
Liam McLoughlin & Stephen Ward
University of Salford, UK

[This is a preliminary draft. Please do not cite or distribute without permission of the authors]

Introduction

The role played by social media in political campaigns during elections has certainly increased in recent years. In 2015, the Conservative party spent £1.2 million on Facebook alone, signifying how digital media has become central to the modern election campaign (Moore, 2016). Yet, even from this early analysis on the use of political adverts on social media, concerns have been raised over the lack of transparency in this campaigning (ibid, 2016). Through the use of microtargeting and dark adverts, political campaigns have been able to segment out voters by their interests and promote inconsistent or even misleading messages (Full Fact, 2018). In response to this perceived issue in political campaigning, several data activists have created projects to collect and publish adverts by political campaigns in elections. These projects have included the New York Times Political Advert collector, Who Targets Me?, and ProPublica’s Political Advert Collector (PAC). However, what has yet to be analysed is how effective these data activist projects have been.

Through a case study approach on each of the above efforts to collect and publish political adverts during elections, this paper uses the lens of Normalisation vs Equalisation to understand if internet technology allows activists the ability to properly meet their aims, or if technology provides more of an obstruction to the collection and publication of this data. Ultimately, this paper finds that the analysis of normalisation and equalisation should also include a measure for the access of real-time data. Analysis from the case studies in this paper found that data access is a major obstacle for their aims of increasing the transparency of political advertising. While each of the advert collectors were able to collect and publish some data during the campaigns, this is hampered by the self-selecting basis of their volunteers. Furthermore, as the technology used to deliver adverts increases in sophistication, being able to publish data on campaigns during the election becomes problematic due to resource restrictions. Hence, suggesting an overall normalisation effect. However, we find that the long-term impacts after elections might after all suggest elements of equalisation, as while the campaigns might not have met their aims
during the election periods, there is inferences that the data collected, and the attention gathered, around the issue of transparency of Facebook advertising might end in legal regulation.

**Dark Adverts and Microtargeting in political advertising**

Over the past decade, the way people consume news has been subject to a significant shift in behaviour, for better or for worse, as more people turn to social networks for current affairs updates (Pentina & Tarafdar, 2014). Indeed, in 2018, 39% of Americans and 27% of Britons used Facebook as a source of news (Newman et al, 2018:11). Social media has also influenced the way citizens participate (Loader & Mercea, 2011; Penny, 2017), not only allowing them the ability to like and share party posts during elections, but also create their own political campaigns (Gibson, 2013; Vaccari & Valeriani, 2016). This shift in communication is not just a bottom-up transfiguration. Top-down changes in communication by elected representatives have adapted the way they communicate with their citizens (Highfield, 2016; Anstead, 2018). However, another area of significant change due to social media has been in the (under analysed) area of the political advertising. To a certain extent, those running political campaigns have sought to adapt to new media landscapes ever since the first paid advert in 1812, when Sir Matthew Ridley spent 50 guineas on a newspaper advertisement (Rosenbaum, 1997). Ever since then, each electoral campaign has come with its own development in advertising – from the introduction of advertising agencies, to the use of the internet to mobilise grass-root supporters and to drive campaign donations - first made notable in Howard Dean’s 2004 presidential bid (Stromer-Galley, 2014:71).

Recent elections such as the 2016 U.S. Presidential election, and the 2017 UK General election have shown how the newest trend, social media advertising, has become a prominent part of political campaigns with vast amounts of money being spent on social platforms. It’s interesting to highlight the winners of both the 2016 U.S presidential election and the 2017 UK general election were also those who had spent the most on Facebook. With Donald Trump’s campaign outspending Hilary Clinton’s by an estimated $16m from June to November 2016, and the Conservatives spending £18.5m to Labour’s £11m on the platform in the 2017 General Election (D’Urso, March 31, 2018; Frier, April 3, 2018). While it’s hard to make the link that this type of online advertising was the main factor behind each campaigns’ victory; some have gone on to argue that, at bare minimum, the use of Facebook adverts was a significant factor in both election campaigns, and the 2016 EU Referendum in the UK wins (Walsh, 2017; Mullen, 2016). Research from U.K 2015 General Election is somewhat more concrete, showing that the Conservative’s outspending of all other political parties on Facebook advertising had a significant effect in the election - particularly in marginal Liberal Democrat constituencies where advertising budgets were in the region of £3,000 per month in each constituency (Wring & Ward, 2015).

The benefits of political advertising on social media platforms for parties lies not just in the size of the audience that can be found there but in the ways they can be targeted with the data sites hold on voters. With a majority of the UK and US citizenry on Facebook, the tools provided to advertisers goes much further than what was once previously available in regard to targeting. Data driven targeting allows political advertisers to know each individual voter likes, dislikes, values, hobbies, and more controversially, data from other sources, such as psychographic data taken from online polls, that can be matched to a user’s profile to target them to a much higher level of precision than previously (Aldrich et al, 2015; Motta & Fowler, 2016; Hughes, 2018:86). This results in adverts that are granularly targeted to convince a citizen how to vote based on the argument most persuasive to that individual voter. For example, during the EU referendum, the Leave campaigns had adverts specifically for voters who where classed as animal lovers, tea lovers, and those concerned about immigration (c.f. DCMS, 2018). During the 2016 US presidential election, the Trump campaign tested more than 50,000 different adverts, per

---

1 Anstead (2018) goes on to argue that there are discourses that over-emphases the impact of new campaigning tech, such as the use of data.
day, to target voters (Beckett, Oct 9, 2017). There has also been some suggestion that similar tactics have been used, not to attract voters, but to put them off voting altogether. For example, evidence that Facebook adverts were used as part of a voter suppression campaign against Clinton’s key voter bases (idealistic white liberals, young women, and African Americans) in the 2016 election (Green & Issenberg, Oct 27, 2016). This type of advertising is often referred to as Microtargeting. In a proposal pack written for Leave.EU by Cambridge Analytica, Microtargeting was described as a “predictive analytics and campaign messaging capacity [that] can help you segment and message the population according to a range of criteria” this criterion included: predicted turnout, psychological clusters, Priority issues, persuadably, partisanship, fundraising potential and most effective channels (Cambridge Analytica, n.d.)

