MICROTARGETING
IN GERMANY FOR THE 2019 EUROPEAN ELECTIONS

Simon Hegelich & Juan Carlos Medina Serrano
Contents

Introduction  page 3
Transparency and political advertising  page 4
The advertising archive  page 4
Political online advertising of the German parties  page 5
Success of microtargeting?  page 6
Reverse engineering of the targeting strategies  page 7
Advertising versus organic reach  page 13
Bibliography  page 14
Authors  page 16
Imprint
Introduction

There are ever more reports worldwide stating that elections are being influenced by microtargeting in social networks (Hegelich & Thieltges, 2019). Since the successful Obama campaign in the United States, there has been an assumption that targeted political advertising on platforms such as Facebook is a powerful instrument to support election campaigns. Especially the Trump campaign made intensive use of this opportunity and spent over $44 million on Facebook advertising (Frier, 2018). This online strategy has repeatedly been mentioned as one of the keys to the success of the Trump election campaign.

"Microtargeting” means communicating targeted advertising to voters based on data analysis. At first glance, this represents a completely normal approach for election campaigns. A politician delivering a speech to the Chamber of Agriculture will likely address different topics than she would if speaking at a labour union meeting. However, targeted advertising in social networks alters this approach considerably and poses fundamental ethical questions. Communication in social networks is fragmented, which means that it is possible for various segments of the public to absorb completely different messages. This leads to the risk that politicians or parties could disseminate contradictory messages, thus obfuscating their true political agendas. The second risk is much more difficult to assess. Spreading information in social networks may provoke unforeseen effects. A cleverly drawn-up campaign might take advantage of this structure and make it efficacious in a completely new way. This is based on the fact that information is spread throughout social networks using two different mechanisms: on one hand when users forward it themselves or embed it in their messages and on the other hand by means of the platforms' algorithms that control which messages are ultimately shown to which users. Because the interaction of users in turn acts as the foundation for these algorithms, situations could arise in which these two mechanisms amplify each other. As of yet, no one knows which effects would then be conceivable. However, it has been clearly documented that social networks can have an impact on politics. Facebook itself already proved in a 2012 study that significantly more people go to the polls when they see that their Facebook friends are also voting (Bond et. al, 2012: 295). Presumably this effect could also be reversed and used for demobilisation. Against this backdrop, it is important to pay close attention to microtargeting during current elections like the European elections.
Transparency and political advertising

In light of the aforementioned risks, transparency is the utmost priority. Only when the online efforts of political stakeholders have been revealed can there be an estimation of whether there is actually a need to act and whether online election campaigns fundamentally differ from what is already happening offline. Given the increasing public and political pressure, the major social media platforms Facebook, Google (including YouTube), and Twitter have agreed on a "Code of Conduct" with the EU Commission, comprising measures including for more transparency (European Commission, 2019).

We must pay attention whenever we discuss social network data given that an interest in transparency invokes a conflict of objectives between the users, advertisers and other actors. Users have an interest in having their personal data. We must pay attention whenever we discuss social network data given that an interest in transparency invokes a conflict of objectives between the users, advertisers and other actors. On the other hand, the parties and other political stakeholders are interested in not completely disclosing their election campaign strategies. However, the question can be posed in this context if those political parties do enough to ensure the same transparency that they in turn demand from the platform operators. This conflict of interests leads to data relating to political advertising being provided in an aggregated form. Facebook and Google now offer application programming interfaces ("APIs") which are available to scientists. These interfaces include advertising considered political, as well as other data such as the customer, date, spread, and advertising costs.

Moreover, aggregated information is also made available regarding where (federal state) which advertising was seen by whom (demographic breakdown by age and gender). These efforts by the platform operators are the first key step in establishing transparency. We have thoroughly analysed these tools, which allows us to first point out the difficulties and shortcomings entailed by these approaches. Secondly, the APIs provide us with insight into paid online political advertising during the European elections in Germany.

