Voter preferences, voter manipulation, voter analytics: policy options for less surveillance and more autonomy

Jacquelyn Burkell

Department of information and media studies, The University of Western Ontario, London, Canada, jburkell@uwo.ca

Priscilla M. Regan

Schar School of Policy and Government, George Mason University, Fairfax, United States, pregan@gmu.edu

Abstract: Researchers in psychology have long known that preferences are constructed in the decision-making process, influenced by choice environments that trigger unconscious biases and heuristics. As a result, choices, including those of voters, can be manipulated by political information. Personalised political messages, designed to influence based on detailed personal profiles, can undermine voter autonomy. We suggest that these practices should therefore be regulated, and discuss policy options and approaches, specifically the appropriate balance between freedom of political speech and privacy rights and interests, the implications of voter analytics for the electoral process, and how and by whom sophisticated voter analytics practices should be regulated.

Keywords: Political autonomy, Micro-targeting, Privacy, Surveillance, Voting behaviour

This paper is part of Data-driven elections, a special issue of Internet Policy Review *guest-edited by Colin J. Bennett and David Lyon.*

Voting (including the decisions of *whether* to vote, and, if so, *which way* to vote) is the cornerstone of the democratic process. A vote (or the decision *not* to vote) is also a *choice*. Central to the democratic value of voting is the ability of the individual to exercise autonomy in making this choice; indeed, the secret ballot recognises and emphasises the need for privacy in order that voters can make an autonomous choice (Evans, 1917). Traditional political advertising is an obvious and very public tactic to influence voter preferences. The impact of such messages can be increased by two forms of personalisation: message targeting (directing messages to selected sub-populations) and tailoring (developing different versions of a message designed to appeal to different people) based on demographic, behavioural, or psychological characteristics (see, e.g., Hirsh, Kang, & Bodenhausen, 2012). Targeting and tailoring have long been used to increase the impact of political messaging, including speeches and broadcast media advertising, in the offline context (see, e.g., Miller & Sigelman, 1978). Recently, *online* targeting and tailoring techniques are being used in new, subtle, and powerful ways to design and deliver political messages that have an even greater potential to influence voter behaviour and voter choices.

Today's political operatives develop highly detailed voter profiles, integrating demographic information, information about the economic, social, and political activities of potential voters, and detailed records of online and even offline behaviour into a rich voter profile that can also reveal, through powerful data analytics, additional insight into thoughts, beliefs, and psychological characteristics (see, e.g., Kosinski, Stillwell, & Graepel, 2013). The resulting voter profiles can be combined with insights from psychological studies to develop persuasive messages that are tailored with respect not only to the content but also the form of the message (e.g., appearance, specific language, timing of the message), designed specifically to appeal or persuade based on specific recipient characteristics (see, e.g., Issenberg, 2016). As Calo (2014) has demonstrated in the context of consumer marketing, these techniques can take advantage of cognitive limitations and vulnerabilities to shape consumer decisions.

Personalised political messages, using the same techniques Calo references, are being employed in the political realm and can shape political decisions - or, as Slovic (1995) argues, these techniques can be used to construct voters' expressed preferences. The techniques go beyond targeting and tailoring messages based on demographic variables (age, gender, party affiliation) and/or social, political and economic activities, to designing messages for and delivering messages to individuals based on psychological variables such as personality characteristics (extroversion, neuroticism, authoritarianism, etc.), attitudes, and interests, and other psychological information that is revealed or can be inferred (see, e.g., Hine et al., 2014). Targeting and tailoring on these and other psychological variables is generally known as psychographic profiling. Effective use of psychographic profiling information includes the manipulation of the form, content and timing of political messages, often using strategies that have been identified in empirical research in cognitive psychology and decision making as increasing message impact. Manipulated messages can be designed to activate implicit attitudes and biases, with effects that are likely to be subtle, and operating at an unconscious level.

It is important to emphasise that these subtle techniques of persuasion and even of manipulation are enabled by equally subtle techniques of surveillance, often taking the form of increasingly sophisticated behavioural tracking techniques. In the consumer marketplace, it may well be the case that the consequences for consumers (e.g., paying a bit less for something or buying on impulse) are relatively small, and the upsides for firms are likewise marginal (Calo, 2014, p. 1002). The effects in the consumer marketplace, therefore, may be of little importance in terms of the number of people whose behaviour is affected and in terms of the impact of those effects on the marketplace; however, it is important to recognise that consumer profiling

practices do have implications beyond consumption. In the political arena, however, affecting the political preferences, decisions, or actions of even a small proportion of voters in a competitive election could be critical to the outcome. Such manipulation of voters raises fundamental issues of democratic theory.

When political communicators have the advantage of deep and detailed knowledge about the public and when they leverage that information to develop and deliver political messages designed to persuade specific individuals based on what is known about their demographics, personality, attitudes, beliefs, etc., *and* when those messages take advantage of persuasive principles drawn from the empirical literature in order to exploit a predictable interaction between individual and message, the result is an unfair system that undermines voter autonomy. We refer to political ads that employ tailoring and/or targeting, manipulating the timing, content and form of messages, and that are employed, not in the interests of informing or even persuading voters, but rather with the goal of appealing to non-rational vulnerabilities as revealed through algorithmic (and particularly psychographic) profiling. We argue that the use of such ads warrants policy intervention, since they have the potential not only to affect individual autonomy and an individual's ability to render her genuinely own voting decision, but also to contribute to the fragmentation and polarisation of the electorate – both results that are antithetical to democratic theory.

Drawing upon our backgrounds in psychology and political science, we previously explored the consequences of the individualised, highly selective, and structured information environment on voter preferences, examining the ways in which personal profiling could be used to manipulate voter preferences and thus undermine voter autonomy (Burkell & Regan, 2019). In this article, we refine that earlier analysis of the civil liberties, privacy and democratic values questions and extend the analysis to focus on psychographic profiling in political advertisements. We argue that this type of profiling in particular should be regulated to protect voter autonomy and mitigate political polarisation. We first discuss the development of personalisation techniques generally and its incorporation into political messaging. Next, we examine the particular issues associated with psychographic profiling as they arose in the commercial space and are now increasingly prevalent in the political arena. Finally, we identify policy options and approaches for regulating sophisticated voter analytics practices that employ psychographic profiling. We should note up front that delimiting psychographic profiling in a way that it can be separated from other forms of profiling is difficult; it is, in effect, only the latest stage in a continuum, as we note below.

ONLINE PERSONALISATION MOVES INTO POLITICS

Personalisation is ubiquitous in the online information environment, and indeed it is a natural technologically-mediated response to the overwhelming amount of information that confronts users online. While online search results *could* return 'everything' relevant to a users' query, some order must be imposed on the results, and personalisation helps to ensure that the information deemed most relevant to users is the information they are most likely to encounter, by placing that information early in the search results. Filtering techniques, including personalised filtering, address what Benkler (2006) has termed the 'Babel objection':

> Having too much information with no real way of separating the wheat from the chaff forms what we might call the Babel objection. Individuals must have access to some

> mechanism that sifts through the universe of information, knowledge, and cultural moves in order to whittle them down to a manageable and usable scope (p. 22-23).

Ranking algorithms for search results (e.g., Google PageRank) obviously and directly address the Babel objection. More subtle forms of information environment shaping that include some degree of personalisation are evident in recommender systems, which select a subset of items to suggest to users, and in online advertisements, which are directed to people who, based on demographic and behavioural information, are most likely to be interested and/or influenced. As the advertising industry realised the economic value of more finely tuned personalisation, and as more activities moved into the online environment, advances in computer modeling and behavioural economics identified more fine-grained methods for identifying individual characteristics that indicated preferences for certain products and services. In order to achieve a more personalised effect, complex algorithms select, sort, and prioritise information about the nature of the items themselves, the characteristics of the user and user interests/needs, and the match between item and user. User behaviour in response to this personalisation is folded into new analytics, feeding new algorithms and giving rise to even better predictions, in an iterative upward spiral of personalisation.

