News by Popular Demand:
Reputation, Ideology, and Issue Attention in Social Media

Natalia Aruguete* Ernesto Calvo † Tiago Ventura ‡

April 21, 2020

Abstract

In social media, activation and propagation of content depend on real-time interactions between users and media outlets selling their news on this market. In this new environment, what does it explain users' decision to share news articles in their social media posts? More important, how do news outlets incorporate users' behavior when producing their content? In this article, we thoroughly answer these questions using a theoretical model where readers "vote" on social media by embedding links to news organizations conditional on user-media cognitive dissonance and the outlet prevalence in the network. Following an extensive literature on spatial models of voting, the model estimates the importance of proximity across the network, the reputation of each news organization and provides a strategy to estimate where in the network each media organization has higher activation. We analyze Twitter data collected from the recent election of the populist leader Jair Bolsonaro in Brazil. The empirical application of our model shows where in the network users' ideological considerations had higher weight, how larger outlets find their sweet-spot in the center of a polarized network, and show how the community of the supporters of the President interact mostly with online conservative political operatives.

*UNQ, Conicet
†GVPT, University of Maryland, Corresponding author: ecalvo@umd.edu
‡GVPT, University of Maryland

1. Introduction

In social media, activation and propagation of content depend of the users' decision to read, click, like, hate, or share the content published in their network. In this environment where users can in real time activate information, indeed determining the number of people exposed to particular content, news outlets have become increasingly attentive to users' preferences when deciding their editorial lines. Through dashboards, ubiquitous in most news organizations, editors monitor the performance of their journalists, their published articles, and every type of media product, from photographs to cartoons, as soon as they go live on their websites. Technology allows editors to compare in real time the performance of competitors, both large and small, and to process a stream of feedback from vendors, clients, and donors. Readership matters for editors' decision.

In our previous work, we model how social media users' decide to activate content shared by their peers conditional on their issue cognitive congruence or dissonance in the network (Aruguete and Calvo, 2018). We build an empirical model in which stronger/weaker activation is measured according to faster/slower *time to retweet* which is explained by users' cognitive congruence/dissonance. Now, in this paper, we move one step forward on this agenda. We provide a model to explain when social media users decide to activate news articles from media organizations instead of focusing on activation only between users.

We provide a model that takes as inputs: i) the ideological congruence/dissonance between the user and the outlet and ii) the prevalence rate - what we call reputation - of the outlets in the network. The model answers three crucial questions. First, it estimates the importance of ideological proximity and reputation in the users' decision to activate content from media organizations. Second, the model assumes issue congruence varies across the network; therefore,

it allows the researcher to identify where in the network proximity has higher weight *vis-a-vis* more or less reputable outlets. Third, the model derives theoretically trough simulations the optimal location of news organizations, and using observational data proposes an innovative strategy to identify the where media outlets have higher activation in the network.

The crucial assumption of the model is that readership matters. In this sense, we build our model assuming outlets want to maximize their readership while voters consider proximity and prevalence in their decision. Under these assumptions, we argue our model has striking similarities with a long lineage of spatial models of voting in Political Science, where politicians maximize their vote-share considering positional and non-positional preferences of voters (Poole, 2005; Merrill and Grofman, 1999; Schofield and Sened, 2006) . The decision to embed a link from an outlet is how users "vote" in social media. Thus, our model begins with the decision by readers to (i) share content published by media organizations and peers. We then explain how editors could use this information to (ii) make editorial decisions conditional on the ideological leaning of readers, the reputation of the news organization, and the behavior of competing news organizations.

We provide an empirical test of our model using social media embeds to news articles in Twitter data. We analyze 2,943,993 tweets published by 162,107 high activity accounts from the Bolsonaro election in Brazil collected from September 26 through October 02, 2018. The election of Bolsonaro in Brazil is an unique political event in many directions: Brazil is one of the largest economies in World, his victory ends with a period of dominance of leftist parties in the country and the continent (Levitsky and Roberts, 2013), and, unfortunately, puts Brazil in the list of countries living the wave of election of far-right populist leaders. As a candidate and during the first months of his mandate, Bolsonaro actively invested in using social media as a strategic tool of campaigning and activation of his supporters, and, more important for this

research, used a vast network of online fake operatives to support his presidential candidacy. The Bolsonaro case works as a crucial test for our model of social media framing and editorship strategic positioning.

2. Proximity and Reputation on Users' decision

Let us consider the following hypothetical case. The user i_l who has more liberal views observes a news article with a similar content published by two different outlets in Brazil: a reputable nation-wide outlet and a small left-leaning online blog. A second user i_r observe receives similar inputs, news from a large newspaper, but instead now a small blog having very conservative views. i_l choose the activate the information of the large newspaper, while the more i_c chooses the right-wing blog. Assume this behavior repeats in the long run with multiple votes being cast following a similar distribution. What does explain the voter i_l decision to share the content from the larger newspaper *vis-a-vis* the voter i_c who decides to go more ideological? What are the consequences of this repeated game, with millions of real-time interactions, to the editors of these three outlets? Assuming editors want to maximize their readership, what should these actors do considering the behavior of the users?

Using our model as a device, we can quickly provide answers to these questions. We assume users' decision to share some news in their walls is derived from two parameters: ideological proximity between the user and the outlet, and the prevalence (reputation) of the outlet in the network. We argue these behavioral parameters both are heterogeneous across the social media users and their communities; in other words, proximity, *vis-a-vis* reputation, matters more heavily for some users compared to others. Additionally, the inputs provide our explanations for where and how each outlet behave in the network.

Returning to our basic example, on the voters' side our model would predict that i_c votes

for the right-wing less reputable blog because from him proximity has a higher weight than reputation; as a consequence, this user prefers to vote an organization closer to his view than supporting one with a broader audience. In the opposite direction, for user i_l reputation has higher weight, then she decides to vote for a larger more centrist outlet, instead of going with the one closer ideologically. On the supply side, the less reputable conservative outlet have weak incentives to moderate their content, while the reputable newspaper uses his reputation bonus to move more to the center of the network maximizing its market-share be receiving votes from centrists and leftists for whom reputation matters more.

The model we describe in the next section formalizes the intuition above. We provide a formal explanation to users' decision to share news content, and from this model we theoretically derive the behavior of news outlets in the era of social media. In doing so, we extend the scholarship on the multi-party competition by Adams et al. (2005), where voters select among multiple candidates in response to their competing policy offers and conditional on their distinct reputations for managerial competence. Different from voting models, however, media markets include a significantly larger set of media organizations. Similarly, users have multiple "votes" to communicate their preferences, as they constantly click and share content they agree with. Different from voting models, therefore, the total "votes" of each user is only limited by the level of interest and the attention span of their online behavior (how much time they spend interacting with content on their social media walls) as well as the number of publications covering an event (how prevalent is an issue on their walls). Therefore, our specification considers a count-model of "votes" (embeds) rather than the decision to opt for one candidate.

The Model

Our model has two actors: a social media user i and organization j. The model uses a utility function which decreases with the cognitive dissonance between the user ideological position, x_i^k, and the news published by an organization, N_j^k. The model is remarkably general, but we consider as organization j only media outlets, such as newspapers, blogs, online journals, among others in this category. The voters' utility also considers the prevalence of the actor j in the network, what we call reputation. Reputation matters in a sense that cognitive congruence is valuable insofar as information is trustworthy. We assume reputation to be fixed in our analysis, therefore, at the moment of publishing news, organizations have a fixed reputation. Consequently, news organizations have a standing reputation and can only maximize readership by altering the ideological leaning of the content they publish. [1]

Our model makes the simple assumption of the utility being negative on cognitive dissonance which means users are more likely to share content that agrees with their preferences. Additionally, reputation produces positive gains in the utility since, all else equal, we assume social media users prefer to interact with more reputable sources. The reputation scores and the ideological leaning of an organization, as observed by voters, could or could not be correlated, but for simplicity, we empirically assume the latter. Both ideology and reputation are also issue-dependent, but here since we empirically test our model on one issue, we propose a shorter version of our model in which issue in constant [2].

Therefore, a Reader i will perceive a utility from "voting" (reading, liking, or reposting news) on issue k by organization j as described in Equation (1):

[1] Future extensions of the model can consider reputation to be dynamic, but we do not go on that road here.
[2] See our working paper ..... to a more general model

$$U_q\left(ij\right) = -\alpha_{iq}\left(x_i - N_{jq}\right)^2 + \beta_{Rj} + \delta_i \tag{1}$$

In Equation (1), the quadratic term $\alpha_{iq}\left(x_i - N_{jq}\right)^2$, describes the disutility of a post that is further removed from the reader's preferred ideological position, x_i. For every unit of increase in cognitive dissonance, the utility of reader i declines by $-\alpha_q$. The parameter $-\alpha$ also has a natural interpretation as the weight that a reader attaches to the ideological leaning of a media organization in this network. For example, for a Brazilian reader, browsing about soccer may weigh less heavily regarding the weight of proximity than in the Bolsonaro network. In this sense, our results here are local using the statistical jargon, and speaks solely with the data in hand.

