Article:

Gorrell, G., Bakir, M.E., Temple, L. et al. (5 more authors) (Submitted: 2019) Local media and geo-situated responses to Brexit: a quantitative analysis of Twitter, news and survey data. arXiv. (Submitted)

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Local Media and Geo-situated Responses to Brexit: A Quantitative Analysis of Twitter, News and Survey Data

Genevieve Gorrell, Mehmet E. Bakir, Luke Temple, Diana Maynard
Paolo Cifariello*, Jackie Harrison, J. Miguel Kanai, Kalina Bontcheva

University of Sheffield, UK; *University of Pisa, Italy

Abstract

Societal debates and political outcomes are subject to news and social media influences, which are in turn subject to commercial and other forces. Local press are in decline, creating a "news gap". Research shows a contrary relationship between UK regions' economic dependence on EU membership and their voting in the 2016 UK EU membership referendum, raising questions about local awareness. We draw on a corpus of Twitter data which has been annotated for user location and Brexit vote intent, allowing us to investigate how location, topics of concern and Brexit stance are related. We compare this with a large corpus of articles from local and national news outlets, as well as survey data, finding evidence of a distinctly different focus in local reporting. National press focused more on terrorism and immigration than local press in most areas. Some Twitter users focused on immigration. Local press focused on trade, unemployment, local politics and agriculture. We find that remain voters shared interests more in keeping with local press on a per-region basis.

2016 saw the UK vote to leave the European Union, and Donald Trump elected as president of the USA, and with this came a rising concern about truthfulness in politics and the quality of information people have available to them (e.g. Rose, 2017). A parallel issue relates to the way attention is drawn and shaped (e.g. "attention economy", Harsin, 2015). As media and public figures compete for attention, we may fail to notice the cost in terms of reduced attention for the practical issues that have more impact on voters' lives. When the UK voted to leave the EU on June 23rd 2016, it was unsurprising that voters on the border between Northern Ireland and the Republic of Ireland voted to remain in the EU. The prospect of a customs border separating Irish people from jobs and family is a significant concern for them, and in voting to remain in the EU they voted in their own clear interests. However, in other regions of the UK the relationship between voting behaviour and local interests is less explicable. Cosmopolitan, affluent areas were more likely to vote to remain in the EU, whereas more rural and less affluent areas voted to leave, despite the importance of trade relationships with the EU[1] and the fact that the majority of EU funding received by the UK goes to support agriculture and development, making them net beneficiaries (Los et al. 2017). The case of Sunderland, in the North East of England, perhaps illustrates the apparent contradiction. It voted overwhelmingly to leave the EU, despite the EU's funding of the university and new aquatic centre[2] (Rushton 2017).

Close ties in offline communities increase resistance to propaganda and media-promoted ideas (Seaton 2016), and might have offered some ballast against the more extreme voices in the Brexit campaign. Yet local newspapers have faced numerous challenges over the last two decades due to concentration of media ownership, prioritization of shareholder interests above readers through a reduction in the numbers of journalists in local newsrooms, a decline in traditional advertising revenue which has resulted in cost cutting and redundancies, competition from companies such as Facebook and Google, which offer locally targeted advertising, the growth of social media as sources of local information, and changing news audience consumption patterns (Ramsay and Moore 2016). The steady loss of traditional local print news sources through the closure or mergers of weeklies in local communities, daily newspapers in towns and cities and the demise of regional newspapers has led to "news gaps" (Currah 2009), raising concerns about the means by which communities can stay informed about relevant local affairs, develop community connections or sustain local identities within a geographic proximity (Hess and Waller 2014). The range and scope of local news coverage contributes to its vulnerability, as local news can be both a vital arm of local democracy, but it can also be of poor quality, routine, trite and trivial, "frequently terrible and yet also terribly important" (Nielsen 2015). Local coverage can range from exposing corruption in local public office to the mundane village fete, with relevance and importance coexisting alongside entertainment and a prurient fascination with local crime, human interest and lifestyle stories. Sources of news in the local setting are generally limited to a "news beat" which can lead to routine engagement with a narrow range of local sources such as local government and business and the police (Harrison 2006), missing the

---

[1] https://tinyurl.com/theconversation-brexit-regions

[2] https://www.nytimes.com/2016/06/28/world/europe/european-union-brexit-sunderland-britain-cameron.html

opportunity for individuals to express their views via traditional local media. Despite this decline, many important stories originate from the journalism that takes place at ground level in urban and rural areas and local news also tends to stay with stories once the national papers have left.

In this context, social media are both a problem and a possible solution. Whilst the success of digital local news and ultra-local start-ups have been patchy and audience take-up limited, the importance of local news has been recognized recently both by Facebook and the BBC. Facebook's Community News Project[3] supports local journalism in the UK by enabling the National Council for the Training of Journalists (NCTJ) to oversee the recruitment of eighty trainee journalists on a two year scheme. The 4.5 million investment aims at increasing local news reporting in areas of the UK that have lost their local newspaper. In 2017 the BBC allocated 8m a year to pay for 150 reporters to work for local news organizations rather than the BBC[45] to cover council meetings and public services, both areas increasingly neglected by the local press, and to share these stories with the BBC. A 2018 OFCOM report of audiences use of news media indicates that 23% of news users use printed local or regional newspapers, showing that despite its decline in recent years, local newspapers still remain an important source of independent journalism about local public affairs. There is also evidence that online social networks have a strong local component, providing a new channel for localized discussion (Bastos, Mercea, and Baronchelli 2018).

In this work, we use a large corpus of local and national news articles to investigate the way in which regional interests were covered in the Brexit debate period. We cross-compare this with a corpus of Twitter data in which Brexit is discussed, considering how interests are reflected and media are amplified. In this corpus, the location and Brexit stance of the tweeters have been ascertained, enabling a novel localized, vote-orientated analysis to be performed. We also use survey data to cross-check our findings with a sample of the electorate. These three corpora taken together allow a new comprehensive picture to be formed. Specifically, our work addresses the following research questions:

- What were the main foci in national and local newspaper coverage in the different regions of the UK in the run-up to the referendum, and how did they differ?

- How did newspaper coverage compare with what individuals from different regions were discussing on Twitter? What does this say about media representation of real interests?

- How do interest patterns and media consumption differ between those who favoured leaving the EU and those who favoured remaining in it in different parts of the country? How do they relate to local and national media?

---

[3]https://www.facebook.com/facebookmedia/blog/facebook-launches-community-news-project-in-the-uk

[4]https://www.bbc.co.uk/news/uk-38843461

[5]https://www.nytimes.com/2018/12/09/business/bbc-local-news-partnership.html

Related Research

Social media are increasing in popularity as a way for people to access news (Newman et al. 2017), and this has the potential for widespread impact on global politics. Social media have been widely observed to provide a platform for fringe views (Faris et al. 2017; Silverman 2015; Barberá and Rivero 2015; Preoţiuc-Pietro et al. 2017), and political asymmetries on social media are a further cause for unease. Research in the US political context (Allcott and Gentzkow 2017; Silverman 2015) raises concerns regarding polarized political debates in other countries. Ferrara (Ferrara 2017) presents findings on the anti-Macron disinformation campaign in the run-up to the 2017 French presidential election, and a body of work (Lansdall-Welfare, Dzogang, and Cristianini 2016; Mangold 2016) has begun to explore Brexit.

