

Understanding the Complexity of Detecting Political Ads

Vera Sosnovik

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
Grenoble, France

Oana Goga

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
Grenoble, France

ABSTRACT

Online political advertising has grown significantly over the last few years. To monitor online sponsored political discourse, companies such as Facebook, Google, and Twitter have created public Ad Libraries collecting the political ads that run on their platforms. Currently, both policymakers and platforms are debating further restrictions on political advertising to deter misuses.

This paper investigates whether we can reliably distinguish political ads from non-political ads. We take an empirical approach to analyze what kind of ads are deemed political by ordinary people and what kind of ads lead to disagreement. Our results show a significant disagreement between what ad platforms, ordinary people, and advertisers consider political and suggest that this disagreement mainly comes from diverging opinions on which ads address social issues. Overall our results imply that it is important to consider social issue ads as political, but they also complicate political advertising regulations.

ACM Reference Format:

Vera Sosnovik and Oana Goga. 2021. Understanding the Complexity of Detecting Political Ads. In *Proceedings of the Web Conference 2021 (WWW '21), April 19–23, 2021, Ljubljana, Slovenia*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3442381.3450049>

1 INTRODUCTION

Social media and the public sphere's digitalization have changed the political campaigning landscape in both good and dangerous ways. While social media are creating new opportunities for engaging citizens in political conversations, they have also raised several risks for the integrity of elections and the political debate. For example, online ads can be tailored to specific groups of people, hence polarizing the voter base and distorting the political debate. Advertisers can buy large amounts of ads to flood people's social media feeds and steer public debates on issues that are of interest to them. Anyone, from political parties to interest groups, and specialized advertising companies such as Cambridge Analytica, can steer the political debate through online advertising.

Ad platforms have put forward several measures to mitigate risks and allow for public scrutiny of ads. Twitter and TikTok decided to ban political ads altogether. Google and Facebook allow political ads, but advertisers are subject to a higher degree of scrutiny and limitations. On Google, advertisers can only use geographic location, age, gender, and contextual targeting to target political ads. Facebook does not restrict the micro-targeting of political ads. Advertisers,

however, need to verify their account (by showing proof of identity or a public listing of their business [14]) and are only allowed to send political ads to users that reside in the same country as them. Moreover, advertisers have to self-declare when their ads are political, and all political ads sent on the platform appear in the Facebook Ad Library where the civil society can further scrutinize them [10].

On their side, governments are also working on legislation to regulate political advertising. For example, the European Commission is working on two pieces of legislation, the Digital Services Act (DSA) [8] and the European Democracy Action Plan (EDAP) [7], that aim to regulate in part online advertising. In a recent interview, Věra Jourová (the European Commission Vice-President for Values and Transparency), declared [23]:

... we are coming in the European Democracy Action Plan with the vision for the next year to come with the rules for political advertising, where we are seriously considering to limit the microtargeting as a method which is used for the promotion of political powers, political parties or political individuals.

Measures from both ad platforms and governments are positive developments. However, all of them implicitly rely on the assumption that *one can reliably distinguish political ads from non-political ads*.

In this paper, we take an empirical approach to test this assumption by analyzing the characteristics of ads deemed political by ordinary people, the characteristics of ads that lead to disagreement, and whether there are differences between what advertisers consider political and what ordinary people consider political. Our analysis is based on a dataset from ProPublica that contains 55k Facebook ads received by U.S. residents, labeled by at least one volunteer as political, and that received three or more votes (Section 2). The dataset was collected by a browser extension that collects the ads users see when they browse their Facebook timeline and allows users to label whether the ads they see are political.

First, we investigate whether ad platforms, volunteers, and advertisers agree on which ads should be considered political (Section 3). All ad platforms agree that ads from or about political actors and ads about elections and voting should be considered political. However, only Facebook and TikTok consider ads about *social issues* (such as climate change or immigration) as political. Our results show that volunteers disagree on whether an ad is political for more than 50% of the ads in the dataset, and only 83% of the ads labeled as political by advertisers are also labeled as political by a majority of volunteers. Hence, the fundamental assumption that we can clearly distinguish political from non-political ads does not hold, since there is no consensus even on what constitutes a political ad, and volunteers and advertisers label different sets of ads as political.

Next, we analyze the characteristics of ads that are labeled as political by volunteers and advertisers in the ProPublica dataset, which can be useful to inform the debate on definitions of political ads (Section 4). To that end, we gathered data about the advertisers

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '21, April 19–23, 2021, Ljubljana, Slovenia

© 2021 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-8312-7/21/04.

<https://doi.org/10.1145/3442381.3450049>

sending political ads and the content of their ads. We hired Prolific users to annotate 2300 ads with the political or social issues the ad is referring to. Our analysis shows that a wide range of advertisers (from political actors to NGOs and businesses) are posting political ads on Facebook and that ads about social issues account for a large fraction of the ads labeled as political; hence emphasizing the importance of including such ads in political ads definitions. Our analysis also shows that the ads labeled as political by volunteers and advertisers are very diverse. We see ads with a clear political message from advocacy groups (e.g., ads addressing abortion issues in the U.S.); but also ads from NGOs that address humanitarian issues and do not seem to directly or indirectly impact U.S. elections or legislation (e.g., ads asking for donations for ending world hunger). As political ads may be subject to higher restrictions, this questions whether it is desirable that the same restrictions apply to both types of ads. More generally, this emphasizes the need to account for the diversity of political ads in devising regulations.

We finally analyze the ads that lead to disagreement among volunteers and between volunteers and advertisers (Section 5). We first observe that advertisers mislabel some ads as either political or non-political (according to the Facebook ToS). Then we find that advertisers seem to underreport ads (that are considered political by volunteers) about social issues, especially the economy and civil and social rights. Volunteers seem to underreport ads (considered political by advertisers) from advertisers such as NGOs and charities, and about social issues, especially civil and social rights and health. Part of the problem may be that the definition of ads about social issues may be too broad and vague, which leads to being interpreted in different ways by people. This also raises the question of whether *all* ads related to social issues should be considered political, and if not, how to filter social issue ads that are not political.

Because of the high volume of ads, enforcement mechanisms need to rely on automated machine learning (ML) algorithms to detect political ads. However, it is not clear how one should train and evaluate such models since there is disagreement on which ads are political (i.e., the positive examples). To investigate that, we train four classifiers with different groups of positive examples (coming from advertisers and volunteers). We test how they perform over various groups of political ads with varying degrees of disagreement (Section 6). While all classifiers achieve high accuracy in detecting ads everyone agrees are political; their accuracy drops on ads that only a few find political.

Another important question is whether (and to which extent) models trained with labels from advertisers would declare as political the same ads as models trained with labels from volunteers (i.e., reliable detection of political ads). Theoretically, if ads labeled as political by advertisers and volunteers are representative of political ads in general, the resulting models should declare the same ads as political. Our results show that the overlap between different models is relatively high (ranging from 83% to 97%), but that discrepancies in the input data transfer to discrepancies in the output data. This suggests that existing labeled datasets are not providing a representative set of political ads needed to build reliable detection schemes.

Overall, our work suggests that, given the complexity of deciding which ads are political, it would be beneficial to have ad libraries that contain *all* ads running on the platform, not only ads deemed political by the ad platform. Following this work, we issued a statement

Table 1: Number of ads in the ProPublica and the AdAnalyst datasets, and percentage of ads with official “Paid for by” political disclaimer.

	All ads	Official political	Official non-political
ProPublica	54.6k	50.8k (93%)	3.8k (7%)
AdAnalyst	9k	2%	98%

together with civil societies asking for “*Universal advertising transparency by default*” that we submitted to the DSA consultation [17]. However, this crucial first step is not enough because political ads are also subject to higher restrictions; hence, we still need to detect political ads reliably. We hope this study can help policymakers to define political speech and decide on appropriate restrictions and ad platforms to set infrastructures for detecting political ads.

2 DATASETS

For our analysis we use the following two datasets of ads that users have received on their Facebook timeline:

ProPublica dataset. ProPublica, an investigative journalism organization, has developed a browser extension that collects the ads users are receiving on Facebook and allows users to label whether the ads they are seeing are political or not [27]. The extension is currently maintained by the NYU Online Political Transparency Project [25]. While ProPublica was not able to make available all the ads it has collected, it shared with us **all the ads for which at least one user has labeled it as being political**, as well as **all the ads that have the “Paid for by” disclaimer** (i.e., the official political ads that have been declared as such by advertisers). This dataset is valuable because it provides us with a unique view of which ads are considered political by “ordinary” people/volunteers. To our knowledge, there are no studies of such data.

