The 2018 Swedish Election Campaign on Twitter

Constanza Sanhueza Petrarca, Maria Tyrberg & Steven Lloyd Wilson

Abstract
How did Swedish parties and candidates use Twitter during the 2018 election? And what topics were discussed the most? Using an original dataset containing over 9.1 million tweets collected over the four weeks of the electoral campaign, this paper explores candidates' campaigning strategies online. We show that candidate's individual characteristics and political career have significant effects on the adoption of and use of Twitter, yet party and district-related factors also explain some patterns in online campaigning.

1. Introduction

In September 2018, general elections were held in Sweden to elect 349 members of the national legislature, the Riksdag. The Swedish political elites have quickly recognized the growing popularity of social media usage in the general population and the potential of using it to communicate with voters. All Swedish parties and sixty-five percent of the elected members of parliament had an official Twitter account in 2018. And, to give a sense of the scale of Twitter usage by Swedes, despite a population of only ten million, Swedish is the 12th most used language on Twitter (GNIP 2018). The study of social media usage in political campaigns is still in its nascent stages, and as such exploratory work is a critical step in documenting the state of play in order to set the stage for causal work. This paper investigates two key features of the 2018 electoral campaign online using an original dataset including over 9.1 million tweets collected for the period of the electoral campaign, which lasted four weeks. We make a descriptive contribution to the literature on electoral campaigning by exploring the usage and popularity of Twitter among parties and candidates and shedding light on the content and congruence of the e-campaign between parties, candidates, and voters.

Constanza Sanhueza Petrarca works at the Democracy and Democratization Unit, Berlin Social Science Center. Maria Tyrberg works at the Department of Political Science, University of Gothenburg. Steven Lloyd Wilson works at the Department of Political Science, University of Nevada, Reno.
E-mail: constanza.sanhueza@wzb.eu; maria.tyrberg@gu.se; stevenwilson@unr.edu

2. Theory

In recent years, social media has become increasingly popular among citizens around the globe. While Facebook remains by far the most popular social network platform, the micro blogging platform Twitter now has more than 130 million active users worldwide. In Sweden, Twitter is very popular. Around 25% of the population has a twitter account, compared to 23% in the United States[1]. Campaigners and politicians have quickly recognized the growing popularity of Twitter and its potential for campaigning and to communicate with voters. Based on data collected in 2013, Larsson and Kalsnes reported that 58% of the Swedish members of parliament had a Twitter account. By the end of 2018, we found that the share of newly elected Swedish deputies on the platform had increased to 65%.

Twitter presents many advantages for candidates running for office. Politicians are able to instantaneously reach a wide audience with one simple click. Moreover, unlike other communication channels, candidates' tweets are less likely to be under centralized control by the party giving candidates the possibility to show their own preferences and views. For researchers, Twitter data is particularly interesting for the examination of electoral campaigns. Unlike Facebook, Twitter provides access via programming interfaces for researchers to systematically collect data, allowing the analysis of campaign strategies of parties and candidates alike. Moreover, as Twitter is one of the most popular social network platforms among candidates, but also in the general population, it is possible to assess as well how parties and candidates compare to the general population in terms of their political activity online. Lastly, Twitter requires users to create personal accounts, and when such accounts become of public interest such as these from politicians and parties, the company verifies the identity of the owner. This one-step verification process makes the data analysis more reliable and less prone to manipulation.

The rapid adoption of social media and its use during electoral campaigns has awakened the interest of political scientists. A stream of research has evaluated individual factors affecting candidates' twitter adoption. Age is an important factor in predicting the level of adoption of social media as individuals of younger age have higher computational skills and are more likely to adopt new technologies. In the U.S., Lassen and Brown (2011) assessed factors affecting Twitter adoption among members of Congress and showed that Twitter users tended to be younger. Similar evidence is found in other countries like the Netherlands (Vergeer & Hermans 2013) and in Sweden (Larsson and Moe, 2012). Besides age, another individual characteristic that is likely to affect the use of social media is gender. Empirical research shows that there are differences in terms of the use of social media by gender, as women are more likely

1 https://www.statista.com/statistics/623048/twitter-users-in-sweden/

to use social media than men (Hargittai 2007). At the same time, there is evidence of a gendered usage of social media. For example, men are reported to use Twitter to a greater extent than women to search for political news online (Abraham, Morn & Vollman 2010) while women use social media more socially than men do (Pujazon-Zazik & Park 2010). Data on the users of different social media platforms reveal important gender differences between the platforms as well: while 52% of Facebook users in the US in 2018 were female, women represented only 34% of Twitter users (Statista 2019). While these gender differences are less likely to be reflected on Twitter adoption at the elite level as a result of the professionalization of electoral campaigns, the persistence of patterns of gendered bias against women – such as the increasing evidence of death threats and harassment against female politicians (Krook 2017) – is likely to affect how female candidates use this platform.

Besides these individual characteristics, candidates' political career and party have been also shown to affect candidates' campaigning strategies online. Newcomers, candidates with less experience, and candidates running in competitive districts can use Twitter to increase their visibility. Candidates whom are ranked lower on the party list may benefit from actively using Twitter to gain visibility during the campaign. Moreover, in the analysis of the 2010 Dutch election, candidates from parties that had lost seats in the general elections of 2006 were observed to be more likely to subscribe to Twitter, suggesting that they sought new ways to reach out to voters (Vergeer & Hermans 2013). This ability of candidates to bypass traditional gatekeepers and address the public directly is a critical and distinct element of social media usage by politicians (Jacobs & Spierings 2016).

Furthermore, there is evidence indicating that parties also influenced the Twitter adoption of their candidates. In a study in the U.S., Lassen and Brown (2011) show that Twitter users in Congress were more likely to have an account if they were urged by their party leaders to tweet. Similarly, in a comparative study of Dutch and British elections in 2010, Graham, Jackson and Broersma (2016) showed that parties that encouraged their candidates to use the platform were also more active on Twitter. In many cases these parties offered their candidates advice and training on its use. At the same time, the authors show that populist left and right-wing parties in the Netherlands and the Conservative party in the UK actively restricted and controlled the communication of individual candidates to avoid scandals.

Further analysis examining the linkage between party ideology and candidates' e-campaigning is inconclusive. Liberal parties have been observed to be early adopters in the use of new technologies and socials media during campaigns (Copsey, 2003). Examining over 30 parties in Europe, Sudulich (2010) shows that left-wing parties used more interactive applications on their websites, which may favor their use of twitter among their candidates. Similarly, in a study conducted in the US in the 2004 and 2008 elections, Williams and Gulati (2013) show that Democrats were more likely to use Facebook. At the same time, other studies do not find a correlation between party ideology and the use of social media by candidates. For instance, Vergeer and Hermans (2013), do not find evidence that party ideology effected the use of Twitter in the Netherlands similar to Larsson and Karlsen (2014) in their analysis of politician's use of Facebook in Sweden and Norway.

A general pattern observed is that candidates use the platform to broadcast information and to mobilize electoral support (Lamarre & Suzuki-Lambrecht 2013). Social media data provides scholars with rich text data to examine politicians' policy positions, which has been proven to provide reliable estimates to conduct micro and aggregate level analysis (Ecker 2017). Building on the research presented above, we explore the activity of politicians on Twitter during the 2018 electoral campaign in Sweden, focusing primarily on the decision to adopt Twitter and the content of tweets.