The equation (1) also indexes the parameter $-\alpha$ by q; here, we allow ideology to have a heterogeneous impact in the network with q representing different places in the network. We generate q by splitting the network into one hundred equally sized squares multiplying the deciles of each of the two dimensions of the network layout. In this format, we allow cognitive dissonance to vary according to where in the network the user is. Finally, equation (1) also shows that news published by a more reputable actor, R_j, increases the utility of reader i.

The choice function for equation (1) describing the likelihood of clicking in a particular news produced by media j out all organizations is described in Equation (2):

$$V_{ijq} = \tau_i \frac{e^{U_{(ij)}}}{\sum_{j=1}^{J} e^{U_k(ij)}} \ \forall \ i,j,q \tag{2}$$

In equation 2, the total number of clicks, likes, or reposts is a function of the probability that a user will select a post by agent j in the network square q, subject to the user's time constraints, τ_i, which describes the number of times a user will vote. That is, some readers may

click, like, or share many news while others may do so sporadically.

The Comparative Statics of the Model

Equation (1) explains the utility derived by readers in social media considering their ideological proximity to a particular media organization and the reputation of the outlet. We model the editor's decision regarding the editorial line of their news organization as a maximization problem that is conditional on the utility function of readers. Subject to Equations (1) and (2), we need to solve for the ideological position of N_j^{k*} that will provide the largest count for editor j by readers, conditional on the $N_{(\neg j)}^{k}$ for all other organizations. This is a difficult numerical optimization problem, but Luckily for us, Adams et al. (2005) provide an algorithm to address this type of numerical optimization problem. For a further discussion about the algorithm, we suggest the reader to consult Adams et al. (2005). From now on, we focus instead on the comparative statics of the model to formulate the hypothesis we test here regarding the location of the editors.

We estimate the comparative statistics mapping variations on three parameters of the model: first, the importance that readers attach to ideological considerations (α) on issue k; second, the importance that readers attach to the reputation of the news organization (β) on issue k; third, heterogeneous values for reputation for the media outlets. From the results, we derive the hypothesis of the paper

Implications of the model

After running equilibrium models for all parameter permutations (approximately 2 million solutions), we post-process the data to assess the effect of users' preferences on the optimal ideological content published by media organizations. In the next few pages, we describe the implications of the model for one particular environment that reassembles the social media context tremendously.

We describe the implications for editorial choice in cases where the distribution of the users vary and ideological attachment and the reputation of media organizations are heterogeneous across the users. Putting in other words, we assess how editors should react to: a) heterogeneous levels of cognitive congruence/dissonance of readers with the ideological leaning of the outlets, in other words, the impact of ideological proximity, b) under different distribution of users' preferences, and c) the presence of media outlets with more or less reputation.

First, we test how organizations react under two different distributions of the users' preferences. We consider both a normal distribution (non-polarized media market) and a bimodal distribution (polarized media market) for users' preferences. Most applications of Adams et al. (2005) consider a voting population with preferences normally distributed. The higher mass of voters in the middle of the distribution drives parties with higher valence to the middle while crowding out smaller parties to the ideological extreme. In our case, we consider not only a normally distributed population of readers but also a mixture distribution with two-modes, where the center has few readers (the median is thin), and where most readers are either on the left or right of political spectrum.[3] As we will show, in polarized political environments there is a stronger centrifugal effect that pushes high reputation organizations to the region that falls between the overall median voter and the high-density regions on the left and right. Therefore, we compare non-polarized and polarized distributions of voters and their effect on the editorial decision of news organizations.

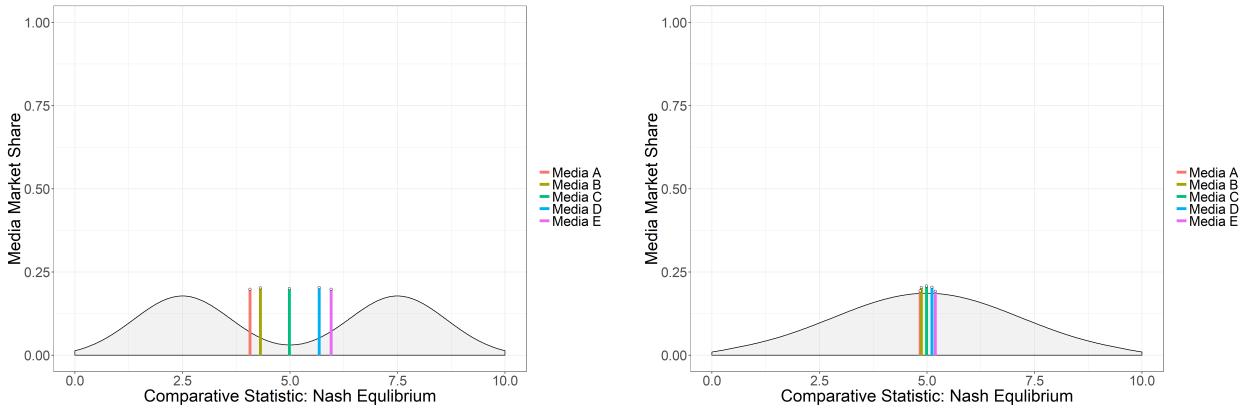
Let us begin by holding the importance that readers attach to ideology and reputation to their median levels, $\alpha = -0.06$ and $\beta = 0.6$. We also allow the $cov(\alpha, \beta) > 0$ to be strictly positive, with readers on the left having higher assessments of reputation for Media A and B, while readers on the right have higher assessments of reputation for Media D and E. We also

[3]We combine two normal distributions, mean centered on the left, 2.5, and the right, 7.5, of the political spectrum.

set reputation values for all organizations to be identical, $R_1 = R_2 = ... = R_5$. Figure 1, therefore, presents the comparative statistics for the equilibrium location of the organizations. Here, we are only describing the effect on readers' polarization of the ideological leaning of each media organization. The plots describe the ideological position of a media organization on the horizontal axis and the corresponding market share in the vertical axis in equilibrium.

Given that readers with different ideological leanings have heterogeneous assessments of each media's reputation, changes in the distribution of the readers' preferences yield changes in the optimal editorial line of media organizations. The left plot shows how polarization among readers pulls media organizations away from the center of the distribution. When the social media environment is not polarized, by contrast, the centrifugal effects on the optimal editorial is weaker.

Figure 1: Comparative Statistics: Impacts of readers' ideological polarization



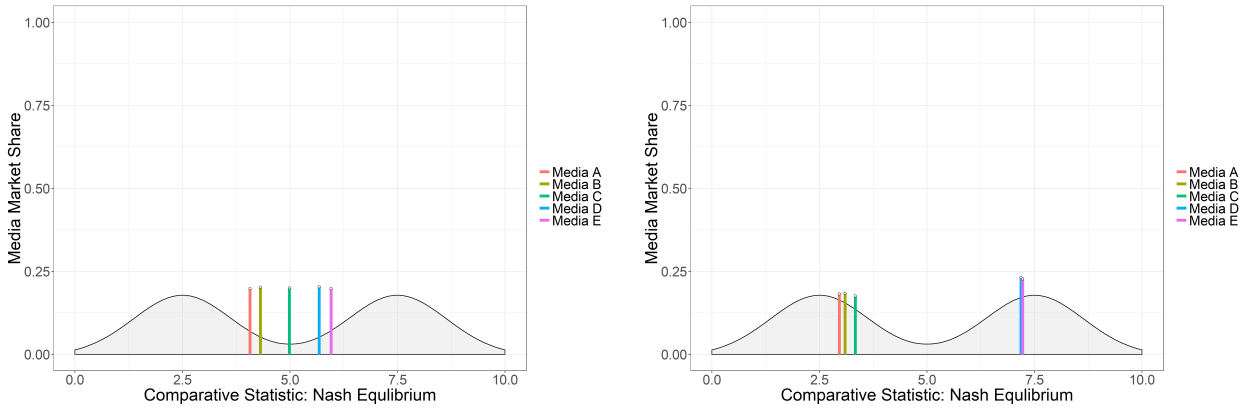
a) Bimodal Normal Distributions b) Normal Distribution

Consider now the situation in which readers increase the weight or importance of ideological concerns when sharing news content on issue k. Figure 2 presents the equilibrium positions for the media organizations, holding all parameters to their median values except for the weight of ideology (*alpha*), which increases from -.06 to -.12. The equilibrium positions indicate that as readers become more ideological, the media organizations move more to the inner hillside of each mode towards the leftist or rightist median voter. That is, towards the local median reader in the

left and right of the political spectrum. Notice that more ideological readers does not mean more extreme readers but, rather, that they care more about cognitive congruence when activating content. In fact, the underlying distribution of readers has not changed in this example. Only the intensity of readers ideological considerations on issue k.

