Journal of Quantitative Description: Digital Media 5(2025), 1–37 10.51685/jqd.2025.024

The Strategic Goals of US Election Ads on Social Media: A Descriptive Study of the 2022 Midterms

MARKUS NEUMANN
Duke Kunshan University, China

SEBASTIAN ZIMMECK
Wesleyan University, USA

JIELU YAO
National University of Singapore, Singapore

ERIKA FRANKLIN FOWLER
Wesleyan University, USA

MICHAEL FRANZ
Bowdoin College, USA

BREEZE FLOYD
Wesleyan University, USA

TRAVIS NELSON RIDOUT
Washington State University, USA

Election campaigns increasingly pursue their strategic goals online, using digital advertising not only to persuade voters but also to raise funds, mobilize supporters, and collect data. These varied objectives reflect how campaigns operate within

Neumann: markus.neumann@dukekunshan.edu.cn
Date submitted: 2025-04-04

broader networks of candidates, parties, and outside groups. The extended party network forms the strategic backbone of modern U.S. campaigns, yet its internal division of labor in paid advertising has remained largely unmeasured. By classifying the goals of over 377,000 Facebook and Instagram ads from the 2022 US elections, we provide empirical evidence about the strategic behavior of candidates, parties, and outside groups. To this end, we build on previous work to extend and refine a taxonomy of nine goals for online election ads: acquisition, contact, donate, event, learn, persuade, poll, purchase, and vote. We analyze how advertisers allocate spending across goals and show that some — notably persuasion, donation, and learn — receive far greater spending than others. We also show that different types of sponsors often pursue different goals. For example, candidate advertisers devote a much higher share of their ads to fundraising than non-candidate advertisers, highlighting the costs of running for political office in the US. We also find that ad goals vary depending on the timing of the ad, the type of sponsor, and the characteristics of the targeted audience. Together, these findings offer the first systematic multi-sponsor account of how digital advertising reflects a functional division of labor within the extended party network that structures contemporary U.S. campaigns.

Keywords: election advertising, election transparency, political advertising, social networks, machine learning, classification

Introduction

Modern campaigns operate within what scholars describe as an extended party network, in which candidates, parties, and outside groups coordinate their activities to achieve shared electoral goals. Still, empirical data on these strategic activities remain scarce. As Nyhan and Montgomery (2015, p. 296) lament, "unfortunately, there is surprisingly little systematic data available on campaign strategy, especially for challengers. One reason is that many campaigns, especially noncompetitive races or those distant from major media markets, provide little direct evidence by which their activities can be evaluated. They receive low levels of media coverage and do not advertise on television."

Nearly a decade later, this limitation remains only partly addressed. Scholars have made progress in describing how modern campaigns operate—often emphasizing their integration with parties and allied groups (Desmarais et al., 2015; Martin and Peskowitz, 2018; Oklobdzija, 2024; Foy-Sutherland, 2025)—but direct evidence of what these actors actually do to pursue their goals is still limited. Traditional sources such as campaign-finance data or media coverage primarily capture spending flows rather than the substantive objectives of campaign communication. Recent work using Federal Election Commission data has improved our understanding of how resources are allocated (Martin and Peskowitz, 2018; Sheingate et al., 2022; Oklobdzija, 2024), but these records reveal little about what campaigns and outside groups aim to accomplish with their ads. In this study, we help fill that gap by analyzing more than 377,000 Facebook and Instagram ads from the 2022 US elections. We classify the strategic goals that underlie paid digital advertising–whether to persuade, raise funds, mobilize supporters, or collect data–offering a large-scale descriptive account of how campaigns use online platforms to pursue their electoral objectives. With these descriptive data in place, new research questions open up for investigation: for example, how certain actors in the party network, outside groups in particular, allocate advertising resources across campaigns and over time. We expand on the potential for future research in the conclusion of the paper.

We note here, however, that as campaign communication has shifted online, the purposes of advertising have diversified. Whereas television ads are almost exclusively intended to persuade (Fowler et al., 2021), digital ads are used for a wider range of objectives: raising funds, mobilizing supporters, acquiring contact information, promoting events, or selling merchandise. Research examining these goals remains limited. Existing studies have focused on the 2012 (Ballard et al., 2016) and 2018 (Ridout et al., 2021) elections, when online ad spending was still overshadowed by television. These studies relied on manual coding, limiting their scope to a few hundred (Ballard et al., 2016) or a few thousand (Ridout et al., 2021) ads from a small number of candidates. They also drew on third-party data sources such as Moat Ad Search and Pathmatics, rather than platform-owned ad libraries. Analyses of ad goals were similarly narrow, focusing primarily on overall distributions, comparisons between two presidential candidates (Ballard et al., 2016), or changes over time (Ridout et al., 2021).

With the volume of online election ads in the U.S. now reaching hundreds of thousands per cycle, manual coding is no longer a viable strategy, underscoring the need for scalable, computational approaches. One effort to classify campaign goals (Zhang et al., 2017) uses machine learning to classify campaign goals, but it focuses on tweets and Facebook posts rather than paid ads and does not analyze the distribution of goals. Critically, no existing work that we are aware of has examined the goals of non-candidate advertisers–such as parties and PACs–who are especially important for several reasons: 1) outside groups are often exempt from many of the transparency rules that apply to candidates, which makes their spending hard to track; 2) the extended party network framework requires a full account of the pro-and anti-candidate electioneering efforts, including those beyond candidate campaigns. Although it seems highly likely that advertising from different types of sponsors will vary based on what we observe on television, television ads have a singular goal: persuasion. Thus, despite the large gap in the literature with respect to digital advertising by groups and parties, in addition to ads from political challengers, we have little existing evidence upon which to build expectations of strategy. Our goal is to provide some initial evidence in service of future theory building. Therefore, we ask one overarching descriptive research question: how do the goals of digital advertising vary within and across different sponsors, by party, over time, and across audiences?

To address this question, we provide the most comprehensive descriptive account to date of the goals that all digital political advertisers pursue in online election ads. We examine all ads that feature federal candidates for office in the US between Labor Day and Election Day 2022. We build on existing work to extend a taxonomy of nine common objectives—including acquisition, contact, donate, event, learn, persuade, poll, purchase, and vote—and apply it to over 377,000 Facebook and Instagram ads from the 2022 US midterm elections, which we acquired directly from Meta's Ad Library (Meta, 2024c). Our dataset includes ads from 649 federal candidate sponsors, as well as down-ballot candidates, parties, and outside groups, representing a total of $275 million in spending. By comparison, Ridout et al. (2021) analyzed 3,943 ads by 24 candidates, accounting for $6.5 million in ad spending. To assign goals to individual ads, we train a set of binary classifiers on hand-coded data, enabling us to identify at least one goal in over 90% of ads. These models, along with the goal taxonomy, represent a resource that can support future work on political advertising at scale.

We conduct a comprehensive analysis of how federal digital ad goals vary across a large number of dimensions, including time, candidate type, political affiliation, incumbency status, audience demographics, and platform. Our findings show that candidate advertisers devote a larger share of their ads to fundraising, while non-candidate groups more frequently focus on persuasion and issue education. Democratic advertisers are more likely to focus on fundraising, devoting a larger share of their ads to donation appeals, while Republican advertisers place greater emphasis on persuasion. Challengers—who, as Nyhan and Montgomery (2015) note, were once difficult to study because they rarely advertised on television—can now be observed much more directly through online ad data. We find that they rely more heavily on persuasion-oriented ads compared to incumbents, who instead concentrate their efforts on raising funds. In terms of audience impressions (which depend on both sponsor targeting and Meta's algorithms), older users are disproportionately shown donation ads, whereas younger users are more often shown get-out-the-vote messages. Finally, although Instagram accounts for a much smaller share of overall ad volume than Facebook, it carries a higher proportion of vote-related ads, suggesting that campaigns may strategically use the platform to mobilize younger or harder-to-reach voters. This work contributes a uniquely large-scale and nuanced account of how political actors use online advertising to pursue diverse strategic goals. In the following section, we situate our study within the broader literature on data-driven political campaigning, where social media and other digital tools have increasingly shaped campaign strategies and voter outreach, and connect it to work on the extended party network that emphasizes coordination among candidates, parties, and allied groups.

Related Work

Recent scholarship no longer conceptualizes campaigns as isolated, candidate-centered efforts, but as components of extended party networks linking candidates, party committees, outside groups, consultants, and allied interest organizations. Foundational work by Koger et al. (2009) and Skinner et al. (2012) shows that these networks coordinate electoral activity through overlapping personnel, organizational ties, and flows of money that circumvent formal party boundaries. Desmarais et al. (2015) demonstrate that integration into these networks benefits candidates–particularly challengers–by connecting them to streams of expertise and funding. Nyhan and Montgomery (2015) trace how consultant networks diffuse tactics and messaging across campaigns, while Martin and Peskowitz (2018) reveal the agency problems and inefficiencies that arise in consultant-driven campaigning. More recent research extends the EPN concept beyond registered committees: Oklobdzija (2024) identifies "dark party" formations of non-profits and vendors that coordinate political work outside disclosure rules, and Foy-Sutherland (2025) details how parties and interest groups collaborate in cooperative, quasi-coalitional relationships. Together, this literature portrays campaigns as interconnected operations that pool resources, professional talent, and strategic expertise toward collective partisan goals.

