

# Different news for different views: Political news-sharing communities on social media through the UK General Election in 2015

Matthew J. Williams<sup>1</sup>, Iulia Cioroianu<sup>2</sup>, Hywel T.P. Williams<sup>1\*</sup>

[1] College of Life and Environmental Sciences,  
University of Exeter, Exeter, UK

[2] Dept. of Politics, University of Exeter, Exeter, UK

## Abstract

Media exposure is a central concept in understanding the dynamics of public opinion and political change. Traditional models of media exposure have been severely challenged by the shift to online news consumption and news-sharing on social media. Here we use network analysis and automated content analysis to examine the interaction between news media and social media around the UK General Election in 2015. We study a large corpus of UK newspaper articles and Twitter content, finding significant temporal correlations between newspaper topic coverage and the content discussed on Twitter. We also identify news-sharing communities around groups of news sources that are ideologically clustered. Analysis of topics covered within each group shows that different communities are exposed to different news content during the election. Our results confirm that ideological bias and selective news-sharing affect patterns of online media exposure in social media.

## Introduction

Media exposure is perhaps one of the most central concepts in social sciences (Prior 2013); in order to understand change (and stability) in opinions and behaviour, it is necessary to measure the information to which a person has been exposed. The web has radically changed the media environment. Individuals now browse and share diverse information from social and traditional media on a wide range of online platforms, creating new patterns of exposure and alternate modes of content production (e.g. user-generated content) (Valkenburg and Peter 2013). The fundamental dynamic of online “communication exposure” (Castells 2007; 2011) involves formation of ties between users and media content by a variety of means (e.g. browsing, social sharing, search). Online media exposure

is thus essentially a process of network formation that links sources and consumers of content (nodes) via their interactions (edges), requiring a network perspective for its proper understanding.

Online media have become an important channel for public debate and opinion formation about many issues. Social media can engage large populations and bring information to the attention of large audiences. However, another feature of online communication is selectivity. The huge variety of online content allows users to easily avoid content they disagree with, leading to biased content exposure. Potential “filter bubbles” arising from content recommendation algorithms have been widely discussed (e.g. (Pariser 2011)), but networked user interactions can additionally create social filter effects when users preferentially share content from favoured sources. A recent study of political news-sharing on Facebook showed that both social and algorithmic filtering contributed to large ideological biases in content exposure (Bakshy, Messing, and Adamic 2015). Political news-sharing on Twitter appears to be similarly affected by partisan bias; one recent study found that 90% of Twitter users only subscribed to news sources from one political leaning, although the study also showed that the diversity of exposure was increased by retweeting of news from alternate viewpoints by friends (An et al. 2014). Analysis of online political networks has repeatedly identified partisan “echo chambers” in which users interact only with like-minded others and are isolated from alternative viewpoints (e.g. (Adamic and Glance 2005; Conover et al. 2011; 2012)). Online echo chambers have also been identified for other contentious issues, e.g. climate change (Williams et al. 2015; Elgesem, Steskal, and Diakopoulos 2015). Selective dissemination and online echo chambers may increase polarisation and promote fragmentation of public discourse, such that existing views become more extreme and consensus is hard to achieve (Sunstein 2007). Such findings appear to contradict the notion of online media as an open “marketplace for ideas” and compromise its potential for cross-constituency public debate.

Here we combine network analysis of news-sharing on social media and content analysis of news arti-

\*Corresponding author: h.t.p.williams@exeter.ac.uk

Copyright © 2015, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

cles to empirically examine aspects of online media exposure around the 2015 UK General Election. The election took place on 7th May 2015 and was primarily contested by the right-wing Conservative Party led by David Cameron and the left-wing Labour Party led by Ed Miliband. A number of smaller parties also fielded candidates, notably the centrist Liberal Democrat Party, the regional Scottish National Party (SNP), the far-right UK Independence Party (UKIP) and the environmentalist Green Party. The election was won by the Conservative Party.

We analyse social media content from Twitter and news articles from a variety of UK national and regional newspapers. An important feature of Twitter is the ability to share links to external content, such as news articles, images and videos. Previous research found that URLs are used more frequently in political communication than in other types Twitter communication and that the use of URLs increases the number of retweets (Stieglitz and Dang-Xuan 2012; Bruns and Stieglitz 2012). After first analysing the topics that were most popular in Twitter and newspaper coverage of the General Election, we next construct networks representing the shared audiences for different newspapers amongst Twitter users, based on patterns of news-sharing. We show that there are statistically significant relationships between Twitter and news media during the election, that news sources covering the election can be grouped by the overlaps between their Twitter audiences, and that different audiences are exposed to different distributions of news topics and political ideology.

