Journal of Quantitative Description: Digital Media 2(2022), 1–61 10.51685/jqd.2022.026

Digital Advertising in U.S. Federal Elections, 2004-2020

ADAM SHEINGATE[1]
JAMES SCHARF
CONNER DELAHANTY
Johns Hopkins University, United States

Digital advertising is now a commonplace feature of political communication in the United States. Previous research has documented the key innovations associated with digital political advertising and its consequences for campaigns and elections. However, a comprehensive picture of political spending on digital advertising remains elusive because of the challenges associated with accessing and analyzing data. We address this difficulty with a unique dataset (*N*=3,639,166) derived from over 13 million expenditure records reported to the Federal Election Commission (FEC) between 2004 and 2020. Employing supervised machine learning to classify expenditures into nine categories, this paper makes four key observations about digital campaigns. First, 2020 was a watershed election in terms of the growth of digital spending. Second, there are clear partisan differences in the resources allocated to digital advertising. Third, platform companies play a central role in an otherwise partisan market for digital ads and services. Fourth, digital platforms and consultants occupy a distinct ideological niche within each party.

[1] Sheingate: adam.sheingate@jhu.edu
Date submitted: 2022-04-18
The authors wish to thank the editors and two anonymous reviewers for their comments on earlier drafts.

Keywords: U.S. elections, campaign spending, political advertising, digital media

Contemporary political campaigns in the United States rely heavily on advertising to mobilize and persuade voters (Fowler et al. 2021). Increasingly, this work takes place online (Fowler et al. 2020a). According to data from the Kantar/CMAG media group, presidential candidates spent $435 million on digital advertising between April 1 and Election Day 2020, nearly a quarter of total ad spending during this period (Ridout et al. 2021). Using data from the Facebook and Google ad libraries, Ridout et al. (2021) identified $743 million in online spending by presidential candidates and their affiliated committees over the course of the entire 2020 election cycle. This represents a significant increase from 2016 when presidential candidates spent an estimated $166 million online (Williams and Gulati 2018).

Despite its increasing importance, reliable data on how much campaigns spend on digital advertising remains difficult to access and analyze (Fowler et al. 2020a). Proprietary data, such as those collected by Kantar/CMAG is not widely available to scholars. Publicly available sources, such as the Facebook and Google ad libraries, only cover spending since 2018 and therefore cannot shed light on the long-term growth of digital advertising (Fowler et al. 2021; Ridout et al. 2021). A complete picture of digital campaigning would also include spending on the consultants who drive and inform strategy (Hersh 2015), something that cannot be ascertained from the ad libraries alone.

Another public source of data is the Federal Election Commission (FEC). All candidates for federal office, political parties, and political action committees (PACs) must file reports with the FEC that include a description of each individual transaction. Although comprehensive in many respects, FEC data presents its own set of challenges, including the need to clean and classify millions of individual records (Williams et al. 2020). This paper addresses the challenges of working with FEC data using a supervised machine learning model that classifies expenditure descriptions into nine categories. The resulting dataset provides the most detailed picture to date of digital political advertising and related services in U.S. federal elections.

Our analysis of digital spending yields four key observations. First, the 2020 election cycle was a watershed in the development of digital campaigning. We estimate that federal candidates and committees spent $2.8 billion on digital ads and services in 2020. To put this sum in perspective, inflation-adjusted spending on digital ads during the 2020 election exceeded the entire amount spent in the previous eight election cycles combined. Second, we find evidence of a Democratic advantage in spending on digital ads and services. Although the electoral benefits of outspending one's rival (online or offline) depends on a variety of factors (Baldwin-Philippi 2017, Sides et al. 2022), our data points to a persistent partisan difference in the resources allocated to digital advertising. Third, a network analysis of shared clients among the top digital firms confirms research showing the central position of Facebook and Google in an otherwise partisan market for digital ads and services (Barrett 2021). Fourth, digital consulting firms and platform companies occupy a distinct ideological niche within each party. Using estimates of candidate and committee ideology derived from campaign contributions (Bonica 2014), we find that clients of firms specializing in digital ads and services are more ideologically extreme compared to the average clients of traditional media firms, fundraisers, pollsters, and consultants.

