and $party_i$ as the independent variables.

As shown on Table 3, all substantive topics are associated with lower positive and higher negative sentiment when comparing to **Greetings**. Additionally, INC and third party politicians demonstrate lower positivity and higher negativity compared to the BJP. This result aligns with prior work which suggests that politicians from non-governing parties collectively spend more effort criticizing the policies and leaders of the current majority ruling party (Jungherr, 2016). Furthermore, previous India-specific analysis has also shown preference for non-controversial tweeting as part of the BJPs online strategy (J. Pal, 2015).

Qualitative Analysis: Finally, we qualitatively analyzed the sentiment appeal of tweets from key politicians: Narendra Modi and Rahul Gandhi¹⁰. While such tweets are a small sample, they are useful indicators of the politician's branding as well as what drives their popularity online. For both politicians, we took one representative viral tweet from each category to compare the differences.

Development: As demonstrated in Figure 5, while Modi talks about development in a tweet with positive valence, presenting a optimistic vision of India in a tweet that congratulates the average Indian citizen, Rahul Gandhi includes a link to a news story alongside a confrontational tweet attacking the government for a development-related fiscal policy. This contrast in styles also underlines the overall difference between the more positive-toned governing party and the more confrontational opposition party approaches.

Technology: In Figure 6, we show a viral tweet from Prime Minister Modi from 2016 urging young Indians to use more digital technology for financial transactions as part of the demonetization effort by the Indian government. The tone is again optimistic, aligning with the techno-optimistic style that Modi has been known for (J. A. Pal, 2016). In contrast, the

¹⁰ Out of the top 200 most retweeted tweets for each category, Modi alone accounted for 41% of the total tweets and Rahul Gandhi's tweets accounted for 37%. This highlights the importance of these two key leaders in the political social media universe.

viral tweet from Rahul Gandhi, shown in Figure 6, includes a link to a story that claims to expose the prime minister's role in a data breach that exposed the personal information of millions of Indians. Unlike the positive tone on technology that Modi has, Gandhi presents a dystopian view of technology, of companies tracking citizens through their mobile devices. While the tweet refers to technology, the underlying intent is to attack the prime minister.

Poverty: On poverty, as shown in Figure 7, Modi highlights the role of his supporters in a political victory for the poor, and has it alongside an image of himself striking a pose. The viral tweet from Rahul Gandhi uses a technical term "real wages" and adds facts to support his attack on the prime minister's economic policies. He then uses wordplay ("Modi Made Disaster" - MMD) in sarcasm.

Inflation: Modi's inflation tweets on average received fewer retweets than his other tweets; his most retweeted message offers discounts on petrol for cashless buyers, which is arguably more about demonetization than the price of gas. As seen in Figure 8, Modi also uses a strategy he employs frequently, of posting a picture of himself alongside the discount announcement. Rahul Gandhi, in comparison, tweets frequently and aggressively about inflation, which has been a hot-button topic for the opposition. Here, referring to Modi as a "king of misinformation", Gandhi uses excerpts from Modi's own speech to discredit his claims on inflation.

Corruption: Modi tweeted significantly on corruption when he was an opposition politician. However, being in power makes that difficult since one would presumably be responsible for the corruption. His highly retweeted message, in Figure 9, was a quasi-greeting to citizens. Following the Rafale arms deal scandal, corruption has been a major topic for Rahul Gandhi. He referred to Modi as "Supreme Leader" on his tweets, suggesting that the prime minister had dictatorial tendencies. The tweet highlighted here is reflective both of the frontal, insulting tone towards Modi, and of Gandhi's own appropriation of the soldier metaphor, borrowed from his opponent the long-successful value of bringing patriotism into an antagonistic exchange.

Greetings: Modi's tweets are overwhelmingly positive and usually around festivals or key moments. While Rahul Gandhi also post positive, non-confrontational greeting tweets, he

has also used congratulations on occasion with sarcasm. As shown, Modi's tweet from early 2018 congratulates the blind cricket team for its victory. His tweet uses upbeat language to present the disabled sportspersons as being a source of inspiration. In contrast, Rahul Gandhi uses combative language and irony in congratulating Narendra Modi's chief ally, the BJP party president Amit Shah, specifically calling out a financial concern that had destructive outcomes for some Indians (Figure 10).

Although these tweets are a small selection from the large body of tweets from both leaders, their virality lends them important symbolic value. The common thread is a tone of combativeness and sarcasm from Rahul Gandhi, compared to the positive-toned tweets from the Prime Minister. It is important to note, however, that all of the viral tweets came from a period when Modi has been in power. If the roles were reversed, we may well see Modi being the more aggressive one.