Microtargeting was only one benefit to political advertising on social media. The other is the ability to create so-called ‘dark adverts’ on Facebook. Dark Adverts2, or their official name, unpublished posts, are posts made by an account, but are not published to a Facebook page. This allowed Facebook page owners to use these posts as adverts with only the target audience being able to see them (Facebook, n.d). This type of advertising is particularly useful for political campaigns where the aim is to segment out the audience to deliver highly targeted ads that contain conflicting information in order to gain the most amount of votes; at worst it allowed campaigns to spread disinformation to particularly vulnerable groups without others knowing (Tambini et al, 2017:16). This type of advertising was deployed heavily in the 2016 EU referendum, 2016 US election for the republican nomination campaigns of Ted Cruz and Donald Trump then later the main presidential campaign, and in the 2017 UK general election (Ellison, June 7, 2017; Wong, March 19, 2018).

---

2 It should be noted, that Facebook disabled the ability to create unpublished posts in late 2017
The use of both microtargeting and dark adverts combined has significant negative issues for democracy. Some have highlighted issues in its use relating to the factual accuracy of the messages displayed during campaigns, lack of regulatory oversight, data privacy issues, underhand campaign tactics, and the use of dark ads to anonymise campaign origin points, along with other important information such as donor information (Moore, 2016; Chand, 2017; Edwards, 2018; Lomas, March 28th, 2018; Worrall, 2018). A report from Full Fact suggests that the use of these type of adverts in political life has multiple layers of negative effects, including disengagement from democracy to risk to life as due to adverts designed to emotionally rile up their audience targeted to members of the public predisposed to violence (Full Fact, 2018). Throughout all these critiques are calls for transparency forcing all adverts, and how they are targeted, to be publicly available (Janetsky, March 7, 2017; Edelson et al, 2019). There are also suggestions for the need for more stringent legal measures. In the UK, a report by the Electoral Commission stated that this type of advertising does, indeed, pose a significant challenge to the organisations aims yet they don’t have the necessary powers to make campaigns fair during the election periods (Electoral Commission, 2018). For instance, while the ICO would go on to fine campaigns for illegal use of data in during the EU referendum, they did not have the powers or the data to monitor infractions during the campaign itself (ICO, 2018). The alarm about this type of campaign activity was not limited to domestic advertisers. There was noticeably uproar following revelations that up to $100,000 worth of dark adverts during the US 2016 Presidential election could be traced back to Russian information operations intended to sway the result of the election (Weedon, Nunland & Stamos, 2017; Leonning, Hamburger & Helderman, Sept 6, 2017). This is all in the context of a time of vast increases of campaign funds being spent on social media (Tambini et al, 2017). Indeed, the Vote Leave campaign director was quoted saying that most of their ad spend during the EU referendum was spent online: simply due to how effective it was (Cummings, Oct 29, 2016). However, it has not only been academics, but pro-transparency bodies, politicians, and institutional bodies that have highlighted the issues raised by the introduction of new forms of political advertising. A number of transparency groups have also tried to actively counter it through data activism – something that will be highlighted later in this paper.

**Normalisation vs Equalization thesis**

One lens in which the deployment of microtargeting and dark ads has yet to be viewed is through the theories of normalisation and equalization. Since the introduction of the internet, each new internet technology or development has been assessed and debated by researchers in how they might alter the political landscape (Bimber & Davies, 2003; Chadwick, 2006). One common theme in how they are assessed is though the Normalisation vs Equalization thesis; these are two competing theories that suggest opposing patterns of effects in the political realm due to technology (Ward & Gibson, 2009:34). On one end is Normalisation: this suggests internet technologies allow dominant political entities to entrench their pre-existing positions through internet technologies (Bimber & Davies, 2003; Hindman, 2008; Stromber-Galley, 2014; Ward et al: 2003). On the other, is the equalisation thesis: this suggests that the features of the internet, such as it’s relatively flat hierarchy and lower costs, allows smaller political entities a platform where they could effectively compete against dominant actors (Barber, 2001; Corrado and Firestone, 1996; Gainous and Wagner, 2011:2014). When these two competing theories were used in the context of political parties, initial research determined that the internet had equalisation effects, allowing smaller parties to compete with larger, more dominant ones. With third parties reporting a greater access to the electorate through websites and email in the UK (Gibson & Ward, 1998). While in America there was evidence to suggest that the internet allowed for a growth of third parties in an otherwise closed two-party political system (Rash, 1997; Morris, 2000). Internet technologies ability to avoid traditional gate-keepers, mixed with low costs, was seemingly levelling the political game. However, subsequent evidence suggests that the internet in the long run has a more normalising effect (Koc-Michalska et al., 2016; Gibson & Ward 2017). As the technology behind the internet has become more complex, so too has the requirements with campaigns now needing more
money for advertising online (see above), and now need web/graphic designers and IT professionals to effectively run online campaigns (Gibson et al, 2003:50; Schweitzer, 2011; Gibson & McAllister, 2015). As technologies have increased in sophistication in the web 2.0 era, smaller parties have found many of these new technologies outside the reach of the low-budget parties or candidates in the UK (Southern, 2015).