The advertising archive

We used the Facebook Ad Library API and the Google Cloud BigQuery API to analyse microtargeting in Germany during the European elections. We were able to analyse the advertising placed on Facebook, Instagram, in Google searches, and on YouTube. We used these interfaces daily from the date when the APIs were provided (March 15, 2019) until a week following the European elections (June 2, 2019) and collected all advertisements from Germany. We retrieved approximately 34,000 advertisements via both these APIs. However, it is not clear which criteria are applied when deciding whether an advertisement is political. While Google refers to known political stakeholders, Facebook attempts to define political advertising by its content. For example, this means that any advertisement in Germany pertaining to the debate on refugees should show up in the Ad Library. Both of these criteria are prone to error. There are advertisements which emerge in the archive which have no political link, and conversely there must also be political ads which are not recorded in the archives. It is quite natural that errors like this occur, especially when automated systems are used to distinguish between political and non-political advertising. As a matter of principle, this does not present any further problems. Admittedly,
the difficulty here is that the whole issue of transparency then recurs on another level. Not only is it unclear which advertising has been placed, but also how the advertisement were identified or which errors occurred during the process. However, only the platform operators themselves possess this information. As a consequence, the utility of the information provided for the purposes of scientific analysis is tremendously questionable since it is impossible to evaluate the reliability of the data relating to online political advertising.

Political online advertising of the German parties

First, we studied which advertising strategies were pursued by the parties in Germany. To do so, we selected all the advertisements placed by accounts of the seven parties represented in the German Bundestag. Each advertisement is associated with a single social media account, but the parties run a series of different accounts. We limited ourselves to the official accounts of the federal parties and their state chapters and omitted advertisements which may have been placed by municipal chapters or individual politicians.

Table 1: Online advertising and spending of the parties

<table>
<thead>
<tr>
<th>Party</th>
<th>Ads</th>
<th>Facebook spending</th>
<th>€ per ad</th>
<th>Ads</th>
<th>Google spending</th>
<th>€ per ad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative for Germany (&quot;AfD&quot;)</td>
<td>48</td>
<td>22,278</td>
<td>464</td>
<td>17</td>
<td>23,400</td>
<td>1,376</td>
</tr>
<tr>
<td>Christian Democratic Union (&quot;CDU&quot;)</td>
<td>17,449</td>
<td>296,801</td>
<td>17</td>
<td>33,120</td>
<td>261,200</td>
<td>8</td>
</tr>
<tr>
<td>Christian Social Union in Bavaria (&quot;CSU&quot;)</td>
<td>27</td>
<td>60,816</td>
<td>2,252</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Free Democratic Party (&quot;FDP&quot;)</td>
<td>5,456</td>
<td>138,762</td>
<td>25</td>
<td>259</td>
<td>32,600</td>
<td>126</td>
</tr>
<tr>
<td>The Greens</td>
<td>7,804</td>
<td>229,451</td>
<td>29</td>
<td>769</td>
<td>140,750</td>
<td>183</td>
</tr>
<tr>
<td>The Left</td>
<td>958</td>
<td>41,526</td>
<td>43</td>
<td>7</td>
<td>3,200</td>
<td>457</td>
</tr>
<tr>
<td>Social Democratic Party (&quot;SPD&quot;)</td>
<td>15,234</td>
<td>283,664</td>
<td>19</td>
<td>90</td>
<td>13,390</td>
<td>149</td>
</tr>
</tbody>
</table>

Table 1 provides an outline of the online advertising activities of the parties. The information relating to the number of advertisements is from the APIs (Facebook, Inc. 2019, Google Developers 2019). The spending figures were generated using the platforms' reporting tools for the same observation period because it was possible to determine the exact spending using the information relating to individual advertisements. In contrast, the API only provides extremely vague spending ranges. The CDU and SPD have evidently taken up microtargeting. Both parties placed more than 15,000 advertisements on
Facebook. The content of these advertisements was often identical, but they were used for different target groups, which can be seen in the low average cost per advertisement. The SPD and CDU only spend around €19 and €17, respectively, per ad placed. This leads to the assumption that these parties have begun experimenting with different targeting approaches. Facebook’s advertising tool allows a budget to be defined for a group of advertisements. Each advertisement is deployed to various target groups, and Facebook automatically registers where the most reactions have been generated and subsequently pushes this targeting. In contrast, the AfD currently does not perform any microtargeting (Serrano et al., 2019). The party placed extremely few ads but provided each of these with a very high budget. This is even more extreme for the CSU, although it must be pointed out that the actual European election campaign was run by the CDU. It is rather remarkable that the CDU is the only party which placed Google advertisements on a large scale and which also pursued a very evident microtargeting approach in doing so. The CDU spent an average of €8 on each advertisement. In comparison, the Greens, the party with the second highest amount of spending for advertising on Google at €140,000, spent an average of €126 per advertisement. This makes the CDU the party which invested the most money by far in online advertising and which pursued the most professional microtargeting strategy – to the extent this can be ascertained from the aggregated data. This dominant position represents a general upheaval for online election campaigns and is in line with the observations gathered at the 2017 parliamentary elections (Serrano et al., 2018). During this election, the CDU attracted attention with its extremely professional tool linking online campaigning and canvassing.