The next section discusses previous research on campaign spending, the growth of digital advertising, and the role of political professionals in contemporary campaigns. This work generates several research questions that guide the present study. Next, we introduce the model we used to classify FEC data on political expenditures. We then describe the growth of spending on digital advertising since 2004, including partisan differences in the resources allocated to digital ads and services. Our attention then shifts to the firms providing digital ads and services, their connections to one another, and the ideological character of the candidates and committees they serve. We conclude with a discussion of directions for future research and the implications of our work for campaign finance and election transparency.

Campaign spending online and offline

Previous research on campaign spending showed there is remarkable consistency in how candidates allocate financial resources. In a study of U.S. House races, Limbocker and You (2020) found very little change in spending patterns from one cycle to the next, even when candidates face an altered strategic context. An important exception to this behavior is that candidates spend more on media when facing a new challenger.

One explanation for both the general consistency in resource allocation and the higher media spending in certain races is the role of political consultants. Martin and Peskowitz (2015) find that consultants tend to work for the same incumbents each cycle, rather than in competitive races where their services might serve the strategic interests of the party. Consultants also charge significant markups on the production and purchase of advertising, although the size of the markup decreases as race competitiveness (and the volume of political advertising) increases (Martin and Peskowitz 2018). In sum, efforts to understand how campaigns allocate resources must consider the role and financial incentives of the political professionals who advise the candidates (Sheingate 2016).

Research on digital advertising suggests how campaigning online might alter resource allocations and the role of political professionals in U.S. elections. Digital advertising is significantly cheaper than television because it can be purchased in smaller increments and targeted at narrow audiences. The lower cost of Facebook or Google ads relative to television makes advertising on these platforms especially attractive for candidates with limited resources, such as challengers and down ballot candidates (Fowler et al. 2021).

The growth of digital advertising has consequences for the political consulting industry as well. Research by Barrett (2021) found that social media platforms such as Facebook and Google occupy a central role in the market for digital ads by virtue of their work on behalf of clients across the political spectrum. In this key respect, platform companies differ from typical political consulting firms that operate as part of an extended partisan network and whose staff is ideologically aligned with the clients they serve

(Kolodny and Logan 1998; Koger et al. 2009; Nyhan and Montgomery 2015; Martin and Peskowitz 2018). Facebook and Google's work with clients from both parties departs from this partisan organization of professional advice.

Scholars have also examined the partisan dimensions of online advertising, especially in terms of the staffing capacities of campaigns (Kreiss and Saffer 2017). Research on the early development of digital campaigning, for instance, identified key innovations among staffers and firms working on behalf of Democratic presidential candidates before diffusing broadly across the political landscape (Kreiss 2012, 2016; Baldwin-Philippi 2015). Other work shows a digital advantage on the Republican side, at least at the presidential level, starting with the 2016 Trump campaign (Williams and Gulati 2018). An assessment of media spending in the 2020 election also found that Trump spent more on Facebook and Google than Biden, although Democrats outspent Republicans overall, on television and online (Ridout et al. 2021). These partisan gaps at the presidential level depart from overall spending patterns where there are typically small differences between the parties. Looking at House races, Limbocker and You (2020) found Democrats allocated slightly more resources to staff, Republicans spent slightly more on fundraising and media.

Research Questions

Research on political spending in the United States finds consistent patterns of resource allocation that reflect established relationships between campaigns and political professionals. Recent studies of online campaigning point to a variety of ways that digital advertising has disrupted these established patterns and relationships. However, a clear picture of both the timing and magnitude of these disruptions is limited by a lack of data (Fowler et al. 2020a; Williams et al. 2021). Tracking the growth of digital media and its implications for candidates and campaigns requires accurate estimates of political spending that extends beyond the snapshot of a single electoral cycle. Similarly, understanding the role of platform companies in an otherwise partisan political marketplace or whether digital consultants serve a distinct clientele from other professionals requires data on the professional firms that populate the ecosystem of American politics.

In sum, detailed descriptive work on political spending is necessary for advancing research on digital advertising and its consequences for the conduct of campaigns. We can summarize these data requirements and their implications for future scholarship with reference to four research questions:

RQ1: How has spending on digital ads and services changed over time?