With the increasing use of social media, a new wave of normalisation vs equalising analysis has been undertaken. With similar claims made that social media would level the playing field by giving politicians an audience that competes with traditional media, while also circumventing the media (Meraz, 2009). However, subsequent analysis suggests that the sites such as Facebook are forces for normalisation rather than equalisation (Samuel-Azran, Yarchi & Wolfsfeld, 2015). The economics of social media is said to be responsible for the normalisation effects from Facebook in the 2012 U.S Congressional elections (Gainous et al, 2017). As social media is financed through advertising, the platform obviously gives prominence to political entities that can afford to spend more. With political campaigns becoming more centred on the advertising opportunities through social media, as argued to be the new normal by Dommett and Temple, one could expect normalisation to be the lens to view this area of online campaigns (Dommett & Temple, 2018). As dominant parties, with more finances for advertising and the resources to plan more advanced and targeted ways to serve adverts, will of course be dominant in this area. However, it seems equalisation isn’t dead on social media, with both Ganious et al (2007) and Southern (2015) suggesting while normalisation effects are prominent on social media, both studies found equalising elements within social media, as third parties have a higher rate of engagement in social platforms where costs are lower. In addition, there is also evidence to suggest social media can foster citizen-initiated campaigns that have the potential to help smaller parties that face a resource deficit (Gibson, 2013). So while normalisation is prevalent on social media, one could come to expect some equalisation effects on social media too. However, the Normalisation/Equalisation thesis lens has not yet been applied to activities outside of direct political campaigns, for instance, data activists. Yet, as will be discussed later, it is a perspective that matches well with some of the activities of projects such as Who Targets Me and ProPublica’s advert collector.

Transparency and Data Activism in Western Liberal Democracies

Data activism is a campaign form based on digital, and often internet, technologies that seek to challenge current power relations through the use of data. In recent times, this form of activity has been said to be moving slowly from a small elite few, to a more active non-skilled form where new tools have allowed the everyday person (or amateur activists) to take part, resulting in a substantial growth of this type of activism (Milan & Gutiérrez, 2015). For instance, groups and companies have sought to lower the technical barrier to data activism by providing tools or services to counter the growing perception of governmental and private privacy violations (Birchall, 2012; Sarikakis & Winter, 2017; Bowcott, Sept 13, 2018). One such example is the proliferation of Virtual Private Networks (VPN) with increases in its use in use in both the UK and US based on the notion of the protection of personal information: at least to those users within the tech industries that understand the implications of such a service (wombat, 2017).

Within this type of activism is the belief that the greater access and publication of data about society and politics has become an important element of levelling the political playing field, enhancing the networked free press, and making politics fair (Winseck, 2013:166). Camaj goes as far to argue that the two concepts of data activism and journalism have become blurred in their aims: to make politics accountable and to show the public what really goes on in politics (2016). The apparent linkages between

---

3 A VPN is a service that allows users to appear to be from a different geo-graphic location. It can be used to access internet content not available to the users own country, to protect a user’s identity, or to appear anonymous on the internet.
transparency and data activism can be explained as both roughly appeared at the same time. As the technology to make data activism possible came around, there was also a resurgence of the transparency movement, one which has reappeared due to the neoliberal idealism of the 1990’s that argued that transparency and accountability was a necessity for a healthy and prosperous democracy, and is a necessity to keep populations informed on their governments, and unable to hide mistakes or corruption (Florini, 1999; Hetherington, 2011). The two movements often intertwine, as digital transparency activists have lobbied for governments freely disclose more information to the public, the creation and publication that data as activists themselves, or so create systems where information can be shared freely without censorship (York, 2013). There is also a political context for why the two movements are placed together. Winseck argued that the rise of data transparency movements can be explained as a reaction to the growing controls and regulation of the internet; the monopolisation of online services; disproportional legal responses to issues of copywrite protection; and the increasing militarisation of the Internet through intelligence agencies (2013).

One of the most prominent examples of this type of activism is the publication of secret documents. The most famous of these is through Wikileaks; an online organisation that publishes classified government information in the form of data dumps. Data releases such as the United Stated Diplomatic cable leak, a release of over 250,000 secret diplomatic cables sent by U.S embassies causing a diplomatic crisis and disclosing the use of diplomats as intelligence gatherers (Leigh, November 28, 2010; Gill & Spirling, 2015). While another notable leak by the same website disclosed the July 12th, 2007 Bagdad Airstrikes video, also known as Collateral Murder. A 39-minute video from U.S. Military forces showing several people, including war correspondents Saeed Chmagh and Namir Noor-Eledeen, being killed by U.S action. The result was, at minimum, a domestic debate in the U.S. about the role the military in Iraq, and as an example of the reality of the war (Marouf, 2012). While this type of data activism is the most visible – it is also one of the most criticised. With ethical issues being raised about data dumps, such as the accidental release of data that violates someone’s privacy, its use as a political tool with unequal or questionable sources, or even endangering the lives of intelligence operatives whose cover had been blown due to leaks (Hansen & Flyverbom, 2014).

However, the range of activity undertaken by data activists varies massively and does not necessarily equate to the use of computer hacking. Data activism projects have also included the use of citizen scientists to the map the radiation levels and other related pollutants from the Fukushima Daichi Nuclear Power Plant disaster in 2011, to provide transparent and accurate data to the general public — importantly out of the hands of those that might want to downplay any radiation issues (Brown et al, 2016). Activists of this type have also sought to capture and publish data of corrupt acts that the state might not be adequately solving. For instance, years before the Arab spring, Egyptians had used the photo sharing website Flickr to create a directory of police brutality called ‘مجموعة الجلادي’ or Piggipedia4 (York, 2013). The ‘Targuist sniper’, another demonstration of this type of an advocacy, filmed police officers in Morocco taking bribes and published them to YouTube leading to the arrest of nine corrupt policemen (Abdo, November 12, 2007). Since 2003 mySociety has sought to provide several services to make the inner workings of government more transparent. Such as their WriteToThem tool, that seeks to make contacting political representatives easier, alongside tracking the response rates of MPs, or their WhatDoTheyKnow, a platform for submitting Freedom of Information requests, and publicises the responses to the wider public (Gibson, CantiJoch, & Galandini, 2014). The use of data activism in society is not limited to the above examples but demonstrates that it’s use in society has had important effects, as a growing number of online initiatives set up by citizens have sought to capture and disclose matters of society that they feel needs addressing (York, 2013).