In the next sections we present our data, empirical strategy and results.

3. Data and Methods

This paper examines the 2018 Swedish electoral campaign on Twitter based on an original dataset containing 21 million tweets collected over 83 days (from June 20th to September 10th). In this paper we focus on the electoral campaign period in the 28 days leading up to the election plus election day itself (from August 12th through September 9th), which accounts for a total of 9.1 million tweets.[2]

We used a complex data gathering technique that collects data from four sources: 1) tweets from accounts of parties and candidates, 2) tweets matching a set of political keywords, 3) geo-coded tweets, and 4) tweets identified as being in the Swedish language and matching a set of 100 Swedish language stop-words. All tweets were downloaded using custom software that accessed the Twitter streaming API in order to download and process tweets in real time. This data included the full text of each tweets along with meta data about the user posting the tweet such as number of followers. The streaming API is a programming interface in which all tweets matching specified search criteria posted to Twitter worldwide can be downloaded in real time by appropriate software.[3] The API provides hooks that allow programmers to pull down subsets of that stream based on queries of two types: geocoding and keywords. We utilize both in this project.

2 Note that the time series graphs also include the day after the election (September 10th) in order to make the expected spikes of activity on election day itself visually distinct.

3 Search results are capped at 1.5% of the overall Twitter stream (the so-called "firehose") at any given time, however none of the queries performed in our project come close to that threshold.

The first source are those tweets that are posted in the official and verified accounts of parties and candidates. We collected all twitter handles for all members of the Riksdag along with those running for office (EveryPolitician 2018), and then scraped from the API all tweets posted from those accounts during the duration of the time period of the study. This amounted to 76,397 tweets.

The second source of tweets have been gathered as they match at least one of the keywords that we have identified were related to the election. Keywords include political slogans, hashtags, and relevant dates such as debates and elections (a detailed list of the keywords used is in the Appendix). These keywords were carefully selected such as to be specific enough to the Swedish elections as to not produce false positives (particularly about entirely unrelated popular issues). This set of data represents some 2.9 million tweets.

The third source include all those tweets that were posted from within Sweden itself. Approximately 1.5% of all tweets are geocoded, which means that a set of latitude and longitude coordinates generally accurate to within two meters is attached to the tweet at the time of its posting. This is generally based on the GPS functionality of a smart phone or similar device used in the posting. By downloading *all* tweets regardless of topic from within Sweden during this period, we gain three things. First, this serves as a proxy of a denominator for the keyword matches. Second, it allows us to investigate whether major political speech is occurring without our identified keywords. That is, are our keywords subject to selection bias of some sort by the researchers. Finally, the geocoded tweets allow us to evaluate whether the keywords selected were well bounded and returning primarily content from Sweden. All keyword matches geocoded from within Sweden should be in both the keyword set of tweets and the geocoded set of tweets. Of the geocoded keyword matched tweets, 85% originated from within Sweden, giving us confidence that our keyword selection was well bounded.

Finally, using a standard "stop-word" list of the Swedish language (i.e. common semantically meaningless words such as prepositions and particles) we downloaded all Swedish language tweets matching such keywords. Nearly all tweets written in Swedish should contain at least one stop word, which means that this set of tweets gives us a rough proxy for the total number of tweets originating from all Swedes. This accounted for 17.1 million total tweets. In addition, this provides an additional robustness check in that 94% of the geocoded tweets in the stop-word tweets did in fact originate from within Sweden, giving us additional confidence in our approach as capturing Swedish communication without foreign false positives.

On average, Swedes posted 249,000 tweets per day, of which 13.8% matched one of our identified political keywords, 26.9% of which included links and 58.6% of which consisted of retweets, that is reposts of original posts (which often contain additional commentary). Figure 1 features the basic characteristics of the Twitter activity during that period. The volume of tweets increased in the week prior to the election, while keyword matches spiked on days particularly salient to the election. For instance, note the spike on election day itself, along with August 14th, the day when car fires occurred at several places across cities in western Sweden. Figure 2 renders a similar pattern, showing only the totals of tweets posted by official politician or party accounts by day.

Figure 1. Number of Tweets per Day - Swedish Language Matches (Solid), Political Keyword Matches (Dashed)

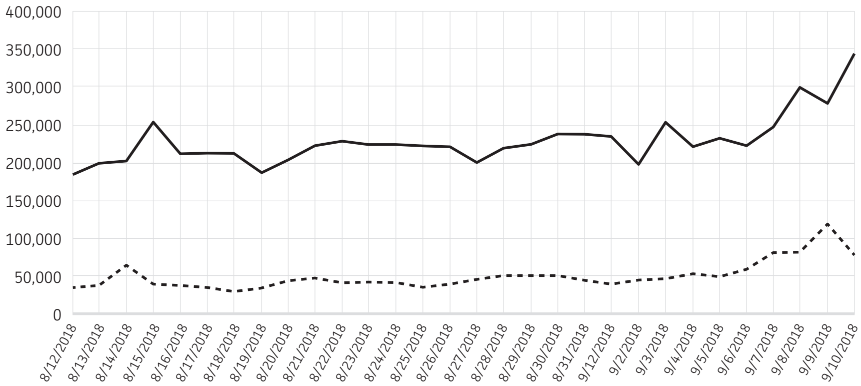
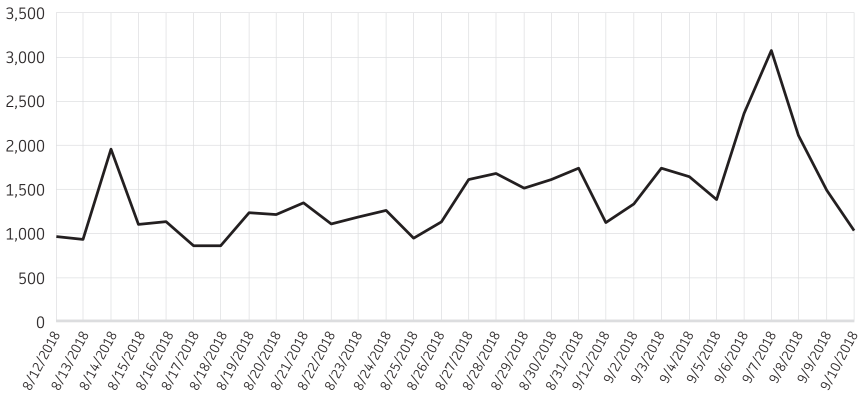


Figure 2. Number of Tweets per Day by Politician (or Party) Accounts



Candidates' characteristics such as age and gender, and their political careers has been gathered from the Riksdag's open data base. We use the Chapel Hill Expert Survey Data to measure party ideology (Polk et al. 2017; Bakker et al. 2015).

The empirical analysis relies on the quantitative examination of the factors that influence a candidate's usage of social media during the campaign and the analysis of the topics discussed online. The first part of the analysis presents the results of logit and negative binomial models (Table A2 in the Appendix presents the summary statistics). The second part uses computerized content analysis in three ways: an unsupervised model to identify the primary dimension of political discussion, an affinity model in which public tweets are statistically matched with political parties, and a dictionary-based approach comparing the political topics focused on by the public and candidates for office.