Chief among those goals is winning office for candidates running under the party label. Many campaign activities are aligned to increase the chances of a candidate's claiming victory. One key campaign activity is persuasion of voters, convincing them to vote for you as opposed to another candidate or party. Persuasion can occur in many different ways. One way to persuade is through informing voters, as they may not know that their positions or values align with yours. This may be especially important in down-ballot races where preexisting knowledge about candidates is sparse. Relatedly, campaigns may want to draw contrasts with their opponents, giving voters a reason to vote for one candidate over another. Another means of persuasion is through agenda setting, that is, telling voters what to think about. Candidates and parties typically want the race to be about those issues that they own (Petrocik, 1996), that is, that they have a reputation for being able to handle better than the opposition. Successful agenda setting can result in priming (Druckman, 2004), whereby voters weigh certain issues more heavily in their minds when evaluating competing candidates or parties. Another key task of a campaign is to establish a candidate's name recognition. Knowing who a candidate is is often a prerequisite to voting for that candidate (Bartels, 1988).

In addition to persuading voters, campaigns and extended party networks must ensure that their supporters turn out to cast ballots. Indeed, there is considerable debate on whether presidential campaigns in the U.S. spend more time focusing on persuasion or more time trying to mobilize voters (Panagopoulos, 2021). Mobilization can take many forms, including knocking on doors, hosting events and rallies, making phone calls and informing people of what they need to know in order to participate, such as the process for registration and where to vote on election day.

Of course, many campaigns activities nowadays are data-driven; they rely on the analysis of several pieces of data in order to inform campaign decision-making (Dommett et al., 2024). Thus, data collection is another important task of a modern campaign. This might involve gathering data while knocking on doors, encouraging people to take "surveys" in which they provide contact information and other data, gathering data from election administrators (such as the voter roll), harvesting digital trace data when people interact with a campaign website or social media page, and purchasing data from commercial providers, to name just a few.

Finally, most of these campaign activities depend on the ability of campaigns to pay for them. Thus, fundraising a key element of the modern campaign. Much fundraising nowadays occurs online, whether at the campaign website—whose splash screen inevitably asks the viewer to "chip in"–through email solicitations or through online ads seeking donations.

In short, today's campaigns and their networks must execute a variety of tasks in order to be successful, and thus we should expect to see many different goals pursued in campaign advertising.

Prior research on the goals of election advertising provides an important foundation for our analysis. Ballard, Hillygus, and Konitzer (2016) hand-coded 840 display ads from the 2012 Obama and Romney campaigns, classifying them into four mutually exclusive categories: persuade (37%), donate (25%), recruit (20%), and vote (18%). While informative, the study was limited to a single election and two candidates. Ridout, Fowler, and Franz (2021) also relied on manual coding, analyzing nearly 4,000 Facebook ads from 24 Senate candidates in 2018. Although their classification scheme included a broader set of 14 goal categories, the scale remained constrained by the labor-intensive coding process. By contrast, Zhang et al. (2017) employed a supervised machine learning approach to categorize a large set of Facebook and Twitter posts—not paid ads—from the 2014 and 2016 election cycles. They classified these posts into six categories: calls-to-action, persuasive, informative, endorsements, ceremonial, and conversational. Their research is primarily a machine learning exercise focused on classifier performance and does not include any descriptive analysis of goal prevalence or co-occurrence. Since the dataset consists of unpaid posts, no spending data is reported. The authors report a micro-averaged F1 score of 0.76. In contrast, our work uses a DistilBERT-based (Sanh et al., 2019) classifier to identify ad goals in 377,721 paid Facebook ads from the 2022 U.S. midterms, achieving a micro-averaged F1 score of 0.90. Our dataset spans 649 candidates and includes a wide range of sponsor types. Beyond identifying goals, we describe how their prevalence varies by sponsor type, party, incumbency, audience characteristics, platform, ad runtime, and timing. Table 1 summarizes key differences across these studies.

In addition to Zhang et al.(2017), a few other studies have moved beyond human coding to classify political messaging. Edelson et al.(2019) developed a rule-based classification approach, described in an arXiv preprint, that labels online political ads based on the links they contain. For example, ads linking to payment processing sites were categorized as donation ads. However, the assumption that links reliably signal an ad's goal has clear limitations, as many ads do not contain links and are thus relegated to an "unknown" category. Stromer-Galley et al. (2021) applied machine learning classifiers to Facebook and Twitter posts by candidates in the 2016 and 2020 U.S. presidential elections. Although their BERT-based (Devlin et al., 2019) approach illustrates the promise of modern NLP in this domain, their primary focus was not on ad goals. Rather, they classified posts into six categories—four types of persuasive messaging (advocacy, attack, image, and issue), and two that describe the focus of advocacy and attack messaging (call to action and ceremonial)—without examining paid ads or campaigns beyond the presidential level. Building on these prior approaches, we turn next to the dataset that forms the basis of our analysis: paid Facebook ads from the 2022 U.S. midterms.

Election Ad Data

Our *inference* dataset of 2022 Meta election ad data is based on the Meta Ad Library (Meta, 2024a) (additionally, we also use 2020 data collected in a similar way for training: see Section 4). Specifically, the population of ads that we intend to make inferences about is all federal ads, as in, all ads by sponsors who run digital advertising that features a federal candidate, which would fit the definition of electioneering. Our choice of dataset is the most inclusive way of defining this dataset, leading to the greatest coverage. See [citation omitted for blind review] for more detail. To identify federal election ads (those we define as featuring federal candidates) on Meta pages we performed daily keyword searches within its "Issues, elections or politics category" using the Meta Ad Library API (Meta, 2024c). Then—with Meta's permission—we scraped all sponsored ads from any page that returned a hit per our keyword searches. We used as keywords a list of Senate candidate names, including variants of legal names, from the

Table 1: Comparison of studies on classification and analysis of ad goals

	Ballard et al. 2012	**Zhang et al. 2017**	**Ridout et al. 2021**	**Our paper**
Election year	2012	2014 / 2016	2018	2022
# creatives	840	43408 / 108605	3943	377721
Ad spend	Unknown	No paid ads	$6.5m	$275m
# candidates	2	78 / 26	24	649
Sponsor types	President	Governor / President	Senate	House, Senate, Downballot, Party, Outside groups
Data source	Moat Ad Search	Twitter / FB	Pathmatics	FB Ad Library
Classification method	Manual	SVM	Manual	DistilBERT
Goal categories	4	6	14	9
Train/test set size	NA	6641	NA	6592
Classifier performance	NA	0.76	NA	0.9
Analysis	Goals by candidate	No analysis	Goals over time	Goals by: – Sponsor type – Incumbency status – Party – Time – Ad runtime – Audience age – Platform

Federal Election Commission (FEC) that we further extended.[1] We also included "senate" and "senator" as general keywords to ensure that we did not miss election ads that do not explicitly mention a Senator's name. To identify US House candidates, save for a few exceptions, we used the keywords "congress" and "representative," which we validated to be sufficient.[2] In addition, to retrieve the spending amounts by sponsor pages we ran daily downloads of the Meta Ad

[1] Our list is publicly available in our replication repository: https://github.com/Wesleyan-Media-Project/ad_goal_paper

[2] One notable exception was Alexandria Ocasio-Cortez, who is known by various aliases. We used "Ocasio-Cortez house," "AOC Ocasio-Cortez new york," and "AOC Ocasio-Cortez NY" to identify her being mentioned in an ad.

Library Aggregate Report (Meta, 2024b). Each page retrieved this way was then hand-coded for who its sponsor was. There are three notable sponsor types: campaign, party, and group sponsors, which are defined as following (see 11 C. F. R. § 100.5 (2025)): Campaign sponsors are pages whose ad spending is paid for by the principal campaign committee of a candidate. Party sponsors are pages whose ad spending is paid for by party committees. Groups are, in effect, all other sponsors, with the exception of some very minor ones (government officials, government agencies, down-ballot candidates and coordinated communications). Pages belonging to federal candidates (down-ballot candidates were excluded) were then linked to their FEC identifier, allowing us to cross-reference them with other variables, such as office, incumbency, and party.

The resulting inference set spans the time period from September 5, 2022, which was Labor Day (generally considered the official kick-off for the general election campaign) to November 8, 2022 (Election Day). Ads that started before September 5, 2022, are included as well. There are 45,640 such ads. Two of them had been running since 2018. The inference set contains 377,721 ads placed by candidates, parties, outside groups, and other sponsors. Because our methodology of selecting election ads is over-inclusive, some ads in the inference set are campaign-related in a broader sense. For example, all ads from a page mentioning a candidate per our keywords will be included even if not all feature candidates. The ads in the inference set are not unique and contain duplicates or near-duplicates, e.g., due to campaigns doing A/B testing with small modifications of an ad. There are 149,413 unique ads with an ad being defined as unique if its text field or combination of text fields is different from all other ads (though these ads often still differ with regard to non-textual attributes, such as spend). 73% of ads in the inference set ran on both Facebook and Instagram, 22% only ran on Facebook, and 5% only ran on Instagram. A small proportion, 0.5% combined, also ran on Facebook Messenger or Facebook Audience Network.