Data and transparency activists have also become active in the area of political campaigns. The first type of this activity can be found in election promise trackers, such as that of electionpromisetracker.in in India,

---

4 The Group can be accessed online at: https://www.flickr.com/groups/piggipedia/
govtracker.co.uk in the UK, the TrudeauMeter at https://trudeaumetre.polimeter.org/, and a number of efforts in the US, including homebrew efforts such as trumptracker.github.io (Mellon et al, 2018). These use election promises made by politicians and parties and track them with real government initiatives to see how they measure up. Traditionally one would have seen this type of action undertaken by members of the press, but in many cases is now undertaken by organisations or activists.

More recently, data activists have noted the concerns behind the use of data, dark adverts, and algorithms to target citizens to deliver political messages. One of the first instances of this was during the 2012 Mexican elections, members of the #YoSoy132 movement began tracking the use of campaign activity that involved the use of online bots to support candidates and the hiring of trolls to attack their opponents (Treré, 2016). However, a new focus for activism was to uncover how exactly political campaigns had used social media; an area that became evidently clear with the introduction of unpublished posts that lead to the formation of dark adverts as discussed above. The issues for democracy have been highlighted several times, including one report by the Campaign for Accountability that found that both Facebook and Google had potentially violated a number of campaign laws during the 2016 U.S. Presidential election by supporting the campaign teams by offering ‘embedds’ Google or Facebook hired employees that would support helping campaign teams (CfA, 2018). The methodology for this report used data from LinkedIn to analyse the activity of Google’s and Facebook employees to see how closely they had been working with both the Democrat and Republican campaign teams (ibid, 2018:18). While other activists have sought instead to collect data on parties campaigns to make available online such as Livingroomcandidate.org that digitally stores and makes available all presidential campaign commercials; or Politicalresources.net that seeks to host all UK party manifestos from 1945 onwards.

In the context of dark ads on social media, a number of initiatives have been set up to collate and publish political adverts served to Facebook users. Of these, three are of interest. The first of which was set up by the New York Times in 2016 (New York Times, April 14, 2016). Seemingly, this is the first attempt at political ad collecting. Two other services has since come out to collect political ads in roughly the same way, Who Targets Me and ProPublica's Political Ad Collector, both install a Brower add-on that automatically collects political adverts alongside a minimal amount of demographic data to determine who is paying for political advertising, who do they seek to target, and with what messages. The issue faced by transparency activists is the highly asymmetrical relationship with those who serve the adverts (i.e Facebook), and thus the political campaigns that use the service. Ultimately, Facebook controls who can access data on advertisers and how campaigns are ran. Some have argued that Facebook chooses not to disclose how advertisers target users to protect advertisers, and thus their revenue streams (Hern, July 17, 2018). Indeed, researchers that have access to targeting data on Facebook adverts have been able to understand, and make public, strategic decisions within political campaigns that some might want to keep undisclosed (Edelson et al, 2019). In January 2019, Facebook went on to change the code in their website to purposefully break political advert collectors (Merrill & Tobin, 2019). This raises some interesting questions in the context of the normalisation vs equalisation thesis. If the internet is supposed to level the playing field across political parties, should it too also level the playing field between all political actors?

**Method**

An investigation of the tracking of microtargeting and dark ads by data activists through the normalisation vs equalisation thesis is an important one given the relative power asymmetry, and recent accusations that Facebook has taken action that stops the collection of political adverts. On one side, you have a seemingly small number of data activists trying to collect and publish adverts and targeting data from political campaigns on Facebook. On the other, you have the Facebook platform with an incentive to encourage political parties to use their platform for advertising, giving campaign strategists the ability to
only show adverts to targeted voters, while not disclosing the wider campaign strategy. These two are seemingly in competition, but at different resource levels: Facebook is seemingly the dominant entrenched party, while data activists are at a much lower level in terms of status and resources. While the normalisation vs equalisation thesis is normally concerned with party positions, in the context of data activism during electoral campaigns, it is one that certainly makes sense to use here. With Facebook essentially replacing the role of the gatekeeper to audiences, but with a financial rather than editorial set of rules for who they promote as seen previously with the media. Ultimately, this paper seeks to understand to what extent these activist groups are capable of counteracting campaign-related dark advertising on social media.

