Facebook Advertising the 2017 United Kingdom General Election: The Uses and Limits of User-Generated Data

Abstract

Despite a growing focus on Facebook advertising in recent elections around the world, there are no large-scale empirical studies analysing these adverts and how they are targeted. This article assesses a new method to tackle this challenge. Working with data collected by the social enterprise Who Targets Me, where volunteers installed a browser plug-in on their computers, we harvested Facebook adverts from users' timelines in order to identify adverts purchased by political parties in the 2017 United Kingdom General Election. In this article, we analyse the resulting dataset, looking at both the content of adverts and who they were targeted at. In addition, we then reflect on the limitations of this type of volunteer driven data-gathering and how methods for researching Facebook political advertising might be improved in the future.

Keywords

Advertising, Facebook, Targeting, Political Communication, United Kingdom, Social Media

Facebook Advertising the 2017 United Kingdom General Election: The Uses and Limits of User-Generated Data

In a number of recent elections in various parts of the world, there has been extensive discussion of the potential dangers that Facebook and other social media platforms may pose to democracy. These concerns have taken various forms, but include renewed criticisms of how digital platforms' algorithmic systems unintentionally create polarizing echo chambers (Sunstein, 2017), allow the malicious circulation of so-called "fake news" (Groshek & Koc-Michalska, 2017) and create possibilities for voter manipulation (Gorton, 2016). Together, it has been argued (Persily, 2017), these developments may deepen political polarization, facilitate misinformation and, as a consequence, foment populism.

While the bases for these concerns are genuine, they are difficult to investigate empirically. In relation to political marketing, in particular, researching Facebook adverts poses significant challenges, as the company does not grant access to its data and adverts are both personalized and ephemeral, appearing briefly on individual users' timelines and then disappearing. This lack of transparency undermines academic research and also efforts to (re)establish transparency in the relationships between political parties and candidates, voters, and interested third parties (Moore, 2016).

This article represent a first attempt to tackle this challenge of increasing transparency in electoral communications.[i] By examining data collected through a browser plug-in created by the social enterprise Who Targets Me (WTM), installed on the computers of more than 11,000 volunteers, it provides empirical insights into Facebook advertising by major political parties in the 2017 United Kingdom General Election as it was experienced by tens of thousands of WTM-cooperating users. We were able to harvest the adverts that appeared on users' timelines, creating a vast – but not statistically representative – repository of political

advert content and data on who saw them (further details of the dataset obtained through this method are provided below).

The aim of this paper is twofold. First, we analyse two datasets constructed through the Who Targets Me plug-in during the 2017 election campaign. The first dataset contains unique adverts placed on Facebook by major UK political parties. This allows us to understand the content and messaging of individual parties and how they compare. The second dataset looks at advert impressions (i.e. the appearances they make on individual users' timelines). We use this dataset to offer some insights into the way messages were targeted at particular voters during the campaign. This allows us to tentatively address some of the concerns about the role that Facebook advertising plays in election campaigning, in terms of WTM volunteers' experience of the social media platform in the months leading up to the election. On the basis of this data, we are able to make a number of observations, including some that suggest fears about Facebook targeting could be overstated. We find no evidence, for example, that campaigns were seeking to send highly targeted but contradictory messages to would-be supporters, with the dominant topics on Facebook adverts reflecting the national campaigns parties were running.

Second, we consider the limitations of our methodological approach and provide suggestions as to how it could be extended in future to provide data representative of the broader national voting population. In so doing, we aim to map out a future agenda for improving the methodological approaches underlying investigations into political communication through social media. We consider this to be a valuable contribution towards the development of approaches that can enable researchers to 'look into' the landscape of social media communications in ways that will allow rigorous, generalizable claims to be

made about the nature of political communication on social media at national and supranational scale.

Existing Research on Facebook in Elections and Research Questions

Facebook is a multifaceted platform that has been used in a variety of ways in election campaigns. Candidates have set up profiles in order to create groups of supporters and communicate with them directly.[ii] Political entrepreneurs not formally affiliated to official campaigns have set up pages to galvanise support of particular politicians. These strategies have been increasingly complemented by a large rise in campaign spending on social media (particularly Facebook) advertising in Europe and the US (Tambini & Labo, 2018). It is this latter area activity that this paper focuses on. While the dynamics of user-generated content and interactions occurring on Facebook have been widely researched, the role played by Facebook as an advertising platform is less well understood.

One approach to understanding the role that Facebook advertising plays in election campaigning is to understand it as part of a broader pattern of targeted political communication based on the segmentation and profiling of electorates based on large scale databases of proprietary information. Recent elections (especially, but not limited to, those held in the United States) have seen significant advances in the ability of campaigns to effectively target members of the electorate through sophisticated tools and services offered by companies such as Cambridge Analytica. Most significantly, this development has involved a shift from segmented targeting to modelling the attributes of individuals and choosing to target them with particular messages, or ignore them all together (Kreiss, 2016; Nielsen, 2012).

While there remains a debate about the effectiveness of data-driven politics (Baldwin-Philippi, 2017; Hersh, 2015), a growing number of researchers and commentators are

concerned about how data-driven political campaigning and message targeting on social media might erode crucial democratic values and good practices. Their concerns include potential violations of privacy (Howard, 2006; Kreiss & Howard, 2010; Cohen 2012; Barocas 2012); a lack of transparency (Kreiss & Howard, 2010; Rowbottom, 2012); and the ability of existing electoral law to maintain a level playing field and thus ensure the legitimacy of elections (Tambini & Labo, 2016; Butrymowicz, 2009). Researchers have also raised concerns as to whether opacity, and expectations of low accountability by political actors, could lead to the increase of divisive negative messages (Barocas, 2012); on 'political redlining', i.e. the ability to direct messaging towards a narrow segment of the electorate and simultaneously purposefully exclude others who are less likely to vote or do not belong to key swing demographics (Hersh, 2015). More broadly, the use of social media in political campaigning has been linked to a general tendency towards more turbulent political environment (Margetts, John, Hale, & Yasseri, 2015). Finally, there exists a concern that the quality of deliberation may be being undermined, and that a diversity of communication during elections is being gradually eroded in favour of targeted messaging otherwise invisible to the public at large.

One concern that links these various claims is the notion that effective targeting may undermine voter autonomy: those voters for whom social media is the dominant source of news and information could theoretically be inundated with a constant stream of political messaging specifically designed for effective manipulation of that person based on their data trail, and able to drown out opposing views and constitute a new powerful propaganda. Many of these concerns came into sharper focus after the shock results of the UK referendum on EU membership and the US presidential elections, both in 2016. A number of commentators started to talk about the risks that social media posed to the democratic process (notably

Cadwalladr, 2017), while others remained sceptical of what they see as "alarmist accounts" that overstate the significance of targeted communications (Tworek, 2017).

At the time of writing, two separate regulatory investigations into the use of targeting during the 2016 UK Brexit Referendum were ongoing: an investigation by the Information Commissioner's office (the UK regulator for freedom of information and data protection) into the use of data for campaign purposes (Denham, 2017); and a separate investigation by the UK electoral supervisor into potential breaches of campaign funding rules, reporting obligations relating to provision of database and targeting services. Whilst international agencies such as the Venice Commission and the OSCE have been relatively slow to respond to the challenge of more sophisticated use of new media, the Council of Europe has carried out a feasibility study for a new recommendation as to how democracies ought to regulate these new practices.

Despite this gathering storm of academic and public debate, there has been a lack of robust, disinterested accounts of how such campaigns actually work in practice. Out of necessity, research into data driven campaigning has tended to rely on interviews (Nick Anstead, 2016; Kreiss, 2016; Moore, 2016), ethnography (Nielsen, 2012) or legal analysis (Butrymowicz, 2009). Despite the valuable insights that such research provides, there has been little analysis of the messages themselves, or how they are targeted. As a result, it is difficult to assess worrying claims that such messaging constitutes new forms of propaganda. The key proposals of the theoretical literature – namely that the legitimacy of elections and referenda may be undermined by these new campaigning tools – have not been effectively tested, and there remains a significant gap between hype (generally of the dystopian variety) and understanding how targeted campaigning on social media has actually been deployed.