## Data collection and methods

**Collection of tweets** Twitter is a social messaging platform with 288 million active monthly users sending 500 million tweets per day (Statista 2015). The Archive.org Twitter Stream Grab<sup>1</sup> collects data from Twitter's public 1% stream and makes it publicly available to support reproducible social media research. We downloaded the original Archive.org tweet corpus covering March-May 2015. We then restricted this corpus to a study period spanning 22nd March to 17th May, to cover the election campaigns, the election, and the immediate aftermath. We then further filtered the retained tweets by matching to hashtags contained in a set of election-related tags `{#ge2015, #generalelection2015, #battlefornumber10, #leadersdebate, #bbcdebate, #bbcqt, #scotdebates, #scotdebate, #walesdebate, #walesdebates}` and a set of party-based hashtags announced by Twitter UK<sup>2</sup> `{#conservative, #labour, #libdems, #ukip, #greens, #snp, #plaid15, #dup, #sdip, #respectparty}`. The original Archive.org corpus contained 267 million global tweets which were filtered to an election-focused dataset consisting of 86,939 tweets made by 52,299 unique users and containing 10,529 unique hashtags. There was a temporary failure in data collection by Archive.org during the period 23rd April to 27th April. We exclude this period from subsequent time series analysis.

<sup>1</sup><https://archive.org/details/twitterstream>

<sup>2</sup><https://twitter.com/TwitterUK/status/586455058264363008>

**Collection of news articles** The database of newspaper articles was built by (Stevens 2016). Articles published by 17 major national and local UK newspapers between the 1st of February and the 30th of May 2015 were downloaded from the LexisNexis database (nex 2015). The newspapers selected were: *The Daily Mail*, *The Daily Star*, *The Express*, *The Telegraph*, *The Sun*, *The Times*, *Western Morning News*, *Daily Mirror*, *The Independent*, *The Guardian*, *The Scotsman*, *Western Mail*, *Yorkshire Evening Post*, *The Evening Standard*, *Financial Times* and *The Daily Record*. A subset of 11,000 articles were human-coded and labeled as either being about the UK General Elections or not. The labeled articles were used to train a supervised classifier in order to identify election articles in the rest of the corpus. Following previous research (Joachims 2002), a linear support vector machines classifier was trained using a stochastic gradient descent algorithm (F-score=0.95), which predicted 21,038 articles to be about the elections. Our study corpus consisted of 13,551 of these articles that were published during the 22nd March - 17th May study period.

**Topic modelling on news articles** Stevens et al (2016) identified the topic composition of their election news articles by estimating a latent Dirichlet allocation (LDA) model using the MALLET (McCallum 2002) implementation of Gibbs sampling. The LDA model was fitted using 30 topics. The 15 most relevant issue topics and 6 topics relating to the major political parties were retained (Table 1).

**News-sharing networks and community detection** To analyse news-sharing communities, we first constructed a bipartite network of users and web domains based on the occurrence of embedded hyperlinks in tweets. Twitter shortens embedded hyperlinks to permit concise tweets. We resolved shortened URLs to full URLs and extracted the primary domain of each URL, removing common leading subdomains, such as *m* (web pages for mobile browsers) and *www*. For example, the shortened URL <https://t.co/p3XS6nd1Rb> resolves to an article in *The Guardian* at <http://www.theguardian.com/politics/2015/may/10/election-2015-exit-polls>, from which we extract the domain *theguardian.com*. Of the 52,299 users in our dataset, 15,152 users tweeted a URL, which after link resolution yielded 2,349 distinct domains.

For all tweets containing URLs, we created network edges between the tweet author and the associated web domain. The resulting bipartite network contained 15,152 user nodes and 2,349 domain nodes connected by 17,855 edges. The giant component contained 1,659 users, 736 domains, and 4,309 edges. We then took a unipartite projection from the giant component of the bipartite user-domain network to create a unipartite network of domains in which each weighted edge represents the number of users who tweeted links to both domains. The domain network contained 736 nodes and 2,989 edges. The original bipartite network captures the pattern of news-sharing amongst Twitter users in our dataset, whereas the unipartite domain network captures the association between domains based on the size of their shared audiences. We analysed community structure within the domain network using the Louvain method for community detection (Blondel et al. 2008), which finds communities as groups of nodes that are densely connected by edges. Each node can only belong to a single community.