Classifying Ad Goals

We designed our ad goal taxonomy as an evolution of Ridout et al. (2021), which in turn builds on Ballard et al. (2016) and Zhang et al. (2017). Ads are classified into, one or more,[3] of the

[3] There is some overlap between the goals. Figure 8 in the Appendix displays correlations between goals in the train/test set as well as the inference set. In particular, persuade and learn goals co-occur in many ads, which is plausible as attempts to persuade often involves exposure to new information.

Table 2: Ad count and spend, by goal. Learn and persuade are the most frequent goals. They also have the highest spend.

	Ad Count	Prop. Ads	Absolute Spend	Prop. Spend
Acquisition	26,457	0.07	17,842,231	0.06
Contact	4,392	0.01	3,812,293	0.01
Donate	58,305	0.15	23,473,289	0.09
Event	6,579	0.02	1,331,839	0.00
Learn	**116,666**	**0.31**	**55,429,749**	**0.20**
Persuade	**194,152**	**0.51**	**103,858,921**	**0.38**
Poll	9,365	0.02	5,573,414	0.02
Purchase	12,967	0.03	7,316,598	0.03
Vote	80,608	0.21	25,620,920	0.09
No goals	39,830	0.11	30,554,635	0.11

following categories:

Acquisition: Acquisition ads have the goal of gathering contact information, e.g., viewers' email addresses. Acquisition ads also involve indirect techniques, such as asking to sign a birthday card for Melania Trump.[4]

Contact: Contact ads ask viewers to contact a lawmaker.[5]

Donate: Donate ads are ads that solicit donations from viewers; explicitly or via a "Donate" button.[6]

Event: Event ads advertise events, such as campaign rallies. They sometimes advertise free tickets to events.[7]

Learn: Learn ads invite viewers to learn more, often linking to pages with additional information.[8]

[4]This category corresponds to the 'Acquisition' aggregate category in Ridout et al. (2021) and 'Recruitment' in Ballard et al. (2016).

[5]We include this category because the ability to contact one's lawmaker has long been considered a cornerstone of the American representative system (Eulau and Karps, 1977).

[6]Asking for a donation is one of the most important goals in Ballard et al. (2016) and Ridout et al. (2021).

[7]Based on the "Attend event" category in Ridout et al. (2021).

[8]Similar to the 'Encourage learning' category in Ridout et al. (2021) and the informative category

Persuade: Persuade ads promote or attack candidates and often include attempts to convince viewers to vote for a specific candidate.[9]

Poll: Poll ads ask viewers to participate in a poll or survey.[10]

Purchase: Purchase ads involve the selling of politics-related merchandise, e.g., Make America Great Again (MAGA) hats.[11]

Vote: Vote (get-out-the-vote) ads call on viewers to vote, provide instructions on how to register, help find a polling place, or make use of absentee and mail-in ballots.[12]

There are several advantages of our goal taxonomy. First, with nine categories, it contains more categories than most previous taxonomies, giving researchers the ability to study variations in the use of ads by goal in considerable detail. Our taxonomy does, however, contain fewer categories than the taxonomy used by Ridout et al. (2021), eliminating some of the goals from that study that rarely occurred in practice. The elimination of these infrequently occurring categories, no doubt, helps contribute to our strong classifier performance. Indeed, the F1=.90 that we obtained is much higher than the F1=.76 reported by Zhang et al. (2017) in a similar study. Moreover, our taxonomy provides good coverage, allowing us to assign a goal to most ads, a point on which we elaborate in our discussion of Table 2. By contrast, Edelson et al. (2019), whose analysis of ad goals is based on the presence of a link, leaves many ads in the "unknown" category because they lack a link.

The 11% of ads without a goal often covered current news, e.g., the Ukraine war, and, thus, would have been challenging to categorize without one-off categories. The proportion of

in Zhang et al. (2017). Some of these ads on Facebook even have a 'Learn more' button.

[9]The question about the extent to which campaign ads focus on persuasion has been one of the most long-running debates in the field (Lazarsfeld et al., 1968) and corresponds to existing categories in Ballard et al. (2016), Ridout et al. (2021) and Zhang et al. (2017).

[10]The ability to directly survey Facebook users is an important advantage of online advertising compared to TV—this category is concurrent with "Answer survey question" in Ridout et al. (2021).

[11]While 'Offer sticker/merchandise' is one of the smallest categories in Ridout et al. (2021), it is one of the interesting idiosyncrasies of online advertising that distinguishes it from TV.

[12]Similar to persuasion, get-out-the-vote efforts are central to any political campaign (Lazarsfeld et al., 1968), and thus also included in Ballard et al. (2016) and Ridout et al. (2021).

Table 3: Frequencies of negative|positive coded instances in the train/test set and double-coded set as well as Krippendorff's alpha values (K's α). Our datasets contained very few contact and event ads, and intercoder agreement for both was poor. For 2020, only 2 event ads were double-coded. The student coders disagreed on both. Thus, we randomly sampled 10 ads in which the classifier identified an event and confirmed that all were in fact event ads. We got the same result for 9/10 contact ads. These results provide some evidence that the precision of the classifiers is good. Merging datasets mitigates poor performance in one.

	Acquisition	Contact	Donate	Event	Learn	Persuade	Poll	Purchase	Vote
Train/Test Set 2020 ($n = 3,671$)	3,409|262	3,622|49	2,972|699	3,641|30	1,802|1,869	1,544|2,127	3,596|75	3,446|225	3,277|394
Train/Test Set 2022 ($n = 2,921$)	2,828|93	2,876|45	2,316|605	2,890|31	1,179|1,742	728|2,193	2,871|50	2,901|20	2,201|720
Test of Train/Test Set 2020 ($n = 735$)	677|58	724|11	592|143	727|8	357|378	308|427	714|21	691|44	664|71
Test of Train/Test Set 2022 ($n = 584$)	561|23	575|9	471|113	574|10	230|354	147|437	577|7	581|3	424|160
Double-coded Set 2020 ($n = 454$)	418|36	450|4	403|51	452|**2**	191|262	171|283	440|14	417|37	411|43
Double-coded Set 2022 ($n = 631$)	610|21	619|12	495|136	624|7	271|360	157|474	620|11	627|4	478|153
K's α Double-coded Set 2020 ($n = 454$)	**0.82**	0.60	**0.94**	0	**0.84**	**0.67**	**0.82**	**0.93**	**0.65**
K's α Double-coded Set 2022 ($n = 631$)	**0.64**	0.39	**0.95**	**0.77**	**0.77**	0.49	**0.69**	**0.75**	**0.63**

ads do not sum to 1 since an ad can have multiple goals. For ads with multiple goals, the ad's total spend is equally distributed between the goals. The proportional spend of only 0.11 on No Goals indicates that the set of ad goals that we conceptualized captures the majority of ads.

To train a model for the automatic classification of these goals, we rely on data from both 2020 and 2022. This larger dataset improves training and, by including two different election cycles, helps the model avoid overfitting to one cycle's idiosyncrasies, enhancing its ability to generalize over time. This is comparable to Zhang et al. (2017), who train on 2014 data and test on 2016 data. The 2020 dataset was collected over the course of the 2020 election cycle, similar to 2022 as described above. We sampled 3,671 ads from the 2020 data (which was originally about 1.4m ads in size, larger than 2022 because it was a presidential cycle) and 2,921 ads from 2022 data (a random subset of the inference set described above), resulting in a set of 6,592 ads (the *train/test set*). We hired nine student coders with experience in political data coding to annotate this train/test set.[13]

[13]Each student coder was paid $15/hour, the regular rate at our institution. We received an IRB

We applied a random 80/20 train/test split.[14] We stratified the split by election cycle to ensure that both cycles are represented equally in the train/test set (2020: 56%, 2022: 44%). A random sample of 454/631 (2020/2022) ads from the train/test set was double-coded by two student coders to evaluate intercoder agreement (the *double-coded set*). Table 3 shows the frequencies of negatively and positively coded instances. It also shows student coder agreement measured by Krippendorff's alpha, reaching values of 0.6 for most ad goals, indicating general reliability of the ground truth data.[15]

Table 4: Weighted F1 scores of DistilBERT and Random Forest classifiers. Generally, the DistilBERT classifiers outperformed the Random Forest classifiers reaching an average weighted F1 score of 0.95 across classifiers. Calculating the micro F1 score for each class and weighting it by the number of instances in each class results in a micro-averaged F1 score of 0.90.

	Acquisition	Contact	Donate	Event	Learn	Persuade	Poll	Purchase	Vote
DistilBERT	**0.964**	**0.991**	**0.986**	**0.997**	**0.872**	0.886	**0.989**	0.991	**0.916**
Random Forest	0.962	0.984	0.979	0.995	0.820	**0.870**	0.988	**0.992**	0.897

We trained one binary classifier for each ad goal. As a baseline we used a Random Forest classifier based on Scikit-Learn (Pedregosa et al., 2011) and TF-IDF features. Comparing this baseline against different classifiers, we obtained the best performance with a base DistilBERT model (Sanh et al., 2019). We fine-tuned the base model with Hugging Face Transformers (Wolf et al., 2020) for 3 epochs using a batch size of 32. Table 4 shows the F1 scores, weighted by label frequency.[16] Overall, our model performs reliably.[17] To process the ads in our fine-

exemption determination from our institution both for the work of the student coders as well as for the collection of data in connection with this study.