Personalisation depends critically on a great deal of information about users and their activities, gleaned through a range of surveillance techniques as well as through inferences, based on those data, about a person's cognitive and psychological styles. Less finely-defined *targeting* relies upon demographic profiling on the basis of (often observable) demographic characteristics like age, gender, religious affiliation, political affiliation, etc. More finely-grained *micro-targeting* relies upon demographic characteristics combined with data about the activities of individuals, including buying patterns, travel destinations, and social interactions. The most finely-defined *psychographic targeting* relies upon psychographic profiling on the basis of personality and behaviour data and inferences: personality (e.g., extroversion/introversion), values, opinions, attitudes, and interests. Some characteristics, such as sexual orientation, may fall in a middle range — one can think of this as being a continuum from externally observable and relatively explicit characteristics (descriptors) to internal psychological characteristics and tendencies.

As these personalisation practices became commonplace, with demonstrated effectiveness, in the consumer arena, they were picked up by political operatives, beginning primarily in the mid-2000s and particularly in the United States (Baracos, 2012; Bennett, 2016; Bodó, Helberger, & de Vreese, 2017; Burkell & Regan, 2019; Rubinstein, 2014; Tufekci, 2014). Issenberg in particular details how these practices were incorporated in political campaigns, culminating in the success of the 2008 Obama campaign with its techniques "that represented an individualized way of predicting human behaviour, where a campaign didn't just profile who you were but knew exactly how it could turn you into the type of person it wanted you to be" (2016, p. 326). Chester and Montgomery (2018) also document how these digital marketing practices evolved in the political arena and how they were used in the 2016 US presidential election including in campaigns closely working with Facebook and Google to target particular groups of voters and to direct ads in real-time and across devices. They quote Brad Parscale of the Donald Trump campaign as crediting these ads for Trump's victory: "Facebook and Twitter were the reason we won this thing" (Chester & Montgomery, 2018, p. 39).

Space does not permit an exhaustive list of examples of such political targeting techniques but a few that demonstrate the realities of those based on psychographic profiling well illustrate the terrain. Digital technologies allow the 'morphing' of two or more faces into a single image.

Bailenson, Ivengar, Yee, and Collins (2008) used digital morphing techniques to create new and individualised versions of candidate faces, subtly altering the candidate images to look more (but only *slightly* more) like the individual to whom the images were presented. Consistent with psychological theory that predicts increased liking of those who are similar to ourselves, viewers who received candidate images morphed with photographs of themselves expressed greater support for the candidates than did those who received candidate images morphed with photos of other people – even though the viewers were *unaware* that the images had been altered. In the 2016 US election political campaigns used Cambridge Analytica's model rating individuals on a five-factor personality model (including: openness, conscientiousness, extroversion, agreeableness, and neuroticism) to develop ads tailored to the vulnerabilities of particular voters (Chester & Montgomery, 2018, p. 23-4)

Although these practices may, as Chester and Montgomery (2017) point out, in essence be classified as micro-targeting or behavioural advertising and the somewhat inevitable result of cross-adoption of behavioural economic insights, sophisticated computer analytics, and online platform and advertiser interest in expanding their markets, we believe their implications are qualitatively different in the political arena (see Turow, Deeli Carpini, Draper, & Howard-Williams, 2012) compared to the commercial arena. Personalisation and targeting practices have evident positive results, including helping to ensure that individuals are directed towards resources, products and services that are of the greatest value to them, and relieving them of the burden of sifting through mountains of irrelevant material. At the same time, there are significant and widely-recognised downsides of personalisation, including its reliance on the collection and repurposing of personal information, surveillance of a greater range of individual activities, discrimination based on selective exposure of information to particular audiences or individuals, and the possibility of manipulation. In the next section, we argue that these downsides, particularly selective exposure to information and possible manipulation, raise distinct and problematic issues in the political arena and that these downsides are greatest when they result from or incorporate psychographic profiling.

PERSONALISATION AND PSYCHOGRAPHIC PROFILING IN POLITICS

In this section, we address two concerns that have emerged in debates about political micro-targeting generally, and that apply even more critically in micro-targeting using psychographic profiling: first, *polarisation* of the electorate in ways that challenge the ability of a democratic polity to understand political questions in similar ways and reach consensus on how to proceed; and second, *manipulation* of voters' decision-making in ways that undermine their ability to act autonomously and develop opinions that reflect their interests. We also close this section with a brief discussion of those who take a more sceptical view of the negative effects of personalised messages.

POLARISATION

One of the main concerns voiced about personalisation in exposure to political information has been about the development of 'filter bubbles' (Pariser, 2011) and digital information 'echo chambers' [1] that effectively stifle information inconsistent with previously expressed (or inferred) interests, opinions, and practices. Concerns arising from surveillance of users and sophisticated algorithmic processing focus on the restriction of content presented to users, and the potential for bias and loss of diversity in the information environment. Some recent

empirical studies, including one measuring exposure to diverse news and opinions on Facebook (Bakshy, Messing, & Adamic, 2015), cast doubt on whether such 'filter bubbles' and 'echo chambers' exist online, though they nonetheless conclude that social media exposure to ideologically different viewpoints is possible and under the individual's control. Many studies reveal a weak 'filter bubble' effect (see, e.g., Bakshy et al., 2015; and Bechmann & Nielbo, 2015) that has the potential to reinforce stronger individual and social information selection mechanisms. Filter bubbles may be more likely to affect some groups, including the politically disinterested who are not avid media consumers (Dubois & Bank, 2018). Regardless, as Lazer (2015) points out, although such selective exposure may not be occurring yet, this remains a potential concern as algorithms become even more sophisticated and opaque sparking subtle changes in behaviour.

In some ways, what have been termed 'digital filter bubbles' are simply a more extreme version of the limited information environment that results from our natural tendency to seek information consistent with confirmed or emerging perspectives (Nickerson, 1988; Sunstein, 2007). However, the multiplicity, ubiquity, and invisibility of the algorithms - their 'black box' (Pasquale, 2015) features - that determine our information environments (Pariser, 2011) will tend to enhance the isolating and fragmenting tendencies we demonstrate spontaneously in our offline information seeking. The technical process of information selection can even catalyse a self-selection process by taking away the choice to avoid or confront dissonant content (Bodó, Helberger, Eskens, & Möller, 2019, p. 2). In fact, it is easy to see how the two processes – personal and technical information isolation – can be mutually reinforcing: "People are diversity averse, and algorithms reduce diversity. Together, users and algorithms create a spiral, in which users are one-dimensional and prefer their information diet to be filtered so that it reflects their interests, and in which this filtering reinforces the individual's one-dimensionality" (Bodó et al., 2019, p. 2).

Whether self-inflicted, technologically mediated, or both, the result of this information isolation in the political arena is that opposing viewpoints are removed, with a consequent negative effect on democratic dialogue. The *particular* negative effect differs under liberal and deliberative view of democracy. In the liberal perspective, in order to make reasonable decisions, citizens should know a range of opinions and options. If information is filtered to them, especially without their consent, that would "violate their autonomy, as it will interfere with their ability to choose freely, and be the judge of their own interests" (Bozdag & van den Hoven, 2015, p. 251). The deliberative democracy perspective puts less emphasis on the loss of autonomy and more on the loss of diversity of opinions and perspectives resulting from targeting information because that will negatively impact the ability of people to deliberate or reason together about issues and candidates (Bozdag & van den Hoven, 2015). Bruns (2019), after reviewing the debate and evidence about 'filter bubbles' concludes that the more fundamental questions are why different groups have come to view information from rather radically but fixed perspectives and how this can be prevented or reversed - "in order to mitigate the very real threat of fundamental polarisation, and even of a complete breakdown of the societal consensus" (Bruns, 2019, p. 10).

Indeed, there is significant public concern about the 'fracturing' or polarisation of the electorate through the creation of a fragmented information environment. A recent article in *The Guardian* highlighted this concern in a quote from Full Fact, the UK fact-checking charity:

> 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...The combination of media buying by computers, and

> adverts being created and personalised by computers, mean that online advertising is not a shared experience any more (Chadwick, 2018, n.p.)