For this study, we only kept ads with at least three votes (either political or non-political) and that were received between June 2018 and May 2020; resulting in a dataset of 54.6k ads coming from 7530 advertisers. The median number of votes per ad after filtering is 5. We call the ads that have the “Paid for by” disclaimer the **official political ads** and the ads that do not have the disclaimer the **official non-political ads**. Table 1 shows the number of ads in the ProPublica dataset as well as the fraction of **official political ads** and **official non-political ads**. Note that this dataset does not contain a representative sample of political ads as they are ads received by people who answered ProPublica’s call for action to install the tool.

AdAnalyst dataset. Similar to the extension provided by ProPublica, AdAnalyst collects the ads users see on their Facebook timeline [2]. The AdAnalyst dataset contains over 500k ads from users in various countries. For this study, we keep only ads in English (detected using text-blob python library [35]) and that targeted users in the US between October 2018 and May 2020. For this, we use information about ad targeting available in the “Why am I seeing this ad?” button and select only ads targeted at people who live in the USA or visited places in the USA recently. The resulting dataset contains 9k unique ads (198 ads with “Paid for by” disclaimer and 8802 without). This dataset does not have votes from volunteers.

Ethical review board and reproducibility. Both data collection by ProPublica and AdAnalyst were approved by the respective ethical review boards. The ProPublica data is available to the public through

a request form [28]. The 9k ads from AdAnalyst, the data collected from the Prolific studies, and other supplemental material is available at <http://lig-membres.imag.fr/gogao/www21.html>.

3 DISAGREEMENT ON POLITICAL ADS

The base to detect political ads reliably is to agree on which ads should be considered political and which ads should not. In this section we look at whether ad platforms, volunteers, and advertisers agree on which ads are political.

3.1 Disagreement across ad platforms

The Terms of Services of different ad platforms provide information on which ads they consider as political. We review the definitions and restrictions for political advertising across four ad platforms.

Facebook defines political ads as: “*Made by, on behalf of, or about a candidate for public office, a political figure, a political party, or advocates for the outcome of an election to public office; About any election, referendum, or ballot initiative, including “go out and vote” or election campaigns; About social issues in any place where the ad is being placed; Regulated as political advertising.*” The social issues are: *civil and social rights, crime, economy, education, environmental politics, guns, health, immigration, political values and government, security, and foreign policy* [14].

Everyone with a Facebook account can be an advertiser if they provide a payment method. However, to be able to send political ads, advertisers need to verify their accounts by providing proof of their identity [15]. Advertisers can only send political ads in the country they reside and need to provide proof of residence. Advertisers need to self-label their ads as political and need to provide a disclaimer about who paid for the ad. This “Paid for by” disclaimer appears on the top of the ad frame, after the advertiser’s name. Finally, Facebook adds the political ads to their Ad Library [10].

Google defines political ads as: “*ads about political organizations, political parties, political issue advocacy or fundraising, and individual candidates and politicians*” [18]. The platform imposes no restrictions on political ads, but the platform expects all political ads to comply with local legal requirements. Google considers *election ads* as a separate category. The definition of election ads depends on the country, but overall it refers only to ads from or about candidates and political parties during an electoral period. Only verified advertisers can run election ads. Election ads can only be targeted by geographic regions (but not by radius around a precise location), age, gender, and contextual targeting options such as ad placements, topics, keywords against sites, apps, pages, and videos.

Twitter defines political ads as “*ads with political content: that references a candidate, political party, elected or appointed government official, election, referendum, ballot measure, legislation, regulation, directive, or judicial outcome; as well as ads of any type by candidates, political parties, or elected or appointed government officials*” [36]. Twitter bans all political ads.

TikTok defines political ads as *ads that promote or oppose a candidate, current leader, political party or group, or issue at the federal, state, or local level — including election-related ads, advocacy ads, or issue ads* [6]. TikTok bans all political ads.

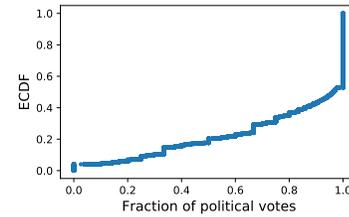


Figure 1: ECDF of the fraction of political votes for the ads in the ProPublica dataset.

Overall there are three categories of political ads: **ads from or about a political figure or political party**, **ads about elections**, and **ads about social issues**. While the precise definition of political ads varies across ad platforms, the most significant difference is that Twitter and Google do not consider ads about social issues as political while Facebook and TikTok do. While it is certainly a debatable question whether or not social issue ads should be regarded as political, the EU Code of Practice on Disinformation mentions both issue ads and political ads as sensitive content. Our results will show the importance of considering social issue ads as political and why they complicate political advertising regulations.

3.2 Disagreement among volunteers

At least three volunteers have labeled each ad in the ProPublica dataset as being political or non-political. The volunteers were given no instructions for what ads they should consider as political, and users were left to decide based on their instinct.

To observe to which extent volunteers agree on what ads are political, Figure 1 plots the distribution of the number of political votes divided by the number of all votes for each ad in the ProPublica dataset. We denote this fraction as fr . A fraction $fr = 1$ means that everyone agrees that the ad is political, while a fraction $fr = 0$ means that everyone agrees that the ad is not political. The plot shows that for more than 50% of the ads, at least one volunteer disagrees with the others (fr is neither 0 nor 1), which shows that *deciding whether or not an ad is political is debatable for more than half of the cases*.

To distinguish ads on which users agree they are political from the rest, we split the ads into four disjoint *ad groups* based on the volunteer votes. We will analyze them separately in the paper. The groups are defined as follows:

- **strong political ads:** ads with $fr = 1$, i.e., where everyone agrees that they are political;
- **political ads:** ads with $0.5 \leq fr < 1$, i.e., where there is some disagreement, but the majority labels them as political;
- **marginally political ads:** ads with $0 < fr < 0.5$, i.e., where there is some disagreement, but the majority labels them as non-political;
- **non-political ads:** ads with $fr = 0$, i.e., where everyone agrees that are non-political.

There are 26k **strong political ads**, 19.7k **political ads**, 7.6k **marginally political ads**, and 1.3k **non-political ads**.

Table 2: Number of ads in different ad groups (based on volunteer votes) and overlap with ads labeled as political by advertisers. † ProPublica was not able to give us access to ads that did not have at least one political vote and that were not labeled as official political ads.

	All	Official pol.	Official non-pol.
strong political ads	26k	96%	4%
political ads	19.7k	93%	7%
marginally political ads	7.6k	74%	26%
non-political ads	1.3k	100%	NA†

3.3 Disagreement between volunteers and advertisers

The ProPublica dataset provides data on whether an ad was labeled as political by the advertiser itself (see Section 2). Table 2 presents the overlap between ads labeled as political by volunteers and ads labeled as political by advertisers (the **official political ads**). The table shows that 96% of **strong political ads**, and 93% of **political ads** were also declared as political by advertisers. Hence, most ads considered political by the majority of volunteers are also considered political by advertisers. There are, however, 4% of **strong political ads** and 7% of **political ads** that advertisers did not label as political.

The more surprising finding is that advertisers label as political a large majority (74%) of **marginally political ads**. Looking the other way around, 83% of **official political ads** are labeled as political by most volunteers. In contrast, 15% of **official political ads** are only labeled as political by a minority of volunteers, and 2% of **official political ads** are not labeled as political by *any* volunteer. Hence, many ads considered political by advertisers are not regarded as political by volunteers. While the reasons can be diverse (this is the subject of Section 5), we conclude that *there is currently a significant discrepancy between the ads labeled as political by advertisers and by volunteers*.

Takeaway: The assumption that we can clearly distinguish political from non-political ads does currently not hold as there are significant disagreements between ad platforms, volunteers, and advertisers on which ads are political. Therefore, it is problematic to apply restrictions on political ads if the decision of whether an ad is political depends on the person labeling it.

4 WHAT GETS LABELED AS POLITICAL

This section provides a general view of ads labeled as political by volunteers and advertisers and analyzes who sends them and what are they talking about. This analysis is relevant for informing the debate on definitions of political ads and understanding the impact of potential regulations. The next section will focus on which ads lead to disagreement.

To interpret the results, we need to know the precise conditions in which the labeling happened. The ProPublica *volunteers* were given no instructions for what ads they should consider as political, and they were left to decide based on their subjective beliefs and background knowledge. However, volunteers could see if an ad was labeled as political by the advertiser itself (these ads have a “Paid for by” disclaimer on Facebook). We present results separately for ads that run with a disclaimer and ads that run without a disclaimer to isolate the potential effect of the “Paid for by” disclaimer.

Advertisers have to self-declare if they send political ads (as defined by Facebook’s ToS). However, there is no public information on how Facebook enforces this policy [33]. Hence, ads labeled as political by advertisers are either a product of their own belief that their ad is political; or the result that Facebook constrained them to label the ad as political to run on the platform (maybe due to false positives in their enforcement algorithms).