To understand if data activists are afforded the ability to compete on a level playing field with advertisers, we selected three elections in which groups have sought to collect and publicise the use of micro-targeting and dark adverts in political campaigns: the 2016 U.S presidential election; 2017 U.K general election, and U.S 2018 Midterm elections. This is undertaken through a case study of the data collection and publication attempts of the New York Times Ad collector, Who Targets Me, and ProPublica’s Political Advert Collector (PAC). Due to the nature of the research, and the limited case studies, this research used elite interviews with individuals who had been active with either of the two projects alongside background collected research such as news articles to build each case study – a method that would seemingly fit with this level of study (Halperin & Heath, 2012:255). The questions during the interview (see appendix) asked about the participants involvement with the project, and the participants history and skills and their interest in the area, the projects history, it’s aims, it’s resources, if they feel adequately resourced, and how effective they judged their project to be. Due to the nature of the interviews, several probing questions were also asked where relevant to drive more information from the user. The research uses the guidelines from both Halperin & Heath (2012) and Hermans (2004) for the creation of this research method. With precautions made for common interview effects by being aware of these issues when communicating with participants before the interview. (so far), this paper uses interviews from two key elites. First is Jeremy B. Merill, who is formally of the New York Times and ProPublica, and had a major role in both the New York Times Political Ad collector and ProPublica’s Ad collector.5 The study also included a search for all news studies using Google News that mentioned either project by name. This was done to create a measure for how successful each project had been in terms of impact. These interviews and the background research were used to create case studies (with supplemental information when available) regarding each attempt to make political campaigns on Facebook more transparent. Each of these cases were then studied in a framework for the analysis of normalisation and equalisation found in Ward and Gibson (2009). The use of this framework will also be useful for testing if the Normalisation/Equalisation lens is a suitable method for understanding the ability of data activists to make data available to the public. The four areas for assessment in the Normalisation–Equalisation hypotheses (Ward & Gibson, 2009:36–37)

**Systemic & Technological:**

- **Media Environment:** How able was each project able to get noticed in the media? Are the campaigns able to overcome traditional issues faced by third parties in regards to media ownership and media gatekeeping
- **Political Environment:** Did the project manage to overcome the entrenchment of institutional frameworks?

**Organisational:**

---

5 (Note, the current draft of this paper does not yet have the full interview with Sam Jeffers from Who Targets Me. It will be updated in due course)
Organisational Capacity: Did the projects face issues with a lack of Staff time, skills, or finances?
Organisational Incentives: How did the campaigns ideology benefit from the ideologies often entwined with the internet? Did they struggle to connect with their target audience? Organisational age – did the age of the project produce any significant issues? Organisational status – did the projects have enough status to be taken seriously, be trusted by voters, or to be noticed by the media?

Ultimately, this research created a number of case studies of by data activists to collect and publish political adverts alongside targeting information and make it accessible to the general public, when it would otherwise only be accessible to those who where served it.


In many ways, the technical use of Facebook by U.S Presidential campaign strategists in 2016 was similar to that of 2012. While 2016 saw an increase in the importance of social media in the campaign, many of the strategies was adaptations to those seen in 2012 (Gualti & Williams, 2013; Persily, 2017). What was new in the 2016 Presidential election was the implementation of behavioural science, data analytics and innovative advertising technology mixed in with the ability to show certain adverts to extremely granularly targeted voters (Bossetta, 2018). While the use of data in elections has been seen since the 1940s (Anstead, 2018), and the statistics modelling used is a variation of what can be found in the 1980’s, Sumpter argues that what differed was the quantity of data used – to levels never recorded from previous elections (2018). However, the true extent of how the campaign was run only truly came out after the election (Madrigal, Oct 12, 2017). Alongside this massive amount of data was the new ability to hide adverts, alongside campaign messages that is said to negatively impact democracy, such as voter suppression campaigns, (Persily, 2017). Following the election, further information has come out regarding the distribution of fake news via Facebook adverts, and at least 3,500 adverts paid for by Russian intelligence operations designed to help the Trump campaign (Dutt, Deb & Ferrara, 2018). What these case studies demonstrate is the negative aspects of the political campaign undertaken through advertising could have been uncovered and acted upon during the 2016 campaign itself if the adverts and how they are targeted were made transparent through either self-disclosure by Facebook, or by a political advert collector by a third party.

In April 2016, the New York Times launched a political advert collector. Coded by Jeremy Merrill, the project sat between a project by journalism and data activism. The system would work by a browser plugin that produced a button over every advert on Facebook on volunteers’ screens, the idea being that users would click the button every time they saw a political advert, and this would send it to the New York Times, adding it into a database (NYT, April 14, 2016). Unlike the other advert collectors mentioned in this paper, the process for users to report the advert was somewhat laborious, requiring users themselves to click whenever they saw a political advert in their Facebook feed (where as other projects did this automatically once the plugin was installed).
The project was initially started after Merrill and his editor discussed ‘politics-linked journalism projects’ and the interest they both had in microtargeting, which at the time was ‘widely discussed but barely understood campaign tactic’.

Our aim was to collect microtargeted political ads and, at minimum, report on any nefarious tactics or obvious trends. We thought candidates might be taking opposite policy positions when talking to different demographic groups, for instance. (This doesn’t appear to be the case, I’ve since learned.) Before we started, we had additional ‘stretch’ goals that were mostly hypothetical: we didn’t know if we could perhaps reverse-engineer the demographic targeting choices used by campaigns or find obvious A/B testing examples that could be shown side-by-side. Once the project got started, I pivoted towards focusing on the ‘Why am I seeing this?’ targeting rationales provided by Facebook. (Merrill)

The project also seemingly had ideological goals too. During the interview Merrill was clear that the project was one based on increasing transparency in politics:

Transparency is key to democracy. While we have (mostly) good norms about transparency in politics conducted offline, online transparency lags – and it poses new risks via microtargeting, etc. So in order to properly hold politicians, corporations and governments accountable in a new digital era of politics, as we had in previous analog eras, we need digital tools to help. (Merrill)

However, the project had minimal organisational support. Merrill stated during the interview that creating the project was mostly a one-person activity with small amounts of help by other members of staff. When the plugin was finished, the New York Times produced one news story for promotion, a few callouts in ‘web NYC stories’ asking people to participate, and a few social media posts. One of the major issues with the project was the amount of resources dedicated to promoting the tool.