**Temporal correlations between news topics and hashtag frequencies** To quantify the relationship between news

Topic	Keys
<b>Tax &amp; Spend</b>	cut spend osborn budget plan tax year chancellor govern public deficit econom fiscal georg billion labour cent tori debt
<b>Polls</b>	poll labour cent vote seat tori parti voter elect conserv win ukip support point result lib dem show lead
<b>Coalition</b>	snp govern parti labour vote scottish deal tori scotland coalit english cameron major mp conserv elect parliament support power
<b>Benefits</b>	work tax cut tori benefit labour peopl manifesto plan govern pay promis wage welfar pledg year cameron conserv bill
<b>EU</b>	eu britain cameron referendum uk european british europ mr defenc minist prime vote countri foreign david membership nuclear govern
<b>Media</b>	elect brand polit bbc show media news russel twitter star channel day sun ed interview tweet david tv daili
<b>Debates</b>	debat leader cameron miliband david bbc parti question farag ed clegg broadcast tv sturgeon prime audienc bst nick elect
<b>Regions</b>	labour seat candid vote mp west north south constitu tori major conserv east local sir ukip green margin elect
<b>Economy</b>	market uk bank elect govern economi year price econom growth cent britain rate financi invest share busi sinc rise
<b>Schools</b>	school educ immigr fee student univers year free tuition migrat polici govern net number children pupil teacher cut fund
<b>Business</b>	busi labour compani tax small britain parti execut letter support firm leader uk chief miliband polici corpor govern employ
<b>Donations</b>	parti donat mp avoid tori lord tax conserv sir donor gmt report fund minist account shapp money labour polit
<b>Housing</b>	hous home wale welsh rent buy properti build council govern plaid labour peopl plan polici associ tenant fund year
<b>London</b>	london johnson mayor bori citi local osborn transport elect manchest north tori region power capit council west rail plan
<b>NHS</b>	nh health servic care patient hospit year gp labour fund doctor nurs privat govern plan extra public staff england
<b>Conservatives</b>	tori cameron campaign parti conserv elect voter david prime minist labour win poll week day messag time leader crosbi
<b>SNP</b>	snp scotland scottish sturgeon labour murphi nicola leader parti referendum ms vote westminist salmond jim uk independ scot elect
<b>Lib Dem</b>	lib dem clegg parti liber nick democrat mr coalit alexand seat leader tori danni secretari conserv elect minist mp
<b>UKIP</b>	ukip farag parti mr nigel leader thanet south elect mp carswel mep immigr campaign candid support seat claim yesterday
<b>Labour party</b>	labour parti union leadership leader shadow mp secretari candid elect unit umunna burnham member miliband support ed back mr
<b>Green party</b>	parti women green vote elect candid mp polit labour campaign femal bennett ms young regist men support peopl group

Table 1: LDA keywords for each topic in news articles.

and social media content we computed temporal correlations between news topics and hashtags. We formed time series of hashtag use  $\{x_{h,t}\}_{t=1\dots N}$  and news topic occurrence  $\{y_{k,t}\}_{t=1\dots N}$ , over  $N$  24-hour time windows, where  $x_{h,t}$  represents the fraction of users tweeting hashtag  $h$  and  $y_{k,t}$  represents the fraction of news articles referencing news topic  $k$  in time window  $t$ . We measured the pairwise temporal association between each news topic  $k$  and hashtag  $h$  using Pearson’s correlation coefficient applied to the two time series  $\{y_{k,t}\}_{t=1\dots N}$  and  $\{x_{h,t}\}_{t=1\dots N}$ . Here we consider only simultaneous correlation between the two time series, while noting that time-lagged correlations may also exist.

## Results

To understand the narrative of the General Election in 2015, we first compared the evolution over time of news topics and hashtag frequencies (Figure 1). The news article topic distribution over time illustrates the complexity of the election discourse as played out in the mainstream news media. Most topics show considerable variability in the level of coverage they receive, although some show distinct periods of high interest amidst lower typical levels; for example, the “debates” topic shows high activity at the times of the various televised leaders’ debates (26th March, 2nd April and 16th April), while the “Labour Party” topic peaks after the election when attention focused on the leadership succession after Ed Miliband resigned. Hashtag frequencies also indicate high attention to the leaders’ debates on Twitter, with related hashtags (*#BBCDebate*, *#LeadersDebate*) showing activity spikes on the relevant days, and to the post-election Labour leadership contest, with *#LabourLeadership* trending after

the election.

To quantify and formalise the association between the social media and news media narratives of the election over time, we calculated the pairwise temporal correlation between the relative frequency of each hashtag and each news topic. Results are shown in Figure 2. These correlations measure the similarity of the temporal profiles of hashtag use and news topic prevalence, and do not imply causal linkage; significant correlations simply indicate co-occurrence of news topics with hashtags. A number of significant positive and negative correlations are observed. For example, the attention given to the leaders’ debates on Twitter is confirmed by positive correlations between the “debates” topic and hashtags (*#NigelFarage*, *#farage*, *#sturgeon*) relating to two of the party leaders who were seen to have performed well, Nichola Sturgeon and Nigel Farage. Overall the high number of significant correlations indicates strong interactions between news media and social media. Out of 1,554 pairwise correlations (comparing 74 hashtags to 21 news topics), 71 were found to be significant; at a threshold of  $p<0.01$  we would expect around 16 significant correlations by chance alone.