To add empirical flesh to these bones, this article seeks to better understand how Facebook was used as a political advertising platform during the 2017 UK general election. In particular, drawing on the questions raised in existing literature and respecting the limitations of our data, we seek to address the following research questions:

1. **What messages were used in the Facebook advertising campaigns of the major political parties?** In particular, what topics were parties focusing on and how did this compare across parties and is there any evidence of messages being tailored in contradictory way for different audiences?

2. **Were these messages more likely to be (i) negative or (ii) personality / leader focused in some parties than others?** The rise of negative and personality-based campaigning has been one of the most significant trends in political communication in recent decades. To what extent does Facebook advertising follow similar patterns to more traditional modes of communication?

3. **Were parties using Facebook to mobilise votes to engage in political actions?** One of the defences that has been made of social media, especially when it is attacked for promoting so-called "Slacktivist" forms of politics, is that it can provide a platform for promoting other, higher threshold forms of political activity, such as donating to political campaigns or mobilising supporters to engage in door-to-door electioneering.

4. **Did voters in marginal constituencies see more political ads on Facebook?** One risk posed by highly targeted advertising is that it might exacerbate existing institutional tensions in election systems. This might especially be true in an election system like the UK's first-past-the-post, where there are incentives for parties to target their campaign messages at the narrow segment of the electorate most vital to their success (i.e. the most persuadable voters in the a few, strategically important marginal seats). Is there any evidence that this is occurring?

Data and Methods

To address these research questions, we explore a dataset gathered by the social enterprise Who Targets Me. The project was created in 2017, in the wake of the Brexit referendum, by two British activists to "monitor the use of dark advertising in our elections" (Who Targets Me, 2018). Because of the lack of publicly available data on Facebook advertising, researchers face substantial difficulties in reaching reliable and valid conclusions in investigating the flows of advertising and other content on the platform and there is significant public debate about the implications of this for democratic communication and trust in elections. One response to this "information gap" has been the development of voluntary, civil society-led, projects that attempt to reverse engineer Facebook's platform in order to generate datasets that open up the 'black box' (Pasquale 2015) of the advertising process. These include for instance investigations by non-profit journalistic organization ProPublica (Angwin and Larson, 2017), which also employed a plug-in.

A Who Targets Me browser plug-in was installed by a total of 11,421 people in England, Scotland and Wales over the course of the 2017 election campaign. When volunteering to install the plug-in, users were asked to give active consent to the data-gathering process. Additionally, they were asked for three pieces of information which could be appended to the data collected from their browsing: their age, gender and the postcode for where they lived (which could be converted into their parliamentary constituency).

The plug-in successfully gathered millions of views of Facebook adverts (what are termed "impressions" - that is, the appearance of an advert in a Facebook users timeline). We then extracted all impressions of political adverts in the overall dataset, totalling 16,109 items.[iii] We used this raw data to generate two datasets for analysis. First, we identified unique adverts in the raw data (n=783). This was necessary for analysing the content of the

adverts that political parties were purchasing. Next, we used geographical data gathered from Who Targets Me users when they installed the plug-in to measure the density of political advertising in individual UK parliamentary constituencies. In order to do this, we developed a metric which we term Political Advert Density (PAD). This is a ratio of political adverts relative to all other adverts being seen in a constituency. This allows for us to control for the different number of users across constituencies, and the differential amount of time those users might be spending on Facebook. This second dataset allowed us to identify the constituencies that parties were targeting the most with adverts within our dataset and examine the extent to which this was related to constituency marginality.

Approach to analysis

Though the Who Targets Me dataset cannot be considered statically representative (a point considered further below), we do not believe that this leaves it devoid of analytical value. We believe that there are three reasons for this. First, given its size and good absolute coverage of key segments of the population, it is reasonable to believe that the dataset provides us with enough evidence to develop arguments on the use of Facebook for political advertising. Second, an understanding of the limitations of this kind of exploratory strategy can help us to better define what future research designs might be able to address these hypotheses in other electoral cycles – as detailed in the conclusion. Third, the significance of the phenomenon we are analysing, coupled with the difficulty of externally assessing it, increases the value of any rigorous insights that can be developed, whatever limitations they may have.

Analysing the data

To measure levels of negativity, personalisation and calls for action, as well as the various topics the adverts covered, we employed a content analysis approach. This method

provides a systematic approach to "the meaning of qualitative material [...] by classifying parts of your material as instances of the categories of a coding frame" (Schreier, 2012: 20). Content analysis has been often employed to make sense of political advertising (see Gunter et al., 2015 and Walter et al., 2013 for recent examples on European election campaigns; see Prior, 2001, for a discussion; our coding frame is included as Appendix 1 to this article). In order to ensure intercoder reliability, we created a random subsample of 100 ads, which were independently analysed by two coders. When tested with Krippendorff's Alpha, all our variables scored in excess of 0.8, a strong indication of intercoder reliability (Krippendorff, 2004: 241). Calculations were done using Freelon's Recal (2013; see results in Appendix 2). In addition to these high scores, coders also reflected on disagreements and made further improvements to our frame. We turn now to the discussion of our findings.

Analysis of Facebook Adverts

RQ1: What messages were used in the Facebook advertising campaigns of the major political parties?

Drawing on the dataset of unique adverts, Figure 1 shows the topics that each parties' adverts concentrated on. It leads to a few observations. First, the adverts produced by the Conservative Party were heavily focused on Brexit (65.6 per cent of all Conservative adverts in the dataset). This is unsurprising, given that the original rationale for calling the election was to provide the government with a mandate to undertake the Brexit process (May, 2017). As might be expected, the Liberal Democrats - the only UK-wide party to actively oppose Brexit and to seek to overturn the 2016 referendum result - also focused on Brexit in many of their adverts (24.7 per cent).

In contrast, Labour barely mentions the issues at all (1 per cent of all Labour adverts are about Brexit). It is necessary to tread cautiously in making interpretations here, given the

unknown representativeness of Who Targets Me's user base. Nevertheless, it does appear striking that so few Brexit-centred adverts were delivered to a demographic for whom this would (in many of the most densely measured constituencies, which remember were largely urban) have been of particular interest. The absence of the topic could be interpreted as Labour side-stepping an issue that was politically problematic to them, due to divisions among the party's voter coalition, made up of regional working class and urban, middle class voters. This is certainly something the party has been accused of, both before and after the 2017 general election (Harrop, 2017).

This finding is potentially important, as it indicates that the claim that Labour refrained from framing the election in Brexit-related terms in their Facebook advertising is more plausible than the claim that they were simply producing different adverts for different constituencies. This is significant as one of the common criticisms of targeted advertising in general and Facebook advertising in particular is that it allows for a form of narrow-casting, wherein contradictory messages can be sent to different audiences. Since these messages are sent and received in a private communication space, there is little or no chance of the contradiction entering into the public space. Yet we find no evidence of anything like this occurring, even on an issue (in this case Brexit) where it would benefit Labour to employ such a strategy. The major focus of policy-based Labour adverts on Facebook include social security (21.6 per cent), education (13.4 per cent) and healthcare (12.9 per cent). These messages build on arguments that the party and the broader political left have been constructing in recent years, particularly focusing on an anti-austerity message (N Anstead, 2017).

These findings are interesting because they point to a Facebook advertising strategy that is not, at least in message terms, highly targeted and differentiated, but rather based on

well-established messages built by parties in the years before the election campaign. This suggests that message fragmentation is defined less by the technological affordances of the platform than by internal and contextual dynamics of the political organization.

RQ2: Were these messages more likely to be (i) negative or (ii) personality / leader focused in some parties than others?

An important question for those seeking to understand how Facebook use in election campaigns is the extent of negative campaign advertising on the platform (Auter & Fine, 2016). Our coding (shown in Figure 2) reveals that the majority of the adverts placed by all four of the major political parties were negative (defined in our coding as naming a specific opponent politician or party). Overall, Labour had the highest proportion of negative adverts (64.1 per cent), followed by the Liberal Democrats (61.6 per cent), the SNP (57.9 per cent) and the Conservatives (56.4 per cent).