I would’ve loved much more! [finances/resources] We could’ve used more resources for promoting people to install the tool (whether ads on Facebook or call-outs in NYT stories). And we could’ve used more reporting resources to write about what we found. (Merrill)

Skills was not an issue with the project, Merrill himself had the ability to code from his background in computer coding and data analysis, and he was also supported by one of the other editors who also had the relevant skillset to assist. However, the experience by Merrill may be a slight oddity, at least in 2016.
Research from media suggests that there is a significant need to teach elements of computer science to journalism students as the changing media landscape requires journalists who could approach information gathering as more of what we do as a society is done online (Spinner, 2014; Chimbel, 2015; Broussard, 2015). So while the New York Times as a journalistic entity might have been able to create a political advert collector, it is hard to judge if other papers could also replicate it without bringing in outside assistance.

The results of the project were limited. During the period it was active, the project collected ‘about 10,000 ads from 1,000 participants’ according to Merrill. Significantly less than the supposed 50,000 different adverts, per day, to target voters by the Trump campaign (Beckett, Oct 9, 2017). The limited results are seemingly due to the lack of resources available to the project to promote it outside the readership of the New York Times, or the organisations availability to partner up with others to help promote it. There was also a contextual element, while people talked about the issues of microtargeting and dark ads, these were not often in a political context. Merrill during the interview claimed that if they knew the types of adverts that would be served during the election, and how they would be targeted with data, then much more resources and public attention would have been given to the project. The output of the project was limited, mostly due to the small number of volunteers, the project resulted in one news story on advertising trends – which did not link to the data collected (c.f. Merrill, Nov 7, 2016). The projects biggest failures one could argue was the failure to detect and publish aspects of the advertising during the 2016 election campaign considered nefarious – such as voter suppression, fake news, or adverts by non-U.S. entities that has only been uncovered by other methods.

**Case Study Two: Who Targets Me? – 2017 General Election**

Who Targets Me was launched by Sam Jeffer and Louis Knight-Webb, both whom have previous software development experience. Sam is also a graduate from the LSE and the Harvard Kennedy School. The project was launched after the announcement of the 2017 GE with the mission statement ‘For the good of our democracy, it’s time to throw some light on dark ads.’ (Who Targets Me, 2017). The project was built in the context and aftermath of the 2016 EU Referendum in which the campaigns were criticised for their use of targeting and data. The project also used a browser add-on for Chrome, that automatically collected any advert a user would see on their screen while on Facebook – in comparison to the much more hands on method seen by the NYT ad collector. Adverts would be sent to the main database, and to the projects partners at the Bureau of Investigative Journalism, and the Media and Communications department at the LSE. The Project also worked with The Guardian, who promoted the project and encouraged readers to download the browser plug-in (Booth, May 3, 2017). The project reported 11,421 volunteers by the final day of the General Election campaign, collecting 783 unique political adverts that appeared 16,109 times in a total of 1,314,004 impressions (Anstead et al. 2018).

Ultimately, no comparison has yet to be made large, or small, of a percentage of total adverts served was captured by the advert collector. Although working from the assumption that only sixteen thousand of a potential 45% of the UK population who are on Facebook used the service it certainly didn’t collect every advert served (OfCom, 2017). However, the number of news articles found that went on to cite the data collected during the election was significant, and much higher than those found for the New York Times advert collector (and from the initial search also more than ProPublica’s too). The Project was widely mentioned in several articles from BBC news, The Guardian, The Metro, The New Statesman, Huffington post, the Times and more (c.f. Ward, June 2, 2017; Booth et al, June 2, 2017; Jackson, June 6, 2017; Demianyk, May 14th, 2017; Gibbons, May 4, 2017). Suggesting that the bulk of

---

6 This section will be updated as interview data is received.
impact from this project should not just be measured by the data it has collected, but also the amount of national interest generated to the issue of dark advertising and microtargeting in political campaigns.

**Case Study Three: ProPublica’s Political Advert Collector – U.S. 2018 Midterm elections.**

Unlike the 2016 U.S Presidential Election, data activists were in a much better position to monitor the 2018 Midterm elections. While during the U.S Presidential Election, issues where wildly talked about, but little understood, the 2018 Midterms had a much better understanding to the issues to democracy faced by the new campaign strategies found on social media. For instance, there were new tools to collect and publish fake news published to Facebook called the *Junk News Aggregator* (Liotsiou, Kollanyi & Howard, 2019). This is just one of the tools launched ahead of the election, others included tools to monitor: foreign interference in the election, user tracking technology on political websites, and in the context of this paper, political advert collectors too (Bossetta, 2018; Patel, 2018). This was also the first election after Facebook released its own political advert archive. The archive was launched after a mix of public pressure following several damming reports on Facebook adverts, and legal pressure after the state of Washington filed a lawsuit after it was found Facebook broke its transparency rules for election campaign advertising (Washington State. Office of the Attorney General, 2018; Mehrotra, Dec 19, 2018; Leathern, May 24, 2018). It should be noted that while the Facebook archive effectively stopped dark adverts, it still did not publish how adverts were targeted: something provided by both Who Targets Me and the ProPublica’s political advert collector. Indeed, researchers who have attempted to use Facebook’s data on political adverts found significant problems in researching it, including vague data, limitations on what data is provided and what can be published, or even being shut out of the data mid research (Edelson et al, 2018).