Larger negative values of *alpha* indicate a sharp decline in the activation of content as the post moves away from a reader. Consequently, in more ideological environments (right plot), media organizations move to the median voter on the left and right of the political spectrum as ideology (cognitive congruence) weights more heavily on the decision of readers to activate content. In conclusion, as ideological attachments increase the global median becomes empty space.

Figure 2: Comparative Statistics: Effect of the weight of ideology, α, on the media's equilibrium positions



a) α = -.06 b) α = -.12

Increased salience on issues that weigh heavily on ideological considerations, therefore, will more forcefully pull media away from centrist positions. When cognitive dissonance matters, media organizations should cater to their ideologies. That is, to the ideologues that already give them high reputation marks for the news they deliver.

Therefore our first hypothesis:

H2: As readers are more polarized or give more importance to ideological considerations, polarization in media Posts increases

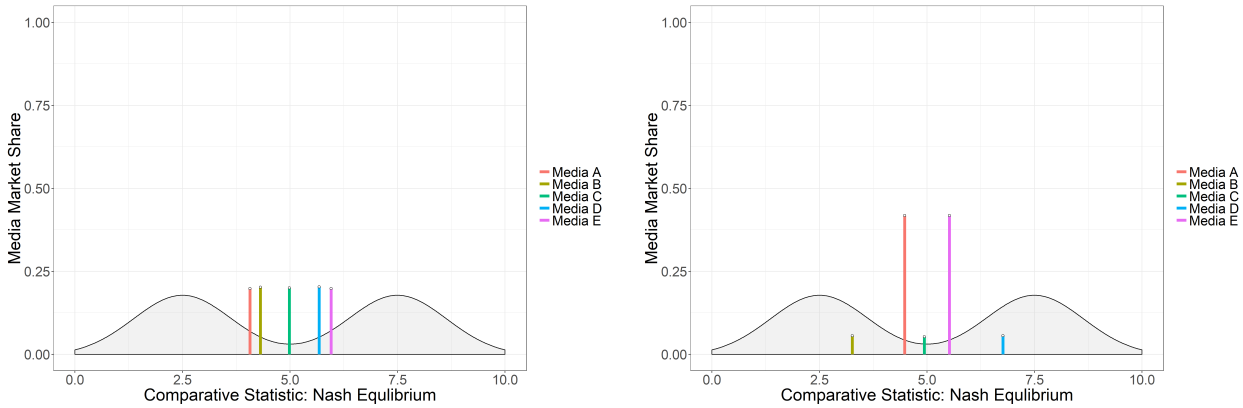
In the previous example, the mean reputation score across media organizations was set to be identical. Therefore, all the work was done by the weight that readers attached to ideological congruence/dissonance. However, both on the left and right of the political spectrum, there are news organizations that are perceived by readers as having a higher/lower reputation than their competitors. Our utility function indeed argues that users decide where to "vote" on a particular outlet by weighting ideological consideration and how reputable each media organization is perceived. Here, we explore how reputation matters for our model predictions.

We simulate an environment where all the parameters are set to their median values, but reputation for each organization varies. The comparative statics of the model shows that organizations that have higher overall reputation scores (the average reputation of Media A for all readers is higher than of Media B, for example) will move centripetally, towards the median reader, while organizations with lower reputation will be crowded out to more extreme ideological locations. Figure 3 presents the equilibrium location of media organizations under parity and asymmetric reputation. In the later, Media A and Media E are recognized as having higher overall quality (Reputation) than Media outlets B, C, and D. Notice that Media A and Media E also have more ideologically extreme readers. The comparative statics of the model provide stronger evidence of the centripetal shift of high-quality outlets.

Our hypothesis suggests that more reputable news outlets maximize readership by taking advantage of their reputation surplus, moving further away from their natural readership (readers with higher assessments of the reputation for that media) towards the overall median reader. In other words, reputation advantages lead to moderation (Calvo and Murillo, 2019). The smaller outlets, on the other hand, are captive to niche readers and, in a polarized environment, move

away from the median reader. The effect of reputation survives in normally distributed media environments, even if there is a stronger drive by all media organizations to gravitate towards the overall median reader.

Figure 3: Comparative Statistics: Impacts of Asymmetric Reputation



a) Parity Reputation

b) Asymmetric Reputation

Therefore our second hypothesis:

H3a: When Reputation is Asymmetric, High Reputation Organizations move to the center

H3b: When Reputation is Asymmetric, Low Reputation Organizations are crowded out to more marginal ideological locations

In the next section, we analyze Twitter data from the #Bolsonaro election in Brazil. We examine the rate at which users embed links to different media organizations and estimate the ideological weight and reputation parameters that explain the centripetal or centrifugal placement of media organizations. We propose a novel way to estimate the latent position of the media j by issue k in the network. Finally, we show how more reputable media occupy the center of the network, and are less dependent on ideological proximity to activate their readers.

3. Embedded links in the #Bolsonaro election

In 1987, when Brazil had a civilian president for the first time after two decades of a brutal dictatorship, Jair Messias Bolsonaro, a captain of the Brazilian Army, won his first election as a local councilor for the city of Rio de Janeiro. His political career started after an interview for the Brazilian Magazine *Veja* speaking out for better wages for the members of the Brazilian Army. In 1990, only four years latter, the former captain was elected a Federal Deputy for Lower Chamber of the Brazilian Congress. During the following five elections, Bolsonaro was reelected for the House until 2018. when Bolsonaro jumped directly to the Presidency of Brazil.

If, in 1986, Bolsonaro won his first election relying on a parochial agenda targeting improvements of the living conditions of the active members of the Armed Forces, in his successful presidential bid a more complex amalgam of interests were in play. This time the elected president ran as a legitimate far-right candidate known around the entire world for bombastic quips demeaning women, gays and people of color, public threats against his opponents, denies of the existence of the global warming, public support for the years dictatorship in Brazil, and a controversial program to fight crime in Brazil based on legal exclusions for police to use lethal force.

From 2014, when Bolsonaro announced his candidacy, to the first months of the electoral campaign of 2018, he was nothing more than an underdog running on a extremely radical, populist agenda. Despite polling consistently in the early surveys, the Brazilian political mainstream and pundits never believed that such a conservative candidate would be able to win the election in the run-off system Brazil adopts. Bolsonaro started with a slim partisan coalition, with only two small parties, reduced campaign donations and official propaganda, choosing a vice-president in the last moment after three of his favorites names denied his invite. Not only a radical, Bolsonaro

was also mostly an unlikely competitor to the presidential race.

How did a far-right underdog candidate win the Presidency in one of the leading economies, political complex, and most diverse countries in the World? The answer to this question is likely to mobilize crucial parts of the political science scholarship in the next few years. In this paper, we disentangle the dynamics of one of the crucial battlegrounds where the electoral dispute was fought: the social media. The election of Bolsonaro has a variety of meanings, yet one of the most important for academic purposes is the centrality of social media platforms for electoral mobilization, framing, and propagation of Bolsonaro's messages. Bolsonaro structured his entire campaign around direct, personal links with his supporters through his Twitter and Facebook accounts, live call, and youtube videos, in particular after the knife attack suffered by him during the first month of the campaign when he was forced to avoid public appearances and debates. In this environment, the Brazilian 2018 election was flooded by false rumors, manipulated photos, decontextualized videos and audio hoaxes in a variety of social media environments, with some ongoing research already giving signs of the overuse this strategy by the winning campaign Tardaguila et al. (2018). We hope that a proper understanding of the episodes involving Bolsonaro's election may help political scientist to disentangle recent changes in the political world and the importance of social media in contemporary politics.

The Data

From September 26 through October 02 of 2018, we connected to the stream API provided by Twitter to collect 5,325,240 tweets that included the characters *"Bolsonaro"*, without the hashtag to ensure that our search captured all tweets that included the string Bolsonaro. Over million of all tweets were posted by users in the primary connected cluster of the #Bolsonaro

network.[4] Using the Fruchterman-Reingold algorithm, we estimated the network's layout, with [x,y] coordinates for each user. We then used the walk.trap algorithm in *igraph* to place each account in their community. The walk.trap algorithm identified two large communities, which we identified by their authorities as aligned with the opposition to Bolsonaro (91,116 users) and aligned with the Bolsonaro campaign (62,289 users). The remaining 8,702 accounts were placed in 1,38 smaller communities weakly connected to the core of the network.