A qualitative examination of the data suggests that these figures can be broken down further, with different types of negative advertising being evident. Conservative negative advertising was strongly focused on attacking Jeremy Corbyn, the Labour leader. In contrast, Labour and Liberal Democrat adverts had a greater tendency to focus on the Conservative Party and its policies more generally. A third category of negative advertising – likely a consequence of the British first-past-the-post election process coupled with multi-party politics – focuses on offering tactical advice to voters, indicating that the party placing the advert were most likely to defeat another party in the constituency. These adverts also attempt to explicitly discourage voters from supporting another challenger party. This latter type of advert is one type of communication where the ability to target geographically is particularly important as, by definition, such adverts are constituency specific, as they necessarily relate to local electoral circumstances.

Although the dataset does not contain many adverts from the Green Party (only 22 adverts in total), the adverts gathered stand out as being part of a relatively positive campaign on Facebook, at least in comparison with other parties. While the Greens did produce some negative adverts attacking their opponents, especially the two largest parties, the majority of the content they placed on Facebook were broader statements of values or attempting to solicit support from activists. This approach is in keeping with communication strategies that the Green Party have adopted in previous elections, seeking to set themselves apart from the larger, mainstream parties (Green Party of England and Wales, 2015).

In any case, the levels of negativity we have found on Facebook targeted ads are remarkably high, in comparison with what studies have concluded on political advertising by British parties in other media during previous elections. Consider, for instance, PEBs (Party Election Broadcasts). Research on these short films, reaching back to the pieces televised in the 1960s, for instance, agree that only rarely the overall proportion of negative appeals surpassed 50% (Walter, 2012; VanHeerde-Hudson, 2011; Hodess et al., 2000; Gunter et al., 2014).

But even Facebook party messages which were not targeted appear to have been much less negative. As Campbell and Lee's (2016) work on the 2015 General Elections suggests, only a small fraction of Facebook "political posters", i.e., "still images posted openly to parties' Facebook pages, rather than distributed as targeted online advertising", was negative. Negativity levels ranged from 27.5% (Conservatives) to 11.1% (Plaid Cymru), they argue. Interestingly, they note that when these posters did not display brand identifiers or the same colour schemes, they were strikingly more negative – particularly among Conservatives and Labour, which "used a noticeable proportion of unbranded and negatively oriented ads" (ibid: 57). When Conservative's posters were branded, only 11.2% of them were negative;

when they were unbranded, this percentage jumped to 83.9%; in the case of Labour, these proportions were of 15.9% and 55.1%, respectively. The suggestion is that parties appear to see in the absence of transparency an invitation to increase their attacks to opponents without suffering potential backlashes. Our study appears to offer further evidence of this phenomenon.

Another difference between the parties is the extent to which they pursued a personality-based / leader-focused campaign. Research on the influence of TV on campaigning since the 1960s has supported a theory that that the medium has been associated with 'personalised' campaigning (Langer, 2011), but it is unclear whether social media campaigns continue the trend. In our content analysis, we define personality-based campaigning as adverts that explicitly mention the party leader. As might be expected, given that from the outset the Conservatives built their campaign around Theresa May in a highly-personalised fashion, the Prime Minister featured heavily (34.5 per cent of Conservative adverts mentioned May explicitly). In contrast, Jeremy Corbyn was not mentioned in a single advert purchased by the Labour Party. This is unsurprising in the context of the start of the election campaign, where received wisdom was that Jeremy Corbyn was a liability to the party. What is interesting however is that Labour did not change its approach over the course of the seven-week campaign, even when their message seemed to be gaining traction and the party's poll ratings were increasing. Therefore, our second observation is that while the opacity of Facebook ads seems indeed to increase negativity levels, it does not necessarily lead to a higher level of personalization of the message.

RQ3: Were parties using Facebook to mobilise votes to engage in political action?

The final codes we applied to the advert dataset examined any appeals made in the advert to encourage action on the part of the reader. We identified several different actions

which appeared across the dataset. What is again notable is that different parties appeared to have adopted quite different strategies in using Facebook. Voting is the only activity requested by Conservative Party advertising, similar to the approach taken by the Scottish National Party. In contrast, the Liberal Democrats and the Greens in particular ask for other activities. In the case of the Liberal Democrats, 19.8 per cent of adverts ask users to sign a petition, likely with the aim of getting more details about would-be voters and supporters, so as they can be more effectively targeted, a tactic that has widely been used in previous elections (Nick Anstead, 2016). The Green Party makes the most diverse use of the medium. 32.1 per cent of its posts ask readers to donate to the party and the same number request that users share posts on Facebook. In contrast only 25 per cent of Green Party posts mention the act of voting.

This discrepancy may not be as strange as it might seem, and may in fact reflect the distinctive needs of the Green Party. Due to the nature of the UK's electoral system, the party struggles to convert its votes into parliamentary seats. As a result, increasing national vote-share (as opposed to increasing vote-share in the few constituencies where the Greens are competitive with the major parties) is of only limited use. Additionally, the Greens have very limited financial resources compared to the other major parties.[iv] Therefore, the content of the Green Party adverts very much seem designed to offset these particular challenges. Interestingly, our findings echo research done on Green Parties in other countries, where social media is used in a similar fashion to galvanise supporters into action (Larsson, 2017).

RQ4: Did voters in marginal constituencies see more political ads on Facebook?

Next, we turn our attention to the targeting of adverts, and the constituencies in which WTM users live. As we have seen, theoretical and public debate about Facebook political advertising has involved multiple claims about targeting. We focus on the claim that

campaigns are focusing resources on marginal constituencies. In this section, we focus on the four parties for which we have adverts that ran candidates across the United Kingdom (excluding Northern Ireland): the Conservatives, Labour, the Liberal Democrats and the Greens.

Of the UK parties, the Conservatives appeared to be most efficient at targeting their Facebook adverts to marginal constituencies, with a correlation 0.320 between the Political Advertising Density (PAD) score and the marginality of constituencies in which WTM users live.[v] The Liberal Democrats also have a correlation, albeit a weaker one (0.102). At the other end of the spectrum, in the case of the Green Party there seems to be very little meaningful relationship between constituency marginality and the propensity to place adverts. However, this is not entirely surprising. As the analysis of advert content above suggests, a large proportion of the content the Green's produced did not seek to mobilise voters, but instead asked individuals to engage with the party as activists and donors.

It is the Labour Party, however, which presents the most puzzling result. There is no measurable correlation between marginality and the PAD for Labour advertisements. Why might this be? Four reasons might provide an explanation. It is important to note that, while a very logical approach to running a campaign, targeting marginal seats might not always be the best campaign strategy. Put another way, the definition of marginal used in this analysis is based on the closeness of the result in the 2015 election. However, changes in the political landscape during the course of a parliament might that parties employ either offensive or defensive strategies. At the outset of the campaign, Labour were a long way behind in the opinion polls, suggesting that targeting adverts at the most marginal seats (i.e. those where they were just behind or ahead of their nearest rival) might not in fact be a rational strategy, as these seats could already be seen as being lost (or at least, this is how it would have

seemed at that moment in time). Labour's low share in early campaign polling also offers a second explanation. During this period, it was widely rumoured in the press that those surrounding Labour leader Jeremy Corbyn were keen to make national vote share a major metric of electoral success, as opposed to seats won and lost (Bush, 2017). It may be that Labour were intentionally targeting resources at seats where it might be possible for them to increase their vote share, even if this would have little effect on the overall number of seats the party would win (although, it should be noted, that this explanation also raises the question of decision making processes within political parties. Prior the election, there was considerable discussion about the estrangement between the officials running the Labour Party campaign and the Corbyn leadership team. It is therefore debatable in whose interests strategic decisions might be made).