[14]Empirical studies have shown that the best results are obtained if 70–80% of the data is used for training and the remaining 20–30% for testing (Gholamy et al., 2018).

[15]Generally, Krippendorff's alpha values above 0.6 indicate substantial agreement (Landis and Koch, 1977). However, any ranges are only guidelines (Artstein and Poesio, 2008). Machine learning classifiers can tolerate data with lower reliability as long as the disagreement looks like random noise (Reidsma and Carletta, 2008).

[16]Appendix, Table 5 shows the F1 scores for the positive instances only.

[17]Model, datasets, and reproducibility instructions are publicly available in our replication repository: https://github.com/Wesleyan-Media-Project/ad_goal_paper. While we do not expect perfect repro-

tuned model we concatenated the text fields in each Facebook ad. Ads on Facebook contain the following text fields: ad creative body, ASR text (we transcribed audio in videos via the Google Cloud Speech-to-Text API (Google, 2024)), OCR text, (we extracted text in images and videos via the Amazon Rekognition API (Amazon, 2024)), page name, disclaimer (which we included because for some sponsors, their primary goal is plainly evident from their sponsor name, e.g., Vote Early 2022), ad creative link caption, ad creative link title, and ad creative link description.[18] Due to the word limit of DistilBERT (we used the default context length of 512 words), we concatenated the text fields in descending order of importance. If the concatenated text would be too long to be fully included, the most important parts come first and are included. We consider ad creative body to be the most important field because it usually contains the primary message of the ad. In most cases, an ad can be understood based on ad creative body alone. For video ads, this is similarly true for the voice-over, which is why we place ASR second. Image and video text also stands out to the viewer, which is why we put OCR third. The other fields contain supplemental information.

We trained and applied the DistilBERT model on a Google Colab GPU instance. Each model took about 7 minutes to train. All other computations were performed on a desktop PC with an Intel i5-13600KF and 64GB RAM and did not take a significant amount of time.

Ad Goal Analysis

After training and testing we applied our model to the inference set of 377,721 ads. We analyzed the frequency of ad goals and their relationship to several other factors, including ad sponsorship, ad runtime, goals over time, and audience. All analyses cover ads run on Facebook, Instagram, or both, unless noted otherwise.

ducibility due to neural model limitations (PyTorch Contributors, 2023), we expect any changes to be small enough not to impact the substantive interpretation of our results.

[18]We conducted robustness tests on how the use of different feature sets would affect model performance by re-estimating the Random Forest based on ad creative body alone, ad creative body + ASR + OCR (the fields that tend to contain the main content), and all fields (which is what we ultimately opted for). Table 6 in the Appendix shows the results. The additional features result in small performance improvements.

Goal Frequency

Table 2 shows ad counts by goal, how much was spent on each goal in total, and the proportions of these quantities. Persuasion is the goal that occurs most frequently. More than half of all ads try to persuade the viewer (51%). Persuasion also receives the largest proportion of spending (38%). The preponderance of persuade ads indicates that getting people to vote for a specific candidate is the major purpose for election advertisers on Facebook and Instagram. By contrast, vote ads receive less than a fifth of the spending of persuade ads. Further of note is how common ads are that invite viewers to learn more (31%). Such ads appear to be the favored strategy to get voters to engage with campaigns, much more so than ads that promote events and ads that contain polls, which would accomplish a similar purpose. From a normative perspective, the relatively high frequency of learn ads ought to be welcome news given that one would not ordinarily expect social media platforms to be a place where voters would go to be educated. To illustrate what ads with particular goals tend to look like, Table 8 in the Appendix shows one example ad for each predicted goal (randomly sampled from candidate ads whose ad creative body was less than 200 characters so they would fit in the table, with probabilities proportional to ad spend).

Ad Sponsors

While some goals, such as persuasion, are considerably more frequent than others in *aggregate*, a look at *individual* ad sponsors reveals a more nuanced picture. Figure 1 shows how much the highest-spending sponsors invest in each goal. This, and all other spend figures show total US dollar amounts based on the mean spend between lower and upper bounds provided by Meta. Spending is added over a set of ads, e.g., by individual sponsors or by all candidate sponsors. If an ad has multiple goals, the spending is evenly distributed between them. For example, if $100 are spent on ad with two goals, $50 are attributed to each goal.

Sandy Hook Promise, as an example, spent a relatively high amount on acquisition ads, while several others frequently ran donate ads—all of them candidates. Figure 2 sheds further light on such differences by considering the sponsor type. Recall there are three notable sponsor types: candidate campaigns, parties, and group sponsors. As we mentioned previously, we do identify a small number of non-group “other” sponsors (i.e., government officials, government

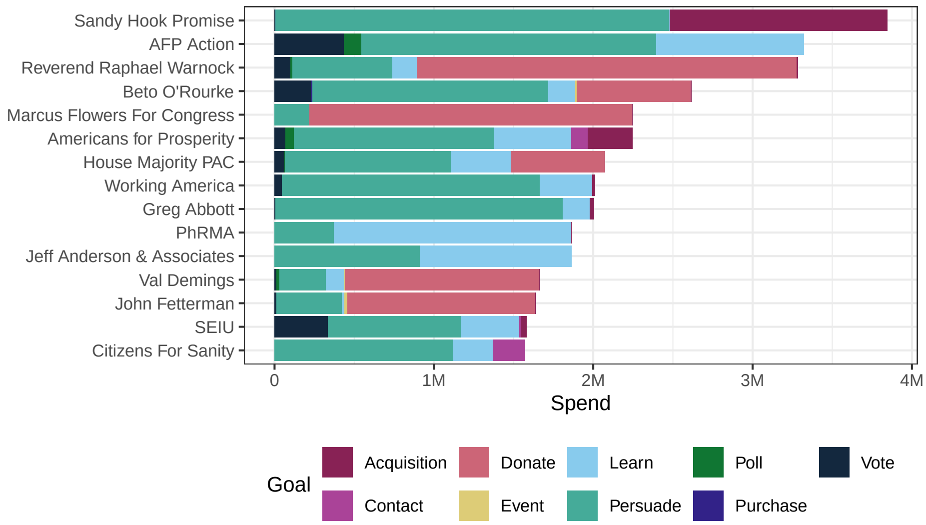


Figure 1. Goal spend by highest-spending sponsors. The values are total spend amounts in US dollars for each sponsor.

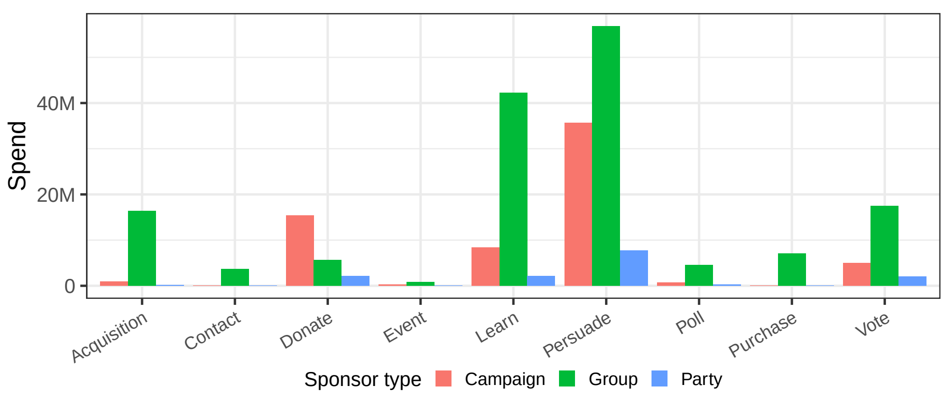


Figure 2. Goal spend by sponsor type. The values are total spend amounts in US dollars for each sponsor type.

agencies, down-ballot candidates and coordinated communications between state and/or local-level parties and candidates). These minor sponsors combined make up only about 1% of all ads, so we omit them from this analysis, with one exception: For the sake of completeness, their ads are included in Table 2.

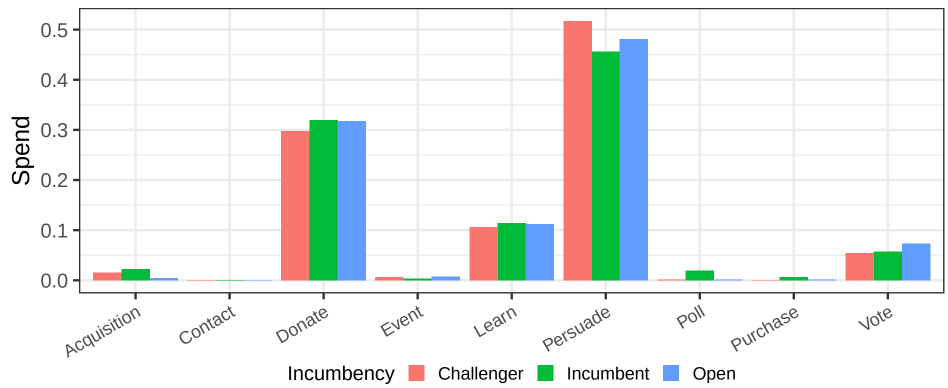


Figure 3. Goal spend by incumbency. The values are proportions for each candidate, averaged by incumbency status.