**Success of microtargeting?**

As described above, microtargeting is already part of the online election campaigns in Germany. This poses the questions of whether this strategy has also had any impact (Papakyriakopoulos, 2017). The ultimate goal of campaign advertising is gaining votes. However, it is imperative not to indulge in the temptation to simply compare advertising budgets with election results because it is obvious that the key factors are not defined by advertising. For example, when you consider the general trends since the parliamentary elections, there is speculation that the Greens would have generated good results even without advertising, whereas the SPD would have delivered poor results even with tremendous advertising spending.

Therefore, we created a new data set to measure the success of online advertising. On Facebook, normal posts can be spread further by means of advertising spending. Thus, the reach of these boosted posts can also be compared with the posts of parties with no additional advertising spending. Using a further Facebook API (CrowdTangle), we were able to see how frequently all the parties’ posts had been seen on Facebook. We identified 11,496 Facebook posts from the parties during the observation period. Because the Facebook Ads API does not include information regarding whether it is a newly created advertisement or a boosted post, we found all posts with the contents of the advertisements and were able to identify 522 posts which had been sponsored. If you now compare the average number of how often a post was seen with or without advertising (impressions), you will realise that sponsored posts are in fact viewed by more people. This means that advertising works at least in the sense that it increases the message’s range. Of course, the effect this has on voters is quite a different matter. This
comparison sounds straightforward although it is relatively sophisticated in terms of methodology since merely determining a difference is insufficient in data analysis. You also want to be sure that this difference was not simply a coincidence. However, the distribution of the number of impressions of social media posts is very inconsistent. Some posts are viewed by very many people, while others nearly get lost in the shuffle. Because of this variation from normal distribution, the differences in mean value cannot be studied using traditional tests (ANOVA). Additionally, there are posts with minimal, moderate, and maximum spending. Consideration must also be given to the reasons why our analysis varies greatly from standard methods. (We make use of a Kruskal-Wallis test with subsequent application of the Wilcoxon signed-rank test with Bonferroni correction.)

Ultimately, we found that boosting posts actually increased reach for all parties, with just the exception of the union parties of the CDU and CSU. The Left and the Greens did see a positive effect, but only for those posts which cost a great deal of money. What are the possible conclusions from this? First, it must be noted that we are only referring to one minor instrument of online advertising here, namely boosted posts. Therefore, this effect cannot necessarily be generalised to the rest of the advertisements. According to our estimations, though, these results are in line with the following image. The CDU in particular placed a great number of advertisements and spent a lot of money on them. However, the question arises whether these resources were also applied purposefully, and the following section will reveal that there is doubt in this context with regard to the CDU.

Whatever the reason may be, the empirical findings (that the party with the greatest efforts generated the least effect) indicate that the impact of microtargeting was relatively weak for this election. At the same time, though, it has become apparent that effects should be anticipated whenever the parties are able to deploy microtargeting in its full breadth and depth.

**Reverse engineering of the targeting strategies**

Microtargeting is taking place and theoretically it can be effective. However, its potential was apparently not yet exhausted during the European elections. Therefore, the question arises which targeting strategy the parties are actually pursuing. There is no straight-forward answer. Platform operators can only share this kind of information to a limited degree because in doing so, data is disclosed relating to both advertisers and users. On the contrary, although parties could easily publish archives online disclosing the target groups for each advertisement, they won’t in order to prevent their election strategies from becoming transparent.