Third, and more broadly, recent British elections have seen a decreased propensity for seats to follow patterns of Universal National Swing (UNS)(Kellner, 2014). This is partially the result of reconfiguration of the underlying social basis of British politics, and partially because of a growing ability for campaigns to reach out to segments of the population who are likely to be most responsive to their messages (Facebook targeting is part of this process). As such, it might make sense to target particular seats that are not marginal in the traditional sense of the term, but contain a large number of voters who might be responsive to a particular message. The Conservative Party employed this approach particularly effectively in the 2015 General Election, for example, when targeting would-be Conservatives in what were assumed to be "safe" Liberal Democrat seats (a strategy that played a role in nearly wiping out their coalition partners and winning David Cameron an outright majority)(Ross, 2015).

Of course, it is also possible that that Labour was, in fact, targeting marginal seats, and the non-random, self-selecting nature of the WTM user base has systematically concealed this pattern in the data that has been obtained. This possibility should cause us to make interpretations with caution, and look at methodological improvements (reflected on shortly) that would allow us to eliminate this possibility from datasets such as this one in future.

We can delve into this data more by examining the ten most targeted seats for each party, revealing what the data suggests as the key battleground constituencies for their campaigns. Perhaps predictably, the evidence here suggests that Labour were more defensive, with four of their five most targeted seats being constituencies represented by a Labour MP. In contrast, the converse is true for the Conservatives with four of their five most targeted seats being constituencies where they were currently in opposition. The Liberal Democrat and Green data is less clearly defined, although the Liberal Democrats most targeted seats include four seats where they are challengers and incumbent's majority is in the four-figure range (a relatively small amount, given the scale of the Liberal Democrats electoral decline in the 2015 election), and the seat of Richmond, which the party had won in a post-Brexit referendum by-election in December 2016.

Brexit most clearly defines the battle lines between the parties in our dataset. Both the Conservatives and Labour focused on constituencies that voted for the UK to leave the European Union. In Labour's case, every single constituency in the top ten voted to leave, seven of them with more than a 60 per cent vote share. For the Conservatives, eight of the top ten constituencies voted leave. In contrast, the Liberal Democrats sought to mobilise voters who had supported remain in the referendum. Five of their top ten constituencies supported remain, including the ultra-strong remain voting seats of East Dunbartonshire (26.87 per cent

leave vote) and Richmond (28.69 per cent leave vote). Thus, while Brexit was often absent from the content of the Facebook campaign (and notably so in the case of Labour), our data suggests it was strongly present in the targeting process, either because parties aimed to take advantage of it or parties feared that their opponents might be able to exploit it.

Limitations of the dataset

Who Targets Me was not designed as research project. As a result, its dataset suffers from limitations that make it impossible for us to claim that it is statistically representative of the UK population. Before explaining why we believe that the data still has (limited) empirical value, let us detail two particularly important issues. First, the group of people who chose to install the plug-in were self-selecting. During the election campaign, Who Targets Me strived to publicise the tool as widely as possible and consciously worked to generate coverage in a range of different types of publications - including national and local media, broadsheets and tabloids - with a range of partisan affiliations. However, it is reasonable to suppose that the people most likely to install the plug-in are those who may have concerns about privacy or an interest in political campaigns and advertising, as these were themes associated with the campaign. Second, the data gathered is limited to those who use Google Chrome as their web browser. While Chrome dominates the UK desktop browser market (Statcounter, 2017. This tool calculates market share on the basis of page views with particular browsers), our dataset necessarily excludes users of other browsers, or – significantly – those browsing through apps and browsers on their smartphones. This is a significant limitation: younger users in particular, are more likely to browse on smartphones[vi].

The self-reported personal data collected during the installation of the plug-in (gender, age and postcode) allowed us to assess how Who Targets Me users compare to the

rest of the UK population. Users were disproportionately likely to be male (78.54 per cent male, as opposed to 50.79 of the UK's population). In terms of age, the plug-in did achieve some coverage across a wide range of ages, and the median age of users (38.36 years of age) is not that different to the UK's overall population (39.08; figure calculated from ONS, 2016). However, these figures mask quite a different age distribution from the rest of the UK population (see Figure 4). A similar problem is present in relation to the geographical location of users: overall, almost all of the UK's 632 parliamentary constituencies in England, Scotland and Wales, had users installing the plug-in, but only a few constituencies accounted for a significant proportion of users (see Figure 5). An examination of the list of constituencies with the most installations quickly reveals a pattern: users were concentrated in areas that are urban, and either in the process of gentrifying or close to major universities (the top five constituencies for installations were Hackney and South Shoreditch; Bethnal Green and Bow; Bristol West; Hackney North and Stoke Newington; and Camberwell and Peckham). This may suggest a user base that disproportionately contains a particular demographic: younger, urban and likely wealthier and more educated than the median voter.

The high number of Who Targets Me volunteers in these types of constituencies also explains the propensity of such users to live in very safe Labour seats. In 2015, for example, the Labour Party won Hackney and South Shoreditch with a majority of 24,213 votes, making it among the 60 safest seats in the country. Nevertheless, it is worth noting that the overall dataset does contain a good cross-section of seats in the UK, ranging from the most marginal to the very safest.

The Who Targets Me dataset is more problematic when it is examined in the context of the European Union referendum of 2016. The types of constituencies where people installed Who Targets Me lived were also the constituencies with the greatest propensity to

vote to remain in the European Union. Figure X shows Who Targets Me installations across, divided into deciles by the votes cast in the referendum in the users constituency (referendum vote share based on modelling in Hanretty, 2016).

An important caveat should be added to this analysis. We can only examine Who Targets Me users' Brexit-preferences at the constituency-level. However, as recent literature on targeting has noted, campaigns increasingly target voters at the individual, not the geographic level (Nick Anstead, 2016; Kreiss, 2016; Nielsen, 2012). In practical terms, this may mean that the users installing Who Targets Me in even the most pro-Brexit constituencies are demographically and politically like users in remain-supporting constituencies. This is an important limitation, and one that needs to be factored in drawing any conclusions about the role played by Brexit in political advertising on Facebook. Lastly, due to how the data was stored by the Who Targets Me, we could not analyse images present in the ads (their content, colour or composition), which entails a significant loss of content.

Conclusion: The Real Challenge of Facebook Advertising

Drawing on the Who Targets Me dataset, this article does offer some findings on the use of Facebook in the 2017 UK election, some of which challenge conventional wisdom about Facebook advertising in election campaigns. One notable finding is that most party political advertising being placed on Facebook is actually in-tune with parties national messaging, communicated through more traditional channels. This was true of both Labour and the Conservative Party. This was also reflected in the geographical targeting of adverts, at least in our dataset, where the most fought over constituencies – at least in the case of the two major parties – were those that had strongly supported Brexit in the 2016 referendum. This is not surprising, given pre-election commentary that it was in these sorts of places that the Conservatives would try to make inroads into Labour's support-base.

That said, and as discussed above, we must be very cautious of over-claiming based on a dataset created by of volunteers installing a browser plug-in. As the metadata from the Who Targets Me dataset indicates, the citizens who chose to install the plug-in disproportionately came from specific segments of the electorate, while other groups were very under-represented.

Being able to study the voting population in the run-up to an election in the manner that Who Targets Me has done could potentially allow researchers to observe communications flows otherwise hidden on Facebook's platform, if only the data could be improved to allow generalizability to the voting population overall. How might the method we deploy in this paper be developed in the future to approach this? Some of the challenges demand technical solutions. For example, a modified collection system would be required to understand advertising exposure on Facebook's mobile applications. To improve representativity, one model might be to recruit (probably with a financial inducement) a statistically representative panel, in a manner similar to the method used by internet pollsters. However, such an approach would not be unproblematic, largely because researchers may lack insights into the Facebook environment required to design a successful sampling strategy. For example, how large would such a panel need to be to ensure good coverage of the various segments of the electorate that political parties are targeting? Given that our dataset suggested that only a fraction of all adverts on Facebook were purchased by political parties (and this, do not forget, with a cohort of users who were likely more politically engaged than the median citizen, so arguably more likely to be exposed to adverts), the panel would likely need to be of a substantial size in gather an effective dataset.