Bastos et al. (2018) present similar work to ours, in which a sample of Twitter users has been classified for Brexit vote intent and situated geographically. They explore network structure in mentioning and retweeting to explore the relationship between "echo chambers" and geography. They find that remainers often had links to people who were further away from them, whereas leavers tended to be more linked to people who were geographically closer to them. While our Twitter corpora are similar, our work differs from theirs in its focus on local and national media influences. We also include evaluation results for our stance and location classification.

Gorrell et al. (2018) explore the media influences that dominated the Twitter Brexit debate. The most linked newspaper is the Guardian, which reflects the majority remain stance on Twitter. However, an aggressive minority of leavers tweeted more media links in total, most notably to mainstream media such as the Express as well as a plethora of alternative media producing materials attracting a strong leave audience. Upheld press complaints centred particularly on the Express and other leave media, supporting others' research regarding issues with truthfulness in the campaign (Moore and Ramsay 2017). Compared with that work, the novel contribution here is the focus on situatedness, as well as the new insights enabled by topic modelling of news text. Locating the users geographically, as well as enabling investigation of how views relate to location, also makes it more possible to give some indicators about whether the impression created by Twitter materials is plausible as a reflection of the attitudes and media-related behaviours of the populace, or whether it is skewed by, for example, "astroturfing" (campaign accounts posing as voters in order to influence) or other distorting influences.

Brexit media research is highly relevant here. Research undertaken by Loughborough University's Centre for Research in Communication and Culture showed the extent of press bias towards the referendum among main UK outlets, with the Financial Times and The Guardian in strong support to remain in the EU, and The Sun and Daily Mail supporting the leave campaign (Deacon et al. 2016). Additional reports from this centre indicate consistency in the issue agendas: both types of outlets covered referendum conduct, economy/business, and immigration as the three most prominent

issues.[6] Previous analysis[7] has investigated attitudes towards particular Brexit campaign topics on Twitter. For example, issues related to employment were discussed much more frequently by remainers than leavers, while issues related to immigration and democracy were discussed much more frequently by leavers. We add to this through drawing connections with local and national news. Moore and Ramsay contrast tabloid and broadsheet coverage (2017), and Matsuo and Benoit [8] investigate differences in the dialogue between leave and remain camps, but little attention so far has focused on the difference between local and national coverage. This is an important novel contribution of our work.

More broadly, this research is positioned in the context of a global "information malaise", and concerns about the integrity of democracy. In the context of emerging geographical research on Brexit and underlying patterns of economically left-behind areas and the discontent of their socially disadvantaged populations *vis-a-vis* metropolitan elitism (Manley, Jones, and Johnston 2017; Dorling 2016), more research is needed to ascertain the degree to which diverse publics were able to comprehend the regulatory complexities of scale and make full sense of its implications for their situated life chances. Los et al. (2017) provide a comprehensive review, and the same authors provide an illustrative graph[9] relating EU exports from a region with their EU membership attitudes. Such research acknowledges a regional aspect to attitudes, for example around Scottish identity, but according to Manley et al. (2017) geographical variation in Brexit attitudes can be elucidated as an expression of differences in population such as age and qualification. We reserve such an analysis for future work.

Corpora

In this section, we describe the three main data sources used in the work. We begin with the news corpus, before discussing the Twitter corpus and the methods we used to ascertain location and Brexit stance. Thirdly we introduce the survey data we draw on. In the following section we present the topic modelling and entity detection approaches we use throughout the work to profile interests and concerns.

Newspaper Corpus

We collected articles from national and regional newspapers that are available in the Nexis[10] database. The selection of national newspapers was based on those with the highest

circulation value that were available in Nexis, comprising a mix of tabloid and broadsheet papers. We collected articles including any of the keywords "Brexit", "EU", "referendum", or "article 50" in the body of the text that were published between February 20th and June 23rd 2016 (the date of the referendum announcement until the referendum itself). The list of newspapers and the number of matched articles are given in the appendix. Unfortunately, the Express was not available for inclusion and could not be accessed in a way that made it suitable for inclusion. Our corpus is large enough to form a fair representation of the UK media landscape despite this, and media material has not been analyzed on a publication-by-publication basis in this work, so the omission does raise problems for the work reported here.

These data have been used to form two region-sensitive divisions. Firstly, regional newspaper articles can be divided according to region of publication. Secondly, both regional and national articles can be divided according to the location that the article is primarily *about*. We identified the location names mentioned in the articles by matching text to a gazetteer list of UK location names extracted from DBpedia[11], a structured encyclopedia derived from Wikipedia and suitable for machine use. The DBpedia location entities are also assigned additional properties such as the coordinates of the locations. We used the coordinate information to identify the level 1 Nomenclature of Territorial Units for Statistics (NUTS)[12] region of locations mentioned in the articles. In this way, a mention of "Edinburgh", for example, would be associated with the NUTS region "UKM", Scotland. The most frequently occurring NUTS region in each article was assigned to it. Not all articles mentioned a UK location, so some articles were unassigned. The appendix gives article counts per region.

Twitter Corpus

Around 17.5 million tweets were collected from 3rd April until 23 June 2016. The highest daily volume was 2 million tweets on June 23rd (only 3,300 were lost due to Twitter rate limiting), with just over 1.5 million during poll opening times. June 22nd was second highest, with 1.3 million tweets. The 17.5 million tweets were authored by just over 2 million distinct Twitter users (2,016,896). The tweets were collected based on the following keywords and hashtags: *votein, yestoeu, leaveeu, beleave, EU referendum, voteremain, bremain, no2eu, betteroffout, strongerin, euref, betteroffin, eureferendum, yes2eu, voteleave, voteout, notoeu, eureform, ukineu, britainout, brexit, leadnotleave*. These were chosen for being the main hashtags, and are broadly balanced across remain and leave hashtags.

Almost half a million of these users were able to be classified by Brexit vote intent, on the basis of tweets authored by them and identified as being in favour of leaving or remaining in the EU. Partisan hashtags such as "#voteleave" at the end of a tweet quite reliably summarize the tweeter's position with regards to the referendum. The methodology used

---

is described in more detail in Gorrell et al. (2018) The end result is a list of 208,113 leave voters and 270,246 remain voters, classified with an accuracy of 0.966.

Users were allocated to a NUTS1 region on the basis of text in the Twitter location field. This is a free text field that users may fill in in any way they choose, or choose not to fill in. As a result of this, users may ignore the field, repurpose it or use it humorously, so only a limited number of locations could be identified reliably. In total, 162,548 user locations were obtained using the same approach as for the newspaper articles, in which text is matched to a gazetteer of UK location names from DBpedia, and the coordinates given are used to assign a NUTS1 region. In addition, a very small number of Twitter users (0.18% of our sample) had consented to location coordinates being added to their tweets. These are too small in number to make an impact on the size of the dataset, but made it possible to tune and evaluate the work to some degree. The evaluation on a test set of 1016 users with location coordinates achieved a precision of 0.82, a recall of 0.67 and an F1 of 0.74. The actual accuracy is probably a little higher than this, since the coordinates may not be especially more reliable than the location field resolution, as users may have moved around.