4. Candidate Campaign Online

How did candidates use Twitter in the 2018 election? And, what factors influence the online campaign strategies followed by candidates? Across countries, Twitter has become a central tool in election and politicians have massively adopted Twitter across Western Democracies. In Sweden, 229 (65%) of the 349 elected parliamentarians elected in 2018 had official Twitter accounts. In order to understand what drives adoption of Twitter among elected candidates the first model presented in Table 1, presents the results of a logit model where the dependent variable indicates if the candidate had an official Twitter account or not. The regression results show that age has a significant effect on twitter adoption, as younger candidates are more likely to have a Twitter account ($p < 0.001$). A unit increase in candidates' age decreases the odds of twitter admission by 0.04. At the same time, we observe that there are not significant differences between females and males. Yet in terms of the political careers, placement in the party list has a negative and significant coefficient ($p < 0.0001$) indicating that for one unit decrease in party list placement measured as a candidates' rank in the list from first to the successive positions, the log odds of twitter adoption (versus not having a twitter account) decreases by 0.04. However, we do not find a significant effect for incumbency, party leadership, opposition, ideology or representing urbanized areas such as Stockholm, Göteborg, and Malmö on twitter adoption.

We find evidence that politicians' use of Twitter varies. Over the 28 days of the campaign, the 228 candidates with Twitter accounts produced over 23,900 tweets and an average of 108.8 tweets per candidate (considering only those that had an account). Jan Ericson, candidate from the Moderate, was the candidate with the largest number of tweets with over 2,300 posts recorded in this period, followed by Tobias Billström (N=1,178), also from the Moderate party and Annika Strandhäll (N=1,173) from the Social Democrats. Among the top five candidates most active in this platform there are two candidates from the Moderate party, one from the Social Democrats and two from the Left party, including the party leader from the Left Party, Jonas Sjöstedt.

Model 2 (Use) explores whether individual characteristics, career, party affiliation and district characteristics influence online campaigning. The model presents the output of a negative binomial regression where the dependent variable is total number of tweets by candidate. The results of the model indicate that the number of tweets was not significantly affected by a candidates' age or gender, incumbency, ideology or type of district. However, challenger candidates, that is candidates that were not in parliament during the legislative period prior to the election, were significantly more likely to tweet than incumbent candidates. The difference in the logs of expected tweets is expected to be 1.01 lower for incumbent candidates, while holding the other variables constant in the model. We also find that party leadership is close to statistical significance, indicating that party leaders were more likely to tweet than other party members.

Our data allows us to count the number of times the official Twitter username of a candidate (Twitter handle) was mentioned during the campaign. The candidates with Twitter account were mentioned over 350,000 times during the campaign that is an average of mentions of 1,549 times by candidate with a twitter account. Out of the five candidates with the highest number of mentions, we find four party leaders including the leaders of the Center Party, the Liberal party, the Left party and the Green party. Interestingly, Hanif Bali of the Moderate Party ranks second in terms of twitter popularity, which may have been influenced by his use of harsh rhetoric when commenting on current political events.

The third model (Mentions) presented in Table 1 explores the relationship between candidates' characteristics, their career, party affiliation and districts on the one side, and the number of mentions they had on Twitter on the other. In this model we control for the total number of tweets posted by a candidate as their activity online is likely to influence their visibility online. The model presents the results of a negative binomial regression, where the dependent variable is the total number of mentions. The coefficient for candidates' age is close to reaching statistically significance while is not significant. Furthermore, most factors associated with a candidate's political trajectory, party affiliation and district do not have an effect on their visibility online. However, we do find that party leader status increases significantly the chances of being mentioned on Twitter as the difference in logs of expected mentions is expected to be 3.8 higher for party leaders compared to other candidates while holding other variables constant in the model. Last, we observe significant and positive effect for the number of tweets that the candidate posted ($p < 0.001$ level) which shows that candidates can also influence their visibility online by being very active on Twitter during the campaign.

Table 1. Candidates' Adoption, Use and Mentions During the Campaign

	Model 1 *Adoption*	Model 2 *Use*	Model 3 *Mentions*
Age	-0.0438*** (-3.67)	0.00414 (0.23)	-0.0332+ (-1.78)
Female	0.275 (1.11)	-0.367 (-0.81)	-0.498 (-1.12)
Incumbent MP	0.0525 (0.19)	-1.010* (-2.15)	-0.142 (-0.34)
Position in Party List	-0.0490*** (-4.04)	0.00323 (0.09)	-0.0119 (-0.39)
Party Leader	0.625 (0.74)	1.947+ (1.71)	3.826*** (3.64)
Candidate Opposition Party	0.444 (1.29)	0.953 (1.54)	0.0244 (0.04)
Party Position Left-Right	0.0646 (0.80)	-0.188 (-1.36)	0.115 (0.89)
Urban District	-0.0343 (-0.11)	0.887 (1.64)	0.275 (0.54)
Total Number of Tweets			0.00792*** (5.09)
Constant	2.086** (3.06)	5.225*** (4.87)	4.877*** (4.08)
N	349	228	228

t statistics in parentheses
Sources: Own Data 2018, Chapel Hill Expert Survey (2017)
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Summarizing, in this section we show that a candidates' age – which reflects a generational digital divide – as well as their rank in the party lists are associated with the adoption of Twitter. Yet, the platform was particularly used by challenger candidates and party leaders to spread their campaign messages across. Last, party leaders received more mentions online, yet as our evidence shows, candidates could also increase their visibility online by increasing the number of messages they posted online.

In other words, we find that candidates' age, party leadership, a candidates' position in the list and the challenger status explains part of a candidate's decision to campaign online. In terms of the effectiveness of these strategies, we observe that party leaders received the most attention on Twitter, yet candidates could increase their visibility online by being active on the platform[4].

4 We also run models including party dummies (Table included in the Appendix). The results do not change substantially when we control for party membership.

Now that we have established the factors that influence the likelihood of campaigning online, in the next section we examine the factors that influence the content of candidates' Tweets.

5. Candidate Social Media Messaging

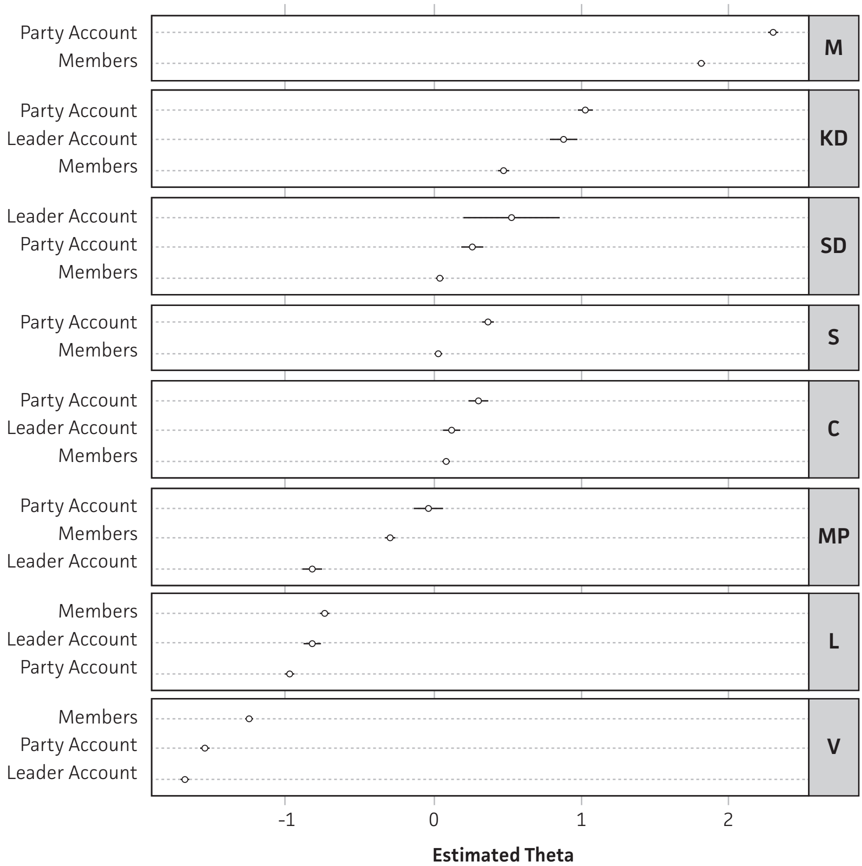
In this section, we examine how parties and elected candidates used social media by leveraging a variety of content analysis techniques to quantify their topical focus and strategies.

