Appendix

This appendix is divided into three parts: in section A1 we provide further details on how some of our variables were operationalised; in section A2 we give full descriptive statistics for our dataset; in section A3 we report on model diagnostics; and in section A4 we report post-estimations from one of our models.

A1 Variable Operationalisation

In this section of the appendix, we provide more details on how our constituency level internet use indicator is operationalised, and the approach we take to constructing and coding our topic models.

A1.1 Constituency Level Internet Use Indicators

Our constituency level internet use indicators are based on the data released as part of a paper by Blank, Graham & Calvino (2017). The paper used small area estimation to provide a measure of internet use at the "output area" level, which is a small geographical area designed specifically for use with the UK census which typically contains around 120 households (there are 227,759 of these output areas in England, Scotland and Wales). The census provides lookup tables which map these output areas to parliamentary constituencies. We use these lookup tables to provide a population weighted average of internet usage by constituency on the basis of the output area measures. In cases where output areas overlapped with two or more constituencies, we assigned these areas to the constituency with which they had the largest overlap.

A1.2 Topic Models

In order to characterise the type of communication activity which political candidates engage in during the campaign period on Twitter, we make use of a series of topic models, a technique which is increasingly used in communication research (Maier et al., 2018). The approach allows us to extract a discrete number of general topics from the textual data within candidate tweets. The advantage of unsupervised topic models over methods such as content analysis and supervised machine learning is that they do not require a large amount of up front data, and therefore they make tractable the process of characterising the communication style of thousands of different political candidates. We will briefly describe the process we followed to produce our topic models here.

The input into a topic model is a 'document-term matrix', which is simply a matrix whereby each row is an individual document (an individual tweet in our case) and each column is a word which appears in the entire corpus of tweets. The entry $d\text{-}t_{ij}$ specifies the amount of times the term t_j appears in the document d_i. We only made use of 'original' tweets in our corpus (i.e. we do not include replies and retweets). We pre-process our corpus of tweets to remove common 'stopwords' (frequently occurring words such as 'and', 'the', 'it' etc. which we assume have little value in classification terms). We also 'lemmatize' all remaining words in the corpus, returning each word to its original base or lemma (such that, for example the words 'angry' and 'angrier' would be reduced to the same term, angry). We then limit our model to the 1,000 most frequently occurring words in the corpus, on the basis that extremely infrequent words are unlikely to be of use in distinguishing topics. Finally, we convert these term frequency scores using tf-idf weighting, which applies a stronger weight to terms which are less common across the corpus of the documents as a whole. From this document-term matrix, two further matrices are estimated: a term-topic matrix (which specifies the likelihood of individual terms appearing in a given topic) and a document-topic matrix (which specifies the likelihood of a given document, or tweet, belonging to a particular topic). These matrices were estimated using the non-negative matrix factorization [NMF] technique (see e.g. Gillis, 2014). We used NMF rather than the slightly more popular latent dirichlet allocation because NMF allows us to work with fractional tf-idf scores.

Table A1

Topic codebook.

Broadcasting Behaviours	Definition
Updating	Posting updates on recent candidate activity, e.g. attendance at events or doorstep campaigning
Promoting	Tweets specifically promoting the skills/ability of the candidate or party
Critiquing	Tweets criticising other parties or candidates
Information disseminating	Dissemination of news reports or informational links
Own / party stance	Tweets where the candidate takes a position on a specific policy area
Interacting beahviours	
Debating	Tweets where the candidate directly debates with opposition candidates or members of the general public
Acknowledging	Tweets where the candidate thanks people or acknowledges support
Organizing / mobilizing	Direct efforts to organise offline or online activity
Advice giving / helping	Candidate efforts to help people in individual constituencies
Consulting	Requesting public input on a given issue

Note. Adapted from Graham et al., 2014, pp.703-707.

A key consideration in topic models is choosing the appropriate number of topics, *k*, which must be specified by the researcher (Maier et al., 2018). The ideal value of *k* is one that allows the full variation of different communication styles to be captured without creating arbitrary divisions between groups of documents that are in practice quite similar. We chose to fit an individual topic model for each party in each year, on the basis that different parties might choose different communication strategies for their campaigns. To select the appropriate number of topics for each campaign year, we made use of the stability analysis technique proposed by Greene, O'Callaghan & Cunningham (2014), which is adapted to the particular case of NMF and involves analysing the extent to which the topic-term matrix is robust to random perturbations of the input data for different values of *k*. For all parties in each of the two waves of our data, we tested all values of *k* between 5 and 15 against 10 randomly drawn samples of 80% of the data. We then picked the value of *k* with the highest level of stability for each party (in the end all values of *k* were in the range 5-7).