RQ2: Is there a partisan difference in spending on digital ads and services?

RQ3: What is the role of platform companies in the market for digital ads?

RQ4: Do digital consultants differ from other types of political professionals?

We attempt to answer these questions using a novel dataset of federal campaign expenditures from 2004 to 2020.

A Method for Classifying Federal Election Commission Expenditures

Although scholars have used FEC reports to examine various aspects of U.S. campaign finance and candidate strategy (Limbocker and You 2020; Martin and Peskowitz 2018; Nyhan and Montgomery 2015; Williams and Gulati 2018), these data present a distinct set of challenges (Williams et al. 2021). First among them is the need to clean and code millions of records. This task is especially vexing because campaigns are inconsistent in how they describe the purpose of expenditures. For instance, spending on media could appear in the data as “television”, “advertising”, or simply “media”. Classification based on keywords can lead to an ever-expanding number of search terms. For instance, Limbocker and You (2020) required 600 keywords to classify a subset of the 3.5 million expenditures they analyzed.

This paper addresses the limitations of keyword searches with a supervised machine learning model that classifies FEC expenditures into nine categories, including digital ads and services.[2] Our dataset is derived from FEC bulk reports of operating and independent expenditures for each of the nine election cycles between 2004 and 2020, 13.7 million records in total.[3] Each record includes a text description of the purpose of the expenditure as reported to the FEC. After removing punctuation and symbols, as well as additional cleaning, there were 620,292 unique expenditure descriptions in the data.

We train and test our model using a random sample (N=9,752) of unique descriptions hand-coded according to the nine categories in table 1.[4] Our categories resemble those used by the transparency organization OpenSecrets, with some important differences.[5] Following work by Limbacker and You (2020), we create separate categories for polling and consulting. In addition, we distinguish digital expenditures from other forms of media.

Table 1: Expenditure Classification Categories

Category	Description
Media	Advertising on television, radio, billboards, and newspapers.
Digital	Advertisements online or through email and data analytics
Polling	Survey research and polling-related expenditures

[2] A full description of the model is in the Appendix.

[3] For a description of the data sources, see https://www.fec.gov/campaign-finance-data/operating-expenditures-file-description/ and https://www.fec.gov/campaign-finance-data/independent-expenditures-file-description/.

[4] Training refers to the process of fitting the model; testing is the process of evaluating the model using a subset of the sample data. Authors Sheingate and Scharf coded the training and testing data using the categories in table 1 (interrater agreement of .82; Krippendorff's Alpha of .78). Sheingate then reconciled differences prior to training and testing the model. Our training data consists of N=7,801 unique descriptions; our testing data consists of N=1,951 expenditure descriptions.

[5] For a list of categories used by OpenSecrets, see https://www.opensecrets.org/campaign-expenditures/methodology

Table 1: Expenditure Classification Categories

Legal	Legal fees, compliance, and accounting.
Field	Canvassing and other forms of direct voter contact.
Consulting	Any type of consultant or consulting-related expenditure.
Fundraising	Direct mail, telemarketing, and other costs of raising money.
Travel	Airline tickets, hotels, car rentals, and other travel expenses.
Administrative	Salaries, food, and other administrative costs.

As reported in table 2, the model's overall accuracy is 93%. Precision refers to the percentage of expenditures labeled correctly. For the category "digital", 96% of observations coded as digital are correct and 4% are incorrect, meaning they belong to another category. Recall refers to the ability to place an expenditure in the correct category. Again, the model classified 96% of digital expenditures correctly and incorrectly classified 4% of digital expenditures as belonging in another category. The F1 score is the harmonic mean of precision and recall.[6] Among the nine categories, the model performs least well classifying fundraising and administrative expenditures. This may be due to the kinds of expenditures associated with these two categories.[7]

Table 2: Model Accuracy (testing data)

	precision	recall	f1-score	support
media	0.96	0.96	0.96	488

[6] In more formal terms, precision refers to true positives as a share of true positives and false positives. Recall refers to true positives as a share of true positives and false negatives. Support is the number of observations by category in the testing data.