Survey Data

In order to add context to this study, we profile both the general Twitter user and those who claim to 'share political information' on the platform by making use of wave 12 data from The British Election Study 2014-2018[13]. The study is managed by a consortium of the University of Manchester, the University of Oxford and the University of Nottingham, and wave 12 was conducted by YouGov between the 5th of May 2017 and the 7th of June 2017. 34,464 respondents participated. The questionnaire covers a broad variety of demographics, as well as politically relevant behaviours and attitudes including social media use.

Methods

Two main natural language processing methods have been used to explore the subject matter in the news and tweet corpora in a quantitative way; a topic view and an entity view.

Topic Modelling

Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003) was used to discover topics discussed in the Twitter and newspaper corpora. In LDA, word frequencies in texts are considered to arise from a weighted mixture of latent topics. Topics are discovered automatically, having manually specified the desired number. Deciding on the number of topics is often done by plotting the coverage provided by the topics against a range of topic numbers; an "elbow" in the graph shows where increasing the number of topics starts to give a diminishing return in covering the data. However, we selected the number of topics heuristically, by trying a few different values and seeing which seemed to capture the

Topic	% of Tokens	Theory
Trade	3.1	Most Localities
Employment	2.6	Most Localities
Local politics	1.0	Most Localities
Steel	0.8	Specific Localities
Agriculture	0.6	Specific Localities
Car pollution	0.5	Specific Localities
Fishing	0.3	Specific Localities
Immigration	1.9	National Interest
Terrorism	1.8	National Interest
Scotland	1.1	National Interest
Northern Ireland/Wales	0.5	National Interest

Table 1: Topics selected for study

most interesting aspects given the research questions of the work.

The topic set was derived from the complete news corpus, as this provided a large enough quantity of material to extract detailed, high quality topics. The topics discovered covered a broad range of relevant subjects, including energy (0.8% of tokens) and the "Queen Backs Brexit" story (0.7% of tokens). Higher token coverage does not necessarily indicate a better topic; the two biggest topics (7.4% and 6.9% of tokens respectively) capture mainly the background language distribution, and it is evident from the examples given that small topics can be very precise. This same set of topics was then applied to Twitter material as well as news subsets as necessary, in order to enable comparisons to be made between the differing datasets.

Of these topics, a number were pre-selected, guided by theory, in order to reduce the possibility of type one errors that would arise in calculating statistical significance of relationships across such a large number of topics. The theory was that local coverage would emphasise practical matters of relevance to local people's daily lives, whereas national coverage would emphasise matters such as national identity. We thus identified a range of topics that were likely to provide a good opportunity to investigate these points, as shown in table 1.

Entities

In the entity view, TAGME[14] (Ferragina and Scaiella 2010) has been used to find mentions of entities in the corpora. An "entity" might be a person, place or organization, or it might be a concept; TAGME matches anything with a Wikipedia page. Any annotation has an associated value $\rho \in [0, 1]$ which estimates the "goodness" of the annotation with respect to the other entities of the input text. We used this value to filter poor annotations, namely those with $\rho < 0.10$ (We used this value as suggested in the documentation [15]).

We then derived the entities associated with a journal as the aggregation of entities found in its articles. Using the correspondence between articles and NUTS1 regions described above, we also derived regional entities as the aggregation of annotations found in articles with same associated

---

region.

In order to effectively extract the "key" topics on both aggregates, we exploited different scoring techniques to rank entities:

- $score(e, A) = articles\_count(e, A) * mean\_\rho(e, A)$
- $score(e, A) = articles\_count(e, A) * max\_\rho(e, A)$
- $score(e, A) = articles\_count(e, A)$

where $articles\_count(e, A)$ is the number of articles in the aggregate A containing entity e, $mean\_\rho(e, A)$ (respectively $max\_\rho(e, A)$) is the mean (respectively max) value of ρ among the occurrences of entity e in A. The final rank is obtained by ordering entities according to their scores. The first two approaches gave similar results, while the last one resulted in a worse performance, perhaps because it does not consider the ρ score associated with entities.

The entity view benefits from TagMe's ability to disambiguate mentions of entities. For example, Theresa May might be referred to as "Mrs. May", or "the UK prime minister". Therefore, counts of mentions can be grouped across different ways of expressing a concept. It also makes it easier to focus on important entities, as only entities salient enough to have a Wikipedia page are extracted. It is like asking closed questions; how do people talk about, for example, Brussels? Trade tariffs? Who talks most about these subjects? The topic view is more like asking open questions; we take our lead from what the texts themselves focus on. This is important as closed questions can miss unexpected trends or more subtle effects. For example, subtleties such as trends toward nationalism in the discourse may be lost in the entity view but might appear as a topic using LDA. It was possible to select entities that matched the pre-selected LDA topics well, enabling various forms of parallel analysis.

Results

We discuss each research question in turn. First we review newspaper coverage, considering both overall emphasis and regional foci, as well as contrasting local and national coverage, and impressions related to particular areas. We then explore reception via the Twitter corpus, considering regional differences. Finally we relate findings on a per-region basis to vote declarations on Twitter as well as referendum outcomes.

RQ1: National and Local Newspaper Coverage

As previously mentioned, topics were derived on the entire newspaper corpus, and used throughout the work. The two largest topics covered background language distributions, enabling other topics to uncover semantically coherent areas in contrast. The third most dominant topic was Brexit itself, which is unsurprising given that the corpus was deliberately comprised of articles mentioning Brexit. The fourth topic is background language relating to the business of government and law, and after that we see topics that give a sense of what the media considered to be relevant issues to Brexit. The economy and the government feature highly, as do aspects of the Brexit campaign, supporting Loughborough's findings (Deacon et al. 2016) about press focus.

Topic	% Tokens	Topic	% Tokens
Trade (UK internat.)	3.1	Refugees (Calais)	1.9
The economy	2.9	Football	1.9
David Cameron	2.8	Business	1.8
Brexit campaign	2.8	Terrorism (Brussels)	1.8
Employment	2.6	UK treasury	1.8
UK politics	2.2	Russia/Ukraine	1.7
Immigration	1.9	EU politics	1.7

Table 2: Most significant topics in newspaper corpus

The top 14 of these are given in table 2. Further background language or unclear topics have been excluded as uninformative.
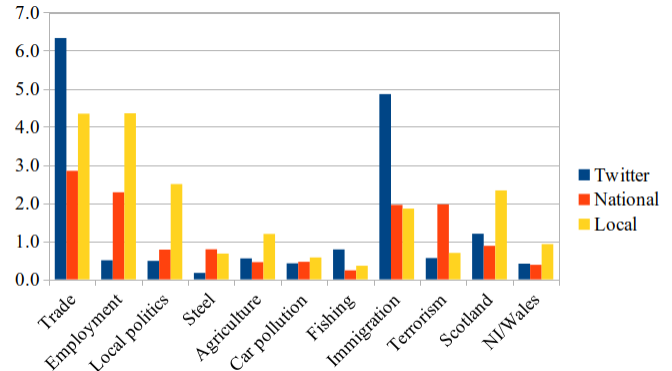

Figure 1: Topics in national and regional press vs. Twitter

In terms of topics, national papers talk about the economy, David Cameron, trade, employment, refugees, terrorism, and immigration. Regional papers discuss employment, trade, football, the economy, UK politics, local politics, and Scotland. Notable differences from national coverage include a greater emphasis on employment, football, local politics and Scotland, and a reduced emphasis on terrorism. Figure 1 shows the extent of representation of our pre-selected topics in national and local news articles, alongside Twitter findings which will be discussed later in the work. All differences between national and regional coverage are significant in a t-test done on a per-article basis (p<0.001) except for steel and car pollution, where the difference is not significant.