The same article noted that important public debate on, and response to, political messaging is undermined when those messages are not universally shared. A previous *Guardian* (Wong, 2018) contrasted the widespread public responses to the 'Daisy' ad of Lyndon B. Johnson's campaign and the divisive 'Willie Horton' ad put forth by George H.W. Bush with the complete lack of debate about ads that were placed by Trump in the 2016 US presidential election. Wong noted that 'no such debate took place around Trump's apparently game-changing digital political advertisements before election day' – because there were 50 to 60 *thousand* versions of those ads each day, effectively ensuring that there *was* no single public representation that could be debated. She also quoted Ann Ravel, a former member of the Federal Election Commission, on her concerns: "The way to have a robust democracy is for people to hear all these ideas and make decisions and discuss... With microtargeting, that is not happening". Sara Bannerman, the Canada Research Chair in Policy and Governance at McMaster University, expresses a similar concern: "On one hand, targeted messaging is similar to the practice of advertising to particular segments of the population in community publications. On the other hand, targeted messaging is completely different because it takes place in 'the dark'... they're visible only to specific selected people and not to a broader public" (Hirsh, 2018, n.p.).

MANIPULATION

A second concern regarding personalisation in the political arena, and particularly personalisation based on psychographic profiling, is the possibility not merely of persuading or influencing, but of *manipulating* voters. There is much written about manipulation, and numerous definitions (see Susser, Roessler, & Nissenbaum, 2019), but Sunstein's definition serves well for our purposes: "An action counts as manipulative if it attempts to influence people in a way that does not sufficiently engage or appeal to their capacities for reflective and deliberative choice" (Sunstein, 2015, p. 443). Advertisement has always been an attempt to manipulate behaviour, but the potential is exacerbated in the online context, and enhanced by increasingly sophisticated algorithms that monitor and respond to user behaviour (Susser, 2019). As Spencer (2019) writes, "the existing infrastructure supporting online behavioural advertising allows for extreme personalisation, enabling marketers to identify or even trigger the biases and vulnerabilities that afflict each individual consumer and tailor content to exploit those biases and vulnerabilities" (p. 4). Also relevant is Zarsky's suggestion of four elements that constitute unacceptable manipulations: 1) they tailor a unique response to every individual based on previously collected data; 2) they adapt the tailored response based on on-going feedback from the user and other peers, rendering the manipulation an on-going process rather than a one time action; 3) they occur in a non-transparent environment; and 4) they are facilitated by advanced data analytics tools allowing insights as to what forms of persuasion are effective over time. (Zarsky, 2019, p. 169). Floridi's (2016) categories of 'structural' and 'informational' nudging also offer some insight into the distinction between acceptable and unacceptable forms of manipulation. According to Floridi, structural nudging alters the choice environment and the courses of action available to the decision maker, and can result in a *de facto* forced choice. Informational nudging, by contrast, changes the information available to the decision maker about the available alternatives, but does not attempt to shape directly the choice itself. The distinction is subtle, but worth careful consideration.

The ability to manipulate individuals has been enhanced by research in psychology, neuroscience, and behavioural economics. Research has demonstrated that social media and

online behavioural tracking information can be used to predict personality characteristics, particularly in the case of extraversion and life satisfaction (Kosinski, Bachrach, Kohli, Stillwell, & Graepel, 2014), and that these predictions are *more* accurate than personality judgments made by friends and family (Youyou, Kosinski, & Stillwell, 2015). The words, phrases, and topics of social media postings are not only highly indicative of age and gender but also with appropriate analysis show strong relationships to the 'big five' personality traits of *extraversion*, *agreeableness*, *conscientiousness*, *neuroticism*, and *openness* (Park et al., 2015; Schwarz et al., 2013). Other researchers have leveraged photos and photo-related activities to successfully predict personality traits (Eftekhar, Fullwood, & Morris, 2014). Advertisements based on cognitive biases or vulnerabilities are difficult for recipients to detect and difficult to counteract, particularly if the effects are small; decades of research in behavioural economics and related fields suggest that these biases are unconscious and persistent (see, e.g., Newell & Shanks, 2014; Tversky & Kahneman, 1974).

The notion that political behaviour is being shaped by leveraging psychological research has been raised in the popular press (e.g., Issenberg, 2016); indeed, John et al. (2013) wrote an entire book examining the use of nudges to shape civic behaviour. It is precisely this concern that is raised by Zittrain in his article entitled 'Engineering an Election' (Zittrain, 2014), and Tufekci raises similar issues under the rubric of 'computational politics' and 'engineering the public' (Tufekci, 2014). At the root of all of these concerns lies the basic truth articulated by Slovic (1995): preferences are *constructed* in the process of political decision-making – and political decision makers can therefore be influenced by the information they encounter in the process of making a decision.

The possibility of manipulation has been discussed more in the commercial than in the political realm (Calo, 2014; Zarsky, 2019; Susser, Roessler, & Nissenbaum, 2018), but persuasive techniques that work to influence consumer purchasing decisions are likely to influence political decisions as well. Consistent with our interest in psychographic profiling, we focus our analysis on political ads that not only use data based on surveillance of one's demographic characteristics and one's social, political and economic behaviour, but that also use sophisticated analysis to draw inferences about one's emotional and psychological inclinations and limitations.

Previously we analysed political targeting generally and the ways in which it challenged the ability of citizens to be autonomous agents in processing the information they receive (Burkell & Regan, 2019). Susser et al. (2019) provide a more general analysis of online manipulation and their conclusions regarding the harms of manipulation to autonomy and the implications for both individuals and society are likewise similar. In situating our concerns about manipulation, particularly in the political arena and as a result of psychographic profiling, Gorton's (2016) argument is prescient and relevant:

> The twentieth century revolution in social science never really made good on its promise of producing theories with genuine predictive power...[but] From the vantage of the twenty-first century...perhaps that magic of prediction and control has at long last arrived, at least in some measures and in certain domains (p. 62).

Gorton notes that the use of social science models and theories have enabled political campaigns to manipulate citizens in their roles as voters, through: 1) precise predictive power, especially when compared to earlier techniques; 2) 'undermin[ing] a healthy public sphere by individualizing, isolating, and distorting political information' (p. 63); and 3) altering the

behaviour of citizens through the use of models of *unconscious* processes of the mind that 'alter voting behaviour and public opinion formation through processes that often completely elude the understanding of their intended targets' (p. 63). All of these capacities in political campaigns raise important and problematic effects but the ability to tap the unconscious processes of decision-making is novel, powerful, and not yet fully recognised. Gorton places responsibility for this capacity on framing theory and focus group research which help campaigns identify words and phrases that "activate certain frames in voters' minds, especially frames that guide their moral thinking" and then use these "to alter voters' beliefs and behaviours by intentionally and precisely targeting their unconscious cognitive processes" (Gorton, 2016, p. 75). Gorton builds upon Lakoff's ideas about the ways in which framing theory affects political discourse and quotes his reasons for why appealing to logic and evidence fails in politics: "not only because the public's mind is mostly unconscious, metaphorical, and physically affected by stress, [but] because its brain has been neutrally shaped by past conservative framing" (Lakoff, 2009, as quoted in Gorton, 2016, p. 76).

Such uses of framing theory are rendered more sophisticated and powerful by ubiquitous digital surveillance and sophisticated algorithms to reveal unique vulnerabilities of individuals; moreover, digital platforms facilitate leveraging insights about individual vulnerabilities into decision-making in something like real-time (Susser et al., 2019, pp. 6-7). Cambridge Analytica's personality model, discussed above, provides a vehicle for these more sophisticated uses. Chester and Montgomery report that Cambridge Analytica compiled a database with thousands of data points per person to identify points on which an individual was 'persuadable' and tailor messages to the vulnerabilities of that individual (Chester & Montgomery, 2018, pp. 23-4). Moreover, research in neuroscience, psychology, and behavioural economics continue to advance more complex understandings of human emotion and behaviour, and ever more complex models to influence individuals.

Calo's research on digital marketing is particularly helpful in identifying the distinctions we think are important. He notes that firms marketing to consumers can "surface and exploit how consumers tend to deviate from rational decision-making on a previously unimaginable scale. Thus, firms will increasingly be in the position to *create* suckers, rather than waiting for one to be born" (Calo, 2014, p. 1018). He argues that the techniques that enable this are distinguishable from previous advertising techniques in two respects – "digital market manipulation combines, for the first time, a certain kind of *personalization* with the intense *systemization* made possible by mediated consumption" (Calo, 2014, p. 1021). Through systemisation, "hundreds of thousands of ads [are matched] with millions of Internet users on the basis of complex factors in a fraction of a second" (Calo, 2014, p. 1021). As discussed above, these same techniques are being employed in the political arena with ads being framed in ways that appeal directly to an individual's decision-making vulnerabilities and at times that they are likely to be most receptive to the message.