[7] For example, an expenditure described as "food and beverages for fundraiser" include keywords for both the administrative and fundraising categories.

Table 2: Model Accuracy (testing data)

administrative	0.87	0.93	0.90	412
digital	0.96	0.96	0.96	359
field	0.92	0.92	0.92	274
fundraising	0.93	0.83	0.88	188
travel	0.93	0.92	0.93	74
consulting	0.92	0.92	0.92	65
legal	1.00	1.00	1.00	46
polling	0.98	0.91	0.94	45
macro avg	0.94	0.93	0.93	1951
weighted avg	0.93	0.93	0.93	1951
accuracy			0.93	1951

To gain a clearer picture of how our model performed, we ran the classifier on the entire set of unique descriptions (N=620,292) and then examined a subset of expenditures that account for 95% of the total spending in each of the nine categories (just over 20,000 unique descriptions between 2004 and 2020). We find our model is slightly more accurate when classifying the top expenditures in each category. For the digital category, precision is 95% and recall is 98%.[8] Our ability to check and correct the top expenditures in each category contributes further to the accuracy of the dataset.

Preparing and cleaning the dataset

Before exploring spending patterns and trends, we carry out additional preparation and cleaning of the dataset.[9] First, using a unique identifier for each committee, we merge the expenditure data with the committee and candidate master files from the FEC for each cycle between 2004 and 2020. Second, we clean each vendor name to correct for small

---

[8] See Table A2.
[9] Details of this procedure are described in the Appendix.

variations in spelling or naming conventions. Third, we designate a party affiliation for each candidate and committee using the FEC committee and candidate files and data from OpenSecrets. For some committees, we impute a party affiliation using FEC data on independent expenditures.[10] Fourth, we correct for double counting and duplicate transactions.[11] Finally, we sum the individual transactions to create a dataset (N=3,639,166) for each unique combination of committee, candidate, vendor, purpose, and cycle.

Although we have high confidence in our model, there are limitations of the data itself. Campaigns frequently report expenditures simply as “advertising” without specifying whether the communication appeared on television or online. Although our model (correctly) classifies these observations as media, we re-code the expenditure as digital if the vendor is a platform company like Facebook or Google. Outside of these exceptions, we cannot be certain that some of our expenditures coded as media were used for online advertising. This means that spending on digital ads and services may be greater than our estimates. In addition, some of the growth in digital spending could be an artifact of changes in reporting practices as committees get better or more consistent at differentiating digital from traditional media in their reports to the FEC.

A second concern with relying on FEC data relates to a lack of transparency in campaign finance reporting. For instance, 501c organizations, often referred to as “dark money” groups because of the undisclosed identity of their donors, report some but not all their political expenditures to the FEC. Dark money groups can also conceal political activities by channeling funds through consultants or other vendors (Massoglia 2021). The potential for dark money groups to circumvent FEC reporting requirements means that our data might underestimate political spending, especially on media. We discuss both limitations and their implications for our results in the next section. In the conclusion, we

[10] Table A3 explains our procedure for imputing party affiliation.
[11] For instance, a committee may list a $5,000 payment to “Chase Bank” and then itemize five, separate $1,000 payments for “airline tickets”. In this example, the sum of reported transactions is $10,000 but the actual campaign outlay is only $5,000.

address potential reforms that would enhance transparency and render FEC data more reliable and accessible to scholars and citizens alike.

Digital Campaign Spending by Cycle

RQ1 asks how spending on digital ads and services changed over time. Figure 1 displays the sum of independent and operating expenditures on digital ads and services for the nine electoral cycles from 2004 to 2020. Totals are inflation-adjusted using constant 2012 dollars. Figure 1 breaks down digital spending into three categories: advertising, services, and consulting. We calculate totals for ads and consulting using a text search of the purpose descriptions for expenditures classified as digital.[12] Services is a residual category (digital expenditures other than ads and consulting) that typically includes spending on data analytics, digital list rentals, and other services associated with digital campaigning. Overall, we estimate an inflation-adjusted total of $2.4 billion in digital spending in 2020, a 268% increase over the 2016 cycle. During the same period, spending on digital advertising grew from $442 million to $1.7 billion, an increase of 285%.