Regional Findings In order to explore how local interests and local foci were reflected in the national narrative, we looked at topic representation in the different regions. Regional variation in interests was found. For example, agriculture was mentioned most in local papers in East Anglia. Steel was mentioned most by local papers in Wales, as well as the North East, and most by national papers in conjunction with Wales. Choropleths for all selected topics are available on the project website[16] and examples of trade and Brexit mentioning in local papers are given in figure 2. Darker green shade indicates more topic representation in

---

[16] http://services.gate.ac.uk/politics/ba-brexit/

that region; note that all values are scaled to the strongest topic/region representation, which is Scottish focus on Scotland. Brexit mentioning pattern in local papers suggests that a focus on Brexit may have disposed toward voting leave, a hypothesis that is tested below under research question 3.

Figure 2: Topics mentioned by local papers published in regions

We now compare topics emerging from regional publications with the overall picture from national papers. It was hypothesized that local interests would be more practical, and that employment, trade and local politics would be more widely mentioned in local publications than nationally. Steel, agriculture, car pollution and fishing were expected to be mentioned more than in national publications only in certain regions. Conversely, it was predicted that immigration, terrorism, Scotland and Northern Ireland/Wales would be mentioned more in national publications (other than, in the case of Scotland and Northern Ireland/Wales, in the regions themselves).

Figure 3: Topic emphasis in local vs. national newspapers

It was confirmed that employment was mentioned significantly more in most regions than nationally, as was trade. Agriculture was found to be of more widespread interest than had been predicted, but was indeed more often mentioned in local publications, as shown in figure 3. In these figures, red shades indicate less local coverage, and green, more; depth of shade indicates statistical significance.

Local politics was also generally mentioned more in regional publications (p<0.001), except in London, Wales and Scotland where it was mentioned less (p<0.001), perhaps because they talk about their regional issues differently, and

in Northern Ireland where coverage was not significantly different to national coverage.

Figure 4 shows that terrorism was widely mentioned significantly less in local papers than national. Immigration is somewhat more of a national topic than a local concern, though the picture is mixed. Steel also presents a varied picture, with significantly more coverage in Wales and the North East (p<0.001), and in Yorkshire and the Humber (p<0.005, Sheffield - in South Yorkshire - has a strong steel industry), and significantly less coverage in the South West (p<0.05) and Scotland and Northern Ireland (p<0.001). This is unsurprising, given steel industry location around the country. Car pollution was mentioned more in the East Midlands (p<0.005), Greater London and the South East (p<0.001) and the South West (p<0.01). It was mentioned less in Scotland (p<0.05) and Northern Ireland (p<0.005). Patterns of interest might be seen as reflecting local concerns.
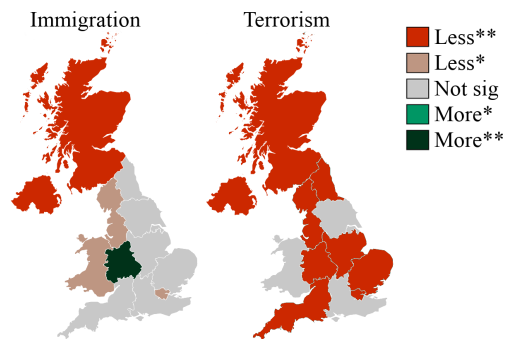

Figure 4: Topic emphasis in local vs. national newspapers; further topics

Topics mentioned less in local coverage are fishing and the regional topics of Scotland and Northern Ireland/Wales, where interest was limited to the regions themselves. Fishing was mentioned generally less in local papers, but significantly so in the East and West Midlands and London (p<0.05) and in Scotland and Northern Ireland (p<0.001 and p<0.005 respectively), despite the fact that Scotland has significant fishing interest and is among the largest sea fishing nations in Europe. Scotland was mentioned generally less (p<0.001) except for in Scotland where it was, unsurprisingly, mentioned more (p<0.001). Similarly Northern Ireland and Wales was mentioned generally less (almost always p<0.001; East of England p<0.05; North East and West not significant) except for in Northern Ireland and Wales where it was mentioned more (p<0.001). All significance testing was performed using two-tailed t-tests.

In summary, findings broadly support our hypothesis that local and national news coverage differs, that local coverage emphasizes different topics, and that regions have their own distinct interests.

As mentioned in the corpora section, national press articles can be grouped according to the region the article is *about*, as ascertained through location mentions in the article. Mentioning regions and the issues that are important

to them is a critical way for national press to show local sensitivity, perhaps even more than by reflecting their interests in the national discourse without mentioning the region. Topic representations in national press articles, divided into regions according to location mentions, were correlated with topics in local press coverage per-region to determine the extent to which emphasis agrees. The strongest correlations were for the topics of Scotland (0.99, p<0.001) and Northern Ireland/Wales (0.96, p<0.001), as might be expected given that the topics would themselves be annotated as location mentions. After that, strong correlations were found for steel (0.83, p<0.001) and immigration (0.71, p<0.001). Local politics correlates significantly (0.79, p<0.01), perhaps because regions with their own topic (Scotland, Northern Ireland and Wales) seem to have a much reduced focus on local politics, so this could be seen as the "negative" of the strong correlations we see for those regional topics.

Aside from the above, we do not see strong correlations for the other topics, which means that national reporting does not show a high degree of reflection of regional differentiation by topic. Agriculture, car pollution and fishing show weak, non-significant correlations, while terrorism, trade and employment show weak *negative* correlations.