Calo argues that it is the "systemization of the personal *coupled with divergent interests* that should raise a red flag" (Calo, 2014, pp. 1022-23). He goes on to say "true digital market manipulation, like market manipulation in general, deals strictly in divergent incentives. The entire point is to leverage the gap between how a consumer pursuing her self-interest would behave leading up to the transaction and how an actual consumer with predictable flaws behaves when pushed, specifically so as to extract social surplus" (Calo, 2014, p. 1023). In the political arena, the divergent interests of voters and the campaign infrastructures are rooted in three factors. The first is the fairly obvious fact that a campaign is interested in promoting a certain candidate or policy position, and interested in persuading a voter to align herself with

the interest of the campaign. The campaign is not interested in providing unbiased information so a voter can judge for herself whether the campaign does indeed represent her interests. Secondly, the digital platforms on which political messages are conveyed are commercial and the platforms are interested in generating as much revenue as possible. The more messages they can display the more revenue and the more they can precisely target an ad and the more accurate in timing the ad, the more they can charge for the ad. Finally, the intermediaries of the ad agencies and political operatives are likewise interested in generating revenue through more sophisticated analytical processing and online outreach.

In the consumer marketplace, Calo points out digital market manipulation can exact economic and privacy harms, as well as damaging consumer autonomy (Calo, 2014, pp. 1024-1034). In the political marketplace of ideas, individual privacy and autonomy will be similarly compromised – and there are very real political harms of fragmentation and polarisation. Additionally, voters arguably incur what could be considered "economic harms" in two respects. The first is that their political message environment is restricted - and if they are challenged by other voters or by confronting counter messages, they take on the costs of reconciling divergent messages. The second is that their vote may not result in the economic or policy results that they anticipated from the messages. As an example, the Trump voters in 2016 may not have benefitted from the tax cut in the way they expected.

SCEPTICAL VIEWS

Some question the need to regulate micro-targeting in the political context. Zuiderveen Borgesius et al., 2018; see also Resnick, 2018) suggest that micro-targeting will have limited and potentially even positive effects on the democratic process, and there is doubt about the effectiveness of micro-targeted ads to change voting behaviour (Kalla & Brookman, 2018; Motta & Franklin Fowler, 2016). Vaccari (2017), evaluating the effectiveness of online mobilisation in three European countries, comes to a somewhat similar conclusion that such mobilisation increases political engagement (Vaccari, 2017, p. 85), however he does not explore the question of whether the engagements actually are in citizen's interests or if they are manipulated. These studies, however, have typically focused on traditional forms of advertising, and may underestimate the impact of more personalised advertising campaigns or psychographic profiling, which can manipulate both advertisement content and advertisement form to achieve maximal persuasion.

The actual impact of micro-targeting as currently practiced may still be an open question but there is every reason to believe that micro-targeting strategies are becoming increasingly sophisticated, based on increasingly detailed profiles and thus potentially more effective. Based on our analysis, the dangers to autonomous decision-making and further political polarisation posed by psychographic profiling tip the scales on the need to regulate. Daniel Kreiss (2017) provides an additional concern about sophisticated targeting that also lends support for some regulation. He takes a more sceptical view of the danger of manipulation of individual voters and emphasises the group basis of politics which leads to his concern about the cultural power of micro-targeting to "create a powerful set of representations of democracy that undermines the legitimacy of political representation, pluralism, and political leadership" (Kreiss, 2017, p. 3) - representations that in effect cause further polarisation. Whether out of concern for manipulation of individual voters or concerns about polarisation of the body politic, some governmental intervention is warranted.

OPTIONS FOR REGULATING/CONTROLLING SOPHISTICATED VOTER ANALYTICS

The first challenge to regulation of sophisticated voter analytics, in particular psychographic profiling, is that political speech is a cherished value in democratic systems and central to a functioning democracy. In the United States political speech is relatively free from regulation. In other democratic countries, governments have imposed some constraints on political speech in order to ensure the rights of voters and to ensure a free and fair exchange of political information so that voters can make informed and autonomous decisions. To date, however, there have been no specific regulations that limit what is generally referred to as "microtargeting" of political messages based on detailed personal profiles. We identify three avenues of response to sophisticated voter analytics and personalised political communication. The first locates the responsibility with voters themselves - what we term *voter responsibility*. The second places the responsibility with the platforms delivering micro-targeted political communications (e.g., Google, Facebook) - what we term *platform accountability*. And the third rests the responsibility with the courts to uphold policies the government adopts to restrict voter manipulation and polarisation of the electorate - what we term *judicial intervention*. We consider all approaches to be important – and the last to be critical.

VOTER RESPONSIBILITY

Some suggest that voters have access to multiple sources of political information and thus need not, and do not, rely solely on political advertisements. These arguments construct the citizen as an active and independent information seeker, capable of gathering and motivated to gather information from a wide range of sources creating an unbiased information sphere. One must consider, however, the difficulty individuals face in recognising that they are the recipients of targeted political advertisements or campaign messages, and their ability to 'step outside' of these selective information environments. Such stepping out may be difficult because, as Just and Latzer point out, "the market for attention – the central scarce resource in information societies – is increasingly being co-produced and allocated by automated algorithmic selection" (Just & Latzer, 2017, p. 239), influencing not only what individuals find or are exposed to but also the reputation of the source and their trust in it (Just & Latzer, 2017, p. 242). This complex interplay affects the ability of individuals, as consumers and voters, to discern the reliability and relevance of information they find or is presented to them. In effect, one's online information reality is largely constructed by algorithmic selection.

In response to users' concerns about the personalisation of messages, Facebook, the largest social media platform and the one at the forefront of current debate in the wake of the 2016 Cambridge Analytica controversy, developed two ad transparency mechanisms: a 'why am I seeing this' button, and an 'Ad Preferences' page. The first explains why a particular user is seeing a specific ad, while the second shows users a list of the information that Facebook has gathered about them and the sources of that information. These mechanisms provide users some insight into personalisation practices, but the mechanisms often offer incomplete, misleading, or vague information and explanations, and thus are of limited effectiveness in promoting ad transparency (Andreou et al., 2018). Moreover, users must be motivated to avail themselves of these mechanisms, and the information they receive only reveals that the advertisements they are viewing are selected specifically for them – they are not mechanisms for accessing unbiased or unfiltered advertisements. To address some of these limitations, Koene et al. (2015) suggest that the 'Internet research community' should develop monitoring tools [2] or

'test kits' that users could deploy to determine if the level of personalisation on a site is acceptable. This approach is consistent with the use 'ad blocker' plug-ins. These tools can provide valuable information to users with the motivation and technical know-how to deploy them, but again they only *flag* the fact that one is receiving a personalised message – the tools do not remove personalisation or message tailoring, nor do they inform others of the targeting and tailoring practices. Other transparency mechanisms such as ad registries that are being offered by or required of platforms (see platform accountability, below) offer solutions that require less technical skill, but still require significant and continued efforts on the part of users, who face a personalised information environment by default. Users can deploy strategies to circumvent personalisation, including deleting cookies, using search engines that do not track, providing false information, or carrying out random online actions such as haphazardly clicking on links (Bozdag & van den Hoven, 2015; Pariser, 2011), but these strategies will undermine the desired as well as undesired effects of personalisation, and they require conscious action on the part of the user. In other words, users must work, and work diligently, to escape and identify the effects of personalised messaging.

Media and information literacy initiatives to improve user skills and knowledge are also important responses, promoted most recently in relation to foreign interference with democratic elections (Tenove, Buffie, McKay, & Moscrop, 2018). In Britain, for example, the House of Commons Digital Culture, Media and Sport Committee highlighted the importance of digital literacy, recommending in its 2018 report that "digital literacy should be a fourth pillar of education, alongside reading, writing and maths" (DCMS, 2018, p. 312), and it suggested that a comprehensive educational digital literacy framework should be funded through a social media company levy. These initiatives seek to empower users by giving them the knowledge and skills to separate 'fake news' and 'alternative facts' from real content (Cooke, 2018). Although these are aimed at the more general issue of digital literacy and training (Stoddard, 2014), they help to increase awareness of the possible dangers in online information flows and may sensitise people to biases in political information. It is important to note, however, that media literacy campaigns will be less effective in protecting audiences against the subtle types of manipulation enabled by psychographic profiling, which often involve small changes to messages to engage processing heuristics and biases that operate below the level of consciousness.