[4]To define the primary connected network, we created a network that included all tweets that were re-tweeted at least once. We then thinned the network to preserve users that re-tweeted at least once or that re-tweeted at least three times. Finally, we selected the primary connected cluster of this network, which left the 196,066 most active users and 2,943,993 tweets.

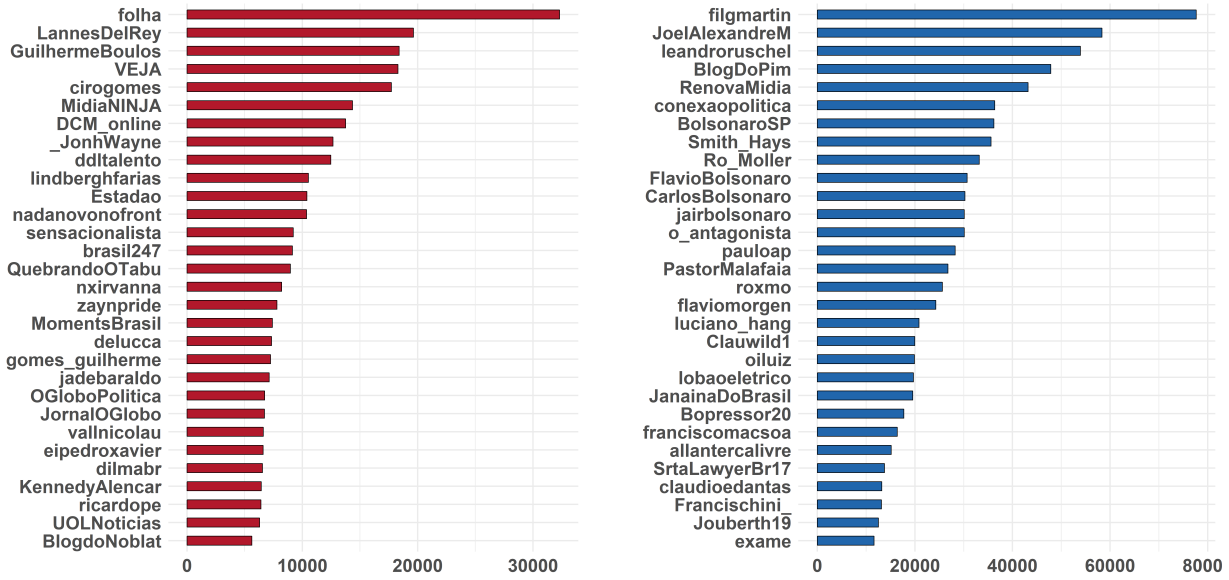


Figure 4: Authorities in the sub-networks aligned with the the Anti-Bolsonaro Communitiy (red) and the Pro-Bolsonaro users (blue)

In Figure 4 we list the most re-tweeted users in the pro-government and opposition communities. In the anti-bolsonaro community, eight of the top ten users had *verified* accounts and included well-known politicians such as two of the Presidential candidates running by the center-left parties Ciro Gomes (@cirogomes) and Guilherme Boulos(@GuilhermeBoulos) in 2018, and the senator Lindbergh Farias from the Workers Party (@lindberghfarias). The list of authorities also includes some left-wing news organization as Midia Ninja (@MidiaNINJA), a network of media activism in Brazil, and *Diario do Centro do Mundo* (@DCMonline), a well-known alternative blog connected to the Workers Party (PT). The surprising finding here relates to the presence of two of the major outlets in the list of authorities in the community in opposition to Bolsonaro: the magazine *Veja* (@VEJA) and the newspaper *Folha de Sao Paulo* (@folha). In the past, these major outlets were rarely activated by users aligned with the Workers Party and other more left-wing communities, as our previous research on the #Dilma crises indicates (Calvo et al., 2016), showing how unique was the election of Bolsonaro and already opening some puzzles of how editors adjust their editorial lines according to each issue.

By contrast, only three of the top ten users in the Pro-Bolsonaro community had *verified* accounts. It calls the attention the presence of well-known *fakes*, *trolls* and anonymous websites in this network, such as @JoelAlexandreM, @conexaopolitica, @RenovaMidia. Among the verified profiles on the top-ten accounts of this network are the accounts of Bolsonaro's sons, the elected senator @FlavioBolsonaro and the Federal Deputy @BolsonaroSP, the journalist @BlogdoPim, who is one of the founders of @Oantogonista, one of the key news outlets aligned with the Bolsonaro community, and @filgmartin as the main authority that is currently the special assistant of Bolsonaro's government for foreign affairs. The comparison between the two communities was striking, while the leading anti-bolsonaro users included institutional accounts of highly visible politicians and journalists, the Bolsonaro's bid was primarily carried out by anonymous political operatives. Later in the campaign, during the run-off round, a front-page report from Folha

de Sao Paulo show vestiges that some Bolsonaro supporters, mostly businessmen, were illegally bankrolling a service of automatic messaging to bombard WhatsApp users with fake news about Bolsonaro's front-runner opponent, Fernando Haddad. Such significant differences provide more evidence for the perception discussed in Brazil and abroad by newspapers and pundits that the spread of false information was a deliberated strategy of Bolsonaro's campaign, which our exploratory analysis indicates that ended up being incredibly effective giving the centrality of non-verified profiles on his community of supporters.

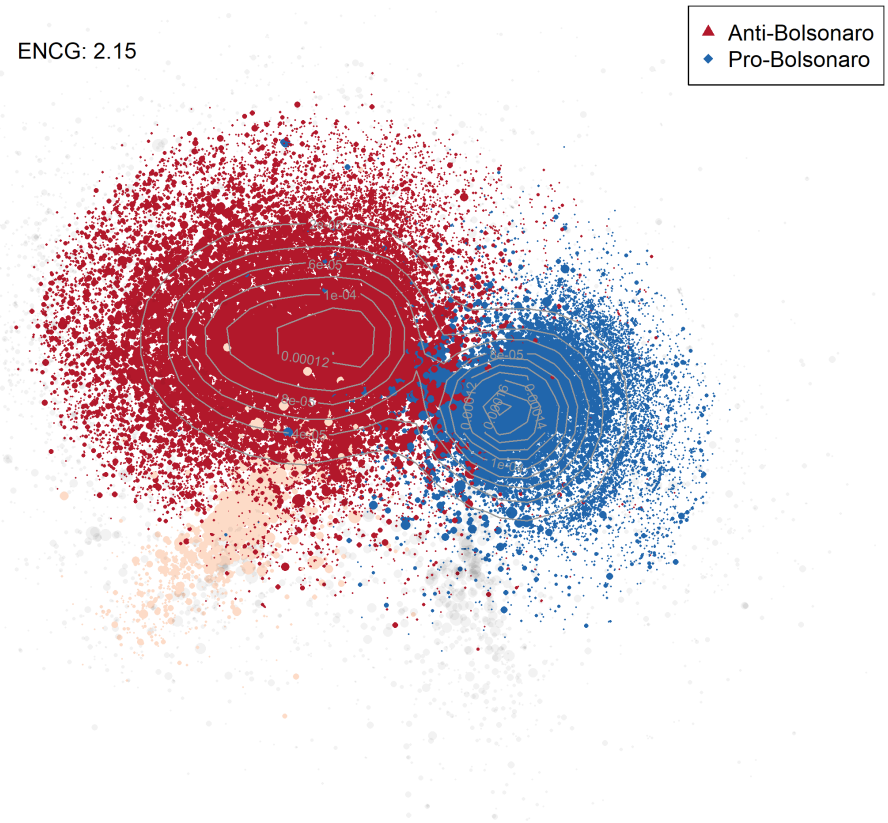


Figure 5: Primary Connected Network of #Bolsonaro. Blue dots describe users aligned with the Bolsonaro. Red dots describe users aligned with the opposition. Layout of users estimated using the Fruterman-Reingold algorithm in *IGraph*. Community detection using Walktrap algorithm in *IGraph*, (Csardi and Nepusz, 2006)

Figure 5 provides a visualization of the #Bolsonaro network, with users aligned with the President-elect in blue circles, users aligned with the opposition in red diamonds, and the rest of the users in gray dots. The size of the nodes is weighted by the in-degree of each user, with larger nodes indicating accounts that were re-tweeted more often. The community of the opposition is 30% larger than Bolsonaro's supporters' cluster. International newspapers and blogs form the small community colored in pink circles in the graph. The interesting caveat here is how the international community is more aligned with the opposition already giving signs of the concern that took place in the international actors with the election of the far-right candidate in Brazil.