We first present the results of a naïve text analysis model that scales all tweets posted in the official party, leaders, and elected members accounts along a single ideological dimension. Wordfish is a unidimensional algorithm that assumes that the underlying texts have a single primary dimension along which they vary and estimates each text's position based on word frequencies. This algorithm has been used to great effect for identifying left/right positions of party speech in a variety of legislative settings (Slapin and Proksh 2008). The produced numbers are unitless and arbitrarily scaled (i.e. they are meaningful in comparison to each other, not in absolute terms). As such, we defined the direction of the scale such that the Christian Democrats were to the "right" of the Social Democrats to have an intuitive orientation. Figure 3 shows the results of this process, grouped by party and ordered from "right" (top) to "left" (bottom), with three different measures per party: the party leader (such as @ BuschEbba, denoted by "Leader Account"), the official party Twitter account (such as @kdriks, denoted by "Party Account"), and the aggregate of the Twitter accounts of all members of the party (indicated by "Members").[5]

The dimension captured is clearly not precisely left to right in a traditional sense, since the Moderates are the furthest along one end of the dimension. However, other measures are intuitive. For instance, the Christian Democrats are essentially on the opposite end of the spectrum from the left wing Left party and the Greens. In addition, the Center Party is fittingly the exact center of the gradient. Yet, some interesting patterns emerge. First, in the case of five of the six parties with a party leader tweeting, the party leader's score is closest to the official party account, reflecting the leader driving the official party dialog. The exception to this is Isabella Lövin, who is significantly different from both the average Green Party member and the official feed itself. This may be reflective of the Green Party's usage of dual spokespeople (unfortunately the other, Gustav Fridolin, does not actively tweet).

5 The Moderates and Social Democrats do not have separate leader accounts in use, and thus only have the other two data points. In addition, we left off the Feminist party as they did not win any seats in the election.

Figure 3. Wordfish Positioning of Party, Leader, and Mean Party Member



The evidence presented above provided some empirical evidence across and within parties during the electoral campaign. Subsequently, we further examine these differences by applying an affinity model to our data. Affinity models are a supervised text classification model that takes as an input a set of texts that have been identified as belonging to different researcher-defined categories. The relative frequencies of words in the texts are used to create a statistical model of each category. These models can then be applied to other texts, classifying them into the modelled categories (Perry & Benoit 2017). We used the tweets from the accounts of parties and party leaders as training data for an affinity model such that each party was its own category, with its own training set [6].

We then applied the affinity model to the 2.8 million election-related tweets that we collected during the campaign, in order to classify to which party's

6 In order to test the accuracy of the affinity model, we used 50% of the over 7,000 tweets posted in the accounts of the parties and party leaders during the campaign and then applied to classify the remaining 50% of the data. When applied to the data, the model showed a 90% accuracy.

speech each tweet was closest. Table 2 shows the results of the affinity model. Overall, we see that there is great variation in the affinity between the tweets of the population and the different parties. We observe that there is greater affinity online with the Green party, the Sweden democrats and the Left party, that is with the parties that have the most extreme positions in the left-right spectrum. At the same time, we find evidence of less affinity with the Christian Democrats and the Moderate Party, while the Center, Feminist, Liberal and Social Democrats have similar scores of around 8 percent.

Furthermore, our evidence reveals important disparities between the vote shares of the parties and the affinities with voters. We find that compared to the percentages of the votes received, the Feminist, Green and Left parties score higher on affinity. This means that despite the actual support in terms of votes (together the three parties obtained around 12% of the votes), citizens reflected online concerns that were also raised by these parties, such as gender equality, the environment and income inequality. For instance, what we observe in relation to the relevance of the environment during the campaign may be explained by the fact that the Swedish election took place at the end of the summer after the country had experienced what is considered the hottest July in over 260 years, sparling heated debates over climate change and the environment. At the same time, it is also worth noting that there are lower affinity levels recorded for the two largest parties, the Social Democrats and the Moderate, which are parties that have broader political agendas. The affinity analysis provides some evidence of strategic voting behavior by citizens, in which there is strong sympathy for the restricted breadth of issues discussed by single-issue parties like the Greens and Feminists, while voters nonetheless cast votes for broader based parties.

Table 2. Affinity and Tweet Reach by Party

Party	Vote Share	Affinity	No of Tweets	Tweet Reach
Centerpartiet (Center)	8.61%	8.64%	38,308	2638
Feministerna (Feminist)	0.46%	7.29%	32,350	6374
Kdriks (Christian Democrats)	6.32%	3.69%	16.366	2641
Liberalerna (Liberal)	5.49%	7.73%	34,292	2496
Miljopartiet (Greens)	4.41%	27.60%	122,246	2932
nya moderaterna (Moderate)	19.84%	5.74%	25,453	2459
Sdriks (Sweden Democrats)	17.53%	16.87%	74,849	1546
Socialdemokrat (Social Democrats)	28.26%	8.26%	36,655	2126
Vansterpartiet (Left)	8.00%	14.18%	62,889	1856

In order to dig into what the patterns of discussion looked like, we developed keyword-based policy areas that allow to examine the political issues discussed during the Swedish electoral campaign. The comprehensive list of keywords by policy area are reported in the Appendix, but the categories examined are: defense, the environment, the elderly, gender, occupation, education, healthcare, law, immigration and the economy.

Figure 4 shows the breakdown of how the population's Twitter activity compared to that of elected candidates during the weeks of the electoral campaign. These figures were arrived at by searching the entire Swedish language database of tweets in order to draw from an unbiased sample of Swedish social media activity. Elected candidates talk about these political topics in general at twice the rate of the population's general discussion, which is to be expected. Therefore, for comparative purposes we scale this discussion by the total number of political topic matches for each group of Twitter users. This allows us to evaluate whether the two groups discuss political topics at equivalent levels, when discussing political issues in the first place.

Looking at the data we observe significant differences in focus between the two groups. Elected candidates are far more likely to speak about environmental issues, healthcare and the state of economy while the population is more likely to speak about employment, law and order, immigration issues and national defense. We do observe, however, a greater level of congruence on topics that are not the most salient ones, like the elderly, gender, and education.