[12] We used the following key words to estimate digital advertising: “advert”, “ad” “ads”, “media”, “advocacy”, “ad buy”, “digital ad”, and “digital buy”. The consulting category includes all transactions with a purpose description of “digital consulting”.

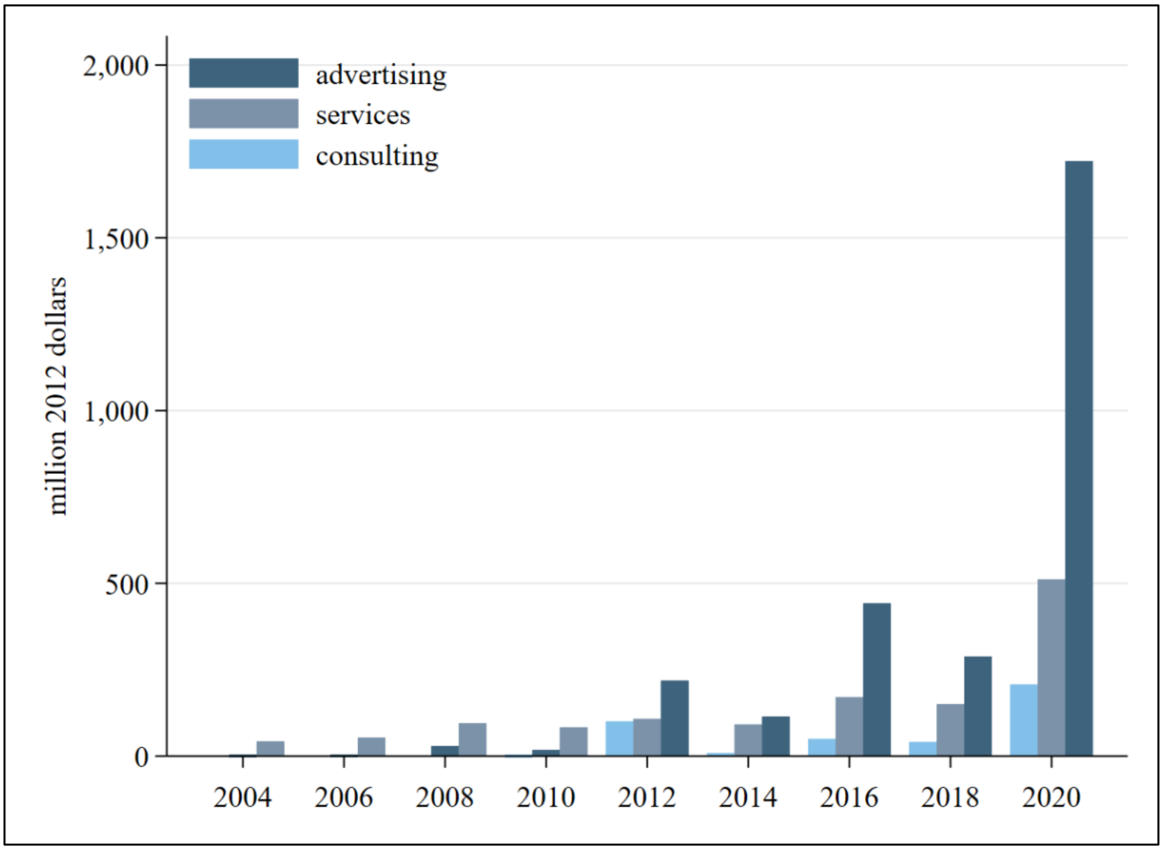

Figure 1: Digital Spending, 2004-2020

Note: includes operating and independent expenditures

To put this growth in perspective, figure 2 compares inflation-adjusted spending on digital ads and services to other categories of campaign spending indexed to a base year of 2016.[13] Again, the increase in digital expenditures is remarkable given the general growth of campaign spending that occurred between 2016 and 2020. During a period when spending on traditional media doubled, expenditures for digital ads and services more than tripled.

[13] See Table A4 for details on spending by category.