ProPublica’s political advert collector was launched in 2017 ahead of the German Federal Election by Julia Angwin and Jeff Larson, with Jeremy Merrill joining before the 2018 U.S. Midterm elections. The project was started following inquiries about the use of Facebook adverts in politics in the 2016 EU Referendum and the 2016 US presidential election campaigns, alongside the misuse of data by Cambridge Analytica which brought significant public attention to microtargeting in political life (Merrill). The aim of the project was to save adverts during elections that were limited to the audience the advertiser has chosen, adverts that are only shown for a limited time, mislabelled adverts, and to stop adverts going narrow, false, and unaccountable (Angwin & Larson, Sept 7, 2017). The project took a much more multinational approach and was active across numerous countries – thus while this research focuses on the 2018 U.S. Midterm elections, the collector also has been used in elections in Australia, Germany, Canada, Finland, Switzerland, Denmark, and Sweden (Valentino-de Vries, Feb 8, 2018). The significant public light on the issue also brought more resources to the project. This allowed the political advert collector by ProPublica to be much more advanced than the previous New York Times advert collector. For instance, the browser plug-in would automatically collect all adverts, without the inconvenience of users having to click (like with the NYT collector), alongside the targeting data provided by Facebook at the point a user is served an advert. The project also received significantly more support from journalistic partners, including being publicised by both local and high authority journalism titles asking readers to download and install the ad-collector (cf. CBC, April 17, 2018; Silverman, July 17, 2018; Cameron, July 11, 2018).

ProPublica’s collection of adverts was much more successful than the previous NYT ad collector, Merrill commented that ‘That project was an order of magnitude larger, collecting over 100,000 ads in the U.S. and generating far more stories’ (Merrill). The collector has been much more successful at not only collecting adverts, but also highlighting adverts not found within Facebook’s own archive and has found interesting targeting mechanisms in place, such as the Taxi company Uber targeting members of the Black Lives

---

7 The whole 2.76GB database of political adverts is available to be downloaded here:
https://www.propublica.org/datastore/dataset/political-advertisements-from-facebook
Matter group, or that Facebook allows political campaigns to target Facebook users labelled as ‘jew haters’ (Angwin et al, Sept 14, 2017; ProPublica, Sept 25, 2018). The project’s not only collected a large amount of data, but the results have also received more media attention than the NYC collector, particularly from journalistic titles that previously promoted the advert collector (c.f. Samuels, Oct 31, 2018; Privacy International., Nov 1, 2018; Vice News. Nov 5, 2018). The data has also gone on to be used in academic research (c.f. Edelson et al, 2019).

**Case study note: Changes in political advert collection since 2018**

The data-collection landscape encountered by these three case studies is now very different to what will now be encountered by similar projects. In January 2019, Facebook implemented new code into its website in an attempt to stop political advert collection by both Who Targets Me and ProPublica (Waterson, Jan 27, 2019). This ultimately caused the ProPublica collection project to go dormant (Merrill), and for Who Targets Me to shift focus (Jeffers, Feb 2, 2018). This leaves Facebook own Advert Library as the only source of political adverts. This comes back to the issue of data and data access. Some research’s attempts to use Facebook’s advert library have been frustrated by the service – some have noted how access to the data is uncertain with the New York University Tandon School of Engineering becoming blocked from the archive after two weeks of access (Edelson, et al, 2018). The data they did receive was much more limited than that from advert collectors, as Facebook only releases vague range data rather than exact figures. Other researchers have found that the advert library actively stops the collection of data after a certain number of access requests (Albright, 2018). The Facebook library also does not contain any targeting information – a large reason for the collection of political adverts.

Website received after a data collection attempt of Political adverts (Albright, 2018).

**Discussion**

Throughout each of the three case studies, there has been a strong ideological drive to provide information and data on the activities of political campaigns during election. This driving ideology can be best describing by a quote found in Full Facts report on misinformation in society:

> When an election stops being a shared experience, democracy stops working … We are used to thinking of adverts as fixed things that appear in the same way to many people. This idea is out of date (Full Fact, 2018).

Each of the case studies has taken very similar approaches to the particular problem of political adverts on Facebook. Each employ a methodology to create tools which allow citizens to help collect the information displayed on their Facebook feeds and send this to the data activists. In all three they sought the use of browser plug-ins, the first iteration of this taking a manual approach with the NYT collector having users click on a button, but with later iterations using more automated approaches to streamline the process.
Using the framework of analysis from the methodology, we can see that the first case study, the New York Times Political Advert collector was far from successful. Overall, the project collected a minimal amount of adverts during the 2016 U.S Presidential election, with very little press alongside the results. This analysis might be somewhat unfair, after all, this research is undertaken with hindsight. However, within the confines of the framework of analysis, the case shows that without prior public knowledge or engagement with an issue, a project that requires citizen input will fail due to the resources (lack of engaged citizens) and status of the project overall. Suggesting without the all-important systematic and organisational incentive factors behind a project, there will be a normalisation as many campaign strategies were able to run their campaigns with impunity at least at the time of the election campaign. The layout of the Facebook advertising platform and the conduct of those using it during the 2016 Presidential election suggests a normalisation effect – with the technologies afforded to the internet significantly benefiting the campaigns. Without the right political environment or incentives for citizens to be concerned about transparency, the project failed to pick up enough media or citizen attention.

ProPublica’s political advert collector and Who Targets Me however suggests a very different landscape. In both of these cases, each project drummed up significant media and citizen attention. They carefully used media and institutional partners for both media stories about what the activists aims are, and to provide a certain level of organisational status. The political context was in their favour following on from the contexts of the 2016 U.S. Election and EU Referendum. Furthermore, both projects had the organisational capacity to deliver the projects. To a degree, both projects met with their aims, and after the elections had happened brought significant attention to the political campaigning strategies. However, through the analysis of the news attention, much of the public’s attention to the issues of transparency in the U.S Midterms was before (to generate volunteers), or after the election campaigns as the projects released reports on the data they had collected. While Who Targets Me was able to generate stories during the U.K General Election campaign, the number of adverts collected by volunteers was much smaller in number (783 compared to the over 100,000 of the U.S midterms). This suggests that in elections where sophisticated campaign strategies are at play, such as the scenario of 50,000 different adverts being served per day, many of these projects do not have the technological capability to process and report on data. So, when the complexity of advertising on the Facebook platform increases, the load also dramatically increases for data activists. This would Indeed suggest an increase of the normalisation effect as the technology behind how adverts are delivered become more sophisticated.