PLATFORM ACCOUNTABILITY

A second avenue for policy responses to sophisticated voter analytics and online personalised and micro-targeted political communication is to require more accountability on the part of internet platforms. In general, such accountability would be in the form of disclosure of who is sponsoring ads and how those ads are being targeted, in effect a form of algorithmic transparency. Since most countries already have some form of disclosure laws, this might be viewed as an incremental change and thus engender minimal opposition. Rubinstein proposes that disclosure of personal information practices, for example, could be required by the candidate and other electoral actors and by the data brokers who make personal information available to electoral actors (Rubinstein, 2014, pp. 910-921). The options we discuss below are instead directed to online platforms rather than to candidates or electoral actors. There appears to be interest in several countries to place more responsibility on platforms.

For example, in December 2018, Canada enacted the Elections Modernization Act, which requires that platforms maintain a record of the political and partisan advertisements they deliver, starting a year before an election, and maintained for two years afterwards (George-Cosh, 2019). Also, in 2018, the Washington State Public Disclosure Commission ruled that the State's political advertising disclosure requirements applied to online advertising. The

requirements included disclosure of: the ad; who or what the ad was supporting or opposing; the name and address of the ad's purchaser; the ad's cost; and, for digital ads, the total number of impressions and demographic information of the audiences targeted and reached to the extent that information is collected by the commercial advertiser (Sanders, 2019). At the national level in the US, the proposed Honest Ads Act is similarly designed to enhance the integrity of the democratic process by extending the disclosure requirements of who has paid for political ads from traditional media to the online environment. With respect to targeted audiences, the bill would require large digital platforms with at least 50,000,000 monthly viewers to maintain a public file of all electioneering communications which "would contain a digital copy of the advertisement, a description of the audience the advertisement targets, the number of views generated, the dates and times of publication, the rates charged, and the contact information of the purchaser." Although the bill has bipartisan sponsorship, it does not have the support of the Republican leadership.

Some companies, such as Facebook and Twitter, have voiced some support for the proposed Honest Ads Act and adopted some of its requirements voluntarily (Newton, 2018a). In May 2018, Facebook required a "paid for" at the top of ads on Facebook and Instagram, with a link to a page with information about the cost of the ad and the demographic breakdown of the intended audience, including the age, location, and gender. This requirement will address targeting generally but not targeting based on more sophisticated voter analytics including psychographic profiling. Facebook has also created an Ad Library with ads for the last seven years and has established a partnership with an academic team to facilitate research about the nature and implications of online political advertising (Newton, 2018b). Twitter has instituted similar rules requiring disclosure of ad sponsors and has established an Ad Transparency Council to provide more detailed breakdowns of ad spending and targeting demographics (Statt, 2018).

It is unclear how effective these self-regulatory initiatives will actually be – and these companies have not been willing to comply with government mandates. For example, in response to the Canadian Election Modernization Act, Google decided to refrain from carrying any political ads rather than comply with legislation designed to support greater scrutiny to online advertising (Dubois et al., 2019). Google and Facebook responded similarly to the Washington State requirements, banning political advertisements rather than following the requirements. The companies argued that the burden of determining whether an ad was political was 'enormous' and that it might be 'technologically impossible' to know what ads are actually running on their platforms (Sanders, 2018). Google, for example, sells advertisement space on web pages through a real-time bidding process that auctions the ad 'slots' visible to a specific viewer who is visiting a web page. The process takes place in a fraction of a second, and the platform (Google in this case) may know only the identity of the successful bidder, and not the content of the ads that were placed by that bidder.

In the Canadian and US contexts, it is also important to note that the accountability required is itself limited, covering only 'official' political advertisements and leaving entirely unregulated other forms of influential online political speech, including bots, influencer marketing, and paid 'audience builders' (Reepschlager & Dubois, 2019). There can be no doubt that political messages, including those constituting foreign influence, will likely slip through the cracks of any system designed to identify them. Platforms, however, have addressed some of these technical issues to disrupt online communications by terrorist groups (Global Internet Forum to Counter Terrorism, n.d.); the same will to act, and the same solutions, could be applied to the identification of online political advertising.

The EU has gone a bit further than Canada and the US in addressing the responsibilities that platforms have with respect to transparency and political advertisements. In April 2018, the European Commission proposed an EU wide policy to counter online disinformation, which was later finalised with input from the major platforms including Google, Facebook and Twitter. By signing this Code of Practice on Disinformation, these platforms are responsible for:

- Ensuring transparency about sponsored content, in particular political advertising, as well as restricting targeting options for political advertising and reducing revenues for purveyors of disinformation;
- Providing greater clarity about the functioning of algorithms and enabling third-party verification;
- Making it easier for users to discover and access different news sources representing alternative viewpoints;
- Introducing measures to identify and close fake accounts and to tackle the issue of automatic bots;
- Enabling fact-checkers, researchers and public authorities to continuously monitor online disinformation (European Commission, 2018).

The policy also provides support for a network of fact-checkers and calls on "Member States to scale up their support of quality journalism to ensure a pluralistic, diverse and sustainable media environment". This is largely a self-regulatory tool, but has also been described as a co-regulatory instrument given the Commission's involvement in its development and oversight (Leerssen, 2019). [3] There appears to be growing support in Europe for efforts such as these. For example, Sofia Karttunen in an LSE Media Policy Blog argues: "Perhaps it would be time for European regulators to take a closer look at the algorithms of social media platforms, which determine which content is displayed to which person and run the risk of creating so-called 'echo-chambers' and 'filter bubbles' that can amplify certain communications over others… and can create social and behavioural change" (Karttunen, 2018). Mittelstadt (2016) proposes 'algorithm auditing' as an 'ethical duty for providers of content personalisation systems to maintain the transparency of political discourse' (Mittelstadt, 2016, p. 4991). He situates this duty in relation to the EU General Data Protection Regulation (GDPR) [4] that requires data processors to explain the logic of automated decision making, and suggests that algorithmic auditing could be carried out by a regulatory body "to oversee service providers whose work has a foreseeable impact on political discourse by detecting biased outcomes as indicated by the distribution of content types across political groups" (Mittelstadt, 2016, p. 4998).

JUDICIAL INTERVENTION

Even if governments impose more effective regulations on internet platforms, political actors, or data brokers, these regulations are likely to be challenged in some countries, especially in the US, on the grounds that they restrict free speech rights. In such cases, it will be up to the courts to determine the validity of the restrictions and the appropriate balance of free speech and other rights and interests that individuals have. Based on our readings of democratic theory and judicial rulings, particularly in the US, we believe that four strains of thinking may provide some justification for restricting micro-targeted voting messages, especially those employing psychographic profiling: a more *expansive view of corruption*; more attention to the *rights of listeners, including the right against compelled listening*; application of the *right to receive information* and an expanded notion of the *rights of voters*. Each of these is discussed briefly below.

CORRUPTION

Since 1976 in *Buckley v. Valeo*, the Supreme Court has consistently held that restrictions on campaign spending are unconstitutional unless there is a compelling interest outweighing the

free speech interest. To date, the court has restricted such a compelling interest to that of corruption, narrowly defined as quid-pro-quo corruption. Walker Wilson (2010) argues for a more expansive view of "corruption" that would address "the relationship between money and potentially manipulative communication strategies" (Walker Wilson, 2010, p. 740), suggesting that "the definition of corruption ought to be expanded to include the potential for distortion in voting behaviour as a result of heavy-handed psychological tactics" (Walker Wilson, 2010, p. 741). As she notes "liberal democracy depends upon a free and willing voting public, and a voting process that is unencumbered by systematic, wide-scale manipulation by any segment of the public, individual candidate, or political party" (Walker Wilson, 2010, p. 742)

RIGHTS OF LISTENERS

The rights of listeners have arguably been under-appreciated, especially in two areas. First is when speakers would prefer not to speak about something that could negatively impact the speaker. An example would be, as Kendrick (2017) points out in product labelling and disclosure, when the public might like to know whether food contains genetically modified ingredients but food producers prefer not to say (Kendrick, 2017, p. 1800). The second area is when courts themselves have shown little attention to the rights of listeners, in part because speakers are the parties invoking free speech claims. Kendrick notes that this has occurred in cases involving net neutrality rules where US courts might have pointed out that such rules served listeners' rights but instead focused on the rights of speakers. Similarly, decisions giving search engines immunity from fair competition laws have not acknowledged that listeners' rights might be furthered by the application of such laws (Kendrick, 2017, p. 1805).