We next consider how social media users are exposed to news content by sharing links to web domains. Within the domain network, which represents associations between web domains based on the links shared by Twitter users (see Data Collection and Methods), we find 11 communities representing groups of domains with similar user audiences (Table 2 and Figure 3). The best-partition modularity score of  $Q=0.283$ ,

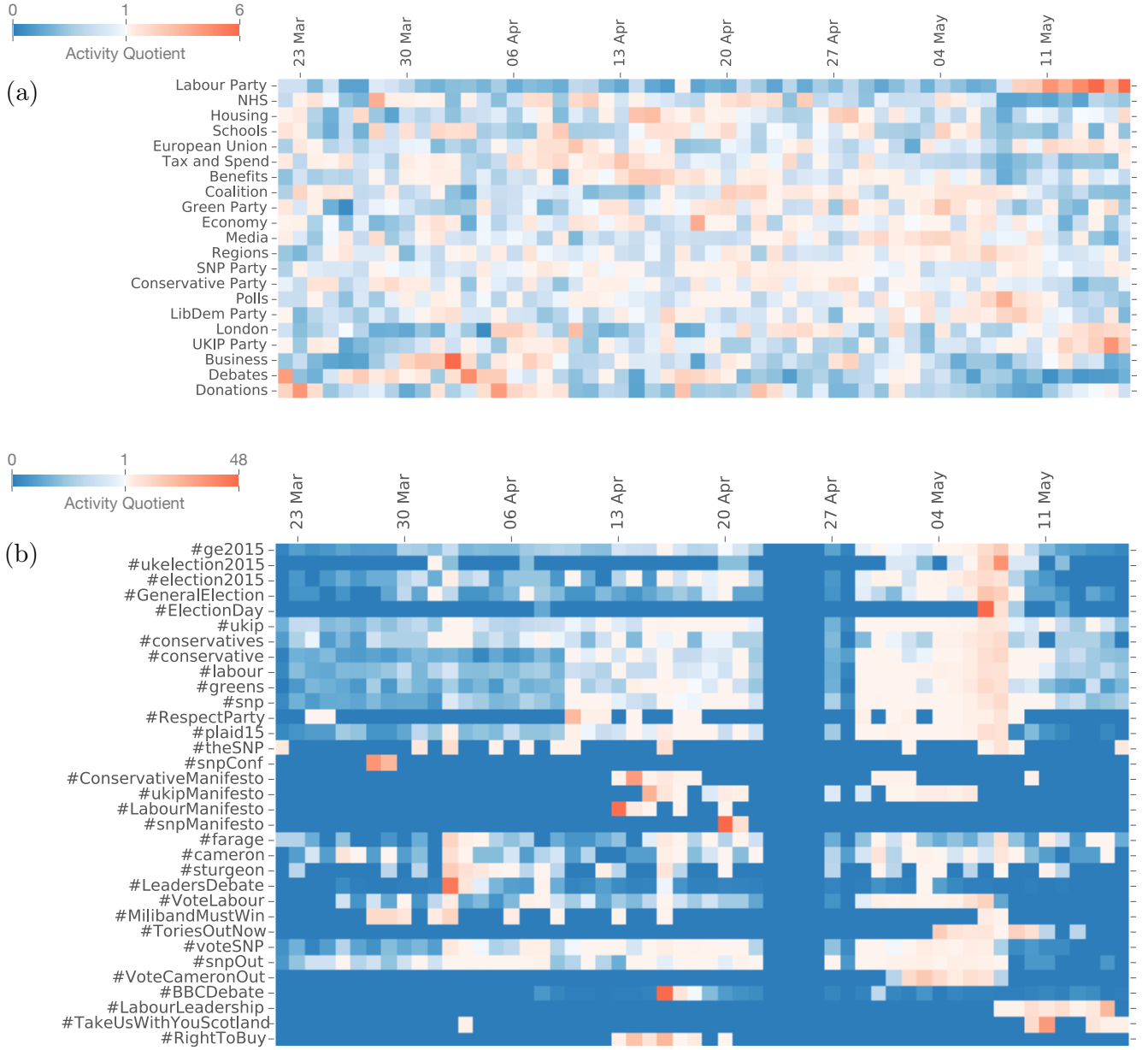


Figure 1: Temporal heatmap showing activity quotients indicating content trends over time for (a) news article topics and (b) Twitter hashtags. Red indicates that a topic was more-commonly used in a given time instant, relative to its overall average use. All 21 LDA topics are shown aggregated over 17 newspapers. Hashtags shown are a subset of those shown in Figure 2 chosen for relevance/interest. There was a temporary failure in data collection by Archive.org during the period 23rd April to 27th April, visible as a universal dip in hashtag use. The activity quotient for a hashtag  $h$  at time  $t$  is given by  $\frac{a_{h,t}}{\langle a_h \rangle}$ , where  $a_{h,t}$  is the number of users tweeting  $h$  in time window  $t$  and  $\langle a_h \rangle$  is the average use of  $h$  over all time windows. Similarly, for the activity quotient of a news topic, we take  $a_{k,t}$  as the fraction of news articles containing topic  $k$ .