4.1 Analysis of advertiser categories

To characterize advertisers we analyze their category. Advertisers need to create a Facebook Page and select from a pre-defined list a category for their page such as “Software Company” or “Political Party” [16]. We use the advertiser’s ids available in the dataset to collect their category using the Facebook Graph API. Some pages no longer exist, we were able to extract categories for 6476 ProPublica advertisers (82%). Figure 2 plots the breakdown of the corresponding advertisers categories for **strong political ads**, **political ads**, **marginally political ads**, **official political ads** and **non-political ads**. We group similar advertiser categories (grouping details can be found in our supplementary material at <http://lig-membres.imag.fr/gogao/www21.html>).

Figure 2 shows that most **strong political ads** come from political actors (58% w. and 48% w/o. disc.), but a significant fraction of ads also come from NGOs (14% w. and 21% w/o. disc.), communities (4% w. and 4% w/o. disc.), and advocacy groups (3% w. and 4% w/o. disc.). In the **political ads** group, a smaller fraction of ads come from political actors (24% w. and 25% w/o. disc.), much more from NGOs (36% w. and 37% w/o. disc.), and we also see more ads from advocacy groups (6% w. and 6% w/o. disc.), news media (4% w. and 4% w/o. disc.), and communities (6% w. and 5% w/o. disc.). In the **marginally political ads** group, only (1% w. and 1% w/o. disc.) of ads come from political actors, the majority (52% w. and 58% w/o. disc.) from NGOs and charity organizations (11% w. and 5% w/o. disc.), some ads come from news media (5% w. and 3% w/o. disc.) and businesses (5% w. and 3% w/o. disc.). In the **official political ads** group, we see a similar diversity in the advertisers labeling their ads as political. *Many countries’ specific electoral legislations only regulate (and impose restrictions on) ads from political actors. However, we see that there is a wide range of advertisers pushing political ads online and that volunteers do label ads from these advertisers as political; hence, prompting for updating legislation.*

Facebook is explicitly exempting news organizations from labeling their ads as political even if they are about political issues [14]; however, yet do seem to consider these ads as political. *This raises the question of whether ads from news media should be treated as political ads. On one side, political journalism is different from political propaganda; on the other side, news media has been used as a tool to manipulate users, and many unauthentic news aggregators are emerging with the purpose of promoting a political agenda [4].*

Table 3 presents examples of political ads from different categories of advertisers such as community, NGO, or business. For each ad, the table shows the fraction of political votes divided by all votes from volunteers and whether the ad was labeled as political by the advertiser itself. The table shows that there is a wide diversity of ads getting labeled as political. For instance, we can see an ad from the ice-cream company “Ben and Jerry” (a business) that is inciting

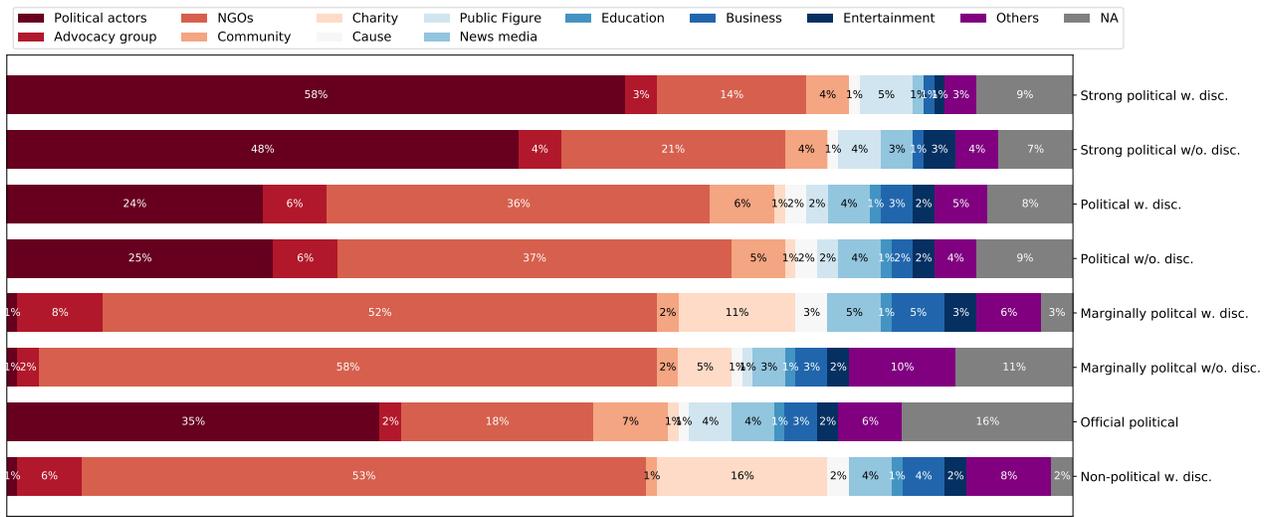


Figure 2: Breakdown of advertisers categories for different groups of ads for ads with and without “Paid for by” disclaimer.

Table 3: Examples of ads from advertisers with different categories.

Advertiser	Text	fr	disc.
Category: Cause			
UnRestrict Minnesota	6% of Minnesotans don’t know the abortion laws in our state.	1	w.
Care2	U.S. Wildlife Services is putting the safety of people: animals at risk in its attempt to control wild predators. Tell them to STOP using taxpayers money to kill wild animals lethally!	0.75	w.
Oregon Forests Forever	Brave men and women from Oregon are helping to fight fires in California	0.37	w.
Home Ownership Matters	Do you want Congress to invest in infrastructure? Click here to sign the petition.	0.33	w.
Category: Charity			
USA for UNHCR	Should America turn away from this child? Not now, not ever. It’s not who we are.	1	w.
World Food Programme	I call on warring parties to allow the constant flow of food for innocent and starving people in Yemen. Add your voice to our petition today.	0.66	w/o.
ChildFund International	She wants a childhood free of worry and a future full of promise.	0	w.
USA for UNHCR	All donations MATCHED for a limited time. People in Syria are still fleeing for their lives. UNHCR needs your help to provide the shelter, food and medicine they need to survive.	0.33	w/o.
Category: Community			
Yes for Washington Elementary Students	Vote YES on the WESD Override to protect full-day kindergarten, music, art, and physical education in our schools.	1	w.
North Carolina Citizens	We have a new survey for North Carolina. Please click the link below to share your thoughts	0.8	w/o.
Healthy Me PA	Workplace violence is 4x more common in the health care industry. Here’s how you can help:	0.3	w.
Protect Coyote Valley	Time and time again, threats of development have been made in Coyote Valley, with some succeeding. We want to see Coyote Valley permanently protected for our wildlife and for our children. All we need is your signature	0.4	w.
Category: Business			
Dissent Pins	Stand for democracy on election day and every day with our Count Every Vote pin.	1	w.
Ben and Jerry’s	Vote YES on 4 and reinstate voting rights for 1.4 million Floridians!	0.96	w.
CREDO Mobile	Help us decide how to allocate our \$50k donation to 5 progressive environmental organizations fighting for climate justice.	0.33	w.
Steady Returns, LLC	Everyone deserves great financial advice!	0.2	w/o.
Category: NGOs			
Democratic Attorneys General Association	Now that we know Joe Biden will be the nominee, we want to know who you think he should pick as his V.P.? Hurry, this round closes soon and we are still missing your response.	1	w.
Pennsylvania Spotlight	Voting from home is easy. By taking thirty seconds to request a ballot, you can fill your ballot out on your couch and mail it in.	0.63	w.
National Audubon Society	Birds and their habitats are under attack, but with your help we can fight back. This Earth Day your monthly gift will go twice as far to protect birds and the places they need..	0.25	w.
FOUR PAWS International	Stray animals are starving in India, will you give them your much-needed support?	0.33	w/o.
Category: Political actors			
Arati Kreibich for Congress	Republicans are suppressing the vote through mass voter purges, polling place closures, and burdensome voter ID laws. Tell the Senate: restore the Voting Rights Act!	1	w.
Bernie Sanders	We are about to make history and I want you to be a part of it. Our campaign is trying to reach 1 million campaign donors faster than any campaign in American politics, and we are VERY close. Can you make a contribution right now to become one of our first million donors?	0.66	w.
Tina Smith	Meet Senator Tina Smith: a big fan of dogs, donuts, and Minnesotans.	0.66	w/o.
Judge Brian Hagedorn	Click here to hear how an adopted daughter changed the Hagedorn family!	0.33	w.

citizens to vote, and an ad from the “Democratic Attorneys General Association” (an NGO) that is asking people who should be the V.P. of Joe Biden. Such ads have a clear association with elections. In the table, we also see many ads, such as the ones from the “World Food Programme” and the “USA for UNHCR” (Charities), that address social issues but do not seem to have any evident association to elections or legislation. *The critical point to recognize is that ads labeled as political can have a very different level of “politicalness”,*

going from straight advocacy messages addressing abortion issues to ads merely asking for a donation to end world hunger.