Out of the 5,325,240 million tweets in the #Bolsonaro network, slightly over 15.3% included hyperlinks to content already published online, 816,694/5,325,240 = .1534. Links to the top 24 media outlets represented 78% of all hyperlinks, 640,595/816,694, with almost a third of them connecting to existing twitter posts and the other two-thirds directing readers to news organizations such as Pagina12, La Nacion, Clarin, and El Destape. While only 15% of tweets included hyperlinks to other media, it is worth noting that 97,160 accounts out of the 196,066 tweeted or re-tweeted content with hyperlinks to news organizations. Therefore, over 45% of the users in the #Bolsonaro network activated content from other sources.

Activation of News Content in the #Bolsonaro Network

Visual inspection of Figure 6 shows that media organizations were activated to a different extent by Pro-Bolsonaro (blue) and opposition users (red). Each plot in Figure 6 describes the region of activation of a different media outlet, measured by the number of times that each user posted or retweeted content from each media source. The figure provides a precise estimation of the relevance of each organization in the twitter network, as well as incorporating heterogeneous

activation per each cluster of users. As shown in the different plots, some links were activated by both Pro-Bolsonaro and opposition users, such as those of Twitter, some agencies connected to the *Folha de Sao Paulo*, *Veja*, and *Globo*. Others, by contrast, were primarily activated by only one of the two communities, as was the case for *Diario do Centro do Mundo*, *O Antogonista*, *Sensacionalista* and *Cidade On Line*, among others.

There are some insightful considerations about the activation maps provided in figure 6. First, larger outlets, as our theory expects, tend to be shared more broadly by both communities. In our theory, reputation concerns might explain this behavior. Users are more likely to embed links from more reputable outlets, moderating their ideological proximity when deciding which news to activate. Second, the figure gives a clear picture of the polarized environment of the Brazilian election. Beyond the more reputable outlets, most of the other source of news, such as *O Antagonista*, *Brasil 247*, and *O Sensacionalista*, are shared exclusively by one of the communities with very littler cross-community exchange of embedded links.

Finally, returning to the issue of activation and propagation of false information in the Bolsonaro's community, figure 6 allows us to make some additional considerations. One of the differences between both communities is precisely the degree to which Pro-Bolsonaro users embed links from anonymous political operatives online. *Conexao Politica*, *Tribuna do Ceara*, *Republica de Curitiba*, *Jornal Cidade On line* are all examples of unknown websites who work, in general attacking progressive social movements and politicians in Brazil, but that in the 2018 election, as ammunition for Bolsonaro's campaign strategy of propagating fake news. We do not observe these operatives with the same centrality in the opposition network of embedded links. In the latter group, the propagation of news comes mostly from left-wing journals and website, who are not anonymous sources, and also from more reputable, well-known outlets.

To be precise about our argument, we do not argue that the use of false information, trolls,

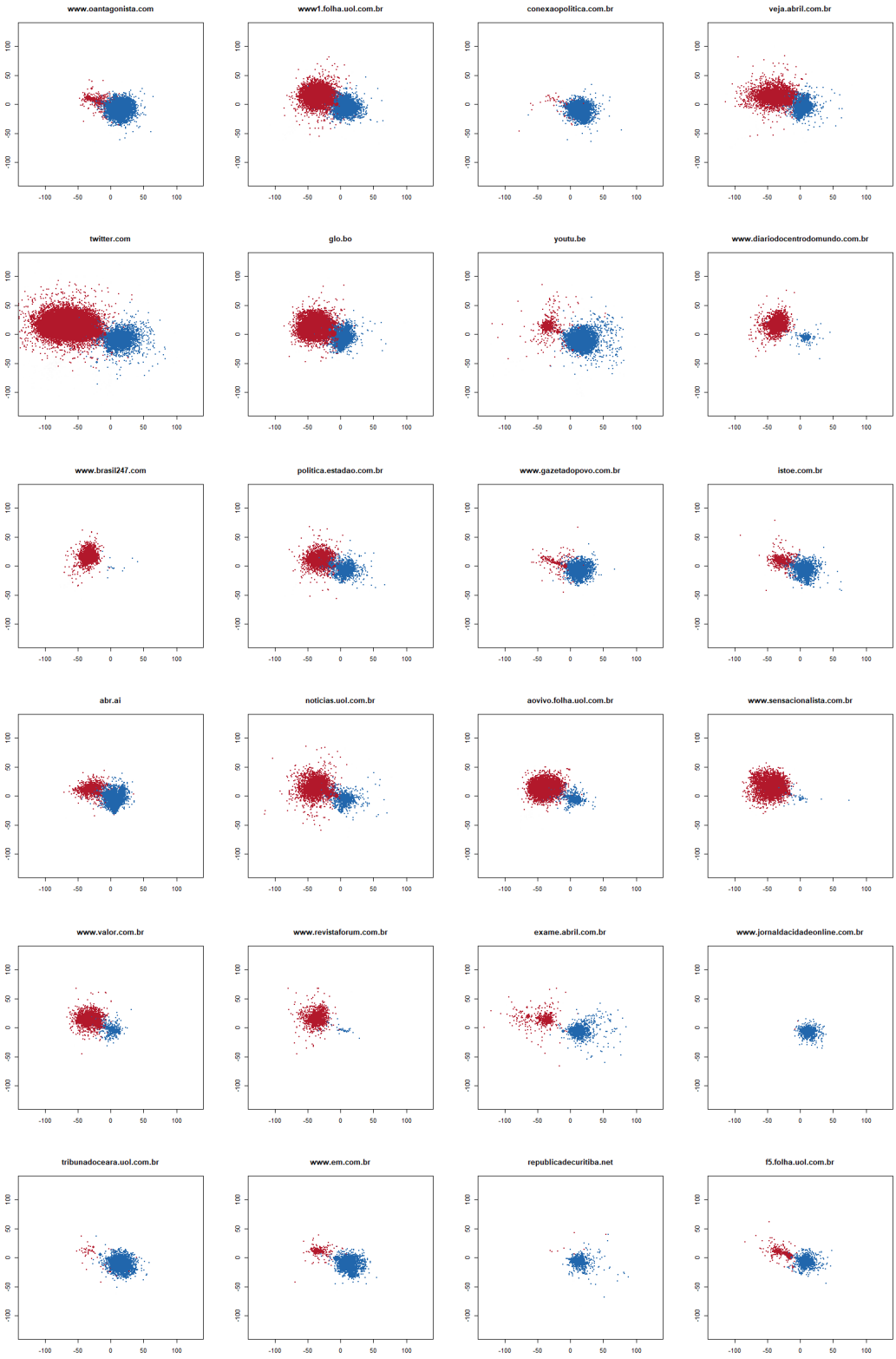


Figure 6: Top 24 Embedded news in the Primary Connected Network of #Bolsonaro. Blue dots describe Pro-Bolsonaro users. Red dots depict Anti-Bolsonaro accounts. The graph represents the activation of each news on both communities

and anonymous websites were a strategy adopted only by one side battling in the electoral dispute; previous research has already shown that, despite being framed in distinct ways, fake news were quite spread among different communities d (Tardaguila et al., 2018). What we do argue is that the relevance of these machines producing daily amounts of fake news is notably higher among the users clustered in the community of supporters to the President-elect Bolsonaro compared to the other groups of our network.

4. Modeling the decision to embed links in #Bolsonaro

We start our analysis discussing how users' decisions of embedding links vary by communities in the network. Figure 7 provides separate Poisson models describing the count of links to each news organization, with pro and anti Bolsonaro accounts dummied out and 95% confidence intervals around the coefficient estimates [5]. In Figure 7, blue dots describe parameter estimates for Bolsonaro's followers and red dots describe parameter estimates for the opposition. Given that larger coefficients describe higher numbers of links embeds, we can readily compare media outlets more frequently linked by Pro-Bolsonaro users, such as *Youtube*, *Tribuna do Ceará*, *República de Curitiba*, *O Antoganista*, *Jornal Cidade On Line*, *Istoé*, *Gazeta do Povo*, *Exame Abril*, *O Estado de Minas (EM)*, and *Abril (abr.ai)*, to media outlets more frequently linked by the opposition, such as *Valor*, *O Sensacionalista*, *Revista Forum*, *Noticias Uol*, *Globo*, *Diário do Centro do Mundo*, *Brasil 247*, and *Folha* .