Thus, only the aggregated data in the ad archive is available to gain insight into the targeting strategies. With this information, it is possible to break down who has seen each advertisement by the percentage of men and women from each age group and each federal state. This data can be used to deduce the underlying advertising strategy. If an advertisement was primarily viewed in one federal state, it was likely shown there deliberately. An ad primarily seen by women was presumably also placed purposefully, although that does not necessarily mean that the parties have differentiated by gender when advertising. For example, if there was a Facebook setting to show an advertisement to those users who liked the page of a pop singer such as Ariana Grande or Taylor Swift, more young women
are likely to see this ad, so we cannot reconstruct exactly what the actual advertising strategy was. But if you regard all of one party’s advertisements, you can recognise how detailed the microtargeting was. If a party systematically addresses completely different target groups, this should become apparent in a great variance in the demographic and regional groups reached with the advertisement.

A variance analysis like this can be represented with the help of data science visualisations. We use Principal Component Analysis (PCA) for this purpose. This algorithm is not trivial mathematically but it leads to a representation which can be interpreted intuitively. For each advertisement, we used various values as a starting point to show the percentages of users which have seen the advertisement and which age range and gender they belong to (e.g. F(emale) 18-24 or M(ale) 25-34). There is also corresponding data for each advertisement relating to the distribution among the federal states. It is not easy to visualise a multidimensional space. The PCA projects the data in a two-dimensional space in such a way that those two dimensions capture as many of the differences as possible in all starting dimensions. This means that data points which are located far apart in this PCA plot are very different.

In a second step, all of the original dimensions are transferred to this plot as vectors. In this dimension, points which move in the direction of an eigenvector like this have higher values. The length of the eigenvectors also shows how important this dimension is for explaining the differences in the data.

This sounds more complicated than it is, as shown in Figure 1.
The plot indicates that there are distinct differences between the number of men and women who have seen an advertisement. The corresponding vectors clearly point in different directions (male to the right – with a slight turn upward and female to left – also with a slight twist). In addition, the vectors with regard to the age groups have a nearly symmetrical form from young (upward) to old (downward) – both also slightly rotated. For example, it can observed here that the Greens reached an above-average amount of women with their advertisements, which suggests that this group was consciously targeted. The AfD and Left predominately reached older men. However, so little data is available for the AfD that this could also reflect the demographic structure of its supporters rather than the result of a targeting strategy. The points at the edges are particularly interesting because the corresponding dimensions have led to vast differences there. Apparently both the FDP and the SPD targeted men and women between 25 and 44 years old and did not show the corresponding advertisements to anyone else. This also stands out because the rather randomly placed points for both parties can be found in the male
segment. Thus, there are several indications that these two parties placed certain advertisements exclusively for women of working age, whereas the CDU advertisements seem to be staggered randomly around the centre of the plot. This is evidence that the CDU carried out very little focused targeting.

We can supplement this insight further with the regional distribution of the advertising (Figure 2).

*Figure 2: Regional targeting*
Most of the federal states actually play no role in explaining the variance of the data, as evidenced by the short eigenvectors. Only the three largest states by area – Baden-Wuerttemberg (BW), North Rhine-Westphalia (NW), and Lower Saxony (NI) – form significant axes. This is not surprising to those who are familiar with Facebook's advertising structure. The more detailed the targeting is, the more expensive the advertisements are. For example, if only FDP supporters in Saarland should be addressed, this could potentially be more expensive than a nationwide campaign. Therefore, an approach like this only makes sense if extremely specific content should be conveyed. Of course, it is questionable if this would make sense for a European election. Nevertheless, the image which was already evident for the demographic targeting recurs here. The SPD has many advertisements on the plot's edges. It is obvious that there were many experiments with targeting here and that some advertisements were purposefully shown only in certain regions. In contrast, the CDU seems to have shown all its advertising more or less nationwide. However, perhaps this “strategy” is the reason that the CDU's advertising appears less effective than that of other parties.