Entities A parallel analysis was performed using entity mentions, on a per-publication basis. Matched entities were used for the pre-selected topics, as mentioned above. Entities were associated with regions in a similar manner to that described above for topics; i.e. by associating articles with regions that are mentioned in them (or places in those regions). However in order to maximize the data, since entity mentions are more sparse, the entire corpus was used to associate entities with locations according to their being mentioned in the article, including regional newspaper articles. Observe that the entity "trade" is widely associated with a variety of regions, confirming the observation for topics, as shown in figure 5. Darker shades indicate a higher incidence of entity mentioning; findings are scaled against the maximum score across all regions and selected entities.
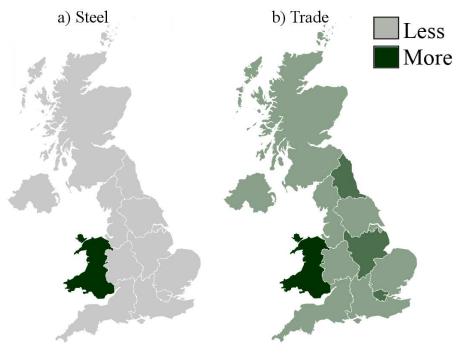

Figure 5: Entities mentioned by national papers in connection with regions

Figure 6 shows example entities discussed in regional papers. We see that "Independence" was an important subject in Scotland. Trade is mentioned mainly in areas that voted

leave, as above. Employment was widely mentioned, suggesting substantiation for our hypothesis that employment was more important in regional papers than national ones. This is further explored in the subsection on research question 3 below. Again, further choropleths can be found on the project website.[17]
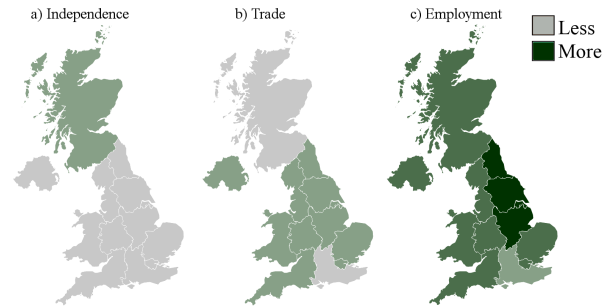

Figure 6: Entities Mentioned by Local Papers in Regions

We now compare local and national papers through the entities lens, in the same way as we did for topics above. For entities, this was done on a per-publication basis to increase reliability (topics were done on a per-article basis). London and East Anglia only have one local paper so it was not possible to calculate statistical significances for them, and generally, entity findings are more subtle. Immigration and terrorism were mentioned generally less in local coverage, as illustrated in figure 7. Scotland was mentioned generally less in local coverage except in Scotland where it was mentioned more and in the North East and Northern Ireland where the difference wasn't significant. Local politics were mentioned generally more in local coverage. Agriculture was mentioned somewhat more in local coverage, significantly so in Yorkshire and the Humber, East Midlands, Scotland and Northern Ireland. The North East mentions employment significantly more in local coverage (p<0.05). No other notable trends emerge; findings for other entities are patchy and generally not significant. Entity findings generally support our predictions and our observations for topics.
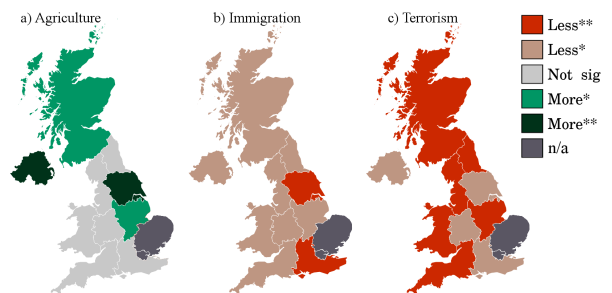

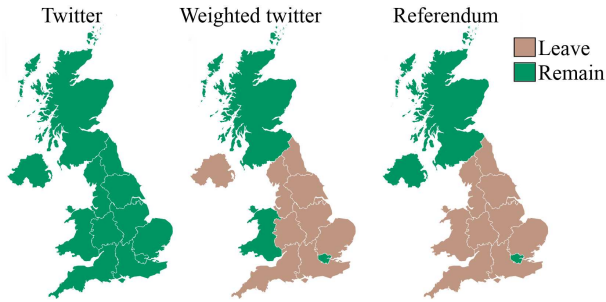Figure 7: Entity emphasis in local vs. national newspapers

RQ2: Twitter Data

Figure 8: Leave/remain balance in Twitter corpus

	Twitter users (1) vs Non-users (0)	Political sharers on Twitter (1) vs Tw. Users (0)
Age	-0.03***	-0.01*
Gender (ref = Male)	-0.11*	-0.26**
Attended University	0.20***	0.13
Class (ref = D/E)		
- C2	-0.16	-0.70**
- C1	0.05	-0.11
- A/B	0.52	-0.19
Retired	-0.30***	0.28
Ethnic Minority	-0.10	0.21
Daily Internet Use (ref = <30mins)		
- Between Half hour and Hour	0.70***	0.80**
- An hour or more	1.13***	1.68***
Ideology (Left 0 10 Right)	-0.79***	-0.18***
Attention to Politics:		
(None 0 10 Great deal)	0.07***	0.41***
Pseudo-r2	**0.12**	**0.17**
n	**11,995**	**3,244**

Table 3: Demographic Characteristics of Twitter Users: Logistic Regression Coefficients

Research profiling Twitter users can be broadly categorized into two strands. The first gleans demographic data from users on the platform itself (Barberá and Rivero 2015; Li, Goodchild, and Xu 2013) whilst the second examines survey data (Mellon and Prosser 2017; Greenwood, Perrin, and Duggan 2016; Duggan 2015). Findings across both approaches suggest Twitter users are likely to be younger, have higher educational qualifications, be more politically engaged, and live in urban areas. Further findings point towards a left-leaning bias on the platform, and a higher prevalence of male users. Twitter activity tends to mirror mainstream politics, spiking at times of heightened political interest, such as campaign debates or scandals (Jungherr 2014; Larsson and Moe 2014). Overall, Twitter is not representative of the wider population and appears to generally reflect existing inequalities in political participation, rather than act as a mechanism to flatten them (Kalogeropoulos et al. 2017).

Figure 8 shows the balance of remainers vs. leavers in the Twitter sample. Across the country, Twitter users tend to be remainers. However, having weighted the sample to bring leavers up to the 52% that actually voted to leave the EU in the referendum, we can see that on a per-region basis our sample resembles the actual referendum outcome. Quantitatively, our weighted sample reflects the referendum outcome per region with a RMSE of 3% and a correlation of 0.89. However, our sample, even having been weighted, over-represents remainers in Wales, and following the weighting, under-represents them in Northern Ireland.

Our finding that Twitter is biased toward remain fits with the picture created by previous research about the Twitter user population (Barberá and Rivero 2015), as well as results from the survey data in Table 3, which indicate a political left bias in keeping with the remain position. Looking specifically at users who share political information in comparison to general users, education and retirement become statistically insignificant, and the left-right result weakens. However, other predictors increase in strength, suggesting gender, attention to politics, and time spent online are key predictors over and above those which predict being on the platform in the first instance. Interestingly, those categorized as C2 class (skilled working class individuals such as mechanics) were significantly less likely to share political information (-0.70, p<0.05), and have been suggested to be one of the more pro-Brexit sections of society (Skinner and Gottfried (2017), though Antonucci et al. (2017) claim that the main leave voters were the "squeezed middle" (those with intermediate levels of education such as A levels). In the table, p<0.01 is signified by "***", p<0.05 by "**" and p<0.1 by "*". The survey data also reveals that Twitter users are more likely to be found in London than all other regions, except for East of England and Scotland which are not significantly different to the capital. This spatial variation is not found for the comparison between general users and those who share politics.