Figure 7 provides similar findings related to the centrality of anonymous outlets among the Bolsonaro's community. The majority of the relevant links shared by this community comes from low reputable, anonymous websites. As mentioned before, *Tribuna do Ceará*, *República de Curitiba* and *Gazeta do Povo* are example of these operatives. Meanwhile, the opposition

[5] The model links from *F5.folha.uol* as the dependent variable ran into separation issues due to lack of variation; therefore, we did not add the parameters to Figure 7

cluster of users tends to rely more frequently on more reputable outlets. The difference of the activation of Youtube is also interesting since this platform works as a repository for videos made to propagate fake news, also more activated by Bolsonaro's community.

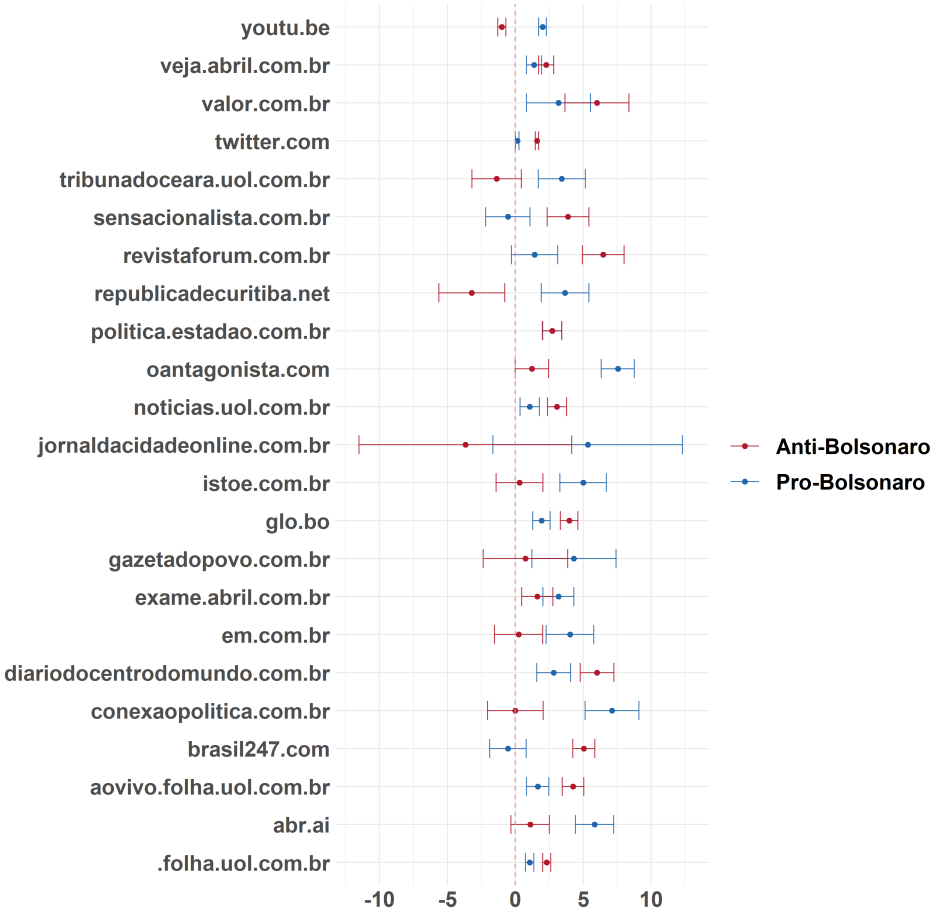
Additionally, Figure 7 provides insights about heterogeneous propagation of outlets across the communities. Take for example the *Revista Forum*, an outlet traditionally more aligned with the opposition to Bolsonaro, and controlled by the Journalist Renato Rovai, a long-age supporter of the Workers Party. The results for the Poisson model show a sharp difference between links from *Revista Forum* embedded by users more aligned with the opposition and the Bolsonaro's community. The same division occurs with *Diário do Centro do Mundo* in favor of the opposition and with the Pro-Bolsonaro operatives mentioned earlier.

An additional finding of these models relates to the comparison between smaller and some more reputable media outlets. For the cases of links from the *Folha's* organization, *Istoé*, *Veja*, and *Estadão*, all more reputable organizations, the gap between the point estimates is not as large as when compared with the smaller outlets and the fake news websites we discussed above. Finally, the figure also predicts how polarization between the communities is more significant among Bolsonaro's supporters; the gap between red and blues dots tends to be larger for outlets more aligned to Bolsonaro than when compared to the gap of outlets more activated by the opposition.

Our theory provides some elements to make sense of these perceived differences. The #Bolsonaro issue in Brazil was deeply ideological resulting in a sharply divided network, and, as we show on figure 5, mostly dominated by the opposition. The comparative statistics for figure 2 hence explain our empirical results quite clearly. When users are polarized, cognitive dissonance has larger impacts on users' decision to activate news organizations. Therefore, one can identify a division between which the outlets activate by each community. As we also discuss theoretically,

reputation moderates the results. More reputable outlets, such as *Folha's*, *Veja*, and *Estadão*, calibrate the impact of cognitive dissonance in their market-share in the network while smaller organizations have a sharper division between the communities.

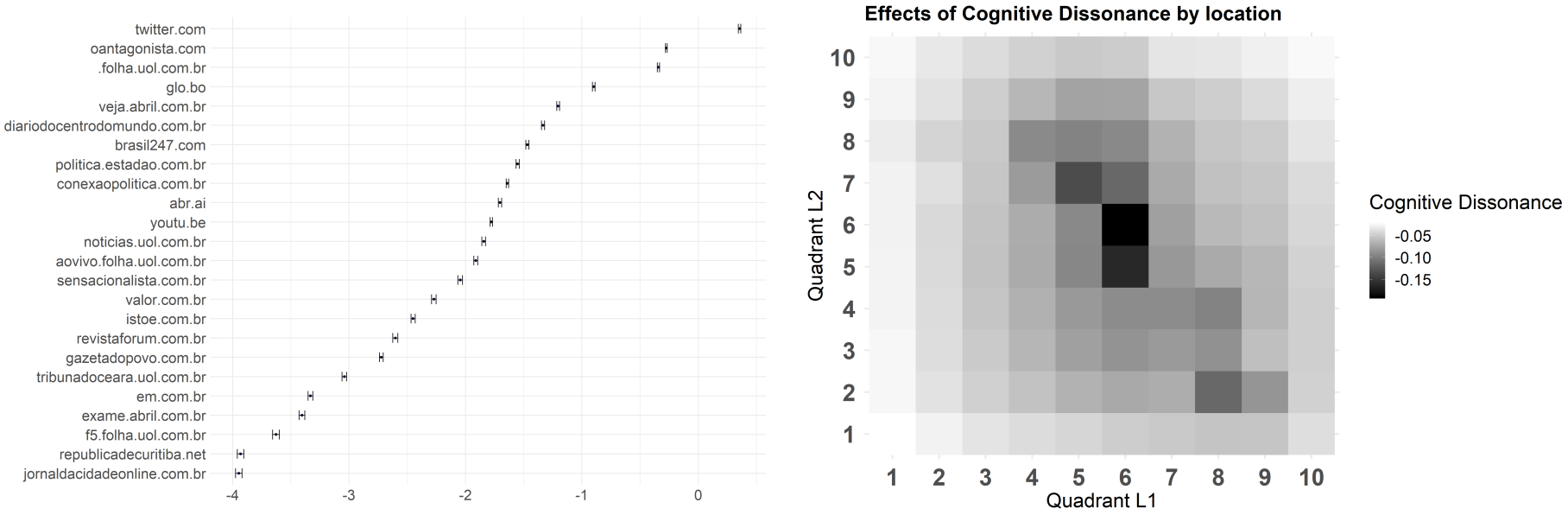
Figure 7: Estimates of Media Embeds (Poisson) for Pro-Bolsonaro and Anti-Bolsonaro users



The basic Poisson model, however, does not allow us to precisely discriminate how much of the count rate is explained by cognitive congruence/dissonance and how much is to be explained by reputation. To better approximate the formal specification in Equations (1) and (2), therefore, we will model the extent to which activation is dependent on the user location in the network space. To model the user decision, we approximate the utility function on equation (1) estimating a generalized linear binomial count model using a logistic transformation. We consider the user decision of embedding the link of news published by a media organization j as a voting decision,

as described in Equation (2), and our data is composed of aggregate counts of those votes by news organizations. We consider links from each user to the different media organizations, resulting in 686,428 observations that report counts of embedded links by 74169 users, and 24 media organizations. As described earlier in this paper, 45.7% of users activated at least one media organization and each of them can be mapped onto the network's layout.