4.2 Analysis of ad messages

To gather grounded information about the topics of ads labeled as political, we took a random sample of 300 ads with a “Paid for by” disclaimer and 300 ads without “Paid for by” disclaimer from each **strong political ads, political ads, marginally political ads,**

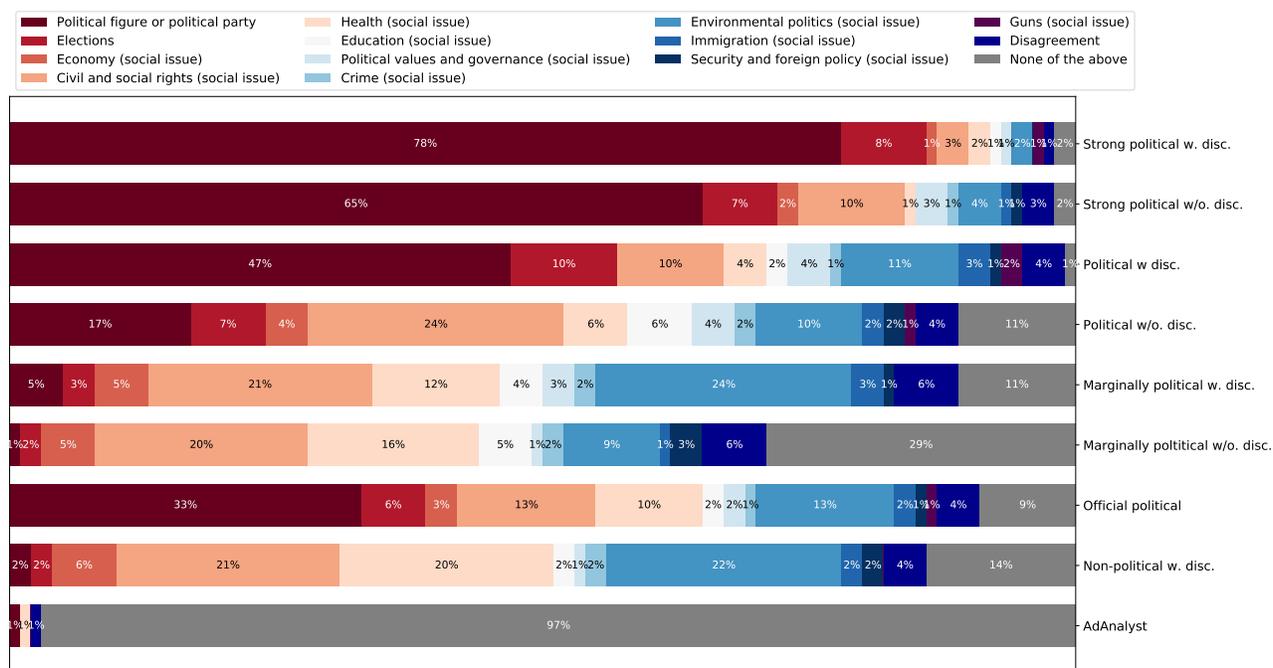


Figure 3: Breakdown of the political and social issues discussed in ads for the different groups of ads with and without disclaimer.

300 ads **non-political ads**, and 200 ads without disclaimer from AdAnalyst. While we picked both ads with and without a disclaimer, we did not show the disclaimer in our surveys. We set up a survey on Qualtrics [29] where for each ad, we ask respondents questions about the ad’s message. We hired workers through Prolific [26], and we redirected them to fill out the survey on Qualtrics. Each worker had to label 20 random ads from the pool of 2300 ads, and each ad was labeled by three workers. We selected workers that are residing in the USA since all the ads used in the experiments targeted people who lived in or visited the USA. The median amount of time that workers spent on the survey was 12 minutes.

Each survey had an instructions page, followed by 20 pages each containing one ad to label. For each ad, we asked the following questions: (1) “Is this ad made by, on behalf of, or about a political actor? (such as a candidate for public office, a political figure, a political party or advocates for the outcome of an election to public office)”; (2) “Is this ad about elections? (such as referendum or ballot initiative, including “go out and vote” or election campaigns)”; and (3) “Does this ad refer to a social issue? (such as civil and social rights, ...)”. Workers were allowed to answer yes to all the questions. If workers selected that the ad is about a social issue, we asked them which social issue: “Which social issue is this ad talking about?” Workers had to choose from the following list: civil and social rights, crime, economy, education, environmental politics, guns, health, immigration, political values and governance, security and foreign policy. We considered these social issues because they appear in the Facebook definition of political ads [13]. Workers were allowed to select multiple social issues if needed.

If workers answered no for all three initial questions (the ad is not about a political figure, election, or social issue), they were asked to choose from a list “What topic describes best the ad?”. We took

the list of 23 topics from the Interactive Advertising Bureau (IAB) categories [11]. Note that we did not ask workers whether the ad is political or not; we just asked them questions about its message.

Figure 3 shows the breakdown of the political or social issue discussed in an ad according to Prolific workers for different ad groups for ads with and without disclaimer. For each ad, we pick the ad topic chosen by the majority of workers or mark it as disagreement if no two workers chose the same ad topic or if two topics had an equal number of votes. We attributed all ads about both a political figure and a social issue or a political figure and election to the political figure group, and all ads about both an election and a social issue to the election group. For clarity, all ads for which the majority of workers chose a (non-political) IAB topic are marked as “None of the above” in Figure 3.

Figure 3 shows that all groups of ads contain most of the ad topics we consider. We see higher fractions of about a political figure or political party and ads about an election in the **strong political ads** (78%+8% w. and 65%+7% w/o. disc.) and higher fractions of social issues ads in the **political ads** (38% w. and 61% w/o. disc.) and **marginally political ads** (75% w. and 62% w/o. disc.). In the **official political ads** group, there is also a high fraction (48%) of social issue ads. The non-political AdAnalyst ads are shown as control. Indeed less than 2% of these ads are labeled as being about a political figure, election or social issue. *Social issue ads are only considered political by Facebook and TikTok, not by Google and Twitter. However, these results tell us that a large proportion of the ads volunteers and advertisers label as political are about social issues. Hence, it is crucial to consider social issue ads as political as well.*

Figure 3 shows that some ads (2% w. and 2% w/o. disc. of **strong political ads** and 1% w. and 11% w/o. disc. of **political ads**) were

not labeled by workers as being about a social issue, a political figure or election. Since there is no expert ground truth, we cannot say whether labels from volunteers or labels from workers are better. Nevertheless, the (non-political) IAB topics that were mentioned the most by workers were society, health & fitness, education and science. *This raises questions on where to draw the line between ads about civil and social right and ads about society; or ads about health as a social issue and ads about health & fitness as a lifestyle.*

One might decide that **marginally political ads** should not be treated as political because only a minority of volunteers labeled them as political. Figure 3 shows that 3%+2% w. disc. and 2%+3% w/o. disc. of **marginally political ads** do contain ads from a political figure or political party or elections. In addition, 21% w. disc and 20% w/o. disc. ads are about civil and social rights, and 24% w. disc. and 9% w/o. disc. are about environmental politics. The numbers look similar for **non-political ads**. Marginally political ads do contain a significant number of political ads as defined by the Facebook ToS. *These results show that marginally political ads should not be ignored because they might contain ads about social issue and ads where only a few people have the right background knowledge to detect them as political.*

Takeaway: Our results show that a large fraction of ads labeled as political are about social issues and do not mention a political actor or elections. Hence, it is crucial to consider ads about social issues as political. Our results also show that a wide range of ads are getting labeled as ads about social issues. Hence, since many legislative projects are considering to severely restrict micro-targeting [23] or ban such ads altogether; we need to decide whether we want ads (with no apparent link to elections and legislation) coming from charities or communities to be subject to the same restrictions as ads that advocate polarizing issues. Such restrictions could hurt a wide range of humanitarian civil organizations.

5 LEARNING FROM DISAGREEMENT

The previous section showed that a very diverse set of ads get labeled as political. This section analyzes the ads that lead to disagreement among volunteers and between volunteers and advertisers. This analysis is relevant for refining political ads' definition and improving the processes and instructions for labeling ads.

5.1 Volunteers vs. advertisers

To understand why advertisers and volunteers disagree on ads being political, we examine separately ads that seem to be underreported by advertisers and ads that seem to be underreported by volunteers.