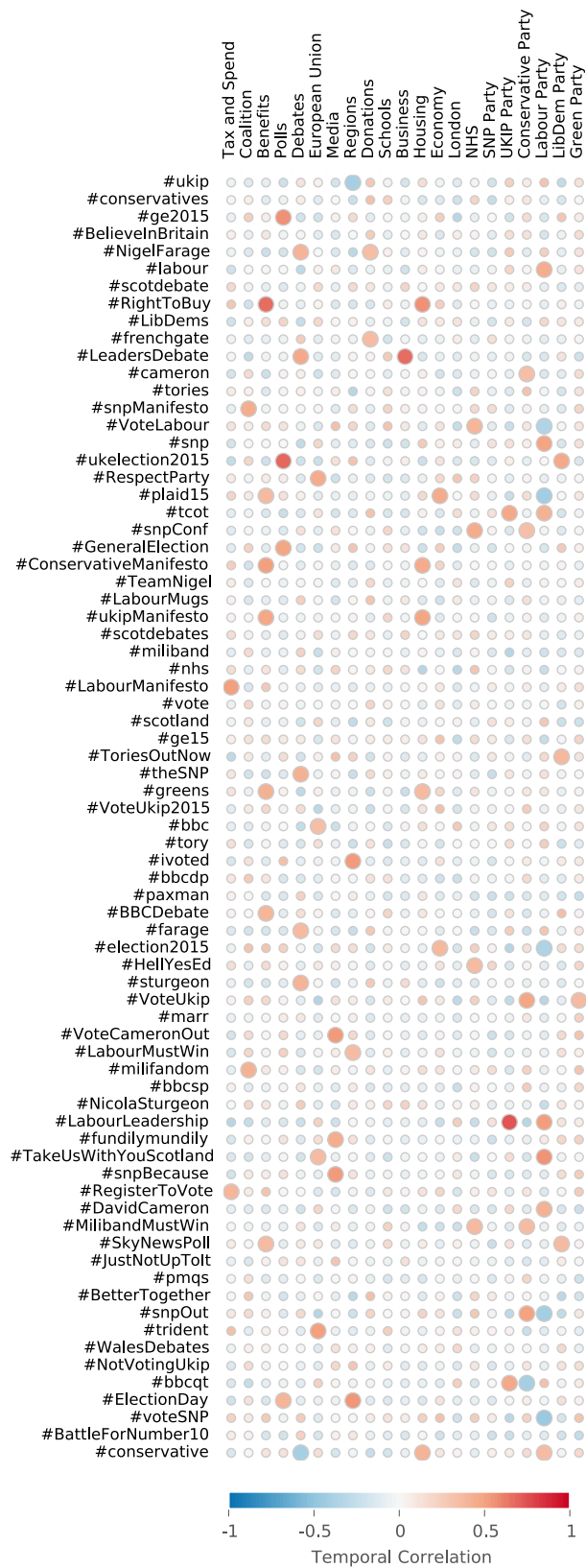


Figure 2: Temporal correlation between news topics and hashtags calculated from 24-hour time windows during the study period. Colour indicates the Pearson correlation coefficient between the temporal profiles of two topics based on time series of hashtag use and news article topic prevalence (see Data Collection and Methods). Large circles indicate significant correlations at  $p < 0.01$ . All 21 LDA topics are shown aggregated across all 17 newspapers. Hashtags shown are all those which appeared in the top-4 ranks by usage on any day during the study period.

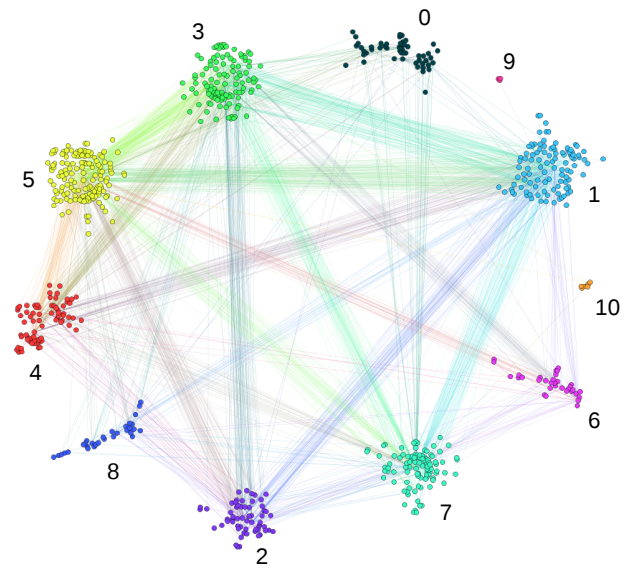


Figure 3: Domain network showing 11 identified communities. Each node is a web domain and each edge represents at least one Twitter user who shared a link to both domains. The network layout was initially created using a force-directed algorithm supplied by the Gephi visualisation package (Bastian, Heymann, and Jacomy 2009), after which the communities were coloured and manually separated to show community-community linkage.