Given that the highest density point of each media organization N_J is unobserved, we need some approximation to calculate the parameter α. We approximate the location N_j using deriving a weighted average of each user i location in the two dimensions of the network by the number of links embedded from each organization j. After estimating these points, we calculate the Euclidean distance between the user location and the highest density point of each media in the network. We also model the cognitive congruence/dissonance of readers with the ideological leaning of the outlets having heterogeneous effects across the network. In other words, we divide the network space in deciles on each dimension of the network space allowing our parameters for ideology to vary spatially. Figure 8 presents the results of the model.



a) Reputation b) Ideology

Figure 8: Estimates for the Media Outlets' reputation and ideological decline in the network space (Binomial Count Model)

Figure 8 presents estimates for the media outlets reputation and the importance of ideology on congruence/dissonance by user. The left figure plots the intercepts for each media in our binomial count model, while the heat-map on the right plots the estimates for the decay function of proximity in each decile of the network. The left plot shows which outlets have a higher reputation on issue k. Larger outlets in Brazil, as we expected, depend less on congruence and dissonance of the ideological preferences from the users. The exception here is the website *O Antagonista* that appears as one of the outlets leading our estimation for reputation. We explain this finding arguing based on the results of 6; *O Antagonista* is the link with broader activation on the Bolsonaro's community receiving high and equally shared attention on most of the space occupied by these users. Therefore, despite not figure as a traditional outlet in Brazil, in the #Bolsonaro network, *O Antagonista* appears as the broadest source of news in one the main group engaging in the debates in this network.

In the other side, smaller news organizations, like independent blogs such as *Jornal da Cidade On Line* and *Republica de Curitiba*, two of the fake news operative highly activate among Bolsonaro's supporters, are on the other extreme of the reputational scale. The results indicate that the latter outlets derive their attention mostly due to the proximity of the network, what we call here cognitive congruence. The right plot indicates that where proximity matters more in the network. Comparing to the basic plot from figure 5, the centers of the Pro-Bolsonaro and, in particular, the opposition community represent the areas where proximity has higher impacts on users' decision to embed links. The core members of each network express more concerns about how close to them the outlet news on the network is observed; in other words, people on the core of the network care more about ideological distance; this finding goes in the same direction of similar research using social media data but using distinct empirical strategies (Barberá et al., 2015) .

To provide a more intuitive understanding of the results on figure 8, we estimate the probability of observing an embedded link in the #Bolsonaro network for each of the top 24 news organizations. Figure 9 and 10 presents the results plotted in a two dimensional space that replicates the network environment. We calculated the predicted probabilities from the estimates on figure 8 and the mean value of the distance of use i to media j in each grid.

For any point in the space, the sum of the probabilities for all twenty-four news organizations adds up to 1. For example, we can see that between a third and a fourth of such probabilities are explained by embeds to other tweets (top-left plot), with small differences in the rates observed in any two points of the grid. The rate of embeds to existing tweets is close to 34% in the periphery of the network, falling to close about 23% at the center of the opposition sub-network. The decay that is explained by ideological reasons, consequently, is relatively small. Of course, Twitter is also more prevalent at the periphery of the #Bolsonaro network, where we find users that are more loosely connected to politics. As we move towards the core of the government and opposition sub-networks, we see a higher number of embeds from news organizations.

More black quadrants on figures 9 and 10 represents news organization's density peak; in other words, the area of the network where the media has higher ideological congruence. Small black grids located in the diagonals of the maps represent outlets with high ideological attention and small reputation. For these cases, their density in the network is concentrated on a small area with a sharp decay as one moves further away. Cases with low variation in the mass of support, and located on the center represent media organizations in which proximity matters less in the users' decisions.

Consider for example the news organization *Revista Forum*, led by the opposition journalist Renato Rovai. At the core of the opposition, *Revista Forum* concentrates close to 8% of all embedded links. However, links to this newspaper rapidly drop to zero as we move to users that

are not closely connected to the opposition. For example, when moving to the quadrants occupies by the cluster of Bolsonaro's supporters, the area becomes purely white indicating a small chance of receiving any attention by users located in this part of the network. The same structure of predicted probabilities appears in the cases of *Diario do Centro do Mundo*, *O Sensacionalista* and *Brasil 247*, all smaller organization more aligned to the opposition and more progressive values in Brasil. In the other side, pay attention to the cases of fake news operatives more popular in the Bolsonaro's community: *Jornal da Cidade On Line*, *Republica de Curitiba*, and *Conexão Política* occupy a quite isolated black area in the extreme of the bottom left quadrants of the network activating the core users of this community solely due to their strong ideological connections.

By contrast, consider now the case of *Globo* and *Folha* [6], the largest TV channel and newspaper in the country, respectively. The most active area in the map for both, where close to 20% of all the embedded links are located, is left to the center of the network. The position is closer to the opposition community, as we would expect, but not exactly in the core of the community, and closer to the median user as our formal model predicts.

Even more important, the rate of decline for both fall at a slower pace, with significant activation rates in the region that extends between the modal reader of the subnetworks and the median reader of the overall network. It goes without say given all the previous findings of this paper, that even for those more reputable outlets, the community of supporters of Bolsonaro is still quite resistant, as one can see by the more red quadrants in the bottom-left part of the figures. However, the graphs indicate that both more reputable news outlets can more easily occupy the center of the network, activate a larger share of users from distinct areas of the network, and optimize its market-share.

[6]The group Folha has some distinct branches as one can see by the different links. Here, I am considering the figure related to link aovivo.folha.uol.com.br that comes directly from the leading newspaper's website

A quick review of all the news organizations described in Figures 9 and 10 shows smaller media organizations in the latter being crowded out to the outer-ladder of the opposition mode. The crowd effect is particularly strong for the fake news operative aligned with Bolsonaro's community. Meanwhile, larger media organizations in Figure 9 tend to be located in the inner-ladder that stretches between the median reader to the core of each subnetwork as well as showing a broader audience of activation. Therefore, our estimation for the hidden location of the media outlets confirms the hypothesis derived from the simulations previously discussed.