Ads underreported by advertisers. These are the **strong political ads** and **political ads** without disclaimer. Table 2 shows that 4% of the **strong political ads** and 7% of the **political ads** are not labeled as political by advertisers. There are several possible (non-exhaustive) explanations: (1) advertisers do not comply with the ToS (e.g., they willingly do not label their ads as political to avoid scrutiny), i.e., volunteers are right; (2) advertisers underreport certain categories of political ads, i.e., advertisers and volunteers have different interpretations of which ads are political; and (3) volunteers misinterpret the ads' message, i.e., advertisers are right.

Figure 2 presents the breakdown of advertiser categories and Figure 3 the breakdown of ad types corresponding to **strong political**

ads, and **political ads** without disclaimer. A significant fraction of advertisers are political figures (48% in **strong political ads** and 25% in **political ads**), and a significant proportion of ads refer to a political figure or political party and elections (65%+7% for **strong political ads** and 17%+7% for **political ads**). *Hence, more than half of strong political ads and political ads without disclaimers are not compliant with Facebook's ToS. These results confirm previous findings that advertisers sometime do not label their ads as political and the need for better enforcement mechanisms [33].*

A large fraction of ads without a disclaimer (23% of **strong political ads** and 61% of **political ads**) are about social issues. Recall that we excluded from this category ads labeled as social issues but mentioning a political figure or elections. Tables 7 and 8 (in Appendix) show some examples of ads about civil and social rights and environmental politics in **strong political ads** and **political ads** without disclaimer. These ads are on topics such as climate change, healthcare, and equal pay, which are very politicized issues in the US, and give valid reasons to volunteers to label them as political.

To understand whether ads about some social issues are less disclosed by advertisers than others, for each ad topic, we compute the fraction of ads that do not have a disclaimer in the **strong political ads** and **political ads** groups. Ads about economy (0.15), civil and social rights (0.28), and security and foreign policy (0.27) have the lowest fraction of ads with a disclaimer. In contrast, ads about political figures (0.6), election (0.57), and environmental politics (0.49) have the highest fractions of ads with a disclaimer. *This shows that advertisers are underreporting ads about social issues, especially if they are about economy or civil and social rights.*

For 2% of **strong political ads**, and 11% of **political ads** w/o. disc. workers did not label them as being about a political figure, election, or social issue; which means that no one besides volunteers labeled them as political. Table 4 shows a few examples of such ads. These ads seem to address some issues but are not clearly related to the social issues provided to workers. This raises an interesting dilemma: if someone labels an ad as political (without being forced or by mistake), can they be wrong?

Ads underreported by volunteers. These are **non-political ads** and **marginally political ads** with disclaimer. There are 1.3k **non-political ads**, and 5.6k **marginally political ads** (74%) labeled as political by advertisers. There are various reasons why advertisers would label their ads as political while all/most volunteers labeled them as non-political: (1) advertisers might be *forced* to label ads as political (even if they are not) because of false positives in the enforcement mechanisms implemented by the ad platform; (2) advertisers might think that disclaimers would bring more attention to their page; (3) advertisers understand better why their ads should be political, and volunteers underreport such ads; etc.

Figure 3 shows that a significant fraction (14%) of **non-political ads** are labeled as not being related to a political figure, election or social issue by workers; meaning that no one besides advertisers are considering these ads as political. Table 4 shows a few examples of such ads. Indeed, the majority of these ads do not seem to be political. *Since substantial restrictions are envisioned for political ads, it is essential to know what enforcement mechanisms are put in place by ad platforms to understand what is the impact of false positives in their algorithms.* Non-political ads mislabeled as political is also

Table 4: Ads underreported by advertisers or volunteers that are not about political figures, elections or social issues (according to workers).

Advertiser	Text	Workers' label	fr	disc.
Ads underreported by advertisers: strong political ads and political ads w/o. disclosure				
Citizens Against Lawsuit Abuse	Frivolous lawsuits are clogging our courts. Want to help tell trial lawyers enough is enough? Join Citizens Against Lawsuit Abuse (CALA) today!	non-pol	1	w/o.
The Young Turks	Support independent investigative journalism while looking fly AF!	non-pol	1	w/o.
Mikey Weinstein, MRFF	MRFF Op-ed: Anti-Theist Airman Memorialized by Air Force Unit with Image of Jesus	non-pol	0.75	w/o.
Voices for Refugees	Torrential monsoon weather has hit Rohingya refugee camps in Bangladesh, destroying 273 family shelters already. Every donation helps us to reach those most vulnerable with emergency support and help to rebuild, reinforce and secure their shelters.	non-pol	0.6	w/o.
Ads underreported by volunteers: non-political ads and marginally political ads w. disclosure				
Grist.org	Lettuce introduce you to the future of your arugula.	non-pol	0	w.
Heifer International	Truth bee told, not everyone can get these 7 questions right. Test your bee smarts and unlock a 50 cent donation for Heifer	non-pol	0	w.
EveryLibrary	Join hundreds of thousands of Americans who love libraries!	non-pol	0	w.
Mercy For Animals	Animals at factory farms suffer in unimaginable ways. They are cruelly confined, abused, neglected, and mutilated. Please support our work to stop this torment.	non-pol	0	w.
The Christian Science Monitor	He need to address corruption in the Arab world is urgent. But if new initiatives are simply politically expedient – as many citizens suspect – they risk only fueling distrust and suspicion.	non-pol	0	w.

problematic when building automated detection methods that use political ads labeled by advertisers to train models. Hence, it is important to look for poisoning attacks when building such models.

Figure 3 shows that the majority of **non-political ads** and **marginally political ads** without disclaimer are related to civil and social rights (21% and 20%), health (20% and 16%) and environmental politics (22% and 9%), while only a few refer to political actors (2% and 1%) or elections (2% and 2%). Figure 2 shows that these ads come mostly from NGOs (53% and 58%), news media (4% and 3%), businesses (4% and 3%), and charities (16% and 5%), while only a few (1% and 1%) come from political actors. Hence, it seems that volunteers underreport many ads about a social issue, especially about civil and social rights and health, and ads from advertisers such NGOs, and charities. Table 7 and 8 (in Appendix) present examples of ads about civil and social rights and environmental politics in **non-political ads** and **marginally political ads**. We see that most of these ads are related to social issues, but volunteers might not consider them as political because there is no apparent association with elections or legislation.

Takeaway: Two main factors contribute to disagreement between advertisers and volunteers: (1) advertisers mislabel ads as political or non-political (maybe to avoid scrutiny; maybe because they are forced to label their ads as political by enforcement mechanisms put in place by ad platforms); and (2) both advertisers and volunteers underreport ads about social issues. Part of the problem may be that the definition of ads about social issues is too broad which leads to different interpretations among people. This raises the question of whether *all* ads related to social issues should be considered political, and if not, how should we filter social issue ads that are not political. For example, one possibility would be to consider as political only ads about social issues that could directly or indirectly impact elections or legislation or that address polarizing issues.

5.2 Volunteers vs. volunteers

To investigate which ads lead to disagreement among volunteers, we check if there is more disagreement on ads coming from specific advertisers and ads about particular political or social issues.

To see if ads from certain categories of advertisers lead to more disagreement, for each advertiser category, we group all corresponding ads (from **strong political ads**, **political ads**, and **marginally political ads**). Figure 4 shows the ECDF of fr for each group. Ads with a fr close to 0.5 have the highest level of disagreement (half of the volunteers label them as political and half as non-political). We

split the analysis on ads with disclaimer and ads without a disclaimer since the disclaimer might have impacted how volunteers voted. We see in Figure 4 that the distributions shift to the right (more political votes) when ads have a “Paid for by” disclaimer. However, we cannot attribute this shift solely to the presence of disclaimers because ads with disclaimers might also have messages that are “more political”.

The plot shows that at least 10% of ads in each advertiser category has $fr = 1$, which means that at least *some* volunteers are not bothered by the fact that the ad is coming from non-traditional political actors. Figure 4 shows that ads coming from political actors and public figures achieve the highest agreement, 85% have $fr = 1$. Besides, ads from communities and advocacy groups tend to be seen as more political, while ads from charities as less political. Ads from other advertisers such as NGOs, causes, news media, education, and businesses are somewhere in between, leading to the highest level of disagreement. To get definite proof if the advertiser category influences the decision (and not the message of the ad), we would need a conjoint analysis that tests the same ad message with different advertisers but our data does not permit such analysis. In any case, *platforms and policymakers should clarify how much consideration should be given to the advertiser when labeling ads as political.*

To see if ads from certain ad topics lead to more disagreement, for each ad type, we group all corresponding ads (from the 1800 ads labeled by Prolific workers in **strong political ads**, **political ads**, and **marginally political ads**). Figure 5 shows the ECDF of fr for each group (we only show groups for which we have more than 20 ads labeled). We can see that the highest agreement is among ads that mention political figures and elections, while, as expected, the highest disagreement is on various social issue ads. We performed a pairwise Kolmogorov-Smirnov statistical tests between the distributions. Ads about elections and political figures are statistically different than the rest; but most of the social issue ads are not statistically different between them.