Figure 4. Topic Discussion by Elected Candidate vs. Population

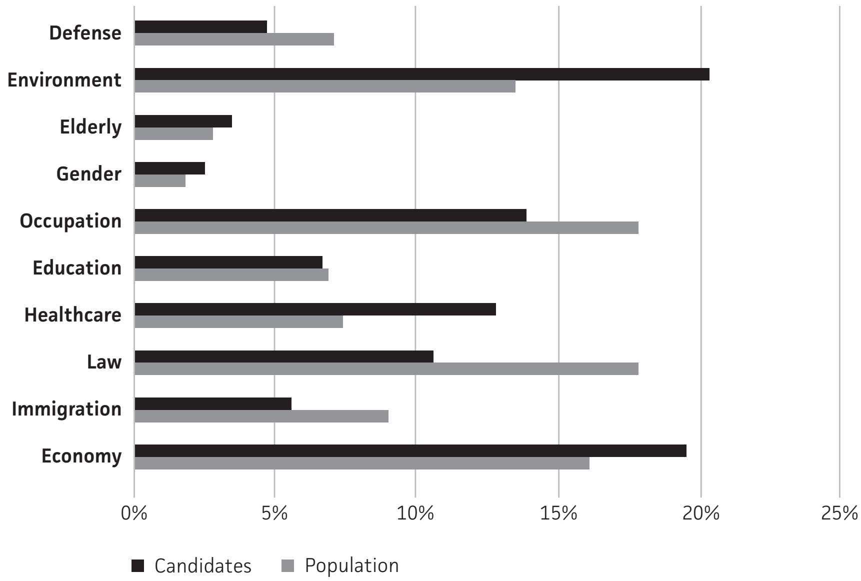


Table 3. Relationship Between Candidate Attributes and Frequency of Tweeting About Topics

	Dependent variable:									
	Defense (1)	Environment (2)	Elderly (3)	Gender (4)	Occupation (5)	Education (6)	Healthcare (7)	Law (8)	Immigration (9)	Economy (10)
Age	0.035***	-0.012	0.023**	0.011	0.009	-0.008	0.015*	-0.008	0.006	0.011*
	(0.010)	(0.010)	(0.010)	(0.011)	(0.006)	(0.010)	(0.009)	(0.006)	(0.007)	(0.006)
Female	0.062	0.831***	0.299	0.703***	0.073	0.348*	0.445**	0.120	0.131	0.241*
	(0.217)	(0.221)	(0.202)	(0.257)	(0.119)	(0.209)	(0.186)	(0.124)	(0.150)	(0.131)
Incumbent MP	0.239	0.223	0.515**	-0.306	0.056	0.424*	-0.132	0.190	0.629***	0.293*
	(0.241)	(0.231)	(0.251)	(0.263)	(0.132)	(0.228)	(0.205)	(0.142)	(0.177)	(0.150)
Position in Party List	-0.025	-0.121***	-0.008	-0.026	-0.041***	-0.026	-0.036*	-0.014	-0.002	-0.025*
	(0.022)	(0.041)	(0.028)	(0.028)	(0.015)	(0.025)	(0.019)	(0.012)	(0.014)	(0.014)
Party Leader	-0.370	0.073	0.216	0.681*	-0.239	0.526	0.179	-0.016	-0.248	-0.350
	(0.466)	(0.501)	(0.397)	(0.405)	(0.248)	(0.419)	(0.434)	(0.267)	(0.302)	(0.306)
Candidate Opposition Party	0.148	0.153	-0.881***	0.139	-0.261	-0.399	-0.019	-0.190	0.408*	0.072
	(0.325)	(0.330)	(0.230)	(0.347)	(0.165)	(0.277)	(0.264)	(0.174)	(0.228)	(0.193)
Party Position Left-Right	0.145**	-0.023	-0.018	0.095	0.051	0.015	0.059	0.114***	0.094**	0.037
	(0.065)	(0.065)	(0.047)	(0.068)	(0.034)	(0.058)	(0.057)	(0.035)	(0.042)	(0.040)
Urban District	0.151	0.511**	0.049	-0.044	-0.230*	0.089	-0.146	0.081	0.560***	-0.058
	(0.246)	(0.245)	(0.219)	(0.271)	(0.136)	(0.232)	(0.223)	(0.137)	(0.164)	(0.155)
Constant	-6.996***	-3.096***	-5.652***	-6.411***	-3.888***	-4.321***	-4.695***	-3.882***	-6.108***	-4.253***
	(0.580)	(0.564)	(0.569)	(0.750)	(0.328)	(0.522)	(0.536)	(0.326)	(0.414)	(0.374)
Observations	349	349	349	349	349	349	349	349	349	349
Log Likelihood	-303.413	-440.968	-214.452	-187.102	-381.746	-285.003	-381.725	-376.242	-278.268	-422.905

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

Sources: Own Data (2018), Chapel Hill Expert Survey (2017)

Table 4. Summary Table of Relationship Between Candidate Attributes and Topical Speech on Twitter

Variables	Defense	Environment	Elderly	Gender	Occupation	Education	Healthcare	Law	Immigration	Economy
Age	+		+				+			+
Female		+		+		+	+			+
Incumbent MP			+			+			+	+
Position in Party List		-			-		-			-
Party Leader				+						
Candidate Opposition Party			-		-	-			+	
Party Position Left-Right	+							+	+	
Urban District		+			-				+	

Finally, we examined how candidate attributes were associated with which topics they tweeted about. Table 3 shows the results of ten regression models, one for each of the primary topics of political discussion in Sweden, using as covariates candidate attributes.[7] The dependent variables are the number of tweets posted in that category by a candidate. As this is count data, we use a negative binomial model, with the total number of tweets by each candidate as an exposure variable, and a zero-inflation step to account for the multidimensionality of excess zero values.[8]

Table 4 summarizes the results strictly in terms of the positive/negative statistically significant relationships between each topic's discussion and attributes of the candidates.

Exploring whether individual and the political trajectories of the candidates affect the topics that the candidates discussed online, we do not find clear patterns.

In terms of the individual characteristics we find that older candidates were more likely to show concern about national defense, elderly issues, healthcare and the economy, while female candidates expressed themselves on issues related to the environment, gender issues, education, healthcare and the state of the economy.

Incumbent candidates discussed the elderly, education, immigration, and the economy more than challengers. However, candidates that were ranked higher in the party lists – who were therefore more likely to be elected – discussed the environment, occupation, health care, and the economy to a higher extent than candidates in the lower ranks in the party lists. Party leaders were more likely to discuss gender issues than other candidates.

Party characteristics also affect the topics of the campaigns online. Being a candidate of opposition and the ideological stances of the parties affected the issues discussed online by individual candidates. As the models show, candidates from opposition parties were less likely to address elderly issues, occupation and education and more likely to discuss immigration than candidates of the governing parties. At the same time, candidates of parties in the right of the political spectrum were more likely to discuss about national defense, law and order and immigration than center-left and left-wing candidates. Lastly, our evidence shows that candidates from urban districts were more likely to

7 It is true that individual candidates belong to parties, and generally in regression analysis involving candidates' behavior in campaigns a multilevel approach is used. In our case however, we believe that a hierarchical model is theoretically invalid: we believe that the social media usage of individual candidates is independent from each other, even if correlated with each other along party lines as candidates are able to circumvent gatekeepers and directly communicate with citizens (see for example Jacobs and Spierings 2016). A hierarchical model is appropriate if the data points are not independent.