URLs found in the tweets were expanded from shortened forms often used on Twitter, possibly following a number of redirects, to result in a target URL. The most common newspaper/media web domains were then counted. Figure 9 shows the location from which the media link-containing tweets originated, normalized by population size of that region. In the figure, four shades indicate where the regions lie on a scale from zero to the most vocal segment (which is London leave-voters). Even after controlling for population size, London still dominates in terms of linking activity in the Brexit Twitter conversation, originating almost twice as many links per capita as any other region among leavers, more than five times as many as Northern Ireland leavers, and around four times as many links at least from remainers as any other region, a result that echoes the survey findings above with regards to location of Twitter users. Figure 9 however illustrates that the leave campaign on Twitter had markedly more engagement in other regions of England than the remain campaign, a finding that still holds for tweet count instead of link count. Further choropleths can be viewed on the project website.[18]

Overall, the Guardian was the most linked paper across regions/voters (60,472 links, of which 67% were from remainers), followed by the Express (56,652 links, of which 99% were from leavers), the BBC (47,577 links of which 59% were from leavers), the Telegraph (36,729 links, of

---

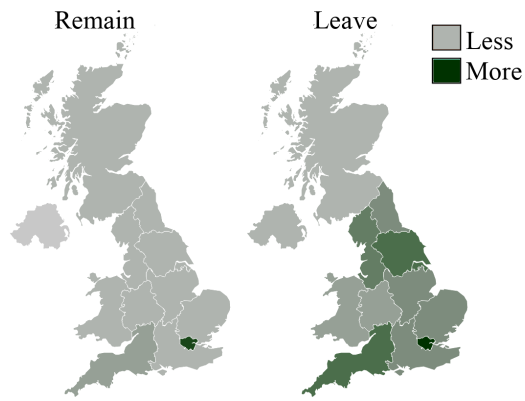[18] http://services.gate.ac.uk/politics/ba-brexit/

Figure 9: URLs (in tweets) per capita

which 83% were from leavers), the Independent (25,645 links, of which 58% were from remainers) and the Daily Mail (23,633 links, of which 91% were from leavers). In most regions, the Express is the most popular link target by some margin. In the north east and south east however the Guardian edges ahead slightly. In London, Wales and Scotland, Guardian is the most popular by some margin. Across the regions, most linked sites are uniformly the Express, the Guardian, the BBC, the Telegraph, the Daily Mail and the Independent in slightly varying orders. Links to local papers are much fewer; the Yorkshire Post received 2,388 links (of which 88% were from remainers) and the Herald, 980 links (of which 54% were from leavers).

Comparing these figures with the readership figures in the survey data reveals an interesting picture, as shown in figure 10. The media that attract the most links on Twitter are not at all the papers that have the highest circulation. The two main online influences of the Guardian and the Express are both attracting proportionally smaller numbers of people who actually state that they are readers, with the disparity in the case of the Express being particularly great. The Daily Mail, on the other hand, is reaching a much wider audience than its Twitter linking figures would suggest. This raises questions regarding the social meaning of linking to a newspaper on Twitter.
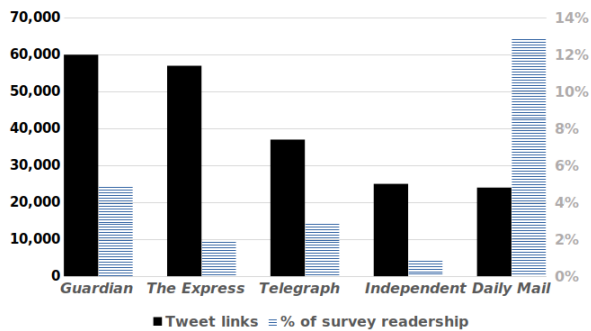

Figure 10: Tweets links from each paper analyzed in study and % of survey readership who read paper

Subjects of general interest in the Twitter corpus (which we recall is selected according to Brexit-related hashtags so can be expected to illustrate subjects mentioned in conjunction with Brexit) include trade, immigration and football. Subjects that attracted more interest on Twitter than in the press as a whole include for example the murder of Jo Cox, the National Health Service (NHS), the rumour that Queen Elizabeth II supports Brexit, and the former UK Independence Party leader Nigel Farage (all $p<0.001$).

We saw in figure 1 that over our selection of topics, Twitter interest levels differ from those in the local and national press. Tweeters are more interested in trade, immigration and fishing than either local or national press. They are less interested in employment, local politics, steel, car pollution and terrorism. Tweeters lie in a middle ground between national (low) and local (high) levels of interest in Scotland, Northern Ireland/Wales and agriculture.

However, due to skewed distributions in some cases, t-tests give different results from the differences suggested by percentage of tokens covered in the corpus by that topic. The most notable example is that in a t-test on a per-article/per-tweeter basis, there is no significant difference between Twitter and national news interest in immigration, despite the apparent Twitter spike. Twitter immigration data has an odd distribution, with an unusual "bulge" of accounts showing medium to high levels of interest in immigration. This means that the histogram in figure 1 shows an elevated score for Twitter immigration that doesn't affect the t-test.

In all other cases the difference in a t-test is statistically significant ($p<0.001$) except for: car pollution, where Twitter does not differ significantly from either local or national news coverage; terrorism, where Twitter and local coverage do not differ significantly; and Twitter and national coverage of Scotland, where the significance level is lower ($p<0.005$). Note however that due to skewed distributions, in some cases the relationships are not as suggested by the histogram, though aside from immigration the differences are small and not especially suggestive. These anomalies are in local politics (national<Twitter), steel (national<local), car pollution (national<Twitter), terrorism (local<Twitter) and Scotland (Twitter<national).

On a per-region basis, the only topics to show a strong correlation between local newspaper reporting and Twitter user focus are the strongly situated topics of Scotland, Northern Ireland/Wales and steel; that is to say, the extent to which these topics were discussed by tweeters in the different regions correlated significantly with the extent to which they were discussed by local papers in those regions, or by national papers in conjunction with those regions ($p<0.001$ for both leavers and remainers, except for steel/leavers where $p<0.01$). Local press showed a sensitivity to local interest levels in fishing ($p<0.05$). Interestingly, employment was reported in the national press in association with those regions where remainers *weren't* talking about it ($p<0.05$), or in fact leavers particularly. The correlation is positive for local press, though not statistically significant.

In summary, Twitter data shows a number of anomalies. It arises from a biased population relative to the general UK electorate, being younger, male, educated, politically

engaged and urban. It shows distortions in media popularity relative to survey readership data, for example in linking the Express far more than the Daily Mail. It shows an unusual level of interest in immigration, with an unusual distribution. However, Twitter users show less interest in terrorism than the national press, and more interest in trade, agriculture and fishing, suggesting that some Twitter use may reflect local concerns.

RQ3: How Leavers and Remainers Differ

In light of the above, we now consider the difference between leave and remain foci and the extent to which each found a voice in the press. Immigration was more discussed by Twitter's leavers than remainers ($p<0.001$). Refugees, terrorism and fishing were more discussed by leavers (all $p<0.001$). Northern Ireland was discussed more by remainers ($p<0.001$). The economy, employment, local politics and steel were more discussed by remainers on Twitter (all $p<0.001$). Scotland was equally discussed by leavers and remainers.