Discussion

In this paper, we first classified a large list of Indian politicians. We then presented a picture of social media behavior of Indian politicians when tweeting various topics. We presented these—technology, development, poverty, corruption, and inflation—as policy-centric, whereas the comparative category of greetings is relatively less policy-centric. An important motivator of this comparison was to contextualize past research Casero-Ripoll's et al. (2017); G. Enli (2017); Evans et al. (2014); Fuller et al. (2018); Hemphill et al. (2013); Kellner (2018); Mudgal (2015) that has suggested a dramatic shift in the media environment for politicians in terms of their choice of subjects of discussion, use of hashtags, as well as their evolving relationship with each other and the professional mainstream media corps.

We observed that politicians do indeed more frequently tweet about greetings and celebrations as compared to policy-relevant issues. This confirmed past research findings that politicians are focused more on building soft, personal ties rather than broadcasting their policies. The mapping of specific politicians with repeat engagement on specific topics including corruption, poverty, and development intuitively suggests that active engagement in

these areas is related to the domains of practice of these specific politicians. Overall, BJP politicians demonstrated a clear dominance in setting the agenda with many of these topical areas in part through their sheer numbers on social media.

We also showed that politicians in opposition tend to have more antagonistic discussions, while politicians in power lean towards the less controversial topics. The messaging of the prime minister—in particular his avoidance of the inflation topic altogether—hints at another important trend. That is, while social media may present the spectre of political leaders constantly communicating with the general public, the lack of professional journalists engaging with political leaders affords them the ability to pick and choose topics to discuss. This may present serious challenges to the free media-led checks and balances process. Furthermore, when we looked deeper at the qualitative findings, we saw both that tweets about what would seem to be "policy" topics such as technology or inflation can sometimes a priori be examples of political performance, or confrontation, rather than substantive discussion about policy. Likewise, we also saw that greetings may not necessarily be benign or free of political implication.

Using network analysis, we observed that politicians demonstrated a higher preference for engaging with news media than with each other when tweeting substantive issues. This implies that politicians are strategically involving news media when discussing actual policies and setting public policy agenda. Thus, rather than professional politics disengaging from the mainstream news process, which has been one of the concerns about the implications of social media use in Global South settings, we are seeing that politicians are still engaging the mainstream media in new ways, often mediated through social media. Using text analysis, we illustrated that policy-focused tweets tend to have much higher language complexity. Further, by comparing the BJP and INC on corruption-related hashtags, we also find that parties in opposition are more likely to use personal or issue-based attacks on corruption, rather than talk about it in a relatively generic fashion. Additionally, we also found a higher use of nationalistic greetings hashtags in the BJP's positive-themed tweets. Overall, politicians' preferential use of hashtags of particular sentiments in each topic clearly depicted party-based collaborative efforts in narrative-building. Finally, a deeper look at the tweets of the two most

important leaders in our sample in terms of following and retweet activity—Narendra Modi and Rahul Gandhi—is instructive on both their individual styles and on the differences in sentiment between governing and opposition parties. While the governing party politician steers clear of controversy and gets widely retweeted for his non-combative, positive tweeting, the opposition politician is rewarded for being aggressive and confrontational on social media.

There are several caveats in our study worth noting here. Not everyone is on Twitter. Given that India still has a substantial share of its population without access to technology or functional literacy, any study of politics and social media is likely to be biased towards those who are technology savvy and belong in higher economic classes De Angeli, Athavankar, Joshi, Coventry, and Johnson (2004). Furthermore, Other technologies such as WhatsApp and Facebook have important implications because of their widespread use among younger populations in India. However, these are largely private. The goal of our work is to study the public discourse of politicians themselves, for which Twitter is a uniquely powerful tool.

While we provided an in-depth assessment of politicians' Twitter priorities and strategies, our paper did not define or address the issue of success with these strategies. Future work can focus on building normative measurements of politicians' success on Twitter. One possible metric could be how many new followers, retweets or mentions politicians gain by, for instance, providing higher emotional appeal in their tweets or engaging more with prominent news media account on Twitter for different type of issues. Furthermore, we also did not delve deep into politicians' social media strategies within different time periods (e.g. before, during, and after elections). Future work that makes this distinction can provide valuable insights into how politicians and politically inclined individuals shift their personal branding with time.

Appendix

Additional Material on the Data Section

Twitter Account Classification: We used features including user follower count, friend count, tweet count, etc. Additionally we also used a *Bag of Words* (Joulin et al., 2016) model, implemented in the *CountVectorizer* package provided by *scikit learn*, to represent the

profile description. Furthermore, we also experimented with 5 different classification algorithms: *NaiveBayes*, *RandomForest*, *StochasticGradientDescent*, *LogisticRegression*, and *SupportVectorMachine*. Additionally, for each classifier, we apply *GridSearchCV* to find the optimal parameters. We find that the *SVM* classifier with an *rbf* kernel is the best. It has an accuracy score of 0.898, recall score of 0.847, precision score of 0.930, and an F1 score of 0.887.