Two points stand out. First, as Figure 2 illustrates, regardless of ad goal, outside groups are the largest spenders. Given that many, such as 501c organizations, do not report their donors, this finding suggests that funding transparency is limited; we do not know who is funding many of the online ads sponsored by groups.[19] Second, groups and campaigns spend their money in different ways. While both run a relatively large number of persuade ads, groups lean more heavily on learn ads. They also run relatively more acquisition and vote ads, which are less common among campaigns. Campaigns, meanwhile, spend relatively more on donate ads indicating the fundraising pressure candidates are under. Party-sponsored ads take up a small share of ad spend and behave similar to groups.

Differences in spending patterns also exist between candidates, although these differences are more modest. While donate and persuade ads dominate among all candidates, incumbents and candidates for an open seat spend more on donate ads than challengers, as Figure 3 shows. Indeed, incumbents are known to spend a lot of their time in office on fundraising (O'Donnell, 2016). Still, the differences across candidate type are not as large as the differences across sponsor type, though challengers devote more resources to persuasion ads than incumbents and open seat candidates.

[19]There is little regulation of online advertising whereas there is much more for television. If ads do not mention federal candidates, federal campaign finance laws do not apply.

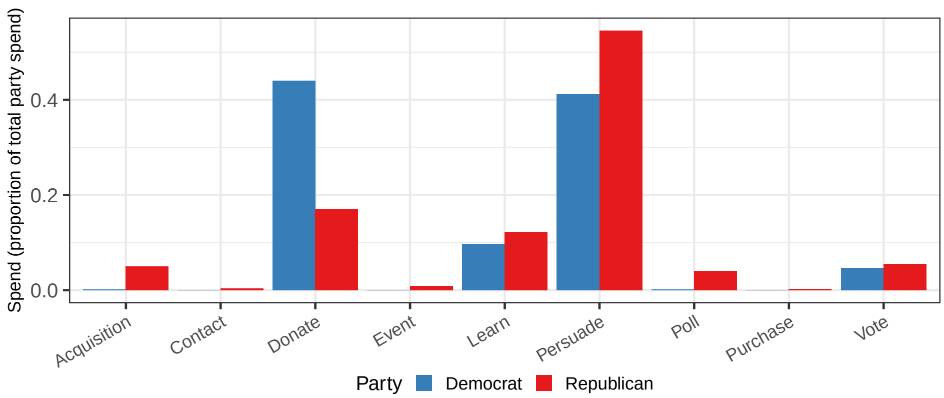
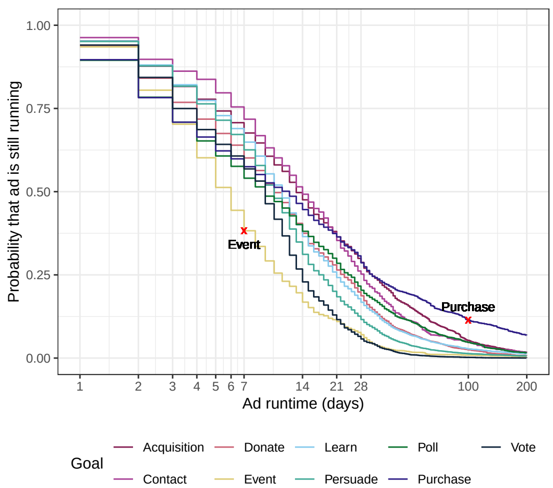


Figure 4. Goal spend by party. The values are proportions of total party-level spend on each goal.

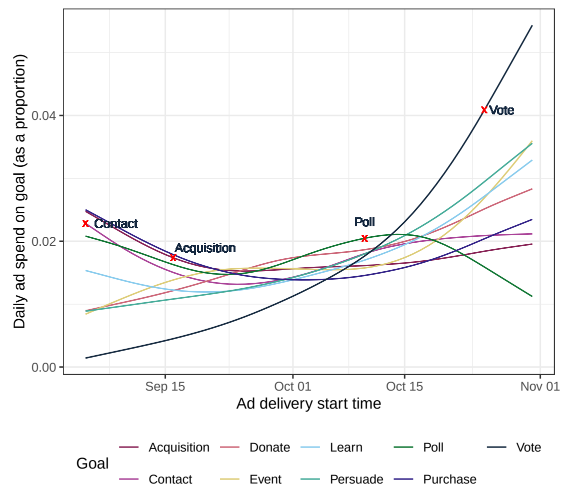
Spending on ad goals also differs by party affiliation. Our categorization of a candidate as a Democrat or Republican is based on FEC data that we match to Meta pages. As candidates register themselves and their party affiliation with the FEC, we can look up their affiliation upon a match of a candidate's name in our data. As Figure 4 illustrates, Democratic candidates spent a much greater proportion on donate ads. On the face of it, this focus looks like Democrats were in greater need of donations. However, the opposite is true. Adding all ad spend in the inference set of Democratic and Republican candidate sponsors, respectively, we find that ad spending by Democrats was 3.75 times higher than by Republicans. The fact that Democrats were able to direct more money towards donation ads may suggest that 'the rich get richer' or that Democrats rely more on small donations. While Republicans spent a greater proportion on learn and persuade ads, their spending was still surpassed by Democrats in these categories in absolute terms (not shown in Figure 4).

Ad Runtime

Another strategic aspect of election ads on social media that can influence the outcome of an election is the duration of ad campaigns. Different goals are associated with different strategies, which are accomplished over different spans of time. To evaluate this point we can treat the ad runtime as a survival process. We calculate the Kaplan-Meier estimates (Kaplan and Meier, 1958) of the probability that an ad is still running after a given number of days. Note that if an ad is stopped on the day it was started, we consider it to have been run for one day. Further,



(a) Probability of an ad still running after a given number of days. For better readability, the x-axis is in log-scale.



(b) Goals over time. The values are the proportion of spend on a particular goal happening on a given day. If an ad has multiple goals, the spend is split between them equally.

Figure 5. Time-based analyses.

our inference set does not have any left-censoring: We only consider ads in our calculation that ran during the period from September 5 to November 8, and the start date of those ads is recorded as September 5 even if the ad ran before that date. We consider ads that run until after November 8 to be right-censored as they do not have any further election relevance.

Figure 5a shows the results. Most ads, regardless of goal, have a short duration. After one week, the chance that an ad is still running is between 54%–72%. An ad's goal generally does not make much of a difference in this regard as the survival curve is approximately equally steep for all goals. However, event ads are a notable exception with only 38% still running after one week.

Generally, short-term oriented ad goals, such as promoting an event (only relevant for a particular event, such as a campaign rally) and vote ads (these ads are typically run directly ahead of the election) only have about a 5% chance of still being online after one month. Purchase ads are particularly long-lasting, and have a 7% chance of running up to and beyond

200 days, a point at which all other ad goals are close to 0%. This finding illustrates the different purposes served by these ads: Purchase ads largely exist to make money off of politics. They do so by appealing to people's partisanship (positive or negative), which does not change much over the course of a campaign or after. Other types of ads react to events occurring at various points in the campaign and are quickly replaced by ads with different goals once the news cycle moves on.

Goals over Time

A related question is the extent to which advertisers' priorities shift over the course of an election campaign. Figure 5b shows spending on new ads that were placed on a given day, by goal. Since some goals receive far more spending than others, which would drown out the smaller goals in the plot, we normalize the data so that the value for a given goal on a given day represents the proportion of total spending on that goal that happens on this day. Due to the high amount of day-to-day spending variability, we apply cubic spline smoothing to the time series to make the trends more obvious. Since Meta implemented a moratorium on new ads the week before the election (while existing ads were allowed to continue running), we only consider new ads until October 31 for our analysis.

A broad trend we observe for almost all goals is that spending on new ads picks up as election day comes closer. This is likely because advertisers make their final arguments and spend any remaining budget right before voters cast their ballots. Furthermore, all ads for the final week have to be in place before the moratorium takes effect. Contact ads make up the highest proportion initially. However, their proportion decreases substantially. The reason for this decrease could be that appeals to contact a legislator to effectuate change become less relevant as a legislator's term approaches its end. Acquisition ads are more frequent closer to the start of the campaign as well. One exception to the general trend are poll ads, which have two peaks—right at the start in September and in the second week of October. Most other goals generally follow an increasing trend towards the day of the election. Unsurprisingly, this trend is strongest for vote ads, which, when compared to all other goals, pick up substantially more towards the end of the period.

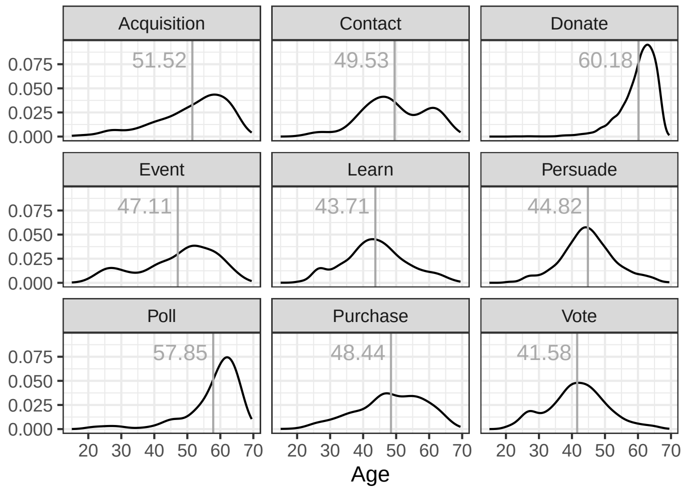


Figure 6. Target group age, by goal. Both densities and means are weighted by spend.