Whilst strong regional correlations were found in the extent of interest from newspapers (local and national) and on Twitter for the topics of the regions themselves (Scotland etc.) and for steel as described above, per-region correlations for other topics were more subtle, but in some cases present for either leavers or remainers. Coverage of local politics per-region correlated with the extent of remainers' interest as determined from the Twitter data (national press, $p<0.05$, regional $p<0.01$), and coverage of car pollution in the national press correlated with the extent of discussion by Twitter leavers in those regions ($p<0.05$).

Recall that figure 2 above shows the pattern of mentioning trade and Brexit in different regions. Trade was found to be correlated with voting leave per-region (again, $p<0.05$). The relationship between talking about trade and voting leave may reflect the initial incongruence that motivated the work; areas most dependent on EU trade voted to leave. Mentioning Scotland in local newspapers correlates with voting remain ($p<0.05$). No other topic among our selection showed a significant correlation, though the regional correlation between mentioning immigration and voting leave was close to significant.

A selection of newspapers was evaluated to assess whether linking to newspapers on a per-region basis correlates with leave stance and/or with remain stance. As a proportion of total links from a particular region, correlating between linking a particular paper and voting on a per-region basis gives predictable results. Linking to the Express correlates with voting leave (-0.63, $p<0.05$), linking to the Guardian correlates with voting remain (0.74, $p<0.01$) and linking the Daily Mail correlates with voting leave (-0.74, $p<0.01$). Survey readership statistics were examined in a similar manner. Percentage of respondents who state that they read each of the following papers was ascertained, as shown in figure 10; the Telegraph, the Independent, the Daily Mail, the Express and the Guardian. This number alone has little meaning however for our purposes, as a 2% readership for paper X could be high in one area but low in another where people do not tend to read a national paper so

much. Therefore we totalled the readership percentages together across the five papers, per region, to give an indication of national paper readership in that area, and then took the percentage for each paper as a proportion of that. This produced interesting results. While, as we saw above, linking to the Express is a strong leave indicator, for readership the correlation is insignificant (-0.03). For the Guardian, however, the correlation is much the same as for linking (0.71, $p<0.01$). For the Daily Mail, the correlation between vote and survey readership becomes stronger (-0.80, $p<0.002$). This might suggest a geographical disconnect between the people who are linking the Express and the people who are reading it, to an extent not found in the other newspapers looked at. Recall that the Express readership is actually low, so a high Express readership compared with other regions is not incompatible with an overall remain inclination in that region. Daily Mail readership is a better indicator of a region's Brexit feelings.

Per-region correlations of local and national media topic coverage against Twitter topic interest were performed, and the resulting baskets of correlations (local vs national) were compared in a t-test. This showed that, across the eleven topics, congruence with local press is significantly greater for remainers ($p<0.05$). This provides evidence in support of the hypothesis that local awareness (or perhaps simply resistance to the national narrative) is connected to a more positive attitude toward EU membership.

Discussion

We have presented findings from a corpus of Brexit-related tweets classified for user location and Brexit vote intent, alongside a large UK news corpus of local and national Brexit-related articles from the referendum period and a sample of survey data. We used topic modelling to ascertain differences in focus between national and local press, and how each relates to topics emerging from the Twitter data. We looked for evidence of regional sensitivity in local and national reporting, or lack thereof, and the relationship between topic focus, regional sensitivity and Brexit stance.

We find that national press emphasizes terrorism much more than local press, and immigration significantly more. Local press emphasizes trade, employment, local politics, agriculture, car pollution, fishing, Scotland and Northern Ireland/Wales. We suggest that local press appear to place greater emphasis on a range of practically relevant issues. National press show a moderate awareness of the issues affecting regions but not a high degree of sensitivity.

Twitter users show a particular interest in trade, which is associated with those wishing to leave the EU. Twitter material appears to show a high level of focus on immigration, but on inspection this is shown to be unrepresentative of Twitter users in general. The unusual distribution of users' interest levels in immigration, with a large "bulge" showing medium to high levels, might be suggestive of campaign activity.

Newspapers popularly linked from tweets show a very different profile from newspapers that survey respondents say that they read. The Daily Mail was most popular among survey respondents, but on Twitter the Guardian and the Express were most popular, demonstrating that their appeal on

that medium is amplified, perhaps in part due to the demographics of Twitter, but maybe also due to a difference in how information is consumed on Twitter, creating a market for a certain type of story. The Express in particular is anomalous in that although those linking to it are almost unanimously leave voters, the number of survey respondents claiming to read the paper per region does not correlate with the overall voting pattern in that region (regions with more Express readers do not have more leave voters). This might suggest that an Express article serves a particular purpose online that it does not serve offline, and this might also be connected to the unusual Twitter activity around immigration–the Express produced an abundance of anti-immigration articles (Ramsay and Moore 2016). The extent of a region's Daily Mail readership is a much better indicator of its likely Brexit stance, and indeed the lack of correlation between Express readership and regional Brexit vote might be explained by the possibility that generally speaking, the country's leavers find the Daily Mail more reflective of their attitudes.

We also find some evidence that remain voters are more aligned with the local press in terms of topic profile than leave voters. This could be seen as supporting Seaton's (2016) suggestion that local press are doing an important job in increasing resistance to "propaganda and media-promoted ideas", and in keeping with Faris et al's (2017) proposition of network propaganda in that to some extent Twitter sharers and some national press co-operated on an immigration and terrorism-focused narrative that didn't cover to the same extent as local press a range of practical issues relevant to people's lives.

Acknowledgments

References

[2017] Allcott, H., and Gentzkow, M. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31(2):211–36.

[2017] Antonucci, L.; Horvath, L.; and Krouwel, A. 2017. Brexit was not the voice of the working class nor of the uneducated-it was of the squeezed middle. *LSE Brexit*.

[2015] Barberá, P., and Rivero, G. 2015. Understanding the political representativeness of Twitter users. *Social Science Computer Review* 33(6):712–729.

[2018] Bastos, M.; Mercea, D.; and Baronchelli, A. 2018. The geographic embedding of online echo chambers: Evidence from the Brexit campaign. *PLoS ONE* 11(13).

[2003] Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3(Jan):993–1022.

[2009] Currah, A. 2009. *Navigating the crisis in local and regional news: A critical review of solutions.* University of Oxford, Reuters Institute for the Study of Journalism.

[2016] Deacon, D.; Wring, D.; Harmer, E.; Downey, J.; and Stanyer, J. 2016. Hard evidence: analysis shows extent of press bias towards Brexit.

[2016] Dorling, D. 2016. Brexit: the decision of a divided country.

[2015] Duggan, M. 2015. Mobile messaging and social media 2015. *Pew Research Center* 19:2015.

[2017] Faris, R.; Roberts, H.; Etling, B.; Bourassa, N.; Zuckerman, E.; and Benkler, Y. 2017. Partisanship, propaganda, and disinformation: Online media and the 2016 US presidential election. *Berkman Klein Center for Internet & Society Research Paper*.