Tweet Classification: We used *gensim*, an open source natural language processing (NLP) toolkit in python, to produce a 300 dimensional vector space representation of our corpus (here, each tweet is a single document). Within this space, each unique word is assigned a numeric vector. Next, for each seed word in a category (e.g. Technology), we first identified the top 25 other words that had the highest cosine similarity scores to the given seed word, using *gensim*'s native function, and then manually selected the ones that made contextual sense. We repeat this process for all categories.

Additional Material on the Analyses Section

Bot Detection: At the time of the writing, Narendra Modi has 45.5M followers and Rahul Gandhi 8.5M. We first queried Twitter API for 1.65M and 975K most recent followers of Modi and Gandhi (we were not able to obtain the entire follower lists due to Twitter API's rate limit). Then, we randomly sampled 5K accounts from each follower list¹¹ and used Botometer (Davis et al., 2016) to determine their bot scores. Out of the 10K accounts, 38.4% had not posted a single tweet and 2.5% had set their accounts to private, therefore we were not able to determine if they were bots. For each of the remaining accounts, Botometer returned a score between 0 and 1 with a higher value indicating greater likelihood of being a bot. Here, an account was labeled as a bot if its score was above 0.5.

Identify News Media Accounts: We identify influential news media Twitter accounts using the following procedure: i) We first determined the entire list of accounts being retweeted or mentioned by politicians, and filtered out all the accounts with less than 1K

¹¹ The sample size for a population of 1.65M with 99% confidence-level and +/-3 confidence interval is 1848. Thus, a sample size of 5K is more than sufficient here.

followers. ii) We then generated a list of keywords that are associated with news media and careers relating to journalism such as "news", "reporter", "blogger" ¹². iii) We obtained a sublist of accounts that contain at least 1 of the keywords in the user profile description field. iv) Finally, we manually examine each account and remove the non-media accounts. A similar approach was used by Priante et al. (Priante et al., 2016) in identifying the occupations of Twitter users.

¹² The complete list of keywords are: news, editor, journalist, columnist, magazine, blogger, anchor, newspaper, journalism, correspondent, coverage, radio, reporter, press, network, commentator.

References

- Adamic, L., & Glance, N. (2005-08). The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd international workshop on link discovery* (p. 36–43).
- Agrawal, B. S. (2006). India public service broadcasting and changing perspectives. public service broadcasting in the age of globalization. *I. Banerjee and K. Seneviratne*, 149–164.
- Ahmed, S., , & J. Cho, K. (2016). The 2014 indian elections on twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics*, 33(4), 1071–1087.
- Ahmed, S., Jaidka, K., & Cho, J. (2016). The 2014 indian elections on twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics*, 33(4), 1071–1087.
- Ahmed, S. M. (2014). My name is khan: the use of twitter in the campaign for 2013 pakistan general election. In *47th hawaii international conference on system sciences (hicss)*, *ieee*.
- Baumer, E. P., Sueyoshi, M., & Tomlinson, B. (2011). Bloggers and readers blogging together: Collaborative co-creation of political blogs. *Computer Supported Cooperative Work (CSCW)*, 20(1-2), 1–36.
- Bennett, W. (2012). The personalization of politics: Political identity, social media, and changing patterns of participation. *The ANNALS of the American Academy of Political and Social Science*, 644(1), 20–39.
- Bennett, W., & Pfetsch, B. (2018). Rethinking political communication in a time of disrupted public spheres. *Journal of Communication*, 68(2), 243–253.
- Bimber, B. (2014). Digital media in the obama campaigns of 2008 and 2012: Adaptation to the personalized political communication environment. *Journal of Information Technology and Politics*, 11(2), 130–150.
- Bøggild, T., & Pedersen, H. H. (2018). Campaigning on behalf of the party? party constraints on candidate campaign personalisation. *European Journal of Political Research*, 57(4),

883–899.

Bonacich, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, 92(5), 1170–1182.

Bureau, Z. (2018). *Pm narendra modi tells bjp mps to have at least 3 lakh followers on twitter*.

Bussell, J. (2012). *Corruption and reform in india: Public services in the digital age*. Cambridge University Press.

Casero-Ripoll's, A., & P. Franch, M.-O. (2017). "the populist political communication style in action: Podemos's issues and functions on twitter during the. *American Behavioral Scientist*, 61(9), 986–1001.

Chadwick, A. (2017). *The hybrid media system: Politics and power*. Oxford University Press.

Chakraborty, S., Pal, J., & D. M. Romero, P. (2018). Political tweets and mainstream news impact in india: A mixed methods investigation into political outreach. In *Proceedings of the 1st acm sigcas conference on computing and sustainable societies, acm*.