Related to the rights of listeners to hear are the rights of listeners not to hear. In the US, as Corbin (2009) points out, the 'captive audience' doctrine has provided some protection for 'unwilling listeners' especially when combined with privacy interests, such as being subject to protesters in front of one's home. According to the captive audience doctrine, private speakers cannot always foist their speech onto unwilling listeners. In order for the government to restrict private speakers, listeners should not be able to easily avoid the message, thus raising listeners' privacy interests. This would be especially true if the speaker follows the listener so the listener suffers repeated exposure (Corbin, 2009, pp. 944-50). One relevant question is whether physical captivity could be compared to online captivity; could, for example, being 'followed' by a message in the online world constitutes captivity similar to that experienced by an unwilling audience that is followed by a speaker in the physical world?

In the EU, freedom of expression entails a right not to listen and a right to refuse information, even if it might be beneficial or valuable, which restricts government involvement in providing a level of information exposure diversity that could infringe individual freedom. However, as Helberger (2012) notes, this interpretation seems to assume that the diversity to which one is exposed is the result of media sources reaching an undifferentiated audience instead of a targeted audience. Moreover, freedom of expression "as a constitutional value, does not only require policy makers to refrain from interferences. It can, under certain circumstances, at least in Europe, create positive obligations to actively protect and promote the realisation of people's right to freedom of expression, part of which is the ability to form one's opinions from diverse sources" (Helberger, 2012, p. 72). Helberger points out that "finding and accessing the kind of diversity that people may seek is also a matter of design aspects, many of which are principally invisible to users" (Helberger, 2012, p. 79). The Council of Europe in 2007 recognised "in particular the importance of transparency regarding the listing and prioritization of information provided by search engines with regard to the right to receive and impart information" (Helberger, 2012, p. 83). More recently, in 2018, the Council explicitly addressed the need for

member states to take measures to "enhance users' effective exposure to the broadest possible diversity of media content" (Bodó et al., 2017, p. 15).

RIGHTS TO ACCESS INFORMATION

Related to the rights of listeners to hear is the right to access information. This has played out primarily with respect to access to government information, as enshrined in freedom of information laws, and with respect to libraries' rights to provide information to the public (Mart, 2003). Language in *Buckley v. Valeo* (1976) reflects the importance of this right to access, noting that the First Amendment: "was designed to secure the widest possible dissemination of information from diverse and antagonistic sources, and to assure unfettered interchange of ideas for the bringing about of political and social changes desired by the people" (*Buckley v. Valeo, 1976,* pp. 48-9). As far back as 1943 in *Martin v. Struthers,* the Supreme Court recognised a constitutional right to receive information, noting that the value to be protected is the "vigorous enlightenment" of the people. In 1969, in *Red Lion v. FCC,* Justice White wrote: "It is the right of the viewers and listeners, not the right of the broadcasters, which is paramount.... It is the right of the public to receive suitable access to social, political, aesthetic, moral, and other ideas and experiences which is crucial here" (Mart, 2003, p. 178). In 1969, In *Board of Education v. Pico* (1982), Justice Brennan in his plurality opinion opined – "the right to receive ideas is a necessary predicate to the *recipient's* meaningful exercise of his own rights of speech, press, and political freedom" (Mart, 2003, p. 181). The right to receive or access information may also provide a basis for legitimate restrictions on the use of sophisticated analytics in targeting political messages.

RIGHTS OF VOTERS

Derfner and Herbert argue that voting should be treated as a fundamental right, protected by the First Amendment as a form of voice and expression (Derfner & Herbert, 2016, p. 485). Indeed, they find an argument for this in *Buckley* itself where the Court stated that "[i]n a republic where the people are sovereign, the ability of the citizenry to make informed choices among candidates for office is essential" (Derfner & Herbert, 2016, p.114) and that the "central purpose" of the First Amendment is to ensure that "healthy representative democracy [can] flourish" (Derfner & Herbert, 2016, p.116). As Derfner and Herbert say, "voters take the information that is put into the marketplace of ideas and ultimately make a decision about which view to adopt and which candidate or political party best represents it" (Derfner & Herbert, 2016, p. 489). If the Court were to recognise more directly that voting itself is an expressive act and that the purpose of the First Amendment is to enable that expressive act, then voting would be brought under the full protection of the First Amendment (Derfner and Herbert, 2016, pp. 489-90). Kendrick similarly argues that freedom of speech can be viewed as derived from right to vote and because individuals have the right to vote, they have a claim to information relevant to voting (Kendrick 2018, p. 1789). Elevation of the rights of voters could provide stronger justification for restrictions on targeting of political speech.

CONCLUSION

Governments in many jurisdictions have demonstrated a willingness to put some limitations on political speech and are increasingly recognising the dangers of highly personalised political messaging. Regulation is of increasing importance because both the sophistication and the penetration of digital marketing techniques has increased in the electoral context (Chester & Montgomery, 2019). The strongest protection of the rights of voters arguably would be to

prohibit micro-targeted or personalised political advertising entirely. This would avoid difficult line-drawing between different types of profiling but would also challenge advocates of political speech. Moreover, the political reality is such that there is likely to be strong pushback by political operatives, including campaigns, political advertising and consulting agencies, and platforms, against a suggestion to prohibit micro-targeted political advertising. Voters themselves may even express, as consumers do, some preference for targeted advertisements; alternatively, they might reject targeted political advertisements consistent with some research that suggests a similar attitude toward targeted advertisements in general (Turow et al., 2009). Instead of a universal ban on personalised political communication, what is needed is clarification of what forms of targeting are problematic. Moreover, arriving at the 'right' policy framework will require multisectoral consultation open to input from *all* stakeholders, including government, platforms, and organised civil society organisations (Marda & Milan, 2018).

Our analysis indicates that micro-targeted political ads based on algorithmic profiling of big data sources about subsets of individuals has the potential to facilitate further polarisation of politics and the manipulation of voters' decision-making capacity. We argue that micro-targeted ads employing psychographic profiling pose the greatest dangers because they are even more opaque, insidious and powerful, as they exploit the psychological vulnerabilities of individuals - in effect, treating citizens as 'suckers'. Although it may be technically difficult to operationalise psychographic profiling and identify ads based on such criteria, we hope we have identified in a meaningful way the dangers of such messaging and outlined possible avenues for regulating these dangers. As democracies begin to grapple with these dangers, the most effective path forward may be through multi-stakeholder or co-regulatory mechanisms, as discussed above with respect to the European Commission's Code of Disinformation Practice.

REFERENCES

Andreou, A., Venkatadri, G., Goga, O., Gummadi, K., Loiseau, P., & Mislove, A. (2018). Investigating ad transparency mechanisms in social media: A case study of Facebook's explanations. *Proceedings of the 2018 Network and Distributed System Security Symposium*. https://doi.org/10.14722/ndss.2018.23191

Bailenson, J. N., Iyengar, S., Yee, N., & Collins, N. A. (2008). Facial similarity between voters and candidates causes influence. *Public Opinion Quarterly*, *72*(5), 935–961. https://doi.org/10.1093/poq/nfn064

Bakshy, E., Messing, S., & Adamic, L.A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science 348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Barocas, S. (2012). The price of precision: Voter microtargeting and its potential harms to the democratic process. *Proceedings of the First Edition Workshop on Politics, Elections and Data - PLEAD '12* (pp. 31-36). https://doi.org/10.1145/2389661.2389671

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven: Yale University Press.

Bennett, C. J. (2016). Voter databases, micro-targeting, and data protection law: can political parties campaign in Europe as they do in North America? *International Data Privacy Law, 6*(4), 261–275. https://doi.org/10.1093/idpl/ipw021

Bodó, B., Helberger, N., & de Vreese, C. H. (2017). Political micro-targeting: a Manchurian candidate or just a dark horse? *Internet Policy Review*, *6*(4). https://doi.org/10.14763/2017.4.776

Bodó, B., Helberger, N., Eskens, S., & Möller, J. (2019). Interested in Diversity: The role of user attitudes, algorithmic feedback loops, and policy in news personalisation. *Digital Journalism*, *7*(2), 206–229. https://doi.org/10.1080/21670811.2018.1521292

Bozdag, E., & van den Hoven, J. (2015). Breaking the filter bubble: democracy and design. *Ethics and Information Technology, 17*(4), 249–265. https://doi.org/10.1007/s10676-015-9380-y

Bruns, A. (2019). Filter bubble. *Internet Policy Review*, *8*(4). https://doi.org/10.14763/2019.4.1426

Buckley v. Valeo, 424 U.S. 1 (1976)

Burkell, J., & Regan, P. M. (2019). Voting public: Leveraging personal information to construct voter preference. In N. Witzleb, M. Paterson, & J.Richardson (Eds.), *Big Data, Political Campaigning and the Law*. Abingdon: Routledge.