To see why for a particular social issues, some ads have higher fr than others, Tables 7-8 (in Appendix) show examples of civil and social rights ads and environmental politics ads for different ad groups. We see that the ads address a wide range of topics (e.g., abortion, wildlife, violence, hunger), they call for various actions (e.g., sign petitions, surveys, donations, call an elected representative) and try to provoke various sentiments (e.g., pride, anger, fear). Ads that address climate change and pollution are seen as more political, while ads about wildlife protection are seen as less political. Besides, ads that refer to problems in the U.S. (ad from NRDC) are seen as more

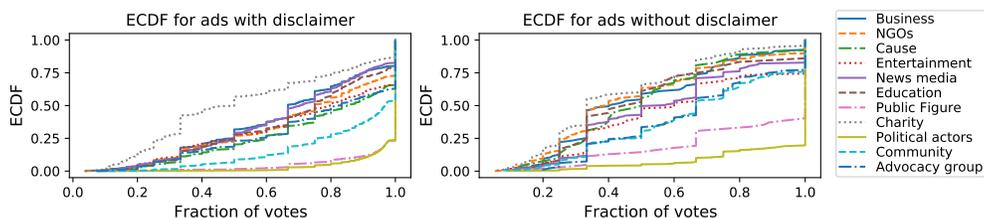


Figure 4: ECDF of the fraction of political votes for ads from different advertiser's categories in strong political ads, political ads, and marginally political ads.

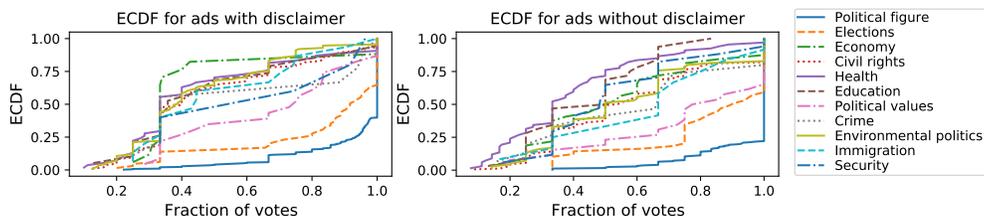


Figure 5: ECDF of the fraction of political votes for ads with different ad topic in strong political ads, political ads, and marginally political ads.

political than ads that refer to problems in other countries (ad from Care2). While these are only anecdotal examples, they emphasize the complexity of deciding which ads are political.

Limitation: There are other reasons for disagreement that we could not analyze with this dataset. For example, the background knowledge of volunteers might impact how they vote (the political nuance of an ad is only recognized by some) or the political ideology of volunteers impacts how they vote. These questions are essential for recruiting moderators, and we leave them for future work.

Takeaway: Ads from NGOs, causes, news media, education, and businesses and ads on social issues lead to the highest disagreement among volunteers. To distinguish better political from non-political ads, we would need policy recommendations that clarify the perimeter of social issue ads. This raises a multitude of questions such as: Should we treat ads about more politicized issues differently than ads about less politicized issues? Should we treat social issues depending on the country? Should we treat ads that call for precise actions differently than ads that just inform citizens? Should we define social issues at a smaller granularity (in both topics and locality) than currently? How should the system adapt to emerging social issues? How much weight should be given to the advertiser's identity (as opposed to just the ad content)?

6 CLASSIFICATION AND DISAGREEMENT

Traditional supervised classification algorithms create models from positive and negative examples that we feed in the training phase. The previous sections showed significant discrepancies between ads labeled as political by advertisers and ads labeled as political by volunteers. Hence, this raises the question of whether classifiers trained on one or the other would result in significantly different models. Intuitively, if the training examples are biased, the models will be different, while if the training examples are representative of political ads in general, the resulting models will be similar. This section

investigates how discrepancies in positive labels from advertisers and volunteers impact the resulting classification models.

For the evaluation we split the ProPublica dataset in two equal-size slices of 28k ads: S_1 and S_2 . We use S_1 as the training and validation dataset and S_2 as the holdout/test dataset. We build four models using four different sets of positive labels but the same negative labels. As negative examples, we took 7.5k ads in English from AdAnalyst without the "Paid for by" disclaimer (see Section 2).

The M_{op} model: the positive labels are a random sample of 8000 official political ads from S_1 . Hence, M_{op} is trained with positive examples from *advertisers*.

The M_{sp} model: the positive labels are a random sample of 8000 strong political ads from S_1 . Hence, M_{sp} is trained with only positive examples where all volunteers agree they are political, $fr = 1$.

The M_{mp} model: the positive labels are a random sample of 4000 political ads and 4000 strong political ads from S_1 . Hence, M_{mp} is trained with positive examples where the majority of volunteers consider the ads as political, $fr > 0.5$.

The M_{1p} model: the positive labels are a random sample of 2600 strong political ads, 2600 of political ads and 2600 marginally political ads from S_1 . Hence, M_{1p} takes as positive examples all ads where there exist at least one user that labeled it as political, $fr > 0$.

To build the different models, we used Naive Bayes. While Naive Bayes is neither new nor sophisticated, it was shown by Silva et al. [33] that it achieves very high accuracy for detecting political ads and outperforms other methods. The classifiers only take as input the ad's text, and as pre-processing, we deleted all HTML tags, stop words, and punctuation. We used Count Vectorizer for text embedding [32].

We performed 10-fold cross-validation for each classifier over its specific training-validation dataset that contains 8000 positive and 7.5k negative examples. Table 5 presents the average accuracy and true positive rate for a 1% false positive rate for the four classifiers. For systems where the fraction of positive examples (political ads) is much smaller than the fraction of negative examples (non-political ads), it is essential to limit the rate of false positives (non-political

Table 5: The average accuracy and true positive rate for a 1% false positive rate for the four models. Each classifier is evaluated over its specific training-validation dataset.

Classifier	Accuracy	TPR for 1% FPR
M_{op} model	96% (+/-1%)	91% (+/-4%)
M_{sp} model	97% (+/-1%)	94% (+/-2%)
M_{mp} model	95% (+/-2%)	91% (+/-4%)
M_{1p} model	95% (+/- 3%)	86% (+/-4%)

ads labeled as political); hence, we are interested in true positive rates for a 1% false positive rate. The table shows that all classifiers achieve high accuracy of over 95%, but only M_{op} , M_{sp} , and M_{mp} achieve true positive rates of more than 90%. The lower true positive rate of M_{1p} (86%) is expected as it has a more challenging task because it is trained and tested with more debatable political ads.

The main challenge in evaluating the classifiers is that we do not have a gold standard collection of political and non political ads. Table 5 only tells us *how good the models are at identifying the same kind of political ads with the ones they were trained on, but not how good they are at identifying political ads in general*. Hence, we look next at how these models perform on detecting other kinds of political ads than those they were trained on.

We use the four models to make predictions for all ads in S_2 . To predict that an ad is political, we took the threshold corresponding to a 1% FRP for each classifier. Table 6 shows how well the four models are at identifying **official political ads**, **strong political ads**, **political ads**, **marginally political ads**, and **non-political ads** in S_2 . As negative examples, we used 1000 ads from AdAnalyst in English that do not have a disclaimer and were not used for training.

Table 6 shows that M_{sp} has the lowest number of false positives, while M_{op} has the largest number. For detecting **strong political ads**, all models detect more than 95% of ads. For detecting **political ads** the M_{op} and M_{mp} perform the best (detecting over 94% of ads). For **marginally political ads**, M_{1p} and M_{mp} , perform equally well (over 85% detection), while M_{op} has a 82% detection. For **non-political ads**, M_{op} labels 85.1% as political while M_{mp} labels 86.8% as political.

The detection rates of M_{op} and M_{mp} are similar across different datasets, with M_{mp} performing better especially on **marginally political ads** and **non-political ads**. Hypothetically if the resulting classifiers would label as political the precise same ads, the input data is representative of political ads, and who is labeling the training data (be it advertisers or volunteers) does not matter. To understand whether M_{op} and M_{mp} label the same ads as political, we computed the fraction of ads labeled by both models as political over all ads for different ad groups in S_2 . The data shows that the two models have an overlap of 97% in **strong political ads**, 94% in **political ads**, 83% in **marginally political ads**, and 84% in **non-political ads**. *These results show that the overlap is relatively high, but discrepancies in the input data do transfer to discrepancies in the output data. Hence, we need to consider how biases in labeling are impacting classification results and whether this may lead to unfairness against certain advertisers.*

7 RELATED WORK

Definitions of what is political: Several works studied what people think is political [12, 19]. Fitzgerald [12] set up several experiments

Table 6: The fraction of ads labeled as political by the four models in different groups of ads in S_2 .