Ad Audience

People are shown different ads depending on their demographics. We gain insight into how age correlates with different ad impressions using Meta's Ad Library (Meta, 2024a). For each ad, we multiply each age group's (13–17, 18–24, 25–34, etc.) mean age by the proportion of an ad's impressions that were displayed to this age group and sum over them. This calculation results in a weighted mean age for each ad. Figure 6 shows the distribution of these mean ages across ads (weighted by ad spend), according to which goals were targeted.

It can be observed that campaigns primarily target donation ads to older voters, who generally have more disposable income. This result is consistent with the resource model of political participation (Brady et al., 1995). Vote ads have the lowest mean age (41.58), consistent with how difficult it is to turn out younger voters. The "bump" in the distribution of vote ads for people between 20 and 30 years old, which is also observable for most other goals, albeit, less pronounced, indicates that campaigns devoted significant resources to countering this problem by speaking to younger voters.

We also consider the platform on which ads are displayed: Facebook and Instagram (the number of ads on Messenger is magnitudes lower and thus omitted here). Figure 7 shows a comparison of the 82,583 ads that exclusively ran on Facebook and the 17,222 that exclusively

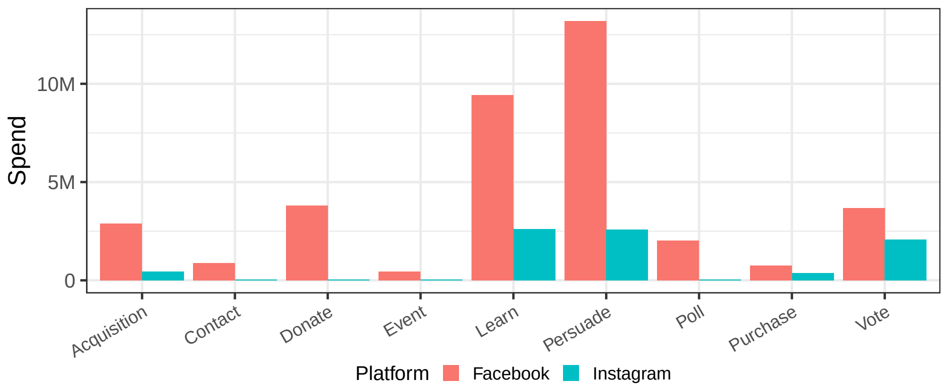


Figure 7. Goal by platform.

ran on Instagram (the vast majority of the remaining ads ran on both Facebook and Instagram). While Facebook makes up the lion's share of ad spending, the amounts are closer for one category: vote ads. It is also noteworthy that donate ads are rare on Instagram. Both of these trends are symptoms of the same dynamic: election advertisers use Instagram to reach younger people, which is why vote ads are so common and donate ads are so rare. Younger people donate little and require more convincing to turn out to vote.

Discussion

Our analysis provides the first comprehensive analysis of digital advertising in federal elections and how the goals of advertising vary within and across different sponsor types. Extended party network theory highlights coordination among candidates, parties, and allied groups, but as Nyhan and Montgomery (2015) note, empirical observation of actual campaign behavior and strategy has long been hampered by a lack of systematic data, especially for challengers and less visible campaigns and sponsors. Through analyzing data from digital ad archives, we make visible the strategic behavior of actors who previously left little trace in traditional data sources.

The EPN literature demonstrates that campaigns operate through a division of labor among coordinated actors. Our findings reveal this division directly in digital advertising: candidate committees focus on fundraising, outside groups on persuasion and informational appeals, and parties play a limited intermediary role. Across networks, Democrats devote more attention to donation appeals while Republicans emphasize persuasion, reflecting both organizational and partisan asymmetries. This specialization aligns closely with the patterns described in prior EPN research.

The temporal distribution of goals further suggests coordination within the network. Fundraising appeals by candidate sponsors peak early in the campaign cycle, while persuasion and vote-related ads from outside groups intensify closer to Election Day. This sequencing indicates that campaign functions are not only differentiated but also temporally aligned. Members of the EPN responsible for fundraising move early, providing resources for the subsequent persuasion efforts of allied organizations. Moreover, because some actors in the network face legal spending limits while others do not (Skinner et al., 2012), unrestricted members likely absorb a greater share of activity once those limits are reached.

Digital transparency data also expose what Oklobdzija (2024) terms “dark parties,” networks of non-party organizations that channel resources and messaging beyond traditional oversight. We show that outside groups are the largest ad spenders, and many top sponsors have opaque funding origins. Their dominance in persuasion and learning ads, contrasted with candidates’ focus on fundraising, reveals the outsourcing of core communicative functions to formally independent actors. Party committees themselves account for a small share of overall spending, an evolution consistent with Skinner et al. (2012)’s argument that new legal and technological environments push parties to incorporate semi-autonomous extensions like 527s. Recent evidence of de facto coordination mechanisms such as “redboxing” (Foy-Sutherland, 2025) further illustrates how nominally independent groups align their efforts with party and candidate strategies.

EPN theory conceives parties as adaptive systems responding to institutional and technological change (Skinner et al., 2012; Sheingate et al., 2022). Our data show such adaptation in the digital domain: Facebook ads are more likely to be shown to older donors and prioritize fundraising, while Instagram ads are more likely to feed mobilization appeals to younger voters. Incumbents emphasize donation appeals; challengers rely on persuasion. These platform- and audience-specific strategies demonstrate the flexibility of those within the extended party network to adapt themselves to digital advertising.

Our data provide a foundation for exploring additional questions and building theory within the EPN framework. Do outside groups sponsor persuasion ads while campaigns work to fill their coffers with donations? Do outside groups and parties shift to persuasion ads when polling flags for a candidate in a competitive race? Do these sponsors abandon or enter races at strategic times, looking for ways to broaden the party's electoral success nationally? Do they focus on acquisition and voter education ads for campaigns with open seat candidates or competitive challengers? The extended party thesis is premised on a wide network of supportive groups "pitching in" to help a preferred national party achieve broad electoral success. This implies divisions of labor not only in where money is spent but how and when.

Beyond EPN theory, segmenting ads by goal also enables other lines of scholarly analysis. In line with the above, assessments of ad impact should consider the relevant subset of ads intended for that goal. For example, do digital campaign ads influence voters' assessments of candidates? To study that question, scholars should subset the digital ad information environment to persuasion ads. A candidate who spends disproportionately on acquisition ads is not trying to influence vote choice. Similarly, one might question the return on investment from donation ads. The correct approach would be to examine donation ads specifically and link those pleas to campaign contribution flows. Without understanding the variation in digital ad goals, the approach of linking ad spending to various dependent variables would lead to imprecise measurements of the relevant information environment.

Limitations

Our approach is subject to various limitations. We analyze data from Meta's social networks (Facebook, Instagram, and Messenger). Thus, we cannot make claims about election advertising on other social networks. Our data relate to one particular time period in the US. While we believe that we would see similar trends for other time periods in the US, we expect that local idiosyncrasies would prevent a generalization of our results to other countries. Several of our analyses report advertiser spending on ads. Meta does not provide the exact amounts spent on each ad; rather, it reports upper and lower bounds, such as \$100–199. We used the average but acknowledge that doing so is an approximation. Finally, the scope of our dataset is determined by Meta's "Issues, elections or politics" category. Because election content can be defined in different ways and we aim to be over-inclusive to capture the broadest universe, our inference set incorporates some advertisers that focus on issue campaigns (e.g. gun control, abortion) rather than candidates in the election. For these advertisers, what the persuade model effectively learns is whether the ad tries to persuade viewers to support an issue rather than a candidate. This is not strictly in line with the election definition of this goal, but it is a fairly sensible thing to do for the model since persuasion is a goal in campaigns more broadly.

Conclusion

Despite the increasing prevalence of election ads on social media and other online platforms, we are just at the beginning of understanding their impact. Our work here shows the feasibility of election ad goal classification on social networks at scale and provides fodder for theory building in literatures like the extended party network. Different types of sponsors (i.e., candidates, parties, and interest groups) pursue different goals in their advertising. We have demonstrated that knowing an ad's goal is an important predictor for how long it runs, where it runs geographically, and whom it targets. Because our approach does not rely on human coding, beyond the coding for testing and training the classifiers, it opens up the study of election advertising to online ads, which number in the hundreds of thousands, if not millions, each year and affords comparisons across candidates, campaigns, and even cross-nationally.

The analysis of election advertising is of civic importance because it reveals the details of campaign strategies that otherwise would remain opaque. For example, that a campaign is focused on fundraising vs persuasion vs mobilization has implications for how people participate in electoral politics and whether they do so in an informed fashion. Although this research cannot speak directly to the relative effects of ads with different goals, the variety of ad goals unearthed suggests that an ad's goal should be considered when trying to evaluate its impact. Other future endeavors could be a longitudinal study over multiple election cycles to track the evolution of election ads online as well as the correlation of online ad spending and content with actual election outcomes.