There is also a broader conceptual problem with any sort of approach that seeks to design a representative sample. What exactly is it representative of? As outlined above, it is

possible to compare some metrics of the sample with the broader population, but this may be misleading. Ultimately, researchers are going to want to understand advert content on Facebook and how it is disseminated. However, without any kind of picture of what that population looks like, it becomes difficult to have complete confidence in how samples are being drawn for any particular dataset.

This takes us back to the major problem, not just with studying Facebook political adverts, but – as, if not more importantly – with regulating elections. Facebook as a political advertising platform remains almost completely opaque. Recent months have seen Facebook publishing data related to the 2016 US Presidential election and in particular, in response to concerns that Russia may have purchased adverts with the aim of subverting the US' electoral process. While transparency around these events is to be welcomed, it underlines rather than challenges the problem: Facebook wholly controls this data and can decide for itself on a degree of transparency that suits its own interests. In the context of electoral politics and regulation, this is not a power that a private company (and certainly one of the scale of Facebook) should wield. Overcoming that problem is a hugely important challenge. Until this is accomplished, researchers will be forced to design of imaginative, best-possible-approach solutions to continue to study this vital area.

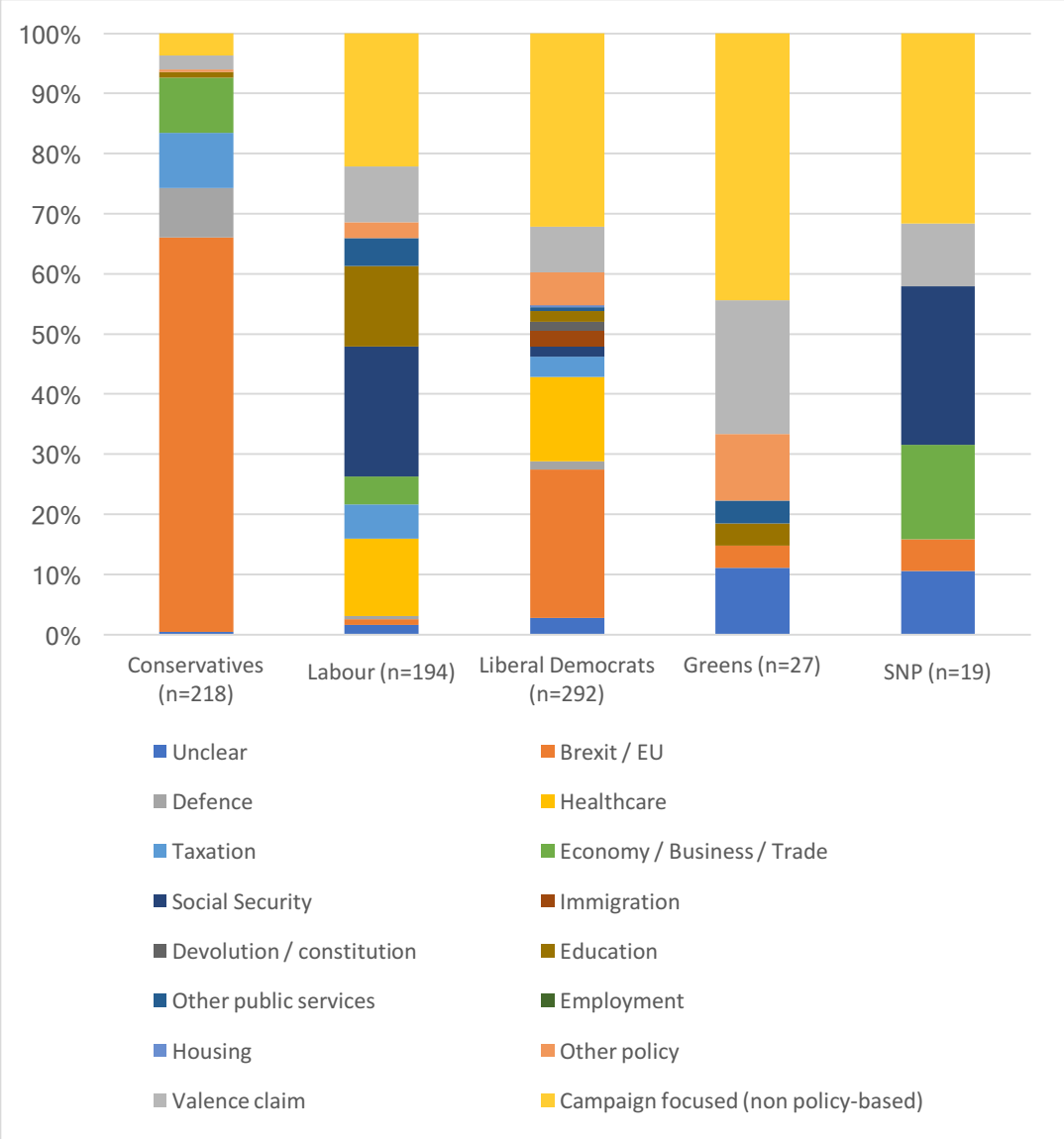

Figure 1: Topics of Facebook adverts for each party (n=750)

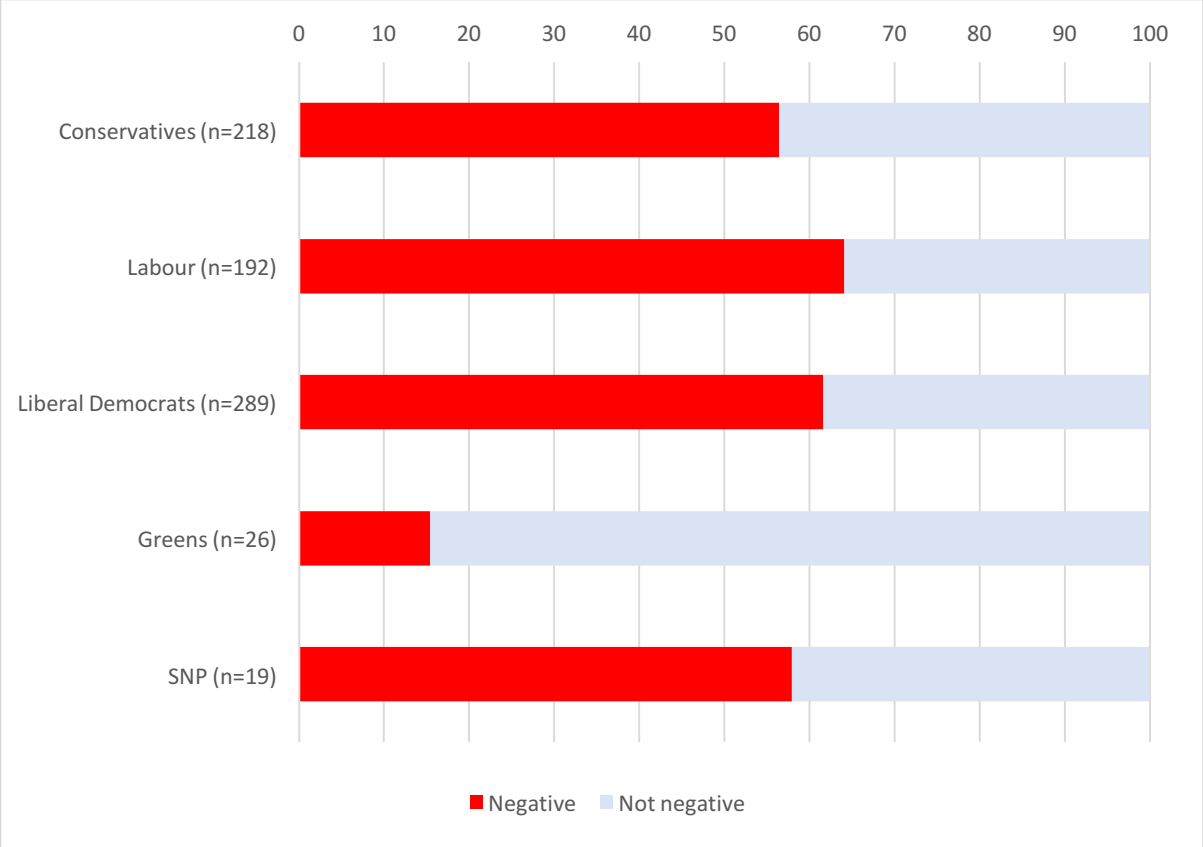

Figure 2: Proportion of negative adverts for each party (n = 744)