Calo, R. (2104). Digital market manipulation. *George Washington Law Review, 82*(4), 995–1050. Retrieved from https://www.gwlr.org/calo/

Chadwick, P. (2018, October 7). This lawless world of online political ads is anti-democratic. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2018/oct/07/lawless-online-political-ads-anti-democratic

Chester, J., & Montgomery, K.C. (2017). The role of digital marketing in political campaigns. *Internet Policy Review*, *6*(4). https://doi.org/10.14763/2017.4.773

Chester, J., & Montgomery, K.C. (2018, September). *The influence industry: contemporary digital politics in the United States*. Berlin: Tactical Technology Collective. Retrieved from https://cdn.ttc.io/s/ourdataourselves.tacticaltech.org/ttc-influence-industry-usa.pdf

Chester, J., & Montgomery, K. C. (2019). The digital commercialisation of US politics—2020 and beyond. *Internet Policy Review, 8*(4). https://doi.org/10.14763/2019.4.1443

Cooke, N. A. (2018). *Fake news and alternative facts: Information literacy in a post-truth era.* Chicago: ALA Editions.

Corbin, C. M. (2009). The First Amendment right against compelled listening. *Boston University Law Review, 89*(3), 939–1016. Retrieved from http://www.bu.edu/law/journals-archive/bulr/volume89n3/documents/CORBIN.pdf

Derfner, A. & Gerald, H. J. (2016). Voting is speech. *Yale Law & Policy Review*, *34*(2), 471–491. https://ylpr.yale.edu/voting-speech

Digital, Culture, Media and Sports Committee (DCMS). (2018). *Disinformation and 'fake news': Final report. Digital culture, media and sport committee* [Final Report]. London: Parliament. Retrieved from https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/179102.htm

Dobber, T., Fathaigh, R. Ó., & Zuiderveen Borgesius, F. J. (2019). The regulation of online political microtargeting in Europe. *Internet Policy Review*, 8(4). https://doi.org/10.14763/2019.4.1440

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, *21*(5), 729–745. https://doi.org/10.1080/1369118X.2018.1428656

Dubois, E., McKelvey F., & Owen, T. (2019, April 10). What have we learned from Google's political ad pullout? *Policy Options*. Retrieved from https://policyoptions.irpp.org/magazines/april-2019/learned-googles-political-ad-pullout/

Eftekhar, A., Fullwood, C., & Morris, N. (2014). Capturing personality from Facebook photos and photo-related activities: How much exposure do you need? *Computers in Human Behavior*, *37*, 162–170. https://doi.org/10.1016/j.chb.2014.04.048

European Commission (2018, April 25). *Tackling online disinformation: Commission proposes an EU-wide code of practice* [Press release]. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_18_3370.

Evans, E. C. (1917). *A History of the Australian Ballot System in the United States.* Chicago: University of Chicago Press.

Facebook (n.d.) Facebook ad library. Retrieved January 13 2020 from https://www.facebook.com/ads/library/?active_status=all&ad_type=political_and_issue_ads &country=US&impression_search_field=has_impressions_lifetime.

Floridi, L. (2016). Tolerant paternalism: Pro-ethical design as a resolution of the dilemma of

toleration. *Science and engineering ethics*, *22*(6), 1669–1688. https://doi.org/10.1007/s11948-015-9733-2

George-Cosh, D. (2019, March 5). Google bans political ads ahead of next Canadian federal election. *BNN Bloomberg*. Retrieved from https://www.bnnbloomberg.ca/

Global Internet Forum to Counter Terrorism (n.d.). *Evolving an institution*. Retrieved January 13, 2020 from https://www.gifct.org/about/.

Gorton, W. A. (2016). Manipulating citizens: how political campaigns' use of behavioural social science harms democracy. *New Political Science*, *38*(1), 61–80. https://doi.org/10.1080/07393148.2015.1125119

Helberger, N (2012). Exposure diversity as a policy goal. *Journal of Media Law*, *4*(1), 65–92. https://doi.org/10.5235/175776312802483880

Hine, D. W., Reser, J. P., Morrison, M., Phillips, W. J., Nunn, P., & Cooksey, R. (2014). Audience segmentation and climate change communication: conceptual and methodological considerations. *Wiley Interdisciplinary Reviews: Climate Change*, *5*(4), 441-459. https://doi.org/10.1002/wcc.279

Hirsh, J. (2018, November 20). Canadian elections can't side-step social media influence. Waterloo, Ontario: *Centre for International Governance Innovation*. Retrieved from https://www.cigionline.org/articles/canadian-elections-cant-side-step-social-media-influence

Hirsh, J. B., Kang, S. K., & Bodenhausen, G. V. (2012). Personalized persuasion: Tailoring persuasive appeals to recipients' personality traits. *Psychological science*, *23*(6), 578–581. https://doi.org/10.1177/0956797611436349

Issenberg, S. (2016). *The victory lab: The secret science of winning campaigns*. New York: Broadway Books.

John, P., Cotterill, S., Richardson, L., Moseley, A., Stoker, G., Wales, C., & Smith, G. (2013). *Nudge, nudge, think, think: Experimenting with ways to change civic behaviour*. New York: Bloomsbury Academic Publishing.

Just, N., & Latzer, M. (2017). Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet. *Media Culture & Society, 39*(2), 238–258. https://doi.org/10.1177/0163443716643157

Kalla, J. L., & Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review*, *112*(1), 148–166. https://doi.org/10.1017/s0003055417000363

Karttunen, S. (2018, September 20). Gearing up for the next European elections: will we see regulation of online political advertising [Blog post]. *LSE Media Policy Project Blog*. Retrieved from https://blogs.lse.ac.uk/mediapolicyproject/2018/09/20/gearing-up-for-the-next-european-elections-will-we-see-regulation-of-online-political-advertising

Kendrick, L. (2017). Are speech rights for speakers. *Virginia Law Review*, *103*(8), 1767–1808. Retrieved from https://www.virginialawreview.org/volumes/content/are-speech-rights-speakers

Koene, A., Perez, E., Carter, C. J., Statache, R., Adolphs, S., O'Malley, C., ... & McAuley, D. (2015). Ethics of personalised information filtering. In T. Tiropanis, A. Vakali, L. Sartori, & P. Burnap (Eds), *Internet Science, INSCI 2015, Lecture Notes in Computer Science, 9089* (pp. 123–132). Cham: Springer. https://doi.org/10.1007/978-3-319-18609-2_10

Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Machine Learning*, *95*(3), 357–380. https://doi.org/10.1007/s10994-013-5415-y

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Kreiss, D. (2017). Micro-targeting, the quantified persuasion. *Internet Policy Review*, *6*(4). https://doi.org/10.14763/2017.4.774

Lazer, D. (2015). The rise of the social algorithm. *Science, 348*(6239), 1090–1091. https://doi.org/10.1126/science.aab1422

Leerssen, P., Ausloos, J., Zarouali, B., Helberger, N., & de Vreese, C. H. (2019). Platform ad archives: Promises and pitfalls. *Internet Policy Review, 8*(4). https://doi.org/10.14763/2019.4.1421

Marda, V. & Milan, S. (2018, May 21). *Wisdom of the crowd: Multistakeholder perspectives on the fake news debate* [White paper]. Philadelphia: Internet Policy Review Observatory, Annenberg School of Communication. Retrieved from http://globalnetpolicy.org/wisdom-of-the-crowd/

Mart, S. N. (2003). The right to receive information. *Law Library Journal, 95*(2), 175–189. Available at https://www2.lib.uchicago.edu/~lar1/LIS450LI/mart2.pdf