8 As with classic zero inflation examples, zeroes in the dependent variable here can be due either to the candidate not mentioning these terms in their tweets, or because they do not tweet at all.

talk about immigration and the environment and showed less concern about occupation than candidates from rural areas.

When taking into account which topics were more relevant for voters online (based on our data, the topics that were most mentioned online by citizens rank-order as follows: 1) occupation, 2) law and order, 3) economy, and 4) immigration) we observe that candidates' individual characteristics, party affiliation and district are more likely to explain the extent to which candidates aimed to campaign on these issues online.

To conclude, in this section we have presented the empirical examination of the topics on which the elected candidates campaigned. We have showed evidence on the position of candidates between and within parties, and in relation to the citizens. Furthermore, we have empirically tested what candidates' factors (individual, partisan, district) affect the choice of topics of their online campaigns. We find that candidates' attributes, party and the districts they represent do influence their choice to discuss specific topics.

6. Conclusion

The study aimed to assess candidates' online strategies during the 2018 Swedish electoral campaign. On the basis of 9.1 million tweets collected over the four weeks of the campaign for all 349 newly elected candidates to the Riksdag. We found that the majority of the elected candidates uses Twitter during the campaign. We observed that candidates' age and political career were related with the adoption and use of Twitter as an additional campaigning tool.

Moreover, using several text analysis techniques, we also explored potential factors that influence the topics mentioned by candidates in their tweets. Our evidence show that individual factors plays an important role on the topics addressed online, as much as partisan and district-related factors.

More generally, the findings of this study show that the ability of Swedish candidates to employ vote-maximize campaign strategies using Twitter may be constrained by candidates' own digital literacy (younger candidates use this tool to a greater extent). Furthermore, we show that individual and career attributes, and the characteristics of a candidates' party and districts affects the choices of topics candidates address in their campaigns online.

In contrast to previous studies conducted in the context of other national elections, we do not find that parties have a strong influence on whether and how Swedish candidates use social media during electoral campaigns as including party dummies yielded similar results than the ones presented in the main analysis.

Future research should focus on the effects of electoral competition and constituency socio-demographics on candidates' online strategies during campaigns.

Acknowledgements

We thank the editors of this issue, the anonymous reviewers, and the participants of the "Political Communication during the 2018 Swedish Election Campaign" Workshop organized by the Swedish National Election Studies for their insightful comments and suggestions on our paper. This research was conducted thanks to the support of the the UNR Wolfcloud computing infrastructure.

References

Abraham, Linda Boland, Mörn, Marie Pauline & Vollman, Andrea, 2010. "Women on the web: How women are shaping the internet". ComScore, Incorporated.

Bakker, Ryan, Catherine De Vries, Erica Edwards, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen & Milada Anna Vachudova, 2015. "Measuring party positions in Europe: The Chapel Hill expert survey trend file, 1999–2010", *Party Politics* 21(1), p 143–152.

Bessi, Alessandro & Ferrara, Emilio, 2016. "Social Bots Distort the 2016 US Presidential Election Online Discussion", *First Monday* 21(11). Available at SSRN: https://ssrn.com/abstract=2982233.

CNN, 2018. "Sweden Twitter Bots Election". Available at https://edition.cnn.com/2018/08/30/europe/sweden-twitter-bots-elections-intl/index.html, cited August 2018.

Copsey, Nigel, 2003. "Extremism on the net: the extreme right and the value of the Internet", p 232–247 in Gibson, Rachel K., Nixon, Paul & Ward, Stephen J. (eds), *Political Parties and the Internet*. London: Routledge.

Ecker, Alejandro, 2017. "Estimating policy positions using social network data: Cross-validating position estimates of political parties and individual legislators in the polish parliament", *Social Science Computer Review* 35(1), p 53-67.

Express, 2018. "Is the Swedish Election Being Hacked?", available at https://www.express.co.uk/news/world/1011297/Sweden-election-2018-is-swedish-election-being-hacked-surge-in-twitter-bots-sweden-news, cited August 2018.

EveryPolitician.org. Database of all politicians in the world. Available at http://everypolitician.org.

Ferrara, Emilio, 2017. "Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election", *First Monday* 22(8). Available at https://arxiv.org/abs/1707.00086.

GNIP, 2018. Languages on Twitter. Available at https://www.mapbox.com/labs/twitter-gnip/languages/.

Graham, Todd, Jackson, Dan & Broersma, Marcel, 2016. "New platform, old habits? Candidates' use of Twitter during the 2010 British and Dutch general election campaigns", *New media & society* 18(5), p 765–783.

Hargittai, Eszter, 2007. "Whose space? Differences among users and non-users of social network sites", *Journal of computer-mediated communication* 13(1), p 276–297.

Jacobs, Kristof & Spierings, Niels, 2016. *Social media, parties, and political inequalities*. Cham: Palgrave Macmillan.

Krook, Mona Lena, 2017. "Violence against Women in Politics", *Journal of Democracy,* 28(1), p 74–88.

Kruikemeier, Sanne, 2014. "How political candidates use Twitter and the impact on votes", *Computers in Human Behavior* 34, p 131–139.

LaMarre, Heather L., & Suzuki-Lambrecht, Yoshikazu, 2013. "Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns", *Public Relations Review* 39(4), p 360–368.

Larsson, Anders O. & Kalsnes, Bente, 2014. "'Of course we are on Facebook': Use and non-use of social media among Swedish and Norwegian politicians", *European Journal of Communication* 29(6), p 653–667.

Larsson, Anders Olof & Moe, Hallvard, 2012. "Studying political microblogging: Twitter users in the 2010 Swedish election campaign", *New Media & Society* 14(5), p 729–747.

Lassen, David S. & Brown, Adam R., 2011. "Twitter: The electoral connection?", *Social science computer review* 29(4), p 419–436.

Murthy, Dhiraj, Powell, Alison B., Tinati, Ramine, Anstead, Nick, Carr, Leslie, Halford, Susan J. & Weal, Mark, 2016. "Automation, algorithms, and politics| Bots and political influence: A sociotechnical investigation of social network capital", *International Journal of Communication* 10, p 4952–4971.

Perry, Patrick O. & Benoit, Kenneth, 2017. Scaling text with the class affinity model. *arXiv preprint arXiv:1710.08963*.

Polk, Jonathan, Rovny, Jan, Bakker, Ryan, Edwards, Erica, Hooghe, Liesbet, Jolly, Seth, Koedam, Jelle, Kostelka, Filip, Marks, Gary, Schumacher, Gijs, Steenbergen, Marco, Vachudova, Milada & Zilovic, Marko, 2017. "Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data", *Research & Politics* (January-March), p 1–9.

Pujazon-Zazik, Melissa & Park, Jane M., 2010. "To tweet, or not to tweet: gender differences and potential positive and negative health outcomes of adolescents' social internet use", *American Journal of Men's Health* 4(1), p 77–85.

Slapin, Jonathan & Proksch, Sven-Oliver, 2008. "A Scaling Model for Estimating Time-Series Party Positions from Texts", *American Journal of Political Science* 52(3), p 705–772.

Shao, Chengcheng, Ciampaglia, Giovanni Luca, Varol, Onur, Flammini, Alessandro & Menczer, Filippo, 2017. "The spread of fake news by social bots." arXiv preprint arXiv:1707.07592, p 96–104.