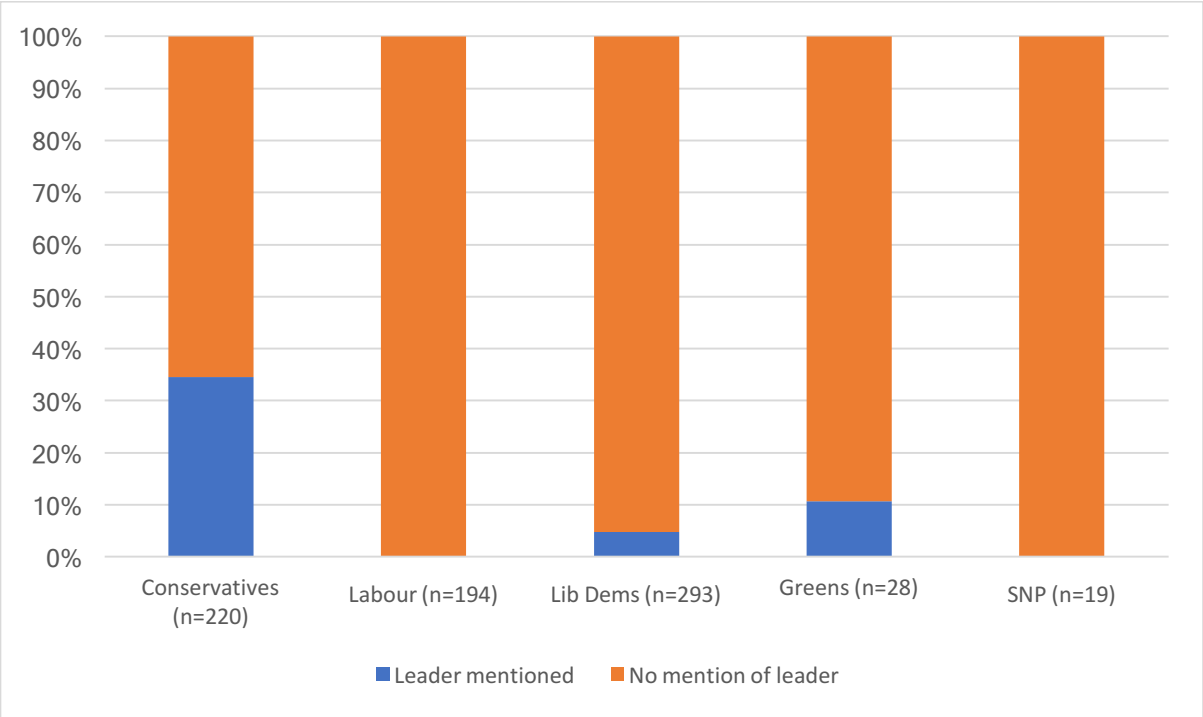

Figure 3: Personality / leader focus of Facebook adverts for each party (n = 754)

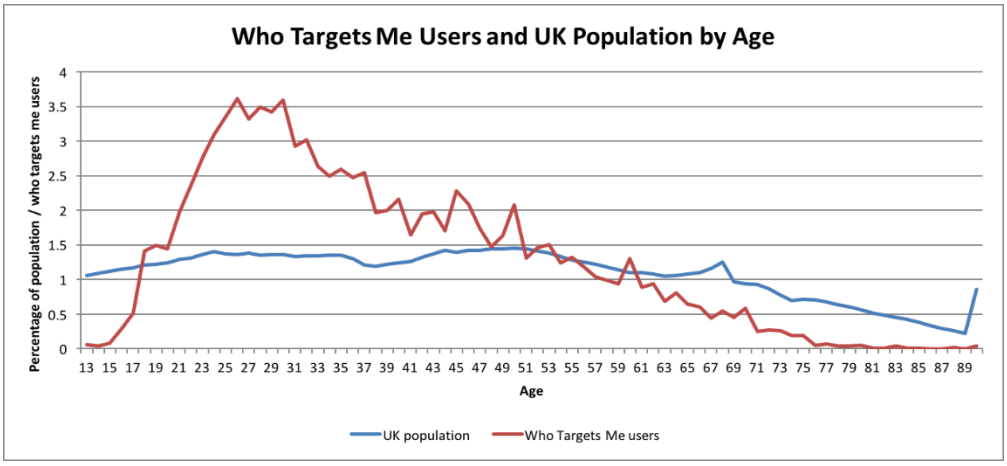

Figure 4: Who Targets Me Users Age, Compared to UK population

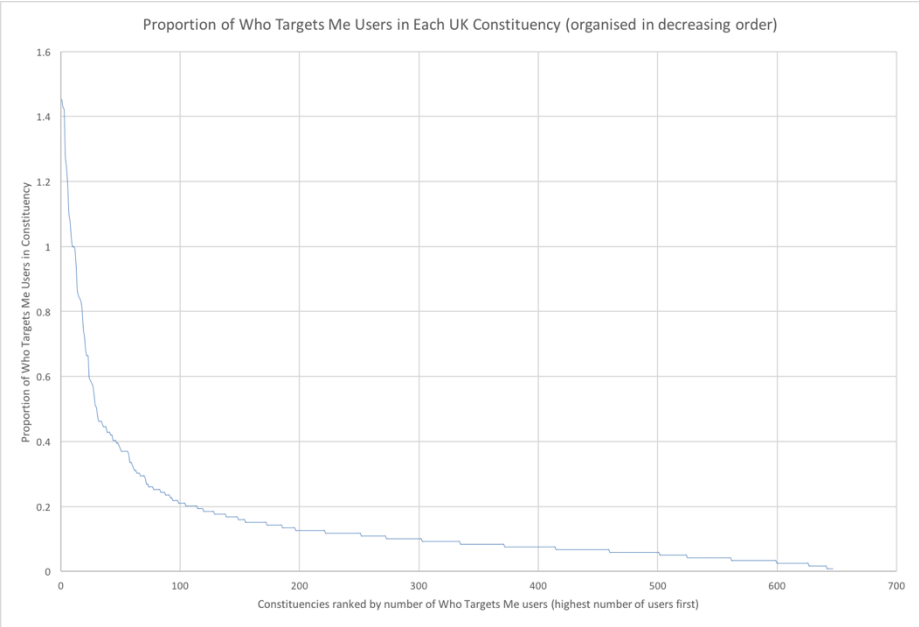

Figure 5: Proportion of Who Targets Me users in individual constituencies, organised in decreasing order