However, this isn’t to say that the projects had ultimately failed – far from it. The data collected has been reported on numerous times and has been the basis of much academic research – much of which would invariably hit the ears of policy makers. For instance, research presented to the U.K Parliament from Who Targets Me as part of its Fake News Inquiry (Knight-Webb, 2019), and that provided to shape the Honest Ads Act in the United States and C86 in Canada (Merrill). Suggesting that while the projects suffer normalising effects from being able to make political adverts transparent during the campaign, with time to analyse the data collected they have long term equalising effects.

One also must consider to what degree these advert collectors were ran by so called outsiders, or third parties. The people who ran these projects either came from established institutions (NYT or ProPublica) or have a history in politics or connections to those who do. It seems, at least on the face of it, that these projects required more than just having the necessary skills to be able to program and code. How these projects became successful (particularly ProPublica’s and Who Targets Me?) could be argued to be due to the project leaders’ connections and ability to convince others to promote the projects. This is in stark contrast to the New York Times advert collector that failed to gain significant traction. This could be compared to other findings such as those by Hindman (2008) who argued that rather than acting as a democratising force, the most popular and read political blogs are by the highly educated, professional, middle class with well honed journalistic skills. To some degree we find a similar finding with our case studies.
The final question of this paper is if the previous model for analysing Normalisation vs Equalisation still apply in the context of data activism? The measures used for addressing equalisation vs normalisation often looks at factors such as skills, finance, systemic factors etc. This paper has highlighted one issue throughout: actual access to real time data. When we come to measure normalisation vs equalisation, we must also consider the access to data as a key measure. Data isn’t always something one can purchase, but must be acquired, often through institutional entrenchment of those collecting the data. While each project managed to collect data, each also suffered from the same issue in that the projects require volunteers, and thus self-selecting. This made the data collected non-representative of the overall states electoral base, and the adverts they would receive. One must question what percentage of the electorate, who might be particularly susceptible to misinformation, would also volunteer to install a browser plugin. Thus, one of this paper’s findings, that can be applicable to the overall theoretical perspective used in this analysis, is that the access to data that should be an additional measure used in analysis. As political campaigning becomes more complex, and more data is used, we can expect that data is another measure in which normalisation vs equalisation can be measured. From this paper’s findings, we expect that data access would favour those with the resources and links to either collect data themselves or be in a position to bargain it from others. For example, data captured by political parties while canvassing, or that offered to political campaigns for targeting. We also found that the level of data a campaign holds is a distinct measure. While it requires both finances and skills to be able to collect and process data, we find there are other factors such as institutional entrenchment that allow campaigns to access it – this is an area in need of further academic research.

Conclusion

To conclude, this paper used the normalisation vs. equalisation lens to measure the success of data activism campaigns during elections. In it, we have highlighted that previous measures to analyse internet technologies are still relevant. With the analysis of the first implementation of a political advert collector finding that the ability and skills to produce a new tool does not automatically mean it will be successful, as the New York Times advert collector still encountered barriers through the political environment and organisational incentives. However, the second and third cases does potentially highlight that a new measure is required to understanding the success of internet campaigns, be it electoral campaigns or activism: Data. Without access to the relevant data, campaigns will face significantly more skills, and resource costs against those with the data such as targeting information. However, as seen from the case studies, data, while linked to skills and finances, is once again distinct, as Facebook has clearly sought to restrict who can access important bits of information regarding the conduct of political campaigns. We further found that within the existing framework of normalisation & equalisation, we can come to think of social media as the new gatekeepers to audiences, except through much more direct mechanisms, being financial rather than editorial. Ultimately, this paper concludes that the advertising platform delivered to political campaigns give strategists a short-term normalising effect – with the more complex a campaign becomes, the harder it is for projects to report in real time. However, this normalisation might be short term. As many of the data advocacy projects have gone on to successfully argue that measures should be introduced to increase transparency to legislatures. Some have argued that the work of political advert collectors has gone on to directly influence new advertising legislation such as U.S. Honest Ads Act and C86 in Canada that might deliver the aims and goals of the projects.


Demianyk, G. (May 14, 2017). Tories Target Jeremy Corbyn In Most Prominent Facebook Election Advert. *Huffingtonpost*. Available at: http://www.huffingtonpost.co.uk/entry/facebook-ads-who-targets-meron_d_5914bef7e00f308c3ebb


Gibbons, K. (May 4, 2017). Facebook tool will show how politicians are targeting you. *The Times*. Available at: [https://www.thetimes.co.uk/article/facebook-tool-will-show-how-politicians-are-targeting-you-jb0fn52c6](https://www.thetimes.co.uk/article/facebook-tool-will-show-how-politicians-are-targeting-you-jb0fn52c6)


Privacy International. (November 1, 2018). An analysis of Facebook’s political advertising in the 2018 US Midterms. Available at: https://privacyinternational.org/examples.pdf


Samuels, A. (Oct 31, 2018). Beto O’Rourke is betting big on Facebook ads: He’s spent more than $6 million in six months. Texas Tribune. Available at: https://www.texastribune.org/2018/10/31/texas-beto-orourke-ted-cruz-facebook-ads-propublica/


Wong, J.C. (March 19, 2018). ‘It might work too well’: the dark art of political advertising online. The Guardian. Available at: https://www.theguardian.com/technology/2018/mar/19/facebook-political-ads-social-media-history-online-democracy