Acknowledgments

We are grateful to the Wesleyan Media Project's undergraduate research assistants for their coding assistance. The work was supported by National Science Foundation grants 2235006, 2235007, and 2235008. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

11 C. F. R. § 100.5 (2025). Political committee. Electronic Code of Federal Regulations. Retrieved October 16, 2025, from https://www.ecfr.gov/current/title-11/part-100/section-100.5.

Amazon (2024). Amazon rekognition. https://aws.amazon.com/rekognition/. Accessed: December 4, 2025.

Artstein, R. and Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Comput. Linguist.*, 34(4):555–596.

Ballard, A. O., Hillygus, D. S., and Konitzer, T. (2016). Campaigning online: Web display ads in the 2012 presidential campaign. *PS: Political Science & Politics*, 49(3):414–419.

Bartels, L. M. (1988). Presidential primaries and the dynamics of public choice.

Brady, H. E., Verba, S., and Lehman Schlozman, K. (1995). Beyond SES: A Resource Model of Political Participation. *American Political Science Review*, 89(2):271–294.

Desmarais, B., La Raja, R. J., and Kowal, M. S. (2015). The Fates of Challengers in U.S. House Elections: The Role of Extended Party Networks in Supporting Candidates and Shaping Electoral Outcomes. *American Journal of Political Science*, 59(1):194–211.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Dommett, K., Barclay, A., and Gibson, R. (2024). Just what is data-driven campaigning? a systematic review. *Information, Communication & Society*, 27(1):1–22.

Druckman, J. N. (2004). Priming the vote: Campaign effects in a us senate election. *Political Psychology*, 25(4):577–594.

Edelson, L., Sakhuja, S., Dey, R., and McCoy, D. (2019). An analysis of united states online political advertising transparency. *arXiv preprint arXiv:1902.04385*.

Eulau, H. and Karps, P. D. (1977). The Puzzle of Representation: Specifying Components of Responsiveness. *Legislative Studies Quarterly*, 2(3):233–254.

Fowler, E. F., Franz, M. M., Martin, G. J., Peskowitz, Z., and Ridout, T. N. (2021). Political Advertising Online and Offline. *American Political Science Review*, 115(1):130–149.

Foy-Sutherland, G. (2025). Working Hand in Hand: Interest Groups and Cooperative Dynamics in Campaigns. *The Journal of Politics*, page 737690.

Gholamy, A., Kreinovich, V., and Kosheleva, O. (2018). Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. Technical Report UTEP-CS-18-09, The University of Texas at El Paso. Accessed: December 4, 2025.

Google (2024). Turn speech into text using google ai. https://cloud.google.com/speech-to-text. Accessed: December 4, 2025.

Jurafsky, D. and Martin, J. H. (2023). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, Upper Saddle River, New Jersey.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481.

Koger, G., Masket, S., and Noel, H. (2009). Partisan Webs: Information Exchange and Party Networks. *British Journal of Political Science*, 39(3):633–653.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174.

Lazarsfeld, P. F., Berelson, B., and Gaudet, H. (1968). *The people's choice: How the voter makes up his mind in a presidential campaign*. Columbia University Press, New York, New York.

Martin, G. J. and Peskowitz, Z. (2018). Agency Problems in Political Campaigns: Media Buying and Consulting. *American Political Science Review*, 112(2):231–248.

Meta (2024a). Ad library. https://www.facebook.com/ads/library/. Accessed: December 4, 2025.

Meta (2024b). Ad library aggregate report. https://www.facebook.com/ads/library/report/. Accessed: December 4, 2025.

Meta (2024c). Ad library api. https://www.facebook.com/ads/library/api/. Accessed: December 4, 2025.

Monroe, B. L., Colaresi, M. P., and Quinn, K. M. (2008). Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4 SPEC. ISS.):372–403. ISBN: 1476-4989.

Nyhan, B. and Montgomery, J. M. (2015). Connecting the Candidates: Consultant Networks and the Diffusion of Campaign Strategy in American Congressional Elections. *American Journal of Political Science*, 59(2):292–308.

O'Donnell, N. (2016). Are members of Congress becoming telemarketers? https://www.cbsnews.com/news/60-minutes-are-members-of-congress-becoming-telemarketers/. Accessed: December 4, 2025.

Oklobdzija, S. (2024). Dark Parties: Unveiling Nonparty Communities in American Political Campaigns. *American Political Science Review*, 118(1):401–422.

Panagopoulos, C. (2021). *Bases loaded: How US presidential campaigns are changing and why it matters*. Oxford University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Petrocik, J. R. (1996). Issue ownership in presidential elections, with a 1980 case study. *American journal of political science*, pages 825–850.

PyTorch Contributors (2023). Pytorch reproducibility. https://pytorch.org/docs/stable/notes/randomness.html. Accessed: December 4, 2025.

Reidsma, D. and Carletta, J. (2008). Reliability measurement without limits. *Comput. Linguist.*, 34(3):319–326.

Ridout, T. N., Fowler, E. F., and Franz, M. M. (2021). The influence of goals and timing: How campaigns deploy ads on facebook. *Journal of Information Technology & Politics*, 18(3):293–309.

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2019). Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Sheingate, A., Scharf, J., and Delahanty, C. (2022). Digital Advertising in U.S. Federal Elections, 2004-2020. *Journal of Quantitative Description: Digital Media*, 2:1–61.

Skinner, R. M., Masket, S. E., and Dulio, D. A. (2012). 527 Committees and the Political Party Network. *American Politics Research*, 40(1):60–84.

Stromer-Galley, J., Rossini, P., Hemsley, J., Bolden, S. E., and McKernan, B. (2021). Political messaging over time: A comparison of us presidential candidate facebook posts and tweets in 2016 and 2020. *Social Media+ Society*, 7(4):20563051211063465.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., and Rush, A. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.

Zhang, F., Stromer-Galley, J., Tanupabrungsun, S., Hegde, Y., McCracken, N., and Hemsley, J. (2017). Understanding discourse acts: Political campaign messages classification on facebook and twitter. In Lee, D., Lin, Y.-R., Osgood, N., and Thomson, R., editors, *Social, Cultural, and Behavioral Modeling*, pages 242–247, Cham. Springer International Publishing.

Appendix

Goal Coding

When coding the ads for their goals, student coders were presented with the following questionnaire:

Which of the following elements are included in the ad?

DONATE Request for donations

CONTACT Request for viewer to contact a legislator

PURCHASE Selling of merchandise

GOTV GOTV efforts

EVENT Advertisement of a event/rally

POLL Asks the viewer to fill out a survey/poll

GATHERINFO Solicitation of viewers email/name/information (other than a poll)

LEARNMORE Invite viewer to learn more

Student coders were instructed to code an ad as [**PERSUADE**] if its primary goal was to promote or attack candidates including attempts to convince viewers to vote for a specific candidate.

Additionally, student coders were provided with the following supplemental guidance for further clarification:

Request for donations

Any time the ad asks you to contribute money. Include ads that have a “donate” button or a link to a donation page, even if there is no other mention of donating in the ad.

Selling of merchandise

Any physical items the ad tries to sell you

- Clothes, mugs, hats, etc.
- Commemorative coins, memorabilia

GOTV (get out the vote)

Any information the ad gives on how/when to vote should be marked as a GOTV element.

- Election date ("Vote Biden on November 3rd")
- How to register to vote
- How to find your polling place
- Information on mail-in/absentee voting

Advertisement of a event/rally

This includes free events or tickets for paid events.

Solicitation of viewers email/name/information (other than a poll)

Any time an ad wants the viewer to give their name, email address, FB information, phone number, etc.

- Signing birthday cards

- Signing petitions
- Filling out forms
- “Add your name”
- Invited you to text a number

*surveys and polls should only be coded as survey/poll, not both.

Invite viewer to learn more

- The “learn more” button at the bottom of of the ad would count here
- Anything like “visit our webpage to learn about ______” or to “read more about ______”

Note that we re-named “GOTV” to “Vote”, “GATHERINFO” to “Acquisition,” and “LEARNMORE” to “Learn” for greater clarity in this paper.

Before starting their coding task, student coders participated in at least four rounds of meetings in which they coded practice sets. The practice sets each included 15–20 ads. After coding a practice set, coders received a written answer key, which they reviewed with one of us either individually or in a group feedback session.

Table 5: DistilBERT and Random Forest classifier performance given as F1 scores of the positive instances. Generally, the DistilBERT classifiers outperformed the Random Forest classifiers, reaching an average F1 score of 0.81 across classifiers.

	Acquisition	Contact	Donate	Event	Learn	Persuade	Poll	Purchase	Vote
DistilBERT	**0.689**	**0.700**	**0.963**	**0.875**	**0.885**	**0.915**	**0.731**	**0.876**	**0.761**
Random Forest	0.651	0.333	0.945	0.800	0.845	0.906	0.681	0.881	0.688

Table 6: Random Forest classifier performance for different feature sets, given as weighted F1 scores.