	# ads	M_{op}	M_{sp}	M_{mp}	M_{1p}
AdAnalyst (non-political)	1000	1.3%	0.6%	1%	1%
official political ads	25k	95%	88.5%	95%	92.6%
strong political ads	13k	97.7%	96%	98.2%	95.3%
political ads	8.7k	94.7%	87.9%	95.2%	90.8%
marginally political ads	3.7k	87%	63.3%	87%	85.6%
non-political ads	0.6K	85.1%	60.4%	86.8%	87%

to identify what topics people consider political and showed that, similar to our study, it is complicated to defined what topics are political. The experiment included 33 different topics such as education, poverty, national parks and space exploration. The participants achieved a 95% agreement that diet pills is not a political topic. However, the mean percentage of respondents who view a topic as political is 42%. Hansford et al. [19] designed an implicit association test featuring the Supreme Court and Congress. The results showed that people perceived the Supreme Court as less implicitly political than Congress. On a more theoretical side, Sartori [31] considered the question of the autonomy of politics. The author concluded that the current situation of politics is reflected in three different ways: outright extinction, autonomy or weakening, which leads to different ways of perceiving, identifying, and defining politics. Warren [38] proposed that the concept of politics should help to clarify normative interests in politics, that the definition of politics should embrace everyday understandings of politics, and serve explanation. He suggested that politics can be define by two attributes: power and conflict.

Analysis of political conversations online: Hersh [20] described how political campaigns changed across time and concluded that social media has a large impact on peoples' decisions. Maruyama et al. [24] showed that Twitter activity could affect people's vote decision. The authors experimented during the 2012 U.S. Senate election in Hawaii. The results showed that people who actively participated in Twitter discussions changed their opinion about their candidate more often than people who did not use Twitter. Kou et al. [22] analyzed the development of public discourses on social media. The authors showed that during the "Umbrella movement", conversations on Facebook (mostly used by Hong Kong citizens) emphasized with protesters, while conversations on Weibo (mostly used by mainland China) emphasized with the government.

Political advertising: Silva et al. [33] created a tool for collecting ads from Facebook and implemented several supervised classifiers to detect political ads during the 2018 election in Brazil. They detected a significant number of political ads that did not have the official "Paid for by" disclaimer. Ribeiro et al. [30] analysed ads send by the Russian interference in the 2016 US elections and found that ads were send to people less likely to report them. Edelson et al. [9] presented a clustering-based method to discover advertisers engaging in a potentially undeclared coordinated activity and proposed recommendations for improving the security of political advertising transparency. Furthermore, Bolden et al. [5] summarized problems with the Facebook Ad Library, such as the lack of clear policies and

data systematicity. Finally, Ali et al. [1] proved that Facebook’s ad delivery algorithms effectively differentiate the price of reaching a user based on their inferred political alignment with the advertised content, inhibiting political campaigns’ ability to reach voters with diverse political views. Our paper focuses on a more foundational question of what should be considered political advertising.

Several other studies have pinpointed problems with the Facebook ad ecosystem without focusing on political advertising such as discrimination [34], lack of transparency [2, 3], and security and privacy problems [21, 37].

8 CONCLUSION

Many agree that online advertising especially political advertising needs to urgently be regulated, but one missing key is how to reliably detect political ads. This paper attempts to dissect some of the complexity of labeling political ads. To our knowledge, this is the first study to show how ordinary people label ads as political, why they disagree and what are the implications for policymaking and enforcement algorithms.

Our paper shows that volunteers seem to underreport ads from NGOs, and charities (that are considered political by advertisers) and advertisers seem to underreport social issue ads (that are considered as political by volunteers). While disagreement can be alleviated through better guidelines to a certain degree, many ads addressing societal and humanitarian issues are intrinsically hard to label. We believe that the community needs a gold standard collection for political ads and to better define the perimeter of social issue ads. We hope our analysis can help policymakers and ad platforms to refine the definitions of political ads and their regulation.

9 ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their helpful comments and Jeremy Merrill for his assistance with the data. This research was supported in part by the French National Research Agency (ANR) through the ANR-17-CE23-0014 and the MIAI@Grenoble Alpes ANR-19-P3IA-0003 grants.

REFERENCES

- [1] Muhammad Ali, Piotr Sapiezynski, Aleksandra Korolova, Alan Mislove, and Aaron Rieke. 2021. Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging. In *WSDM*.
- [2] Athanasios Andreou, Márcio Silva, Fabrício Benevenuto, Oana Goga, Patrick Loiseau, and Alan Mislove. 2019. Measuring the facebook advertising ecosystem. In *NDSS*.
- [3] Athanasios Andreou, Giridhari Venkatadri, Oana Goga, Krishna Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Investigating ad transparency mechanisms in social media: A case study of Facebook’s explanations. In *NDSS*.
- [4] Priyanjana Bengani. 2020. As election looms, a network of mysterious “pink slime” local news outlets nearly triples in size. *Columbia Journalism Review* (2020). <http://bit.ly/3der7V1>
- [5] Sarah E. Bolden, Brian McKernan, and Jennifer Stromer-Galley. 2020. *Facebook Political Advertising Transparency Report*. Retrieved 14 February 2021 from <https://bit.ly/37er8F7>
- [6] Blake Chandlee. [n.d.]. *Understanding our policies around paid ads*. Retrieved 14 February 2021 from <https://newsroom.tiktok.com/en-us/understanding-our-policies-around-paid-ads>
- [7] European commission. [n.d.]. *Public consultation: Commission seeks citizens’ views in preparation of new European Democracy Action Plan*. Retrieved 14 February 2021 from https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1352
- [8] European commission. 2021. *The Digital Services Act package*. Retrieved 14 February 2021 from <https://ec.europa.eu/digital-single-market/en/digital-services-act-package>
- [9] Laura Edelson, Tobias Lauinger, and Damon McCoy. 2020. A Security Analysis of the Facebook Ad Library. In *IEEE S&P*.
- [10] Facebook. [n.d.]. *Facebook Ad Library*. Retrieved 14 February 2021 from <https://www.facebook.com/ads/library/>
- [11] Facebook. [n.d.]. *Facebook advertisers’ categories*. Retrieved 14 February 2021 from <https://www.facebook.com/pages/category/>
- [12] J. Fitzgerald. 2013. What does “Political” Mean to You? *Political Behavior* (2013).
- [13] Facebook for Business. [n.d.]. *About social issues*. Retrieved October 16, 2020 from <https://www.facebook.com/business/help/214754279118974>
- [14] Facebook for Business. [n.d.]. *Ads about social issues, elections or politics*. Retrieved 14 February 2021 from <https://www.facebook.com/business/help/167836590566506>
- [15] Facebook for Business. [n.d.]. *Become authorised to run ads about social issues, elections or politics*. Retrieved 14 February 2021 from <http://bit.ly/379NffN>
- [16] Facebook for Business. [n.d.]. *Best practices to choose a category for your Facebook Page*. Retrieved 14 February 2021 from <https://www.facebook.com/business/help/376650512904346>
- [17] European Partnership for Democracy. 2020. *Universal Advertising Transparency by Default*. Retrieved 14 February 2021 from <https://epd.eu/wp-content/uploads/2020/09/joint-call-for-universal-ads-transparency.pdf>
- [18] Google. [n.d.]. *Google: Political content*. Retrieved 14 February 2021 from <https://support.google.com/adspolicy/answer/6014595?hl=en>
- [19] Thomas Hansford, Chanita Intawan, and Stephen Nicholson. 2018. Snap Judgment: Implicit Perceptions of a (Political) Court. *Political Behavior* (2018).
- [20] Eitan D Hersh. 2015. *Hacking the electorate: How campaigns perceive voters*. Cambridge University Press.
- [21] Aleksandra Korolova. 2010. Privacy violations using microtargeted ads: A case study. In *IEEE ICDMW*.
- [22] Yubo Kou, Yong Ming Kow, Xinning Gui, and Waikuen Cheng. 2017. One social movement, two social media sites: A comparative study of public discourses. In *ACM CSCW*.
- [23] Natasha Lomas. 2020. *Europe to put forward rules for political ads transparency and beef up its disinformation code next year*. Retrieved 14 February 2021 from <http://tcm.ch/3b4MANV>
- [24] Misa T Maruyama, Scott P Robertson, Sara K Douglas, Bryan C Semaan, and Heather A Faucett. 2014. Hybrid media consumption: How tweeting during a televised political debate influences the vote decision. In *ACM CSCW*.
- [25] Damon McCoy. [n.d.]. *Online Political Ads Transparency Project*. Retrieved 14 February 2021 from <https://engineering.nyu.edu/research/online-political-ads-transparency>
- [26] Prolific. [n.d.]. *Revolutionizing people research*. Retrieved 14 February 2021 from <https://www.prolific.co>
- [27] ProPublica. [n.d.]. *Facebook Political Ad Collector*. Retrieved 14 February 2021 from <https://projects.propublica.org/facebook-ads/?lang=en-US>
- [28] ProPublica. [n.d.]. *Political Advertisements from Facebook*. Retrieved 14 February 2021 from <https://www.propublica.org/datastore/dataset/political-advertisements-from-facebook>
- [29] Qualitrix. [n.d.]. *QualtricsXM*. Retrieved 14 February 2021 from <https://www.qualtrics.com>
- [30] Filipe N. Ribeiro, Koustuv Saha, Mahmoudreza Babaei, Lucas Henrique, Johnnatan Messias, Fabrício Benevenuto, Oana Goga, Krishna P. Gummadi, and Elissa M. Redmiles. 2019. On Microtargeting Socially Diverse Ads: A Case Study of Russia-Linked Ad Campaigns on Facebook. In *FAT**.
- [31] Giovanni Sartori. 1973. What is “Politics”. *Political Theory* (1973).
- [32] Scikit-learn. 2007. *Scikit-learn: Machine learning in Python*. Retrieved 14 February 2021 from <https://scikit-learn.org/stable/>
- [33] Márcio Silva, Lucas Santos de Oliveira, Athanasios Andreou, Pedro Olmo Vaz de Melo, Oana Goga, and Fabrício Benevenuto. 2020. Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook. In *The Web Conference*.
- [34] Till Speicher, Muhammad Ali, Giridhari Venkatadri, Filipe Ribeiro, George Arvanitakis, Fabrício Benevenuto, Krishna Gummadi, Patrick Loiseau, and Alan Mislove. 2018. Potential for discrimination in online targeted advertising. In *FAT**.
- [35] TextBlob. [n.d.]. *TextBlob: Simplified Text Processing*. Retrieved 14 February 2021 from <https://textblob.readthedocs.io/en/dev/>
- [36] Twitter. [n.d.]. *Twitter political ads*. Retrieved 14 February 2021 from <https://business.twitter.com/en/help/ads-policies/ads-content-policies/political-content.html>
- [37] G. Venkatadri, A. Andreou, Y. Liu, A. Mislove, K. P. Gummadi, P. Loiseau, and O. Goga. 2018. Privacy Risks with Facebook’s PII-Based Targeting: Auditing a Data Broker’s Advertising Interface. In *IEEE S&P*.
- [38] Mark E. Warren. 1999. What is Political? *Journal of Theoretical Politics* (1999).