Once we had created each party-year level topic model, we then labelled each topic in terms of its style of communication: whether it could be conceived of as largely one-way communication (*broadcasting*) or whether it attempted to engage in some form of two-way communication (*interacting*). Our definitions of these two communication types come from the codebook proposed in Graham et al. (2014, pp.697-698), where they label a number of different types of communication as falling into one of these two overarching categories. The definitions we use are set out in table A1.

The coding process itself was based on both the term-topic matrix and the document-topic matrix. For each topic, we extracted the 10 most probable terms and the 5 most probable tweets. One of the authors of the study then used the content of these terms and tweets to make a coding decision. A second author independently performed the process, producing a Krippendorrf's alpha of 0.69 (percent agreement of 88%). Having labelled all topics as either *broadcasting* or *interacting,* we were then able to label all tweets in the corpus as either *broadcasting* or *interacting* as well, on the basis of selecting the most probable topic for each tweet.

A2 Descriptive Statistics

Table A2

Descriptive Statistics

	2015		**2017**		**2015 - 2017 Panel**		
Categorical variables	**freq**	**%**	**freq**	**%**		**freq**	**%**
Total observations	3,172	100	2,826	100		1062	100
Incumbent MP	540	17%	597	21%	Inc. 2015, Chal. 2017	30	3%
Challenger MP	2,632	83%	2,229	79%	No Change	875	81%
					Chal. 2015, Inc. 2017	157	16%
Conservative Party	631	20%	631	22%	Conservative Party	359	35%
Green Party	567	18%	459	16%	Green Party	105	10%
Labour Party	631	20%	631	22%	Labour Party	285	31%
Liberal Democrats	631	20%	629	22%	Liberal Democrats	191	17%
Plaid Cymru	40	1%	40	1%	Plaid Cymru	3	0%
Scottish National Party	59	2%	59	2%	Scottish National Party	46	4%
UK Independence Party	613	19%	377	13%	UK Independence Party	73	4%
Had Twitter	2,398	76%	1,786	63%	Had Twitter	822	77%
Did not have Twitter	774	24%	1,040	37%	Did not have Twitter	240	23%
Numeric variables	**mean**	**sd**	**mean**	**sd**		**mean**	**sd**
Vote Share (%)	19.73	17.79	22.17	22.20	Share 2017 - 2015	2.79	8.31
Votes	9,358	8,665	11,012	11,291	Votes 2017 - 2015	2,185	4,735
Internet Use in Cons. (%)	77.56	5.61	77.56	5.61			
Numeric variables (those with Twitter account only)							
Tweets	169.65	242.72	152.94	248.53	Tweets 2017 - 2015	-18.87	183.33
Advertising Spend	3,242	4,125	3,951	3,846	Advertising 2017 - 2015	717.0	4,366
Staff Spend	647	1,181	809	1,285	Staff 2017 - 2015	255.7	1314.99
Party Av. Share	19.73	13.63	22.17	18.68	Party Av. 2017 - 2015	2.97	6.51
Original Tweets	46.95	76.20	44.19	84.40	Orig. 2017 - 2015	1.45	71.47
Replies	37.52	70.91	23.82	58.05	Replies 2017 - 2015	-15.51	52.61
Retweets	85.19	154.94	84.92	167.86	Retweets 2017 - 2015	-4.81	124.24
Broadcast Tweets	48.21	74.76	43.63	83.07	Broadcast 2017 - 2015	-1.41	68.21
Interacting Tweets	2.98	10.33	7.88	17.40	Interacting 2017 - 2015	4.50	11.98

A3 Model Diagnostics

In this section of the appendix we report on the diagnostic checks undertaken on our models. All fitted models were analysed with standard goodness of fit diagnostic tests for OLS models. In particular, variance inflation factors were inspected for evidence of multicollinearity, Cook's distance was calculated for all observations in order to identify high influence data points, plots of variables versus fitted values were inspected to check the general assumption of a linear relationship between dependent and independent variables, and a Breusch-Pagan test for heteroscedasticity was conducted.

These tests highlighted two main potential concerns with the fit: evidence of heteroscedasticity and a group of around 150 high influence observations (the exact amount varying depending on the model being fitted). The majority of these observations were parliamentary candidates who were campaigning in "safe" parliamentary constituencies, which are characterised by very high levels of concentration of voters for one party. In these safe seats, parliamentary candidates can win a lot of votes with little campaigning effort. Hence, these observations deviated considerably from the overall fit of the model. This impact was sometimes compounded when an MP retired and a new candidate took their place. In this situation, the new candidate is not treated as an "incumbent" by the model, but has many of the advantages of being an incumbent.