Statista, 2019. Users of Social Media. Available at https://www.statista.com/statistics/828092/distribution-of-users-on-twitter-worldwide-gender/.

Sudulich, Maria Laura, 2010. "Cyber-space Oddity? An Analysis of Political Parties' Websites and Online Campaigning" (Doctoral dissertation, Trinity College Dublin).

Sveriges Radio, 2018. "Dramatic Increase of Social Media Bots". Available at https://sverigesradio.se/sida/artikel.aspx?programid=2054&artikel=7031695, cited August 2018.

Vergeer, Maurice, Hermans, Liesbeth & Sams, Steven, 2013. "Online social networks and micro-blogging in political campaigning: The exploration of a new campaign tool and a new campaign style", *Party Politics* 19(3), p 477-501.

Vergeer, Maurice & Hermans. Liesbeth, 2013. "Campaigning on Twitter: Microblogging and online social networking as campaign tools in the 2010 general elections in the Netherlands", *Journal of Computer-Mediated Communication*, 18(4), p 399–419.

Williams, Christine B. & Gulati, Girish J., 2013. "Social networks in political campaigns: Facebook and the congressional elections of 2006 and 2008", *New Media & Society* 15(1), p 52–71.

Wilson, Steve and Metzger, Megan, 2019 (forthcoming). "RT and the American Midterms on Social Media". Washington Post, Monkey Cage Blog.

Zamora-Medina, Rocío & Zurutuza-Muñoz, Cristina, 2014. "Campaigning on Twitter: Towards the "personal style" campaign to activate the political engagement during the 2011 Spanish general elections", *Communication & Society* 27(1), p 83–106.

Appendix

Table A1. Keywords Used to Gather the Keyword-Based Data

Category	Keyword
Party leader	stefanlofven
Party leader	stefan lofven
Party leader	stefanloefven
Party leader	stefan loefven
Party leader	stefanlöfven
Party leader	stefan löfven
Party leader	annielööf
Party leader	annie lööf
Party leader	annieloof
Party leader	annie loof
Party leader	jimmieåkesson
Party leader	jimmie åkesson
Party leader	jimmieakesson
Party leader	jimmie akesson
Party leader	jonassjöstedt
Party leader	jonas sjöstedt
Party leader	jonassjostedt
Party leader	jonas sjostedt
Party leader	ulfkristersson
Party leader	ulf kristersson
Party leader	janbjörklund
Party leader	jan björklund
Party leader	janbjorklund
Party leader	jan bjorklund
Party leader	ebbabuschthor
Party leader	ebba busch thor
Party leader	gustavfridolin
Party leader	gustav fridolin
Party leader	isabellalövin
Party leader	isabella lövin
Party leader	isabellalovin
Party leader	isabella lovin
Party leader	gudrun schyman
Party leader	gudrunschyman
Party leader	gitanabavi
Party leader	gita nabavi
Debate	13 juni
Debate	13 juni 2018
Debate	29 augusti
Debate	29 augusti 2018
Debate	7 september
Debate	7 september 2018
Debate	partiledardebatt
Debate	partiledardebatten
Election	riksdagsval
Election	riksdagsval2018
Election	val2018
Election	valet2018
Election	svpol
Election_date	9september
Election_date	9september2018
Parties	socialdemokraterna
Parties	socialdemokrat
Parties	centerpartiet
Parties	kristdemokraterna
Parties	kdriks

Category	Keyword
Parties	sverigedemokraterna
Parties	sdriks
Parties	nyamoderaterna
Parties	nya moderaterna
Parties	moderaterna
Parties	liberalerna
Parties	vänsterpartiet
Parties	vansterpartiet
Parties	miljopartiet
Parties	miljöpartiet
Parties	feministerna
Parties	feministisktinitiativ
Parties	alliansen
Parties	rödgröna
Slogan	framåt
Slogan	framåt!
Slogan	likaföralla
Slogan	lika för alla
Slogan	klimatetkanintevänta
Slogan	Nu
Slogan	klimatet kan inte vänta
Slogan	ettsverigeföralla
Slogan	ett sverige för alla, inte bara de rikaste
Slogan	frihetmåsteförsvaras
Slogan	frihet måste försvaras
Slogan	försvara friheten
Slogan	förändringpåriktigt
Slogan	förändring på riktigt
Slogan	trygghet och tradition
Slogan	trygghetochtradition
Slogan	stoppavinstjakten
Slogan	Ett starkare samhälle. Ett tryggare Sverige
Slogan	För ett samhälle där alla tar ansvar
Slogan	välfärdslöftet
Slogan	du ska kunna lita på sverige
Slogan	klartvikan
Slogan	klart vi kan

Category	Keyword
Newspaper	@dagensnyheter
Newspaper	@SvD
Newspaper	@Aftonbladet
Newspaper	@Expressen
Newspaper	@metrosverige
Newspaper	@sverigesradio

Table A2. Descriptive Statistics

Variable	Obs	Mean	Std. dev.	Min.	Max.
Adoption	349	0.63	0.47	0	1
Number of Tweets	349	71.4	235.7	0	2306
Mentions	349	1012.1	6341.3	0	73741
Age	349	45.1	10.6	22	85
Female	349	0.46	0.49	0	1
Incumbent	349	0.69	0.46	0	1
Position in Party List	349	5.38	11.7	1	63
Opposition Party	349	0.66	0.47	0	1
Left-right Party Placement	349	5.7	2.12	1.7	7.7
Urban District	349	0.17	0.38	0	1
Defense	349	4.48	4.67	0	50
Environment	349	5.25	17.31	0	164
Elderly	349	0.97	3.97	0	46
Gender	349	0.59	2.21	0	26
Occupation	349	3.61	9.66	0	93
Education	349	1.39	4.1	0	36
Care	349	2.86	10.14	0	124
Law	349	3.39	8.53	0	67
Immigration	349	1.71	4.86	0	53
Economy	349	4.85	16.07	0	201

Sources: Own Data (2018), Chapell Hill Expert Survey (2017)

Table A3. Candidates' Adoption, Use and Mentions During the Campaign – With Party Dummies

	Model 1 *Adoption*	Model 2 *Use*	Model 3 *Mentions*
Age	-0.0474*** (-3.80)	0.00839 (0.45)	-0.0312 (-1.42)
Female	0.127 (0.49)	-0.579 (-1.16)	-0.550 (-1.21)
Incumbent MP	0.0805 (0.28)	-1.346** (-2.60)	0.0778 (0.17)
Position in Party List	-0.0330* (-2.49)	0.0167 (0.41)	-0.0174 (-0.51)
Party Leader	0.460 (0.53)	2.214+ (1.87)	3.596** (3.21)
Candidate Opposition Party	0.692+ (1.82)	1.245+ (1.91)	0.284 (0.47)
Party Position Left-Right	0.153 (1.33)	-0.0996 (-0.60)	0.0948 (0.57)
Urban District	-0.198 (-0.60)	0.728 (1.29)	0.418 (0.83)
total Number of Tweets			0.00754*** (4.83)
Constant	1.842* (2.40)	4.935*** (4.10)	4.402*** (3.55)
		(19.67)	(19.25)
N	349	228	228

t statistics in parentheses, party dummies excluded from table
Sources: Own Data 2018, Chapel Hill Expert Survey (2017)
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A4. Categories of Discussion and Associated Keywords