A APPENDIX

Table 7: Civil and social rights ads in different ad groups.

Advertiser	Text	fr	disc.
Strong Political			
AFSCME 3299	Stand up for immigrant families. Tell UC to cancel its contracts with ICE collaborators now!	1	w.
SEIU	Too many of us are still paid less for the same work. That's why we need a union.	1	w.
Fight Back	We're fighting for better healthcare and equal pay.	1	w/o
ACLU	Just a few weeks ago, the Supreme Court ruled that the First Amendment forbids religious hostility by the government. If only it applied that standard to the president and his Muslim ban.	1	w/o.
Political			
CREDO Mobile	"We can transcend the darkness of this moment by joining the struggles of past and future freedom fighters. That is how, when we reach the end of our lives and look back on these heady moments, we will find peace in the knowledge that we did our best." – Ady Barkan.	0.82	w.
Granite State Progress Education Fund	Stop anti-abortion shame, stigma and hate from New Hampshire politicians. Sign the petition to support abortion access for all Granite Staters!	0.83	w.
Physicians for Human Rights	Doctors and nurses are standing up against human rights abuses across the world. Join our community and learn more about our work	0.75	w/o.
International Rescue Committee	Women and girls in crisis zones face discrimination, violence, and a lack of equal opportunities. Learn how we're working to change that.	0.66	w/o.
Marginally Political			
Boston Rescue Mission	It's tragic to be all alone and hungry. Your gift can bring hearty, nutritious meals to men and women who struggle with homelessness.	0.2	w.
No Kid Hungry	Giving Tuesday is coming, and you can help end childhood hunger in America. Our partner, Citi, will match all donations up to \$100,000!	0.3	w.
International Rescue Committee	Yemen is facing the largest humanitarian crisis of our time: millions of children are at risk of starvation and a deadly cholera epidemic remains a serious threat. And it's about to get worse if we don't step up our efforts now.	0.2	w/o.
World Food Programme	Authorities in Yemen are blocking aid. Millions are suffering the consequences. Add your name today to keep aid flowing into Yemen	0.33	w/o.
Non-political			
Save the Children US	There's still time to give during the 48 Hours of Giving! Your gift in support of the Center for Girls will be matched 2x by an anonymous donor – but the match ends at midnight Saturday	0	w.
United States Holocaust Memorial Museum	It's more important than ever that people understand the dangers of unchecked hatred. In this time of growing antisemitism at home and abroad, we all have a responsibility to keep the history of the Holocaust alive. Can we count on you?	0	w.
Covenant House International	TRIPLE your impact on precious young lives. Give now to help ensure that Covenant House keeps its pledge to welcome ALL homeless youth who come through our doors and love them unconditionally	0	w.
Nashville Rescue Mission	Water can be life-saving when summer's heat is at its worst and there's no escape. Helping is easy—and it won't cost you a thing	0	w.

Table 8: Environmental politics ads in different ad groups.

Advertiser	Text	fr	disc.
Strong Political			
National Parks Conservation Association	No organization has won more victories for the national parks over the past century than NPCA - but we can't do it without you. Please donate to protect our nation's magnificent public lands.	1	w.
Conservation Northwest	ACTION ALERT: Yesterday, the U.S. House of Representatives passed a budget bill that would block funding for grizzly bear restoration in the North Cascades. Use the links below to send your elected representatives a quick message to ensure Congress provides the funding bears need!	1	w.
Ocean Conservancy	Offshore oil spills can harm marine life, devastate ocean environments and risk the livelihoods of coastal communities.	1	w/o.
Coloradans for Responsible Energy Development	Colorado's first-in-the-nation oil and gas regulations work to protect our communities.	1	w/o.
Political			
Care2	Botswana is considering lifting the ban on hunting elephants. We must act NOW and convince Botswana to maintain their stance on protecting these endangered elephants from poachers!	0.55	w.
American Bird Conservancy	The Endangered Species Act is under attack. Despite the fact that 99% of species shielded by the Act — including Bald Eagles and California Condors — have avoided extinction, opponents in Congress are threatening to undermine this bedrock environmental law. Add your name to ABC's petition and tell the government to help protect endangered birds now	0.83	w.
NowThis	Women are equally impacted by climate change, and it's critical that we have them equally involved in the solution	0.66	w/o.
NRDC	Plastics never break down. And that's becoming a real problem for those of us that depend on the Gulf of Mexico and Mississippi River	0.8	w/o.
Marginally Political			
Defenders of Wildlife	The support from our donors has helped us win many battles for wildlife, but there is always more to be done. Our love of animals is endless, so we are ready to fight tirelessly for imperiled wildlife that can't speak for themselves. Support Defenders today and help us continue the fight for wildlife!	0.33	w.
National Audubon Society Action Fund	Climate change threatens the birds we love. Sign up and we'll alert you to actions you can take to protect birds and the places we all need.	0.17	w.
Potomac Conservancy	Trees are nature's Brita filters! For just \$33, we'll plant a native tree along the Potomac River to help filter out water pollution. Plant a tree today!	0.36	w/o.
Climate Reality	Last year, 39 million people tuned in to 24 Hours of Reality to learn what climate change is doing to our planet and how we can solve it with the solutions in our hands today. Help us make 2018's show even bigger!	0.33	w/o.
Non-political			
National Park Foundation	Working together, you can help us have a powerful impact on our spectacular national parks. Your support right now will go to work immediately to protect the places that matter most for future generations.	0	w.
The Nature Conservancy	The challenges facing our natural world are growing every day. Please, make a tax-deductible gift to give nature and wildlife a future.	0	w.
Oceana	Sea lions are drowning in mile-long "walls of death" off the California coast. Let them die... or help us save them	0	w.
IFAW	IFAW protects animals and the places they call home. With your help, we can continue to make a difference. Let's get to work.	0	w.