	Acquisition	Contact	Donate	Event	Learn	Persuade	Poll	Purchase	Vote
Ad creative body	0.952	0.983	0.965	0.994	0.803	0.851	0.983	0.988	0.874
Ad creative body + ASR + OCR	0.959	0.980	0.968	0.993	0.800	0.864	0.985	0.988	0.892
All fields	**0.962**	**0.984**	**0.979**	**0.995**	**0.820**	**0.870**	**0.988**	**0.992**	**0.897**

Additional Model Performance Results

Table 5 provides F1 scores for the positive instances only, complementing Table 4, which shows weighted F1 scores.

Ad Text

To demonstrate the face validity of classified data, Table 7 shows the keywords most associated with each goal, using the method described by Monroe, Colaresi, and Quinn (2008), which assigns z-scores to keywords for two categories (in this case, a given goal vs. not that goal) using a Dirichlet prior. See also Jurafsky and Martin (2023) (chapter 25) for an introduction to this method. The table shows that keywords associated with each goal are generally what is expected.

For donate ads, the top keywords are unsurprising as they relate directly to donating. ActBlue is the primary donation platform of the Democratic party. A little further down are the keywords "race" and "deadline," indicative of a common strategy in donate ads of instilling a sense of urgency in audiences in order to get them to donate (similar to, for example, "only 1 item left" in online shopping platforms).

Contact ads are rare, so the keywords here are very directly related to the few campaigns

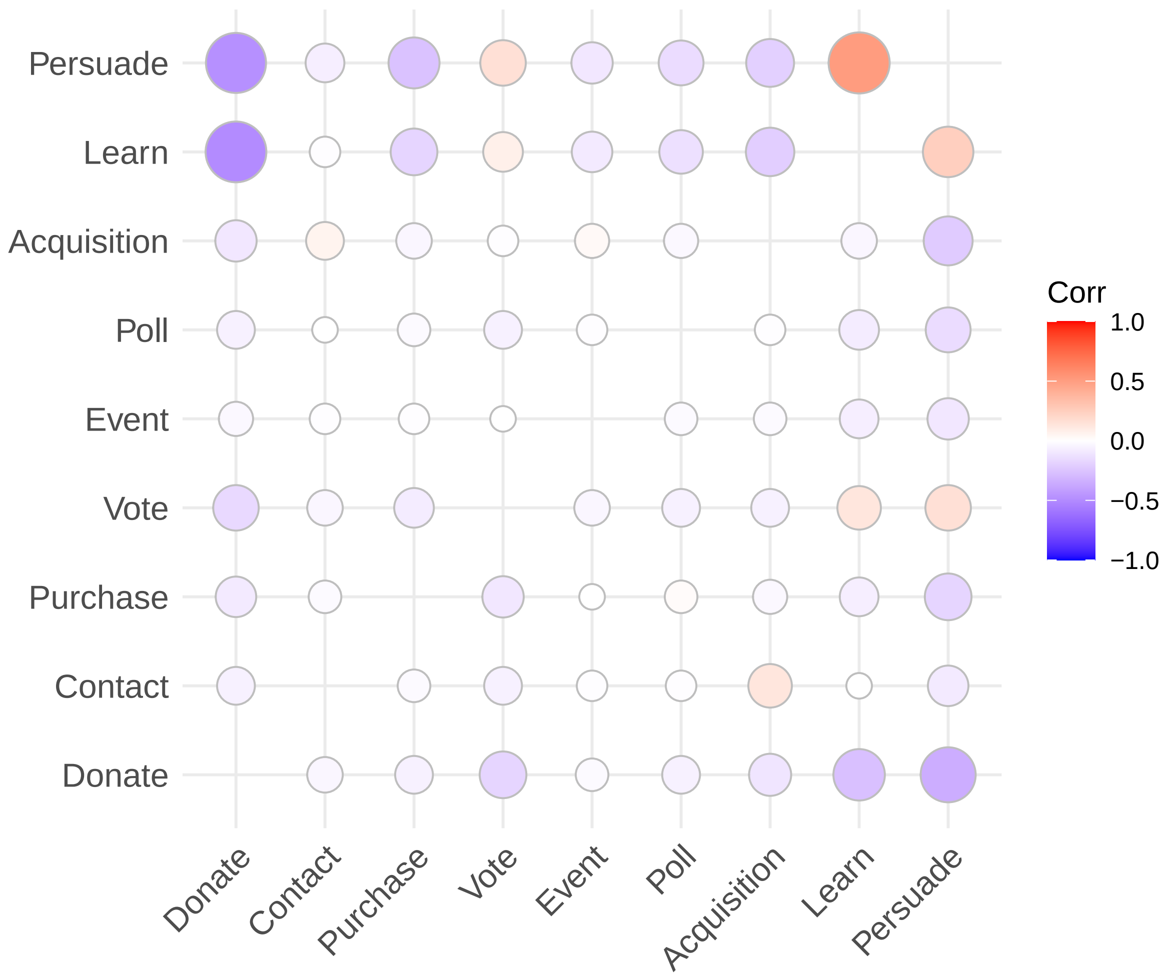


Figure 8. Overlap between goals. The top left triangle shows correlations between goals in the train/test set. The lower right triangle shows correlations between goals in the inference set.

that do run them, especially, Senators Hassan and Warnock. Purchase ads primarily relate to selling MAGA-themed swag. Unsurprisingly, GOTV ads focus on voting. Noteworthy here is how high "early" voting is on the list. This is likely related to the Democratic party's strategy of focusing on early and mail-in (mail is the 8th word) voting in the context of Donald Trump's voter fraud claims. "November 8th" and "Tuesday" are on the list too, though, indicating a focus on election-day voting as well.

The top keywords for events ads are largely unsurprising, and the same is true for poll ads. Acquisition ads contain the keywords "sign" and "petition," which indicates that asking someone to sign a petition is a common tactic used to get people to provide their contact information. Ads that invite the viewer to learn more have several Spanish words among their top keywords, indicating that reaching out to Spanish-language speakers is important for many campaigns. Persuade ads are similar to vote ads but are more targeted towards a specific candidate.

Table 7: Top 20 keywords associated with a given goal.

Acquisition	Contact	Donate	Event	Learn	Persuade	Poll	Purchase	Vote
sign	tell	donate	rsvp	vote	state	prageru	trump	vote
name	sen	help	wing	early	paid	poll	free	november
add	act	now	sunday	para	vote	trump	twins	voting
petition	nation	chip	veep	por	district	survey	tuttle	8th
action	inflation	can	west	que	county	take	limited	voter
join	stop	secure.actblue.com	oct	los	representative	presented	america	early
climate	hassan	democrats	7pm	action	8th	america	2024	election
pledge	liberals	today	9th	voting	house	prageru.com	flag	register
care2	call	senate	event	las	candidate	learning	gift	ballot
demand	project	race	join	learn	city	jill	edition	nov
fb.me	pipeline	win	october	prosperity	november	2024	books	day
markowicz	energy	across	tickets	energy	authorized	owens	boyack	tuesday
karol	taxes	donation	saturday	aarp	council	mobile	cameo	online
para	pass	abrams	rally	staceyabrams.com	police	candace	hat	registration
humane	prevention	campaign	tour	climate	judge	download	facebook	georgia
sexualization	evers	seat	meet	con	michigan	app	poll	mail
society	nrdc	urgent	september	una	board	biden	wexner	plan
que	reckless	country	courier	más	arizona	follow	patriots	make
tell	senator	need	now	votar	taxes	simonian	ame	voted
worker	feeding	deadline	events	planned	mayor	facebook	save	easy

Table 8: Example ads for each predicted ad goal (randomly sampled from candidate ads with probabilities proportional to ad spend). The ad classified as Contact contains further information in the OCR text ("TELL BIDEN AND PELOSI TO STOP HARASSING THE MIDDLE CLASS"), though it likely isn't a sincere call to contact them. The Donate ad contains a link to the GOP's donation platform (winred.marcorubio.com) and its ASR text goes on to say "I'm reaching out to as many people as I can to try and help raise the funds that we need to fight back against these lies."

Goal	Page Name	Ad
Acquisition	Kevin McCarthy	I AM DONE. Pelosi MUST GO. I am calling on ALL GOP Patriots to join my call to FIRE PELOSI. 1 hour left to add your name:
Contact	D'Esposito for Congress	Biden, Pelosi, and the Radical Left just passed a new law spending $80 billion to hire 87,000 new IRS agents!
Donate	Marco Rubio	I just had to take a huge risk. I really hope it was worth it.
Event	Colin Schmitt	Great afternoon in Port Jervis for fall harvest fest with Tom Faggione and Karl Brabenec! —
Learn	Lauren Boebert for Congress	Aspen Adam Frisch is just like Pelosi = he makes everything more expensive! Aspen Adam voted to increase electricity rates for residents living in affordable housing units
Persuade	Ron Johnson	With record-setting inflation, Americans need lower taxes. What's Mandela Barnes' plan? Increase spending AND raising taxes - he's just another dangerous liberal
Poll	Elise Stefanik	I want to know what matters most to my constituents. Will you take my District Priorities Survey?
Purchase	Marjorie Taylor Greene	Joe Biden just declared every MAGA Republican an Enemy of the State. Claim your exclusive gear and wear yours proudly: https://bit.ly/3TCtuUF
Vote	Josh Harder	Ballots are heading to mailboxes—don't wait to vote! Voting by mail is easy! Make sure you vote for Josh Harder on page 2, then sign, and date your envelope. Return your ballot ASAP—no stamps needed!