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *Icwsim*, 133, 89–96.

Correa, J. J. (2017). "ideological consumerism in colombian elections. *Cyberpsychology, Behavior, and Social Networking*.

Curato, N. (2017). Flirting with authoritarian fantasies? rodrigo duterte and the new terms of philippine populism. *Journal of Contemporary Asia*, 47(1), 142–153.

David, Y. C. (2018). Reframing community boundaries: the erosive power of new media spaces in authoritarian societies. *Information, Communication and Society*, 1–18.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016). Botornot: A system to evaluate social bots. In *Proceedings of the 25th international conference companion on world wide web* (pp. 273–274).

De Angeli, A., Athavankar, U., Joshi, A., Coventry, L., & Johnson, G. I. (2004). Introducing atm's in india: a contextual inquiry. *Interacting with Computers*, 16(1), 29–44.

Dâzheer, E., & Verdegem, P. (2014). Conversations about the elections on twitter: Towards a structural understanding of twitter's relation with the political and the media field.

- European Journal of Communication*, 29(6), 720–734.
- Enli, G. (2017). "twitter as arena for the authentic outsider: exploring the social media campaigns of trump and clinton in the 2016 us presidential election. *European Journal of Communication*, 32(1), 50–61.
- Enli, G. E. (2013). Personalized campaigns in party-centred politics: Twitter and facebook as arenas for political communication. *Information, Communication and Society*, 16(5), 757–774.
- Evans, H. K., Cordova, V., & Sipole, S. (2014). Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns. *PS: Political Science & Politics*, 47(2), 454–462.
- Flaounas, I., Ali, O., Lansdall-Welfare, T., De Bie, T., Mosdell, N., Lewis, J., & Cristianini, N. (2013). Research methods in the age of digital journalism: Massive-scale automated analysis of news-contentâ€Topics, style and gender. *Digital Journalism*, 1(1), 102–116.
- Fuller, G., , & C. Fisher, A. (2018). Malcolm turnbull's conversational career on twitter: the case of the australian prime minister and the nbn. *Media International Australia*, 167(1), 88–104.
- Gayo-Avello, D. (2017). Social media won't free us. *IEEE Internet Computing*, 21(4), 98–101.
- Hemphill, L., Otterbacher, J., & Shapiro, M. (2013). What's congress doing on twitter? In *Proceedings of the 2013 conference on computer supported cooperative work* (pp. 877–886).
- Howard, P. (2006). *New media campaigns and the managed citizen*. Cambridge University Press.
- Hutto, C., & Gilbert, E. (2014-06). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international conference on weblogs and social media (icwsm-14)*. Ann Arbor, MI.
- Jaffrelot, C. (2015). Narendra modi and the power of television in gujarat. *Television and New Media*.

- Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.
- Jungherr, A. (2014). The logic of political coverage on twitter: Temporal dynamics and content. *Journal of Communication*, 64(2), 239–259.
- Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of information technology & politics*, 13(1), 72–91.
- Kaur, R. (2015). Good times, brought to you by brand modi. *Television & New Media*, 16(4), 323–330.
- Kellner, D. (2018). Donald trump and the war on the media: From election '16 into the trump presidency. In *The trump presidency, journalism, and democracy*, routledge (p. 19–38).
- Kou, Y., Kow, Y. M., Gui, X., & Cheng, W. (2017). One social movement, two social media sites: A comparative study of public discourses. *Computer Supported Cooperative Work (CSCW)*, 26(4-6), 807–836.
- Kreiss, D. (2016). Seizing the moment: The presidential campaigns' use of twitter during the 2012 electoral cycle. *New Media and Society*, 18(8), 1473–1490.
- Kumar, S. (2015). Contagious memes, viral videos and subversive parody: The grammar of contention on the indian web. *International Communication Gazette*.
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on world wide web* (pp. 591–600).
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010-04). What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on world wide web* (p. 591–600).
- Lee, C. (2017). Facebooking to power: The social media presence of malaysian politicians. *ISEAS Perspectives*, 74.
- Lotan, G., Graeff, E., Ananny, M., & Pearce, D. (2011). "the arab spring| the revolutions were tweeted: Information flows during the. *International journal of communication*, 5, 31.