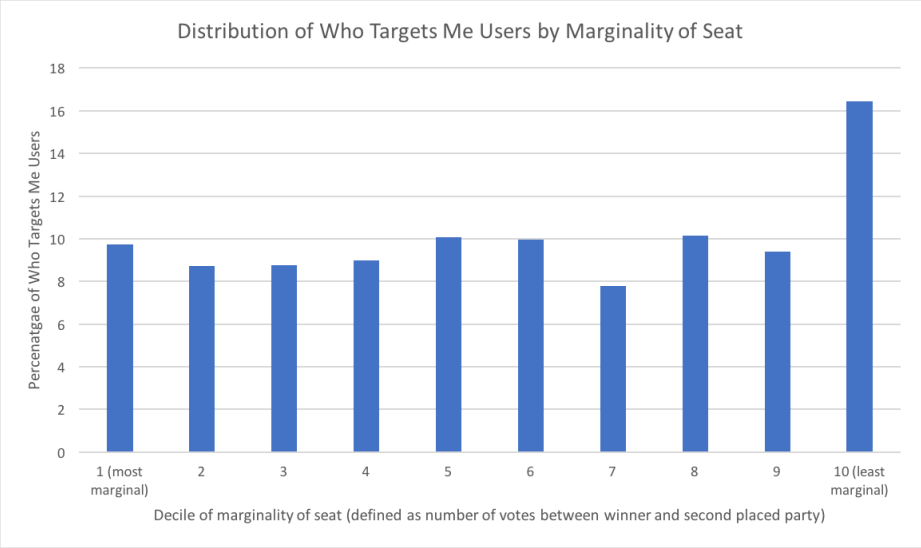

Figure 6: Distribution of Who Targets Me Users by Marginality of Seat

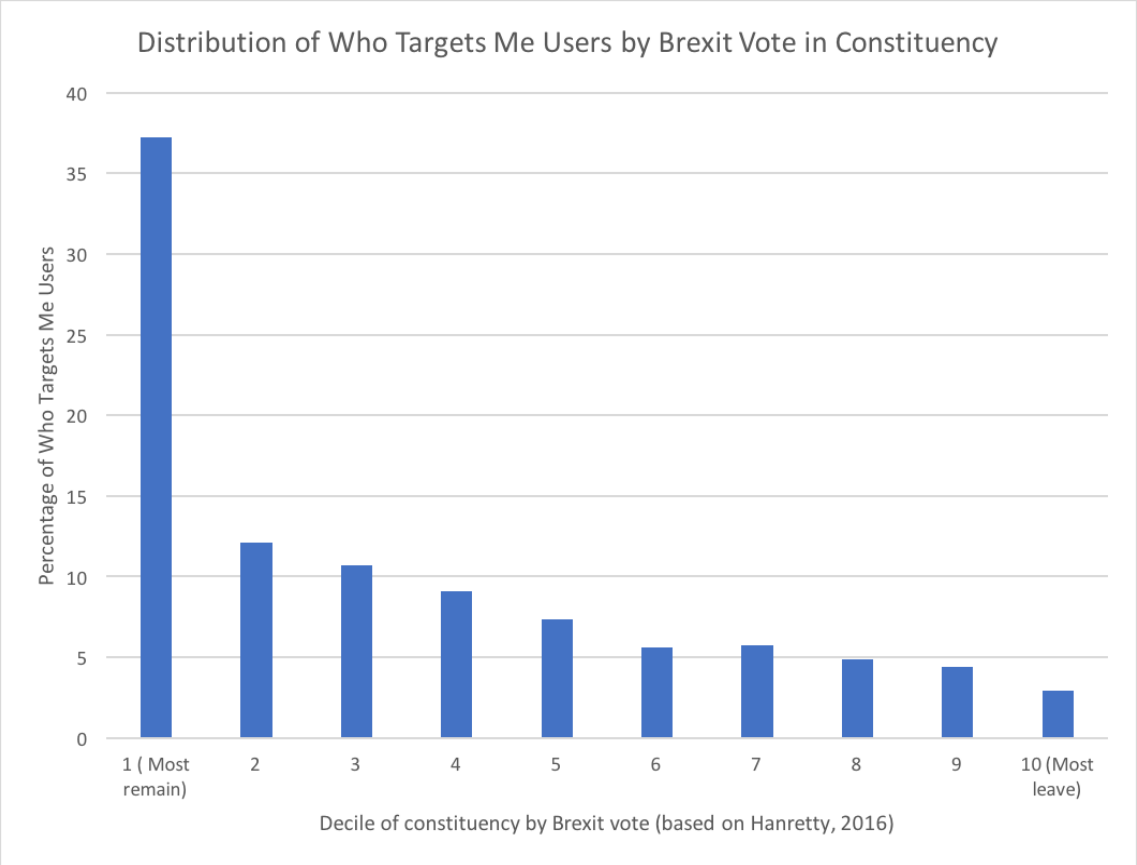

Figure 7: Distribution of Who Targets Me Users by Brexit vote in Constituency

Party		Constituency	PAD Score	Marginality (+/-)	Br
Lab	1	Stoke-on-Trent Central	0.114035088	5179	64
	2	Birmingham, Perry Barr	0.079365079	14828	51
	3	Middlesbrough South And East Cleveland	0.061643836	2268	65
	4	Peterborough	0.05785124	-1925	61
	5	Heywood And Middleton	0.047368421	5299	62
	6	Dover	0.044776119	-6294	63
	7	Penistone And Stocksbridge	0.044642857	6723	60
	8	Erewash	0.042372881	-3584	63
	9	Derbyshire Dales	0.042168675	-14044	51
	10	Mid Bedfordshire	0.03652968	-23327	52
Con	1	Wirral South	0.121621622	-4599	46
	2	Birmingham, Northfield	0.037037037	-2509	46
	3	Copeland	0.023323615	-2564	59
	4	Enfield North	0.019441069	-1086	61
	5	Camborne And Redruth	0.017763845	7004	58
	6	Jarrow	0.017094017	-14880	61
	7	Bolton West	0.016666667	801	55
	8	Keighley	0.013192612	3053	53
	9	Barrow And Furness	0.013081395	-795	57
	10	Dagenham And Rainham	0.012048193	-7338	70
Lib Dem	1	North West Norfolk	0.096385542	-23054	65
	2	East Dunbartonshire	0.046511628	-2167	26
	3	Cheadle	0.042589438	-6453	42
	4	Macclesfield	0.035175879	-22221	47
	5	Heywood And Middleton	0.031578947	-19319	62
	6	Hazel Grove	0.030769231	-6552	52
	7	Cheltenham	0.028498511	-6516	42
	8	Hemel Hempstead	0.026717557	-23843	55

	9	Erewash	0.025423729	-18978	63
	10	Richmond Park	0.024399399	-23015	28
Greens	1	Na H-Eileanan An Iar	0.066666667	-8662	43
	2	Leicester East	0.057692308	-27918	54
	3	Peterborough	0.033057851	-17466	61
	4	Sutton Coldfield	0.028846154	-26356	51
	5	Jarrow	0.025641026	-20154	61
	6	Vale Of Glamorgan	0.023752969	-22553	52
	7	Kingston Upon Hull West And Hessle	0.023255814	-14703	67
	8	Middlesbrough South And East Cleveland	0.020547945	-18133	65
	9	North West Durham	0.020408163	-18507	55
	10	Brigg And Goole	0.019138756	-22031	55

Notes: Data on seat marginality calculated from Norris (2017). Brexit voting figures are based on statistical generated by Hanretty (2016).

Appendix 1: Coding frame for content analysis of Facebook advert content

Apply all codes across the whole ad (i.e. coding unit is combination of headline, subtitle and text)

Topic	**In the event of an advert containing more than one issue, code for the dominant issue.** 0 Unable to say / unclear 1 Brexit / European Union 2 Defence / Military / Security 3 Healthcare 4 Taxation (including references to the dementia tax) 5 Economy / Business / Trade 6 Social Security 7 Immigration 8 Devolution and other constitutional issues 9 Education 10 Other public services 11 Employment 12 Housing 13 Other policy issue (record what issue this is) 14 Valence issue (i.e. claim to competance, ability to deliver, no substantive policy disagreement) 15 Other party political / campaign topic with no policy content (i.e. activist mobilisation, donate money etc. Also includes calls to vote).
Negativity	0 Not negative 1 Negative about specific opponent
Neg_Subject	Record the name of the target of the negativity (this could be either a person or a party. In the event of multiple names appearing choose the dominate target of the ad)
Brexit_position	**Only code in the event of topic code being 1 "Brexit"** 0 Unable to say / not clear 1 Anti-Brexit 2 Pro-Brexit 3 Valance (i.e. claim to competence)

Action codes

Action - vote	0 Not present 1 Present
Action - donate	0 Not present 1 Present
Action - petition	0 Not present 1 Present
Action - join party	0 Not present 1 Present

Action - share (FB)	0	Not present
	1	Present

Appendix 2 – Intercoder reliability test

	Percent Agreement	Scott's Pi	Cohen's Kappa	Krippendorff's
Variable 1 (cols 1 & 2)	86	0.82117767	0.82147411	
Variable 2 (cols 3 & 4)	98	0.95920033	0.95921697	
Variable 3 (cols 5 & 6)	95	0.92094862	0.92101106	

Bibliography

Anstead, N. (2016). The Data-Driven Campaign in the 2015 UK General Election. *International Journal of Press/Politics*.

Anstead, N. (2017). The Idea of Austerity in British Politics, 2003-2013. *Political Studies*.