These problems were not resolved by log transformation of variables, so two further steps were taken in response. First, coefficient estimates, measures of statistical significance and estimates of adjusted R^2 were all computed by bootstrapping (R=5,000). Second, robust linear regressions were also estimated for all models. The results of the robust regressions were largely the same as the original OLS estimates, hence we reported the original estimates to facilitate interpretation. Separate cross sectional models for 2015 and 2017 were also produced, which again largely support the original interpretation. We also duplicated the models produced using a multilevel approach, with individuals nested in constituencies. The results were the same.

The following differences between our robust and cross sectional regressions and the original results were noted:

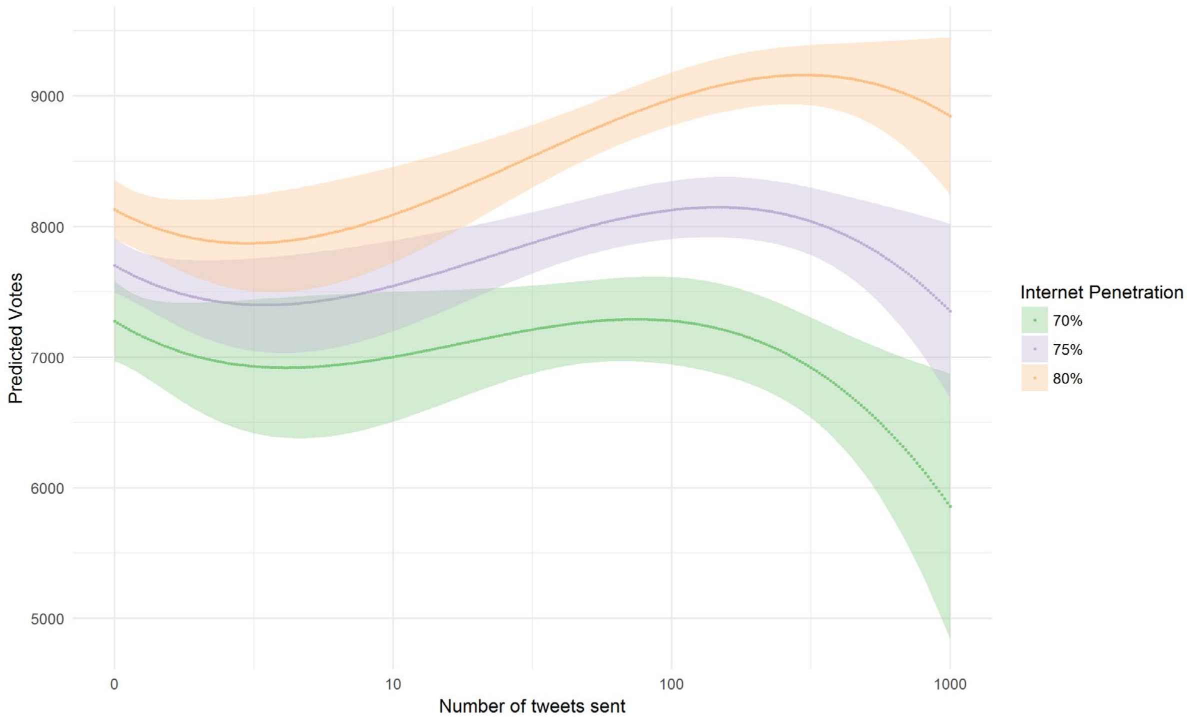
- The term for original tweets is only significant at the 0.1 level in the robust version of our first difference regression.
- The term for broadcasting tweets is only significant at the 0.1 level in our 2015 cross section and in the robust version of our pooled time series model. It is furthermore not significant (though still pointing in the same direction) in the 2017 model.
- The interaction term with incumbency status is also not significant in the 2017 cross section.

A4 Semi-parametric post-estimations

In order to further explore our results, we produced a semi-parametric estimation of model 2.3 (with three knots). This allows the regression line to curve and thus allows us to find where the effect of sending tweets starts to diminish, according to our data. Post-estimations from this model are presented in Figure A1 below.

Figure A1

Post Estimations from Model 2.3



The figure shows that while increasing tweets sent from 10 to 100 provides a clear benefit, beyond that point the effects start to decrease (with the exact point of decrease depending on the internet penetration level of the constituency). According to the model, the 'ideal' number of tweets to send was just over 250 in an average constituency.