Community	Domain Nodes	Edges	Audience Size	Newspapers
0	52 (7%)	166	62	0
1	104 (14%)	362	426	7
2	60 (8%)	104	252	2
3	119 (16%)	239	561	2
4	73 (10%)	123	203	2
5	155 (21%)	273	895	1
6	33 (4%)	66	80	1
7	100 (14%)	164	337	0
8	31 (4%)	61	80	1
9	3 (0%)	3	2	0
10	6 (1%)	7	3	0
Whole Network	736 (100%)	2,989	1,659	16*

\*An extra newspaper, Birmingham Mail, was not among the pruned domains.

Table 2: Statistics for the domain network and identified communities. A community's audience size is the number of users who have tweeted a link to a domain in that community. A user may belong to the audience of multiple communities if they tweet multiple links. These shared audiences can be observed as inter-community edges in Figure 3. The identified communities contained 16 of the 17 UK-based national or regional newspapers in our dataset.

found by the community detection algorithm, indicates moderately strong community structure and the network diagram in Figure 3 shows considerable residual linkage between the identified communities.

Inspection of the communities identified (Figure 4) reveals contextual information that supports their coherence. Different communities appear to have ideological (left/right leaning) or regional (Scottish, Welsh) themes in the domains they contain. Community 1 includes the UKIP and Conservative websites, as well as five right-leaning newspapers (*The Telegraph*, *Daily Mail*, *The Sun*, *The Times* and *The Express*) and two centrist/independent newspapers (*The Daily Star* and *The Morning News*). Community 2 includes Labour's press office and the largest pro-Labour political blog (*LabourList.org*), as well as two left-leaning newspapers (*The Mirror* and *The Independent*). Community 3 is centered around left-leaning *The Guardian* newspaper - which endorsed Labour in the 2015 elections, but supported Liberal Democrat candidates when they were the main opposition to the Conservatives - and the Liberal Democrat website, but also the main Scottish newspaper, *The Scotsman*. Community 4 is centred around Wales politics and contains the website of the Welsh national party, Plaid Cymru, and the Welsh regional newspaper, *Western Mail*. However, it also includes the *Yorkshire Evening Post*. Community 5 is centered around Twitter and the BBC and contains a number of other broad interest political sources, including the SNP website and the London-based *Evening Standard* newspaper. Community 6 seems to be a Scottish community which includes the *Daily Record*. Multiple domains related to the Green Party are represented in Community 7, which does not have a large newspaper associated with it. Community 8 occupies the political centre-right and includes the *Financial Times*, as well as other publications which focus on economics and finance, such as *The Economist*, but also various domains linked to the Liberal Democrats.

We used the LDA topic vectors for the 17 newspapers in our dataset to create topic vectors for each community (Figure 5). Differences between community topic vectors show variation in the balance of news topics to which users following the associated web domains are likely to have been exposed during the election period. The dominant topics associated with each community show good correspondence with political parties which are represented within its member domains, based on party-associated issues that have previously been identified in the political science literature (Green and Hobolt 2008). For example, right-leaning Community 1 shows prominent topics including "tax and spend" (an issue normally associated with the Conservative Party) and "UKIP". Left-leaning Community 2 shows prominent topics including "benefits" (i.e. social welfare payments) and "Conservatives" (presumably in the form of criticism). The prominent topics in newspapers in left-leaning/Scottish Community 3 include "SNP" (the Scottish National Party) and

"coalition", which reflects discussion of whether the SNP would join a left-leaning coalition with the Labour Party if Labour failed to win an outright majority.

## Discussion

Here we have examined the complex relationship between online news media and news-sharing on social media around the UK General Election in 2015. We find significant temporal correlations between topics covered in news media and content discussed on Twitter, indicating a strong coupling between the two media types, although there are also substantial differences in content. We identified distinct groups of online news domains based on similar patterns of news-sharing by Twitter users. These groupings have clear ideological leanings and we extrapolate from newspaper topic distributions to infer that user audiences clustered around the different domain communities were exposed to different news content during the election.

One important limitation to our study is the quality of our social media dataset. We used a public archive of the 1% Twitter stream available at Archive.org. While this data has the advantage of being publically available (making any analysis repeatable), it also suffered from collection failure during our study period, missing several days of data. More significantly, because the archive is a sample of the complete Twitter feed and is not focused on our study area of UK politics, we were only able to retrieve a relatively small sample of relevant data for this study. However, while the data we were able to retrieve is limited in volume, it is unbiased and our results retain validity. Another limitation arises from our assumption that the content exposure for users associated with the domain communities we identify is determined by the topic distributions of the newspapers found within the communities. Since the newspapers were a small number of domains amidst much larger communities, and since users associated with the domain communities are also likely to receive content from other sources, this assumption must be validated in future work. We hope to confirm the relationship between news media content and social media audience exposure with further study working with a larger tweet sample and restricting the topic analysis to news articles shared within each community.