Auter, Z. J., & Fine, J. A. (2016). Negative Campaigning in the Social Media Age: Attack Advertising on Facebook. *Political Behavior, 38*(4), 999-1020. doi:10.1007/s11109-016-9346-8

Baldwin-Philippi, J. (2017). The Myths of Data-Driven Campaigning. *Political Communication, 34*(4), 627-633. doi:10.1080/10584609.2017.1372999

Barocas, S. (2012, November). The price of precision: Voter microtargeting and its potential harms to the democratic process. In Proceedings of the first edition workshop on Politics, elections and data (pp. 31-36). ACM.

Bush, S. (2017). Why Len McCluskey's idea of general election "success" for Labour matters. *New Statesman.* Retrieved from https://www.newstatesman.com/politics/staggers/2017/05/why-len-mccluskeys-idea-general-election-success-labour-matters

Butrymowicz, D. W. (2009). Loophole.com: How the fec's failure to fully regulate the internet undermines campaign finance law. *Columbia Law Review* 109(7), 1708-1751.

Cadwalladr, C. (2017). The great British Brexit robbery: how our democracy was hijacked. Retrieved from https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy

Cohen, J. E. (2013). What Privacy Is For. *Harvard Law Rev2*, *126*, 1904–1933.

Denham, E. D. (2017). *The Information Commissioner opens a formal investigation into the use of data analytics for political purposes*. Retrieved from https://iconewsblog.org.uk/2017/05/17/information-commissioner-elizabeth-denham-opens-a-formal-investigation-into-the-use-of-data-analytics-for-political-purposes/

Freelon, D. (2013). ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, *8*(1), 10–16. Retrieved from http://www.ijis.net/ijis8_1/ijis8_1_freelon_pre.html

Electoral Commission. (2016). *UK Parliamentary General Election 2015: Campaign spending report*. Retrieved from London: http://www.electoralcommission.org.uk/__data/assets/pdf_file/0006/197907/UKPGE-Spending-Report-2015.pdf

Green Party of England and Wales. (2015). Change The Tune. Retrieved from https://www.youtube.com/watch?v=PPgS7p40ERg

Gorton, W. A. (2016). Manipulating Citizens: How Political Campaigns' Use of Behavioral Social Science Harms Democracy. *New Political Science,* 38(1), 61-80.

Groshek, J., & Koc-Michalska, K. (2017). Helping populism win? Social media use, filter bubbles, and support for populist presidential candidates in the 2016 US election campaign. *Information Communication and Society, 20*(9), 1389-1407. doi:10.1080/1369118X.2017.1329334

Gunter, B., Saltzis, K., & Campbell, V. (2015). The Changing Nature of Party Election Broadcasts: The Growing Influence of Political Marketing. *Journal of Political Marketing*, *14*(3), 229–250. https://doi.org/10.1080/15377857.2012.693060

Hanretty, C. (2016). Final estimates of the Leave vote, or "Areal interpolation and the UK's referendum on EU membership". Retrieved from https://medium.com/@chrishanretty/final-estimates-of-the-leave-vote-or-areal-interpolation-and-the-uks-referendum-on-eu-membership-5490b6cab878

Harrop, A. (2017). Where Next for Labour? *The Political Quarterly, 88*(3), 395-399. doi:10.1111/1467-923X.12406

Hersh, E. D. (2015). *Hacking the Electorate: How Campaigns Perceive Voters*: Cambridge University Press.

Howard, P. N. (2006). *New media campaigns and the managed citizen*: Cambridge Univ Pr.

Kellner, P. (2014). Uniform swing is now worse than useless – it is positively misleading. Retrieved from https://yougov.co.uk/news/2014/12/01/uniform-swing-rip/

Kreiss, D. (2016). Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy: Oxford University Press.

Kreiss, D., & Howard, P. N. (2010). New challenges to political privacy: Lessons from the first US Presidential race in the Web 2.0 era. *International Journal of Communication, 4*, 19.

Langer, A. I. s. (2011). *The personalisation of politics in the UK : mediated leadership from Attlee to Cameron*. Manchester: Manchester University Press.

Margetts, H., John, P., Hale, S., & Yasseri, T. (2015). *Political turbulence: How social media shape collective action*: Princeton University Press.

May, T. (2017). Theresa May's general election statement in full. Retrieved from http://www.bbc.co.uk/news/uk-politics-39630009

Moore, M. (2016). Facebook, the Conservatives and the Risk to Fair and Open Elections in the UK. *Political Quarterly, 87*(3), 424-430. doi:10.1111/1467-923X.12291

Nielsen, R. K. (2012). *Ground wars: Personalized communication in political campaigns*: Princeton University Press.

Norris, P. (2017). British General Election Constituency Results, 2010-2017, V1.2. Retrieved from https://www.pippanorris.com/data/

Office for National Statistics (2016). Population estimates. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates
Persily, N. (2017). Can democracy survive the Internet?. *Journal of democracy*, *28*(2), 63-76.

Prior, M. (2001). Weighted content analysis of political advertisements. *Political Communication*, *18*(3), 335–345. https://doi.org/10.1080/10584600152400374

Ross, T. (2015). *Why the Tories Won*. London: Biteback.

Rowbottom, J. (2012). Lies, Manipulation and Elections—Controlling False Campaign Statements. *Oxford Journal of Legal Studies*. doi:10.1093/ojls/gqs016

Schreier, M. (2012) Qualitative Content Analysis in Practice, London: Sage.

Statcounter. (2017). Desktop Browser Market Share United Kingdom. Retrieved from http://gs.statcounter.com/browser-market-share/desktop/united-kingdom

Sunstein, C. (2017). #Republic: Divided democracy in the age of social media. Princeton, NJ: Princeton University Press.

Tambini, Damian and Labo, Sharif (2016) Digital intermediaries in the UK: implications for newsplurality. *Info*, 18 (4), 33-58.

Tworek, H. (2017). Cambridge Analytica, Trump, and the new old fear of manipulating the masses. Retrieved from http://www.niemanlab.org/2017/05/cambridge-analytica-trump-and-the-new-old-fear-of-manipulating-the-masses/

Walter, A. S., van der Brug, W., & van Praag, P. (2014). When the Stakes Are High: Party Competition and Negative Campaigning. *Comparative Political Studies*, *47*(4), 550–573. https://doi.org/10.1177/0010414013488543

Who Targets Me (2018). The project. Retrieved from https://whotargets.me/en/about/

---

[i] As far as we know, no previously published article has gathered and analysed Facebook advertisements.

[ii] Popular examples in the recent UK elections include pages for the Labour Party's Jeremy Corbyn (https://web.facebook.com/JeremyCorbynMP/), the Conservatives' Theresa May (https://web.facebook.com/TheresaMayOfficial/) and Nick Clegg of the Liberal Democrats (https://web.facebook.com/nickclegg/)

[iii] In this article, we employ quite a restrictive definition of what constitutes the political, by only examining adverts that were purchased by accounts named for national political parties and party leaders. There are at least three ways this definition could be expanded. The first would be to include adverts purchased by constituency parties for their candidates. The second would be to broaden our definition of politics to include non-party but formally constituted political actors, such as trade unions and pressure groups. The third and broadest definition would be to define politics by the content of the advert alone, rather than who purchased it. Needless to say, operationalising such an approach would be complex.

[iv] In the 2015 election, for example, the Greens were able to spend £1.13 million nationally. This compares with a Conservative spend of £15.59 million, Labour spend of £12.1 million and a Liberal Democrat spend of £3.53 million (Electoral Commission, 2016: Chart 3).

[v] For these correlations, we do not report p-values. This is because the p-value is a measure of statistical representivity. As we discuss later in the paper, there are serious problems with claiming representivity based on the data gathered by the plug-in.

[vi] According to Ofcom's Communication Market Report for 2017 more than 60% of 16-34 year olds surveyed reported that smartphones were their most important device for internet access compared to only 13% of those aged over 55 (see https://www.ofcom.org.uk/__data/assets/pdf_file/0021/105438/uk-internet-online.pdf)