Milan, S., & Agosti, C. (2019, February 7). Personalization algorithms and elections: Breaking free of the filter bubble. *Internet Policy Review*. Retrieved from https://policyreview.info/articles/news/personalisation-algorithms-and-elections-breaking-free-filter-bubble/1385

Mittelstadt, B. (2016). Auditing for transparency in content personalisation systems. *International Journal of Communication*, *10*. Retrieved from https://ijoc.org/index.php/ijoc/article/viewFile/6298/1809

Motta, M. P., & Franklin Fowler, E. (2016). The content and effect of political advertising in US campaigns. In *Oxford Encyclopedia of Politics*. Oxford: Oxford University Press https://doi.org/10.1093/acrefore/9780190228637.013.217

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, *37*(1), 1–19. https://doi.org/10.1017/s0140525x12003214

Newton, C. (2018a). Congress roasted Facebook on TV, but won't hear any bills to regulate it. *The Verge*. Retrieved from https://www.theverge.com/2018/6/7/17387120/congress-facebook-tv-regulation-bills

Newton, C. (2018b, May 24). Facebook disclosure requirements for political ads take effect in United States today. *The Verge*. Retrieved June 7, 2019, from https://www.theverge.com/2018/5/24/17389834/facebook-political-ad-disclosures-united-states-transparency

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Pariser, E. 2011. *The filter bubble: What the internet is hiding from you*. New York, NY: Penguin Press.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., ... & Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of personality and social psychology*, *108*(6), 934–952. https://doi.org/10.1037/pspp0000020

Pasquale, F. (2015). *The black box society*. Cambridge, MA: Harvard University Press.

Reepschlager, A., & Dubois, E. (2019, January 2). New elections laws are no match for the Internet. *Policy Options*. Retrieved from https://policyoptions.irpp.org/magazines/january-2019/new-election-laws-no-match-internet/

Resnick, B. (2018, March 26). Cambridge Analytica's "psychographic microtargeting": what's bullshit and what's legit. *Vox*. Retrieved from https://www.vox.com/science-and-health/2018/3/23/17152564/cambridge-analytica-psychographic-microtargeting-what

Rubinstein, I. S. (2014). Voter privacy in the age of big data. *Wisconsin Law Review,* 2014(5), 861–936. Retrieved from http://wisconsinlawreview.org/wp-content/uploads/2015/02/1-Rubinstein-Final-Online.pdf

Sanders, E. (2018, September 26). Big tech is fighting to change Washington's pioneering rules on election ad transparency, *The Stranger*. Retrieved from https://www.thestranger.com/slog/2018/09/26/32825020/big-tech-is-fighting-to-change-washingtons-pioneering-rules-on-election-ad-transparency

Sanders, E. (2019, January 2). As 2019 begins, so does Facebook's ban on local political ads in Washington state. *The Strange*. Retrieved from https://www.thestranger.com/slog/2019/01/02/37628091/as-2019-begins-so-does-facebooks-ban-on-local-political-ads-in-washington-state

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open vocabulary approach. *PLOS ONE, 8*(9), e73791. https://doi.org/10.1371/journal.pone.0073791

Slovic, P. (1995). The construction of preference. *American psychologist*, *50*(5), 364–371. https://psycnet.apa.org/doi/10.1037/0003-066X.50.5.364

Spencer, S. B. (2019). *The Problem of online manipulation*. https://doi.org/10.2139/ssrn.3341653.

Statt, W. (2018, May 24). Twitter reveals new guidelines and disclosure rules for political ads. *The Verge*. Retrieved from https://www.theverge.com/2018/5/24/17390156/twitter-political-

advertising-guidelines-transparency-rules

Stoddard, J. (2014). The need for media education in democratic education. *Democracy and Education*, *22*(1). Retrieved from https://democracyeducationjournal.org/home/vol22/iss1/4/

Sunstein, C. R. (2007). *Republic.com 2.0*. Princeton: Princeton University Press.

Sunstein, C. R. (2015). The ethics of nudging. *Yale Journal on Regulation, 32*(2), 413–450. Retrieved from https://digitalcommons.law.yale.edu/yjreg/vol32/iss2/6

Susser, D., Roessler, B., & Nissenbaum, H. (2018). Online Manipulation: Hidden Influences in a Digital World. https://doi.org/10.2139/ssrn.3306006

Susser, D. (2019). Invisible Influence: Artificial Intelligence and the Ethics of Adaptive Choice Architectures. Retrieved from https://philpapers.org/archive/SUSIIA-2.pdf

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, *8*(2). https://doi.org/10.14763/2019.2.1410

Tenove, C., Buffie, J., McKay, S., & Moscrop, D. (2018). *Digital threats to democratic elections: How foreign actors use digital techniques to undermine democracy* [Report]. Vancouver: Centre for the Study of Democratic Institutions, University of British Columbia. Retrieved from https://democracy2017.sites.olt.ubc.ca/files/2018/01/DigitalThreats_Report-FINAL.pdf

Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday*, *19*(7). Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/4901/4097

Turow, J., Delli Carpini, M. X., Draper, N. A., & Howard-Williams, R. (2012). *Americans roundly reject tailored political advertising—At a time when political campaigns are embracing it* [Departmental Paper]. Philadelphia: Annenberg School for Communications, University of Pennsylvania. Retrieved from https://repository.upenn.edu/asc_papers/522

Turow, J., King, J., Hoofnagle, C. J., Bleakley, A., & Hennessy, M. (2009). *Americans reject tailored advertising and three activities that enable it* [Departmental Paper]. Philadelphia; Berkeley: Annenberg School for Communication, University of Pennsylvania; Berkeley School of Law, University of California, Berkeley. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1551&context=asc_papers

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Vacari, C. (2017). Online mobilization in comparative perspective: digital appeals and political engagement in Germany, Italy, and the United Kingdom. *Political Communication, 34*(1), 69–88. https://doi.org/10.1080/10584609.2016.1201558

Walker Wilson, M. J. (2010). Behavioral Decision Theory and Implications for the Supreme Court's Campaign Finance Jurisprudence. *Cardozo Law Review, 31*(3), 679–748.

Warner, M. (n.d.). The honest ads act. Retrieved January 13, 2020 from https://www.warner.senate.gov/public/index.cfm/the-honest-ads-act.

Wong, J. C. (2018, Mar 19). 'It might work too well': the dark side of political advertising online.

The Guardian. Retrieved from
https://www.theguardian.com/technology/2018/mar/19/facebook-political-ads-social-media-history-online-democracy

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, *112*(4), 1036–1040. https://doi.org/10.1073/pnas.1418680112

Zarsky, T. Z. (2019). Privacy and manipulation in the digital age. *Theoretical Inquiries in Law, 20*(1), 157–188. https://doi.org/10.1515/til-2019-0006 Available at
https://www7.tau.ac.il/ojs/index.php/til/article/viewFile/1612/1713

Zittrain, J. (2014). Engineering an election. *Harvard Law Review Forum, 127*, 335–341. Retrieved from https://harvardlawreview.org/2014/06/engineering-an-election/

Zuiderveen Borgesius, F. J., Möller, J., Kruikemeier, S., Fathaigh, R. Ó., Irion, K., Dobber, T., … de Vreese, C. H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82–96. https://doi.org/10.18352/ulr.420

FOOTNOTES

1. Sometimes these two terms are conflated and used interchangeably. However, Bakshy et al. (2015) distinguish 'echo chambers' as when "individuals are exposed only to information from like-minded individuals" and 'filter bubbles' as when "content is selected by algorithms according to a viewer's previous behavior" (p. 1130) and Bruns (2019) distinguishes 'echo chambers' as a group choosing "to preferentially *connect* with each other to the exclusion of outsiders" and 'filter bubbles' as a group choosing "to preferentially *communicate*" (p. 4).

2. For example, researchers at the University of Amsterdam have developed ALEX to unmask the functioning of personalisation algorithms on social media platforms. See: https://algorithms.exposed (Milan & Agosti, 2019)

3. Such a co-regulatory approach may be particularly well-suited as a governance mechanism as demonstrated also by Mard and Milan (2018) with respect to content regulation and fake news.

4. The GDPR places other restrictions on data collection and processing, as well as individual rights, that also limit micro-targeting (see Zuiderveen Borgesius et al., 2018 and Dobber et al., 2019). Dobber et al., for example, point out that data regarding people's 'political opinions' falls within the category of sensitive data.