Social media users who actively discuss politics and share links to related news articles are likely to be a minority of the user population. These strongly engaged users are likely to act as "opinion leaders" (Katz 1957) that disseminate relevant web content produced outside social media (e.g. news articles, reports) to their less-engaged followers, whose exposure to political content depends critically on this association. Thus the effects of selectivity are amplified by the effect of the engaged user group on the content exposure of their followers; since the engaged users who share links to news media are the entry point for news content into the "Twittersphere" the par-



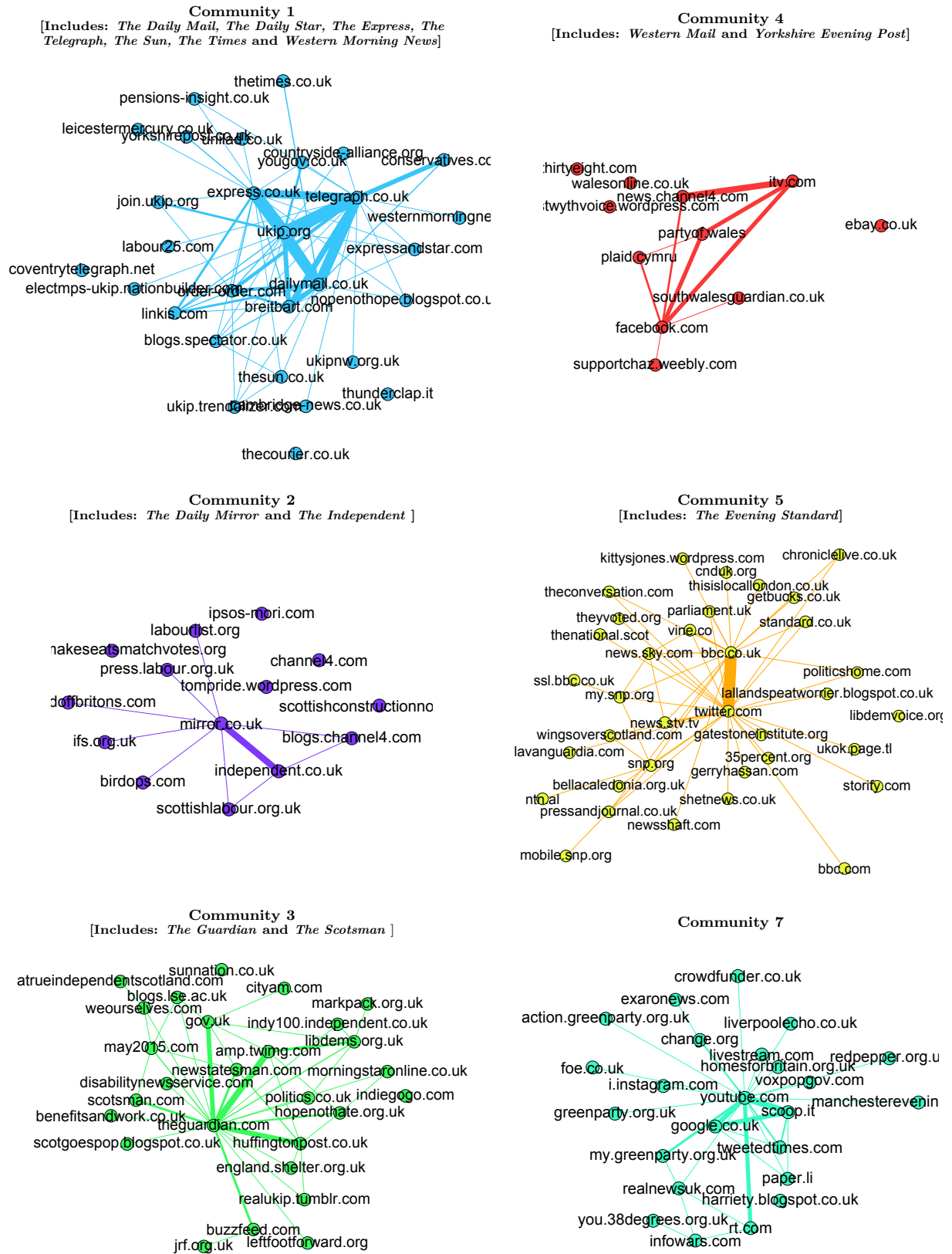


Figure 4: Structure of the six largest domain communities. Node labels give domain names, edge thickness indicates the number of users who tweeted URLs linking to both domains. Captions give the newspapers within each community which were used to calculate the mean topic vectors plotted in Figure 5.