[2010] Ferragina, P., and Scaiella, U. 2010. TAGME: on-the-fly annotation of short text fragments (by wikipedia entities). In *Proceedings of the 19th ACM Conference on Information and Knowledge Management, CIKM 2010, Toronto, Ontario, Canada, October 26-30, 2010*, 1625–1628.

[2017] Ferrara, E. 2017. Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday* 22(8).

[2018] Gorrell, G.; Roberts, I.; Greenwood, M. A.; Bakir, M. E.; Iavarone, B.; and Bontcheva, K. 2018. Quantifying media influence and partisan attention on twitter during the UK EU referendum. In Staab, S.; Koltsova, O.; and Ignatov, D. I., eds., *Social Informatics*, 274–290. Cham: Springer International Publishing.

[2016] Greenwood, S.; Perrin, A.; and Duggan, M. 2016. Social media update 2016. *Pew Research Center* 11:83.

[2006] Harrison, J. 2006. News. abingdon. *MA/Oxon, England: Routledge*.

[2015] Harsin, J. 2015. Regimes of posttruth, postpolitics, and attention economies. *Communication, Culture & Critique* 8(2):327–333.

[2014] Hess, K., and Waller, L. 2014. Geo-social journalism: Reorienting the study of small commercial newspapers in a digital environment. *Journalism Practice* 8(2):121–136.

[2014] Jungherr, A. 2014. Twitter in politics: a comprehensive literature review.

[2017] Kalogeropoulos, A.; Negredo, S.; Picone, I.; and Nielsen, R. K. 2017. Who shares and comments on news?: A cross-national comparative analysis of online and social media participation. *Social Media+ Society* 3(4):2056305117735754.

[2016] Lansdall-Welfare, T.; Dzogang, F.; and Cristianini, N. 2016. Change-point analysis of the public mood in UK Twitter during the Brexit referendum. In *Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference on*, 434–439. IEEE.

[2014] Larsson, A. O., and Moe, H. 2014. Triumph of the underdogs? comparing Twitter use by political actors during two Norwegian election campaigns. *Sage Open* 4(4):2158244014559015.

[2013] Li, L.; Goodchild, M. F.; and Xu, B. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and geographic information science* 40(2):61–77.

[2017] Los, B.; McCann, P.; Springford, J.; and Thissen, M. 2017. The mismatch between local voting and the local economic consequences of Brexit. *Regional Studies* 51(5):786–799.

[2016] Mangold, L. 2016. Should I stay or should I go: Clash of opinions in the Brexit Twitter debate. *Computing* 1(4.1).

[2017] Manley, D.; Jones, K.; and Johnston, R. 2017. The geography of Brexit–what geography? modelling and predicting the outcome across 380 local authorities. *Local Economy* 32(3):183–203.

[2017] Mellon, J., and Prosser, C. 2017. Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics* 4(3):2053168017720008.

[2017] Moore, M., and Ramsay, G. 2017. *UK media coverage of the 2016 EU referendum campaign*. King's College London.

[2017] Newman, N.; Fletcher, R.; Kalogeropoulos, A.; Levy, D. A.; and Nielsen, R. K. 2017. Reuters institute digital news report 2017.

[2015] Nielsen, R. K. 2015. *Local journalism: The decline of newspapers and the rise of digital media*. IB Tauris.

[2017] Preoţiuc-Pietro, D.; Liu, Y.; Hopkins, D.; and Ungar, L. 2017. Beyond binary labels: political ideology prediction of Twitter users. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, 729–740.

[2016] Ramsay, G., and Moore, M. 2016. Monopolising local news. *Centre for the Study of Media, Communication and Power, Kings College London*.

[2017] Rose, J. 2017. Brexit, trump, and post-truth politics.

[2017] Rushton, P. 2017. The myth and reality of Brexit city: Sunderland and the 2016 referendum.

[2016] Seaton, J. 2016. Brexit and the media. *The Political Quarterly* 87(3):333–337.

[2015] Silverman, C. 2015. Lies, damn lies and viral content. Technical report, Tow Center for Digital Journalism.

[2017] Skinner, G., and Gottfried, G. 2017. How Britain voted in the 2016 EU referendum.

Appendix

Newspaper	Region	Articles
The Northern Echo	UKC	405
Sunderland Echo	UKC	113
Manchester Evening News	UKD	169
Liverpool Echo	UKD	321
Lancashire Evening Post	UKD	135
The Bolton News	UKD	106
Hull Daily Mail	UKE	180
Yorkshire Post	UKE	873
The Star (Sheffield)	UKE	290
Yorkshire Evening Post	UKE	192
York Press	UKE	358
Bradford Telegraph and Argus	UKE	67
Leicester Mercury	UKF	248
Derby Telegraph	UKF	313
Nottingham Post	UKF	219
Stoke The Sentinel	UKG	415
Birmingham Evening Mail	UKG	362
Coventry Evening Telegraph	UKG	225
Eastern Daily Press	UKH	336
The Evening Standard (London)	UKI	1571
Oxford Mail	UKJ	207
Hampshire Chronicle	UKJ	24
Bristol Post	UKK	218
The Plymouth Herald	UKK	204
Bournemouth Echo	UKK	258
Daily Post (North Wales)	UKL	619
South Wales Evening Post	UKL	269
South Wales Echo	UKL	267
Aberdeen Press and Journal	UKM	509
Herald (Glasgow)	UKM	1524
Aberdeen Evening Express	UKM	151
Evening Times (Glasgow)	UKM	154
Edinburgh Evening News	UKM	190
Belfast Telegraph	UKN	665
Irish News	UKN	657
Total		12814

Table 4: Regional Newspapers Article Numbers

NUTS1	NUTS1 Name	Papers
UKC	North East (England)	2
UKD	North West (England)	4
UKE	Yorkshire and The Humber	6
UKF	East Midlands (England)	3
UKG	West Midlands (England)	3
UKH	East of England	1
UKI	London	1
UKJ	South East (England)	2
UKK	South West (England)	3
UKL	Wales	3
UKM	Scotland	5
UKN	Northern Ireland	2
Total		35

Table 5: Regional Newspaper Counts

Newspaper	Articles
Daily Mail and Mail on Sunday	2041
Daily Star Daily Star Sunday	328
Financial Times (London)	2464
The Daily Telegraph (London)	2719
The Guardian(London)	9709
The Independent (United Kingdom)	6833
The Mirror and The Sunday Mirror	1353
The New Statesman	116
The Observer(London)	587
The Spectator	181
The Sun (England)	3050
The Sunday Telegraph (London)	542
The Sunday Times (London)	1336
The Times (London)	4070
Total	35329

Table 6: National Newspapers Article Numbers

NUTS1	NUTS1 Name	Articles
UKC	North East (England)	313
UKD	North West (England)	485
UKE	Yorkshire and The Humber	1384
UKF	East Midlands (England)	648
UKG	West Midlands (England)	548
UKH	East of England	265
UKI	London	2464
UKJ	South East (England)	368
UKK	South West (England)	584
UKL	Wales	649
UKM	Scotland	1104
UKN	Northern Ireland	180
Total		8992

Table 7: Regional Mention Article Counts