Economy	Education	Environment
Ekonomi	Skola	Klimat
Skatt	Friskola	Miljö
Inkomst	Segregation	Utsläpp
Tillväxt	Betyg	Flyg
	Skolval	Bil

Immigration	Occupation	Defence
Invandring	Sysselsättning	Försvar
Flykting	Jobb	Nato
Integration	Arbetslös	Hot
Asyl	Arbete	Militär
Nyanlända		Säkerhet
Arbetskraft		

Law and order	Gender equality
Lag	Jämställdhet
Ordning	Samtyckeslag
Brott	Kvotering
Straff	Föräldraförsäkring
Kriminalitet	Me too
Trygghet	

Health care	Elderly issues
Vård	Äldreomsorg
Omsorg	Pension
Kömiljard	Garantipension
Vinster	Premiepension
Sjukvård	

Figure A1. Public Discussion of Political Topics Over Time

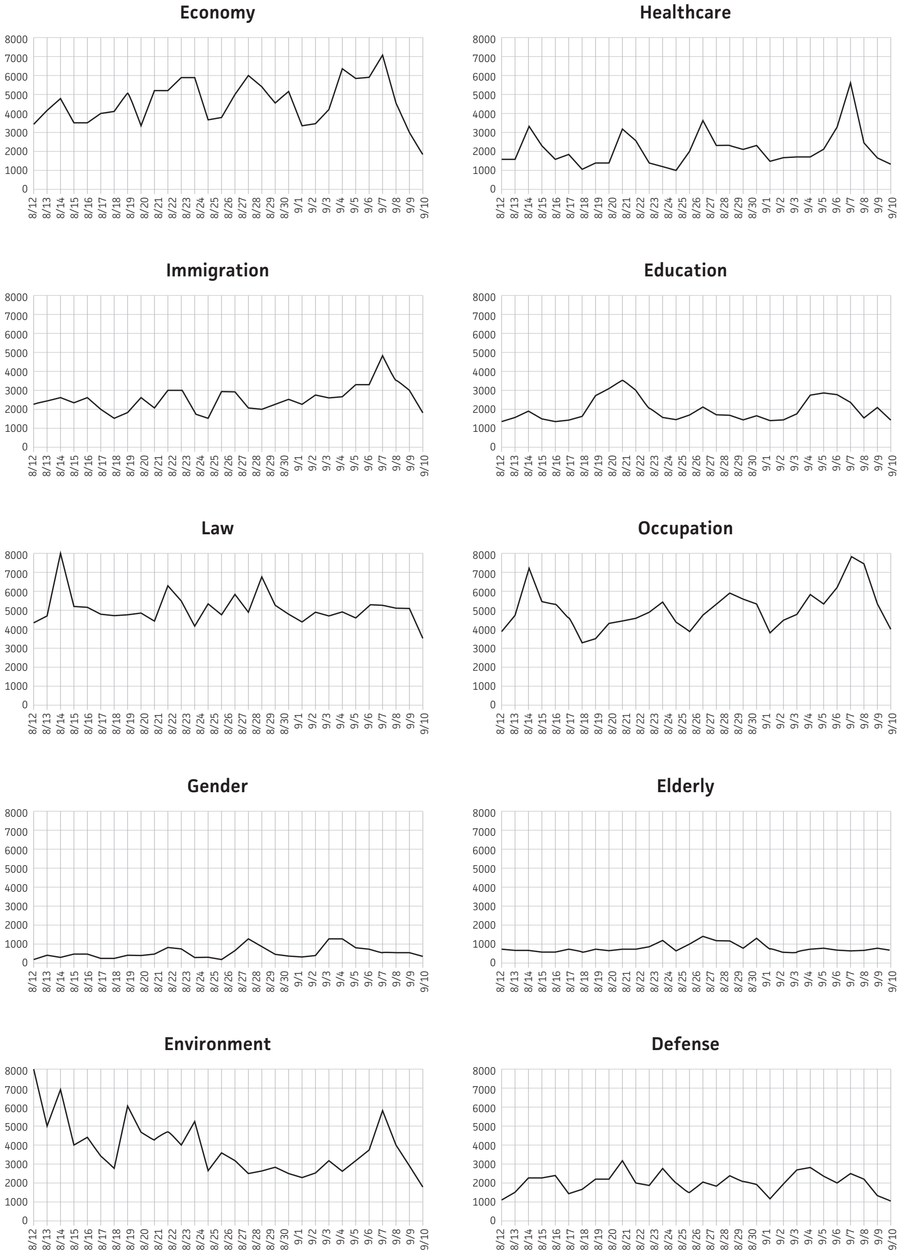


Table A5. Relationship Between Candidate Attributes and Frequency of Tweeting About Topics with Party-Dummies (dummy coefficients not shown)

	Dependent variable:									
	Defense (1)	Environment (2)	Elderly (3)	Gender (4)	Occupation (5)	Education (6)	Healthcare (7)	Law (8)	Immigration (9)	Economy (10)
Age	0.029***	-0.010	0.012	0.010	0.002	-0.014*	0.010	-0.008	0.007	0.006
	(0.010)	(0.009)	(0.010)	(0.009)	(0.005)	(0.008)	(0.007)	(0.006)	(0.007)	(0.006)
Female	-0.068	0.564***	0.149	0.883***	0.040	0.251	0.535***	0.029	0.046	0.007
	(0.223)	(0.210)	(0.192)	(0.169)	(0.112)	(0.172)	(0.165)	(0.123)	(0.149)	(0.123)
Incumbent MP	0.178	0.574***	0.360	-0.055	0.045	0.100	0.016	0.087	0.515***	0.128
	(0.234)	(0.222)	(0.242)	(0.204)	(0.124)	(0.195)	(0.184)	(0.137)	(0.177)	(0.137)
Position in Party List	-0.020	-0.051*	-0.012	-0.017	-0.024	-0.020	-0.010	-0.015	0.001	-0.008
	(0.023)	(0.030)	(0.032)	(0.025)	(0.016)	(0.024)	(0.019)	(0.012)	(0.013)	(0.014)
Party Leader	-0.047	0.011	0.580	0.582**	-0.257	0.356	-0.039	0.163	-0.262	-0.115
	(0.459)	(0.442)	(0.382)	(0.241)	(0.224)	(0.342)	(0.364)	(0.258)	(0.297)	(0.265)
Urban District	0.156	0.315	0.007	0.196	-0.326***	-0.061	-0.005	0.035	0.443***	-0.205
	(0.242)	(0.222)	(0.209)	(0.208)	(0.124)	(0.187)	(0.183)	(0.133)	(0.157)	(0.137)
Constant	-5.994***	-4.382***	-4.940***	-6.209***	-3.291***	-3.482***	-4.259***	-3.305***	-5.799***	-3.536***
	(0.537)	(0.520)	(0.530)	(0.520)	(0.283)	(0.417)	(0.389)	(0.301)	(0.374)	(0.308)
Observations	349	349	349	349	349	349	349	349	349	349
Log Likelihood	-294.261	-413.648	-197.858	-163.828	-363.200	-254.823	-349.211	-367.801	-266.408	-394.146

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

Sources: Own Data (2018), Chapel Hill Expert Survey (2017)