We also conducted a YouTube experiment to obtain further insight into the parties' targeting strategies. We generated 16 YouTube accounts (Google accounts) with identical demographic features (a 26-year-old man with a common German name). We subsequently created a list with 13 popular but non-political YouTube channels from the areas of news, sports, and culture. In addition, we sought out the YouTube channels of the parties, as well as 45 bloggers who actively support one of the parties. For each party there was then one user who only followed the political channels of his party and one who followed these channels as well as the non-political channels. Two control accounts only followed the non-political channels. The accounts were managed completely automatically and initially “watched” at least 20 videos from the corresponding party each day (these videos were started and ran through to the end). Then, a randomly selected video suggested by YouTube was started. Every two hours the accounts than started five further videos. During this process, we recorded the advertisements that were shown between the videos. After two weeks, each of the accounts had viewed 480 videos and we were able to record a total of 2,025 advertisements, which we broke down into five categories: politics, travel, entertainment, trade, finance, and miscellaneous.

Overall, we stumbled upon very few political advertisements: six from the AfD, two from the CDU, and two from the Left. There were a few additional advertisements from parties which are not represented in the Bundestag, but only 1.4% of all ads were political. The AfD advertisements were shown to the account only following the Greens, both accounts following the CSU, and both non-political accounts. The underlying data of this experiment is insufficient to derive conclusions. Nevertheless, interesting trends became obvious. Figure 2 indicates the percentage distribution of the advertising categories among the various accounts. Contrary to our expectations, the non-political accounts saw the largest percentage of political advertising.

It is also interesting to shift perspectives: commercial targeting is occurring at the same time as political microtargeting. Political views are apparently also a key signal for this commercial targeting. At least in the fields of travel and finance, the share of the advertising shown varies greatly depending on the political stance. In this specific case, this would most likely not have an impact on the political discourse. However, in general this raises the question of whether people with different political views aren’t
progressively active in various 'versions' of the Internet, which would not just make a subtle impact on the world view conceivable, but also lead to an ever increasing fragmentation of the public (Shahrezaye, 2019).

Figure 3: YouTube advertising and political accounts
Advertising versus organic reach

Paid advertising only represents a small part of the activities of parties in election campaigns. In addition, all the parties are using social media platforms more frequently to spread certain messages to the nation, especially leading up to elections. Often the content of organic posts does not differ from the content of paid advertising. The difference is that these posts are free of charge for the parties and that they cannot pursue targeting when disseminating them. This blurs the distinction between advertising and organic political content. This has the direct consequence that there is no regulation at all in place for organic election campaigning on social media. Moreover, the idea of the parties’ digital influence shifts sharply when the reach of advertising and the organic reach are contrasted (see Table 2).

Table 2: Advertising and organic reach

<table>
<thead>
<tr>
<th>Party</th>
<th>Facebook/Instagram Impressions</th>
<th>Facebook Interactions</th>
<th>Google/YouTube Impressions (1%)</th>
<th>YouTube Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>from (1%)</td>
<td>to (1%)</td>
<td>from (1%)</td>
<td>to (1%)</td>
</tr>
<tr>
<td>Alternative for Germany (“AfD”)</td>
<td>17,600</td>
<td>39,019</td>
<td>2,632,650</td>
<td>27,600</td>
</tr>
<tr>
<td>Christian Democratic Union (“CDU”)</td>
<td>143,780</td>
<td>639,673</td>
<td>256,493</td>
<td>148,600</td>
</tr>
<tr>
<td>Christian Social Union in Bavaria (“CSU”)</td>
<td>23,750</td>
<td>55,109</td>
<td>197,422</td>
<td>0</td>
</tr>
<tr>
<td>Free Democratic Party (“FDP”)</td>
<td>169,850</td>
<td>576,462</td>
<td>268,756</td>
<td>74,200</td>
</tr>
<tr>
<td>The Greens</td>
<td>141,900</td>
<td>474,408</td>
<td>262,415</td>
<td>49,900</td>
</tr>
<tr>
<td>The Left</td>
<td>59,050</td>
<td>158,404</td>
<td>253,981</td>
<td>11,100</td>
</tr>
<tr>
<td>Social Democratic Party (“SDP”)</td>
<td>177,540</td>
<td>638,163</td>
<td>245,865</td>
<td>32,900</td>
</tr>
</tbody>
</table>

In order to contrast advertising and organic reach, we have to compare two different data sources which each have different information. We obtained the impressions (how frequently the ad was seen) for each advertisement from the Facebook Ad Library. However, a range (0-9, 10-99, 100-499, …) is specified for these impressions rather than an exact value. For all of the advertisements of each party, we added the lower limit (from) and the respective upper limit (to). Using the CrowdTangle API, which we employ to collect information relating to the organic posts, we do not obtain any information about impressions. Instead, we have data about the interactions, meaning how frequently users interacted.
with posts on the pages. In order to make these figures easier to compare, even though they do not signify the exact same thing, we divide the impressions by 100. Generally, a click-through rate of 1% is often assumed, which means that for every 100 people who have seen an advertisement, one person reacts to it.