- Lufkens, M. (2016). Twiplomacy study 2016. *Burson-Marsteller, Editor, Geneva*.
- Maheshwari, S. C. (2018). Political elites and journalistic practices in india: A case of institutionalized heteronomy. *Journalism*.
- Mascaro, C. M., & Goggins, S. (2010). Collaborative information seeking in an online political group environment. In *International workshop on cis of cscw conference*.
- McGregor, S. (2018). Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. *new media and society*, 20(3), 1139–1160.
- Meeks, L. (2016). Gendered styles, gendered differences: Candidates' use of personalization and interactivity on twitter. *Journal of Information Technology and Politics*, 13(4), 295–310.
- Morozov, E. (2009). Iran: Downside to the "twitter revolution". *Dissent*, 56(4), 10–14.
- Mudgal, V. (2015). Framing the 2014 elections: The curious absence of development. *Television and New Media*, 16(4), 354–360.
- Ohm, B. (2015). Organizing popular discourse with and against the media: Notes on the making of narendra modi and recep tayyip erdogan as leaders-without-alternative. *Television and New Media*, 16(4), 370–377.
- Ott, B. (2017). The age of twitter: Donald j. trump and the politics of debasement. critical studies in media. *Communication*, 34(1), 59–68.
- Pal, J. (2015). Banalities turned viral: Narendra modi and the political tweet. *Television and New Media*, 16(4), 378–387.
- Pal, J., Chandra, P., & VydiswaranâŽ, V. (2016). Twitter and the rebranding of narendra modi. *Economic and Political Weekly*, 51(8), 52–60.
- Pal, J. A. (2016). *Political social media in the global south. conference on e-business, e-services and e-society*. Springer.
- Park, M.-J. J. (2000). *De-westernizing media studies*. Psychology Press.
- Priante, A., Hiemstra, D., van den Broek, T., Saeed, A., Ehrenhard, M., & Need, A. (2016). #whoami in 160 characters? classifying social identities based on twitter profile descriptions. In *Proceedings of the first workshop on nlp and computational social science* (p. 55–65).

- Rahman, A. (2014). The problems with reimagining public media in the context of global south. *Stream: Inspiring Critical Thought*, 6(1), 56–65.
- Rao, U. (2010). *News as culture: Journalistic practices and the remaking of indian leadership traditions*. Berghahn Books.
- Rodrigues, U. (2014). Social media's impact on journalism: a study of media's coverage of anti-corruption protests in india. *Global media journal: Australian edition*, 8(1), 1–10.
- Rodrigues, U. M. (2017). Social media as a platform for incessant political communication: A case study of modi's "clean india" campaign. *International Journal of Communication*, 11, 23.
- Schäfer, M. T. (2016). Challenging citizenship: Social media and big data. *Computer Supported Cooperative Work (CSCW)*, 25(2-3), 111–113.
- Sharma, P., & Moh, T. (2016-12). Prediction of indian election using sentiment analysis on hindi twitter. In *2016 ieee international conference on big data (big data)* (p. 1966–1971). IEEE.
- Si, L., & Callan, J. (2001). A statistical model for scientific readability. In *Proceedings of the tenth international conference on information and knowledge management* (pp. 574–576).
- Soroka, S., Young, L., & Balmas, M. (2015). Bad news or mad news? sentiment scoring of negativity, fear, and anger in news content. *The ANNALS of the American Academy of Political and Social Science*, 659(1), 108–121.
- Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: a social media analytics framework. *Social Network Analysis and Mining*, 3(4), 1277–1291.
- Sundaram, R. (2015). Publicity, transparency, and the circulation engine: the media sting in india. *Current Anthropology*, 56(12), 297–305.
- Tusa, F. (2013). "how social media can shape a protest movement: The cases of egypt in. *Arab Media and Society*, 17, 1–19.
- Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint*

arXiv:1703.03107.

Wilkerson, J., & Casas, A. (2017). Large-scale computerized text analysis in political science: Opportunities and challenges. *Annual Review of Political Science*, 20, 529â€“544.

Zittrain, J., Faris, R., Noman, H., Clark, J., , & R. Morrison-Westphal, C. (2017). *The shifting landscape of global internet censorship.*

Table 1

Top 10 Contributors (Followers>=100K) by Absolute Tweet Count for Each Category of Tweets