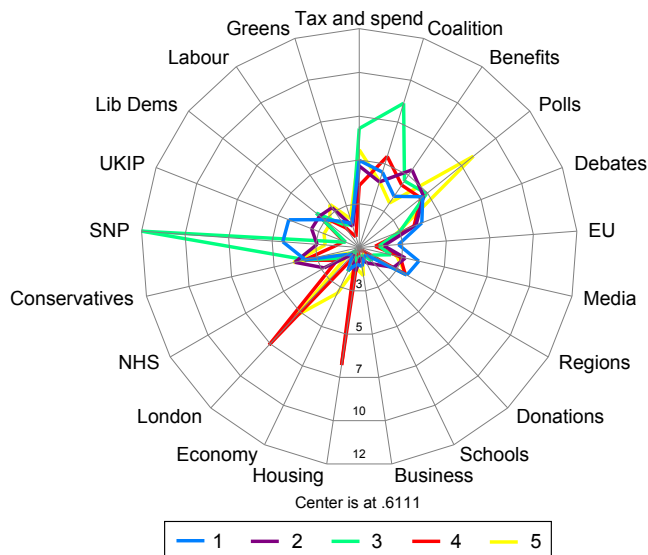


Figure 5: Topic distribution vectors for five of the largest domain communities. The topic vector for each community was calculated as the average of the topic vectors for the newspapers found in each community; the newspapers found in each community are given in Figure 4.

tisan filtering they apply may affect a wider population. While we do not study the political leanings of individual users here, we do observe that the domain communities we identify show clear political leanings. There is a clear analogy with political echo chambers found elsewhere in social media (Adamic and Glance 2005; Conover et al. 2012); we find both left-wing and right-wing domain communities created by users predominantly sharing content from one side of the political spectrum.

**Acknowledgements.** This work was supported by the UK Economic and Social Research Council ES/N012283/1.

## References

Adamic, L. A., and Glance, N. 2005. The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery (ACM)*, 36–43.

An, J.; Quercia, D.; Cha, M.; Gummadi, K.; and Crowcroft, J. 2014. Sharing political news: the balancing act of intimacy and socialization in selective exposure. *EPJ Data Science* 12.

Bakshy, E.; Messing, S.; and Adamic, L. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239):1130–1132.

Bastian, M.; Heymann, S.; and Jacomy, M. 2009. Gephi: an open source software for exploring and manipulating networks. In *Proc. International Conference on Weblogs and Social Media (ICWSM 2009)*.

Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; and Lefebvre, E. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008(10):P10008.

Bruns, A., and Stieglitz, S. 2012. Quantitative approaches to comparing communication patterns on twitter. *Journal of Technology in Human Services* 30(3-4):160–185.

Castells, M. 2007. Communication, power and counter-power in the network society. *International Journal of Communication* 1:239–266.

Castells, M. 2011. *The Rise of the Network Society: The Information Age: Economy, Society, and Culture. Vol. 1*. London, UK: John Wiley & Sons.

Conover, M.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Menczer, F.; and Flammini, A. 2011. Political polarization on twitter. In *Proc. International Conference on Weblogs and Social Media (ICWSM 2011)*.

Conover, M. D.; Gonçalves, B.; Flammini, A.; and Menczer, F. 2012. Partisan asymmetries in online political activity. *EPJ Data Science* 1(1):1–19.

Elgesem, D.; Steskal, L.; and Diakopoulos, N. 2015. Structure and content of the discourse on climate change in the blogosphere: The big picture. *Environmental Communication* 9(2):169–188.

Green, J., and Hobolt, S. B. 2008. Owning the issue agenda: Party strategies and vote choices in british elections. *Electoral Studies* 27(3):460–476.

Joachims, T. 2002. *Learning to classify text using support vector machines: Methods, theory and algorithms*. Kluwer Academic Publishers.

Katz, E. 1957. The two-step flow of communication: an up-to-date report on an hypothesis. *Public Opinion Quarterly* 21:61–78.

McCallum, A. K. 2002. Mallet: A machine learning for language toolkit. <http://mallet.cs.umass.edu>.

2015. Nexis uk. Part of LexisNexis, [www.lexisnexis.co.uk/products/nexis](http://www.lexisnexis.co.uk/products/nexis).

Pariser, E. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. London, UK: Penguin.

Prior, M. 2013. Media and political polarization. *Annual Review of Political Science* 16:101–127.

Statista. 2015. Global social networks ranked by number of users (March 2015). Technical report. Date of access: 12th June 2015.

Stevens, Dan; Banducci, S. C. T. K. G. 2016. The 2015 uk general election coverage in newspapers. In: *Media in Context and the 2015 General Election: How Traditional and Social Media Shape Elections and Governing*. Economic & Social Research Council ES/M010775/1 (2015-2016). Available at <http://hdl.handle.net/10871/19310> or <https://mediaeffectsresearch.wordpress.com/research-output/>.

Stieglitz, S., and Dang-Xuan, L. 2012. Political communication and influence through microblogging—an empirical analysis of sentiment in twitter messages and retweet behavior. In *System Science (HICSS), 2012 45th Hawaii International Conference on*, 3500–3509. IEEE.

Sunstein, C. R. 2007. *Republic.com 2.0*. Princeton, NJ, USA: Princeton University Press.

Valkenburg, M. P., and Peter, J. 2013. Five challenges for the future of media-effects research. *International Journal of Communication* 7:197–215.

Williams, H. T.; McMurray, J. R.; Kurz, T.; and Lambert, F. H. 2015. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change* 32:126–138.