It is evident that although the AfD is the party with the least advertising reach, it is the party with the furthest organic reach by far. According to our calculation, only between 17,600 and 39,000 people interacted with the AfD’s advertisements, whereas there were 2.6 million interactions on the AfD’s Facebook pages during the observation period, which is ten times more than for any other party. The AfD also has the most reactions on YouTube.

We could now simply congratulate the AfD on their successful online strategy since these numbers have nothing to do with microtargeting. However, we consider this development to be tremendously alarming. On one hand, this is due to the fact that the AfD in particular has many posts whose designs look like political advertising. In light of their reach, it should be examined whether a political regulation should also apply for these posts, even if no money was paid for them. On the other hand, the deviations between the AfD and the rest of the parties is so large that it can be assumed that this represents non-authentic user behaviour (Serrano et al., 2019). Hyperactive accounts are seen systematically interacting with each post on the AfD pages. Elsewhere (Papakyriakopoulos, 2019), we were able to verify that these hyperactive users distort the online political discourse. Furthermore, it is also not discernible if money for these activities might also come from channels other than traditional campaign financing.

It is plausible that party supporters generate these interactions for free because of their political convictions. However, an unknown party may also have made investments in an infrastructure to generate social media interactions.

Especially for Facebook, the main issue is that this strategy of artificially driving up the number of interactions is obviously working. Facebook has set up its news feed to use “meaningful interaction” as the main indicator which the algorithm attempts to optimise. Specifically, this means that if posts from the AfD generate ten times more reactions, Facebook also classifies this content as ten times more important and shows it to more people in their news feeds without being asked. There must be further studies regarding the extent to which this effect actually occurs and the scope and dynamics of organic advertising.
Bibliography


The Authors

Prof. Dr. Simon Hegelich has been a Professor for Political Data Science at the School of Governance at the Technical University of Munich since 2016. He studied political science at the University of Münster, where he also received his doctorate and acquired his post-doctoral teaching qualification (habilitation). His research focuses on the connection between political science and computer science and political data science. He is interested in the political dimension of the ongoing digital revolution as well as in answering traditional political science questions with the help of machine learning, data mining, computer vision, and simulations.

Contact: simon.hegelich@hfp.tum.de

Juan Carlos Medina Serrano is a data scientist working at the Bavarian School for Public Policy of the Technical University of Munich. He was born in Mexico and studied Physics Engineering at the Tecnológico de Monterrey. He later moved to Germany, where he obtained two Masters degrees: Computational Science and Data Engineering and Analytics. Juan Carlos has worked as a data scientist in companies like BMW, Bosch, and Siemens. Currently, he is pursuing a PhD and his research involves the spread of misinformation in social media and its effects on the political landscape.

Contact: juan.medina@tum.de
IMPRINT

Publisher:
Media Authority of North Rhine-Westphalia
Zollhof 2
D-40221 Düsseldorf
T +49 211 77007-0
F +49 211 727170
info@medienanstalt-nrw.de
www.medienanstalt-nrw.de

In cooperation with the Bayerische Landeszentrale für neue Medien (BLM), the Berlin-Brandenburg Media Authority (mabb), and the State Media Authority of Rhineland-Palatinate

Project Managers:
Dr. Tobias Schmid (Director of the Media Authority of North Rhine-Westphalia)
Siegfried Schneider (President of The Bayerische Landeszentrale für neue Medien (BLM))
Dr. Anja Zimmer (Director of the Berlin-Brandenburg Media Authority (mabb))
Dr. Marc Jan Eumann (State Media Authority of Rhineland-Palatinate (LMK))

Coordination and Content Support:
Dr. Meike Isenberg (Media Authority of North Rhine-Westphalia)

Published:
October 2019