Category	Top10 CONTRIBUTORS
Technology	DIPP India(other 2073); Ravi Shankar Prasad(bjp 1977); Amitabh Kant(other 1696); Dr. Harsh Vardhan(bjp 1550); Suresh Prabhu(bjp 1252); Dharmendra Pradhan(bjp 1075); Jayant Sinha(bjp 908); Narendra Modi(bjp 896); Manoj Sinha(bjp 861); data.gov.in(other 796)
Poverty	Agriculture INDIA(other 1991); Yogendra Yadav(other 1297); Radha Mohan Singh(bjp 1182); Jitu Patwari(inc 877); Shivraj Singh Chouhan(bjp 828); Raghubar Das(bjp 785); Narendra Modi(bjp 573); UP Congress(inc 565); Narendra Singh Tomar(bjp 557); BJP Delhi(bjp 536)
Greetings	Narendra Modi(bjp 2762); Kailash Vijayvargiya(bjp 2647); Sudarsan Pattnaik(other 2125); Raghubar Das(bjp 2091); Mukhtar Abbas Naqvi(bjp 2012); Sidharth Nath Singh(bjp 1784); Nand Kishore Yadav(bjp 1776); Ashwini Kr. Choubey(bjp 1741); Vasundhara Raje(bjp 1687); Yogi Adityanath(bjp 1617)
Development	Raghubar Das(bjp 2180); Suresh Prabhu(bjp 1922); Agriculture INDIA(other 1852); Dharmendra Pradhan(bjp 1485); Radha Mohan Singh(bjp 1308); Narendra Modi(bjp 1259); Vasundhara Raje(bjp 1229); CMO Chhattisgarh(bjp 1228); Piyush Goyal(bjp 1118); Dr. Pankaj Shukla(bjp 1078)
Corruption	Sanjay Jha(inc 1277); Tejaswi Yadav(other 833); Subramanian Swamy(bjp 765); BJP Delhi(bjp 679); Jitu Patwari(inc 655); Sanjay Singh AAP(other 644); Sanjay Nirupam(inc 638); Sambit Patra(bjp 607); Amit Malviya(bjp 576); Randeep Singh Surjewala(inc 552)
Inflation	Dharmendra Pradhan(bjp 839); Jitu Patwari(inc 307); Sanju Verma(bjp 277); Sanjay Nirupam(inc 244); Alka Lamba(other 238); Randeep Singh Surjewala(inc 229); Harish Rawat(inc 215); Sanjay Jha(inc 202); Sitaram Yechury(other 159); Yogendra Yadav(other 154)

Table 2

Regression Results for BJP and INC Politicians' Likelihood of Using Various Subtypes of Greeting-related Hashtags

<i>Dependent variable:</i>	
	had_used_hashtag
type_hindu	-0.172*
	(0.093)
type_national	-0.405***
	(0.110)
type_party	-0.623***
	(0.107)
type_sports	0.073
	(0.100)
partyinc	-0.383***
	(0.114)
followers_count	0.107***
	(0.037)
friends_count	0.092*
	(0.051)
statuses_count	0.283***
	(0.052)
type_hindu:partyinc	-0.077
	(0.181)
type_national:partyinc	-1.357***
	(0.335)
type_party:partyinc	-0.700***
	(0.254)
type_sports:partyinc	-0.300
	(0.208)
Constant	-6.373***
	(0.240)
Observations	168,597
Log Likelihood	-6,680.171
Akaike Inf. Crit.	13,386.340

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3

Regression Results for Readability and Sentiment with Respect to Party and Tweet Category.
Readability is Measured by Flesch, Gunning, and Smog Index. Sentiment is Measured Using VADER.

	<i>Dependent variable:</i>				
	Flesch	Smog	Gunning	Pos Emo	Neg Emo
	(1)	(2)	(3)	(4)	(5)
partyinc	0.620*** (0.085)	0.371*** (0.067)	0.367*** (0.101)	-0.030*** (0.003)	0.043*** (0.002)
partyother	0.115 (0.083)	0.111* (0.065)	0.075 (0.098)	-0.027*** (0.003)	0.020*** (0.002)
categoryCorruption	0.724*** (0.116)	0.735*** (0.092)	1.144*** (0.138)	-0.299*** (0.004)	0.109*** (0.003)
categoryDevelopment	2.039*** (0.112)	1.616*** (0.089)	2.506*** (0.134)	-0.259*** (0.004)	0.022*** (0.003)
categoryInflation	-0.322*** (0.124)	-0.106 (0.098)	-0.169 (0.147)	-0.303*** (0.004)	0.072*** (0.003)
categoryPoverty	1.343*** (0.114)	1.033*** (0.090)	1.377*** (0.136)	-0.286*** (0.004)	0.104*** (0.003)
categoryTechnology	1.593*** (0.112)	1.183*** (0.089)	1.939*** (0.134)	-0.252*** (0.004)	0.016*** (0.003)
friends_count	-0.286*** (0.060)				
followers_count	0.260*** (0.047)	0.336*** (0.036)	0.469*** (0.054)	0.013*** (0.002)	-0.007*** (0.001)
statuses_count	0.108 (0.066)	-0.009 (0.046)	-0.276*** (0.069)	-0.018*** (0.002)	0.006*** (0.002)
Constant	11.193*** (0.282)	10.819*** (0.214)	13.199*** (0.323)	0.449*** (0.010)	-0.00001 (0.007)
Observations	8,116	8,116	8,116	8,116	8,116
R ²	0.080	0.072	0.073	0.519	0.254
Adjusted R ²	0.079	0.071	0.072	0.518	0.253
Residual Std. Error	3.077 (df = 8105)	2.431 (df = 8106)	3.668 (df = 8106)	0.109 (df = 8106)	0.082 (df = 8106)

Note:

*p<0.1; **p<0.05; ***p<0.01

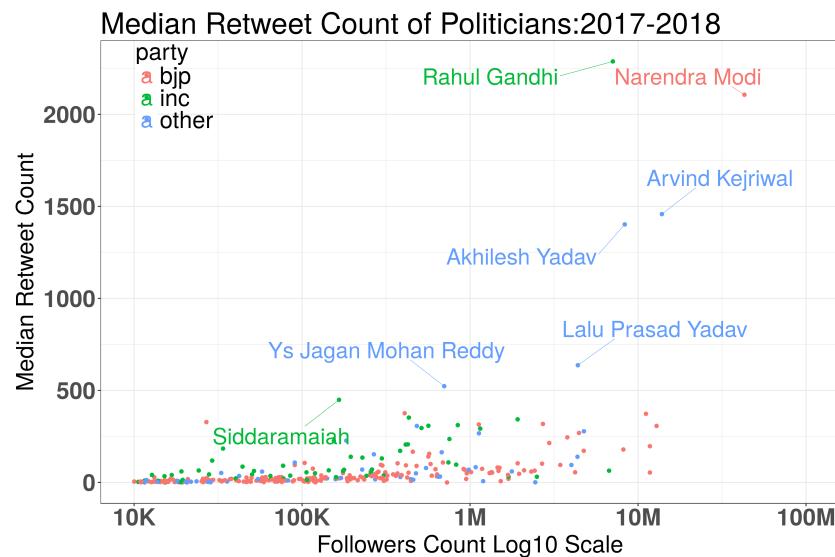


Figure 1. Median Retweet Count for Individual Political Account

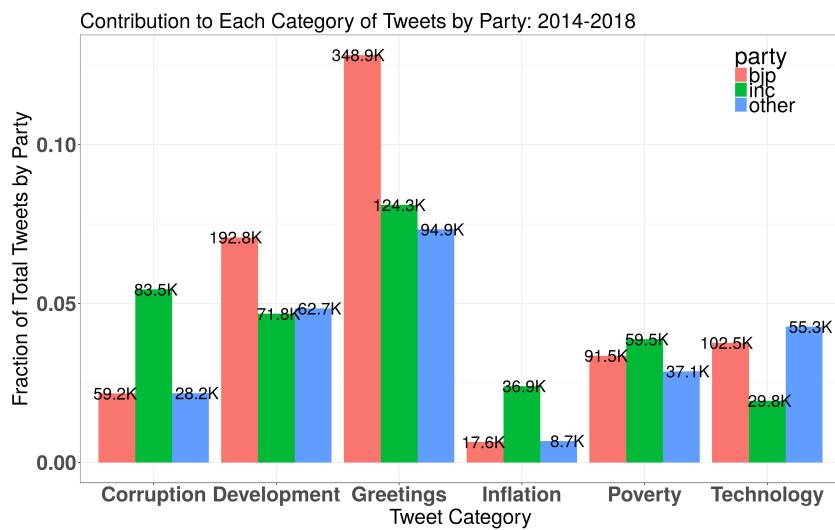


Figure 2. Party Contribution Distribution by Tweet Categories

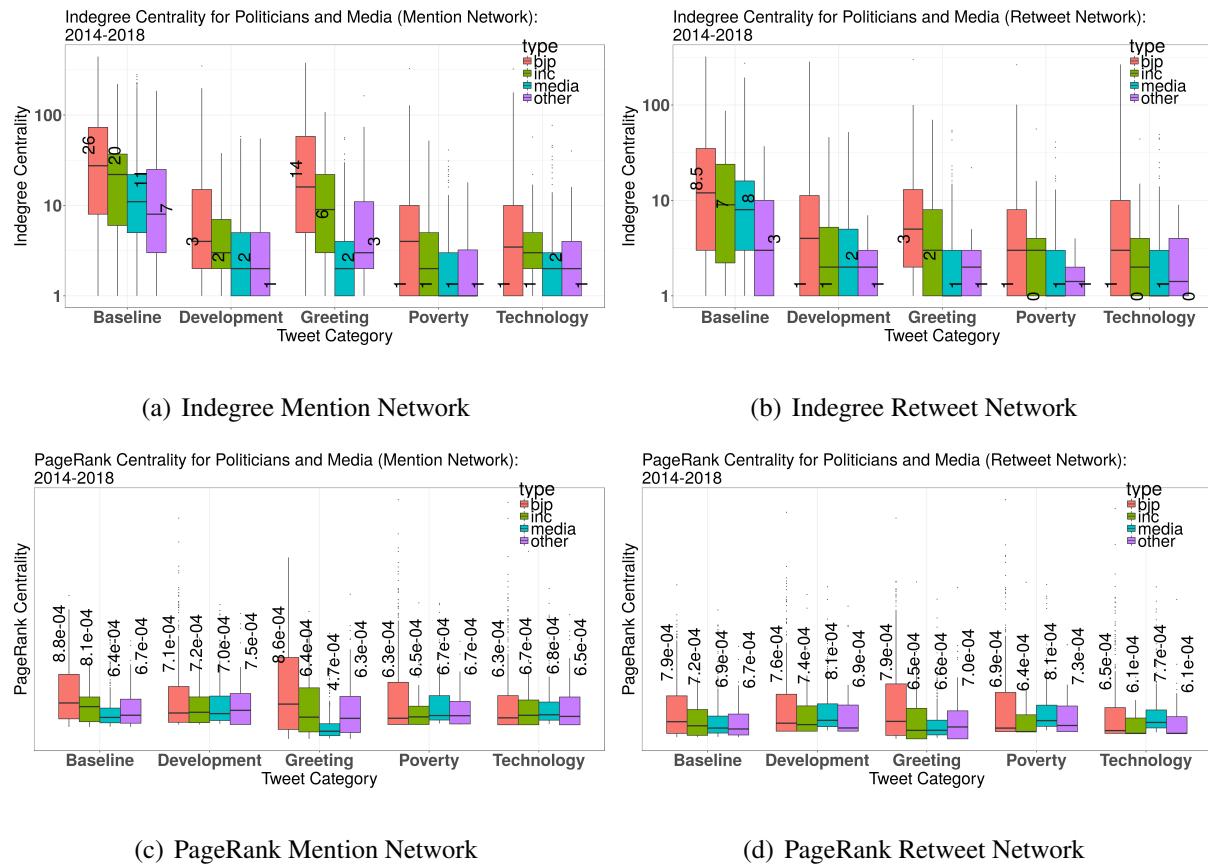


Figure 3. Indegree and PageRank Distribution for Retweet and Mention Networks

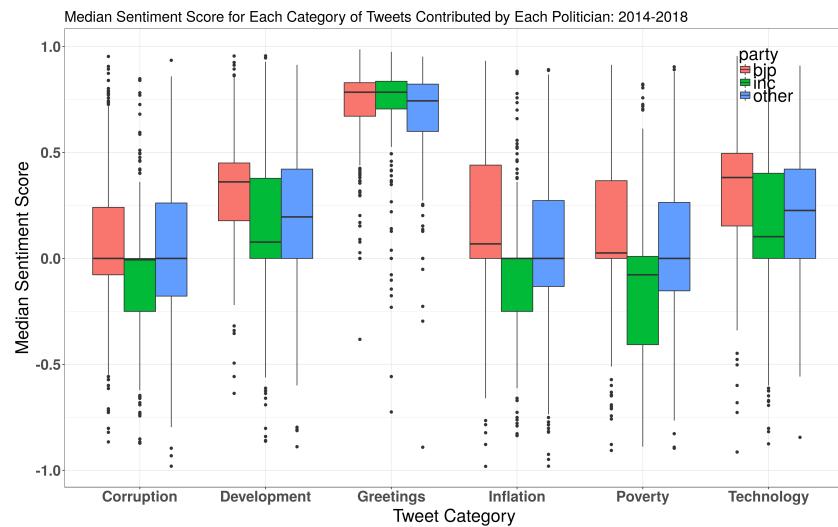


Figure 4. Overall Compound Sentiment Score for Each Category of Tweets

 **Narendra Modi** 
@narendramodi

Follow ▾

A new India is emerging, which is being powered by the strength & skills of 125 crore Indians. This India stands for development.

10:52 PM - 11 Mar 2017

7,607 Retweets 28,000 Likes



2.0K 7.6K 28K

 **Rahul Gandhi** 
@rahulgandhi

Follow ▾

FM Jaitley's genius combines with Mr Modi's Gross Divisive Politics (GDP) to give India:

New Investments: 13 year  Bank credit Growth: 63 year  Job creation: 8 year  Agriculture GVA growth: 1.7%  Fiscal Deficit: 8 year  Stalled Projects 



GDP Growth Seen Slowing To 6.5% In 2017-18, From 7.1% In 2016-17
GDP growth recovered to 6.3 per cent in the September quarter, higher than a three-year low of 6.7 per cent recorded in the April-June quarter.
[nitir.com](#)

7:16 PM - 5 Jan 2018

9,280 Retweets 20,146 Likes



2.7K 9.3K 20K

Figure 5. Narendra Modi and Rahul Gandhi on development



Figure 6. Narendra Modi and Rahul Gandhi on technology



Figure 7. Narendra Modi and Rahul Gandhi on poverty



Figure 8. Narendra Modi and Rahul Gandhi on inflation



Figure 9. Narendra Modi and Rahul Gandhi on corruption



Figure 10. Narendra Modi and Rahul Gandhi on greeting