Figure 9: Predicted Rate of Embed by Media

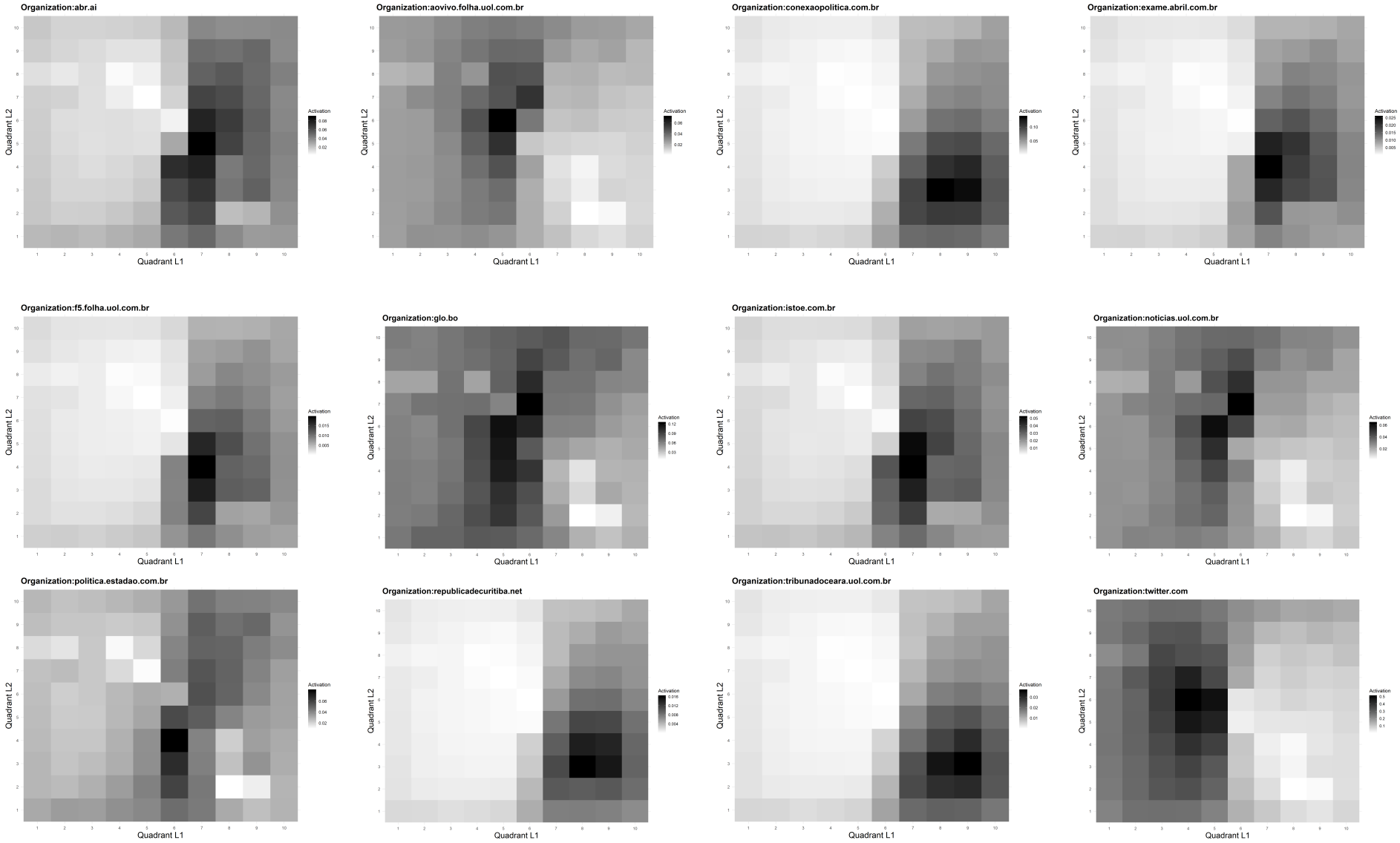
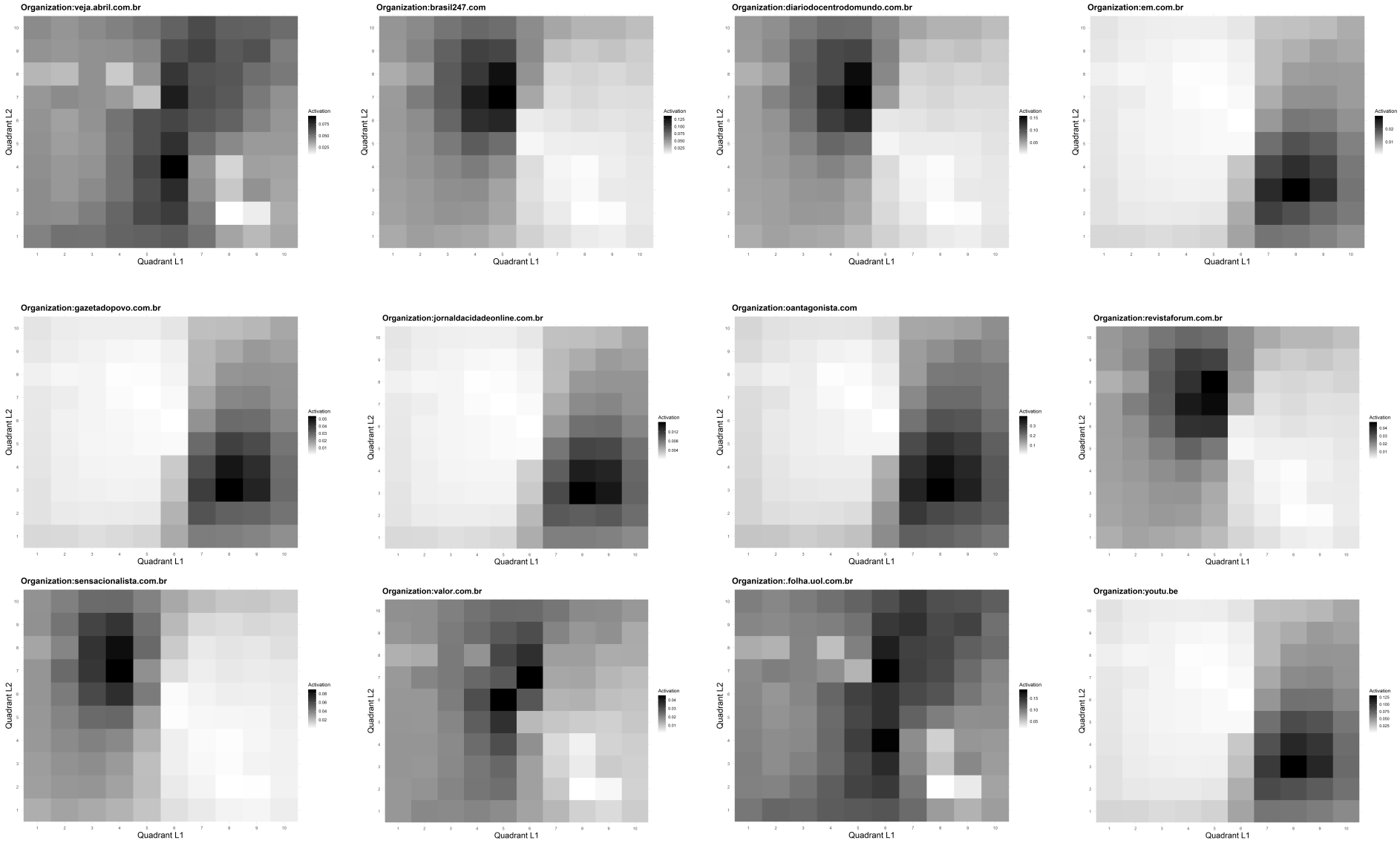


Figure 10: Predicted Rate of Embed by Media



5. Concluding Remarks

How should editors incorporate the preferences of social media users when making editorial decisions for their news organizations? What should be the optimal editorial line if editors are only interested in maximizing online readership? In this paper, we develop a theory that answers both questions. We describe how news organizations should behave when voters are ideologically motivated and have heterogeneous assessments of the quality of different news organizations (reputation). We estimate the determinants of social media embeds in Twitter during a major political event in Brazil, the presidential election of Jair Bolsonaro, highlighting how our model provides new insights and explicit hypotheses to be tested empirically.

Theoretically, our paper extends spatial voting models common in political science to shed light on an exciting communication's problem. The logic of the model relates to ideologically motivated readers and media organizations with distinct journalistic reputations. We show that the users' decision to embed links to these news organizations should affect the extent to which these organizations moderate or radicalize their editorial lines.

In our framework, the reader's decision to share content with friends is explaining by the (1) cognitive congruence/dissonance of readers with the ideological leaning of a post; as well as the benefits of relying on information from high reputation outlets. As we posit that reputations are difficult to alter in the short term, we describe the challenges of an editor as an optimization problem where they can only adjust the ideological leaning (editorial) or the topic (gatekeeping) of the news they publish. Editors that hope to maximize readership, consequently, are pushed towards more centripetal locations if their reputation is already high among readers; or they are pushed towards more extreme positions when their reputation among readers is low.

So far, our model describes media organizations that have no ideological preferences of their

own. Many spatial models in political science take into consideration the policy preferences of politicians. In the communication's literature, we need to consider both the existence of partisan media as well as the economic benefits of higher reputations, which increase the returns that media organizations perceive from vendors and donors. Future extensions of our model will incorporate discount functions for these economic considerations. It will also incorporate the potential benefits of endorsing politicians that could facilitate the expansion of an organization income even at the expense of suboptimal editorial positions.

References

Adams, J., Merrill, S., and Grofman, B. (2005). *A unified theory of party competition : a cross-national analysis integrating spatial and behavioral factors*. Cambridge University Press, Cambridge ; New York.

Aruguete, N. and Calvo, E. (2018). Time to #Protest: Selective Exposure, Cascading Activation, and Framing in Social Media. *Journal of Communication*, 68(3):480–502.

Barberá, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., and González-Bailón, S. (2015). The Critical Periphery in the Growth of Social Protests. *PLOS ONE*, 10(11):e0143611.

Calvo, E., Dunford, E., and Lund, N. (2016). Hashtags that matter: Measuring the propagation of tweets in the dilma crisis.

Calvo, E. and Murillo, M. V. (2019). *Non-Policy Politics: Richer Voters, Poorer Voters, and the Diversification of Electoral Strategies*. Cambridge University Press.

Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, 1695:1695.

Levitsky, S. and Roberts, K. M. (2013). *The resurgence of the Latin American left*. JHU Press.

Merrill, S. and Grofman, B. (1999). *A unified theory of voting : directional and proximity spatial models*. Cambridge University Press, Cambridge, U.K. ; New York.

Poole, K. T. (2005). *Spatial models of parliamentary voting*. Cambridge University Press, Cambridge ; New York.

Schofield, N. and Sened, I. (2006). *Multiparty democracy: elections and legislative politics*. Political economy of institutions and decisions. Cambridge University Press, Cambridge ; New York.

Tardaguila, C., Ortellado, P., and Benevenuto, F. (2018). Fake News Is Poisoning Brazilian Politics. WhatsApp Can Stop It.