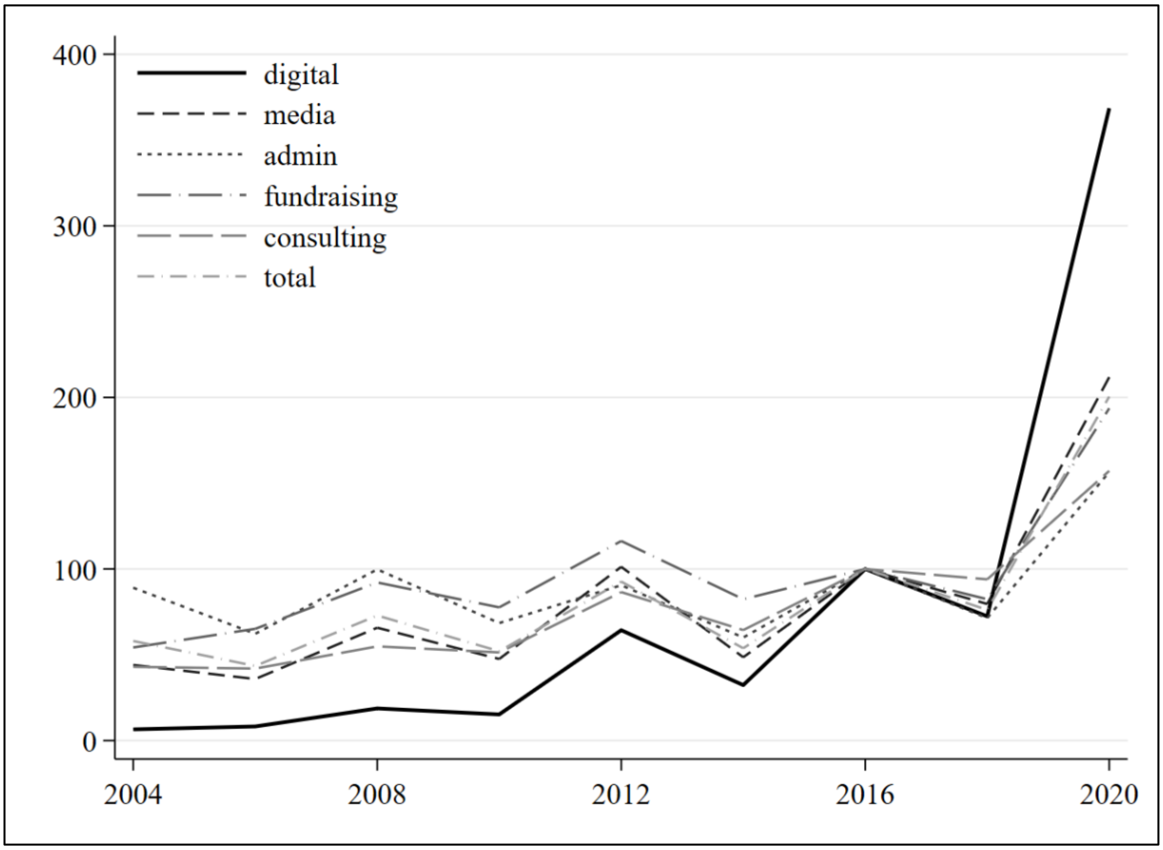

Figure 2: Index of Political Expenditures (2016=100)

Note: includes operating and independent expenditures

The extraordinary growth in digital campaigning represents an important change in how political entities spend money. Figure 3 illustrates this with a violin plot of spending across more than 16,000 committee-year observations since 2004.[14] The data in figure 3 includes major party candidates, parties, and political action committees with positive spending in both the digital and media categories. Although most candidates and committees typically spend more on traditional media than on digital ads, these differences narrowed considerably in recent cycles. In fact, median spending on digital ads and services

[14] Violin plots combine density plots and box plots. The density plot shows the distribution of spending with the width of the curve corresponding to the frequency of data points in each region. The box plot displays the median and interquartile range between the 75th and 25th percentile.

in 2020 slightly exceeded median spending on traditional media.[15] Figure 3 also shows that digital spending at the upper range of the distribution increased significantly. Between 2008 and 2020, the 90th percentile of digital spending increased from $50,000 to $600,000. Over the same period, the 90th percentile of spenders in traditional media grew from $1.2 to $1.6 million.

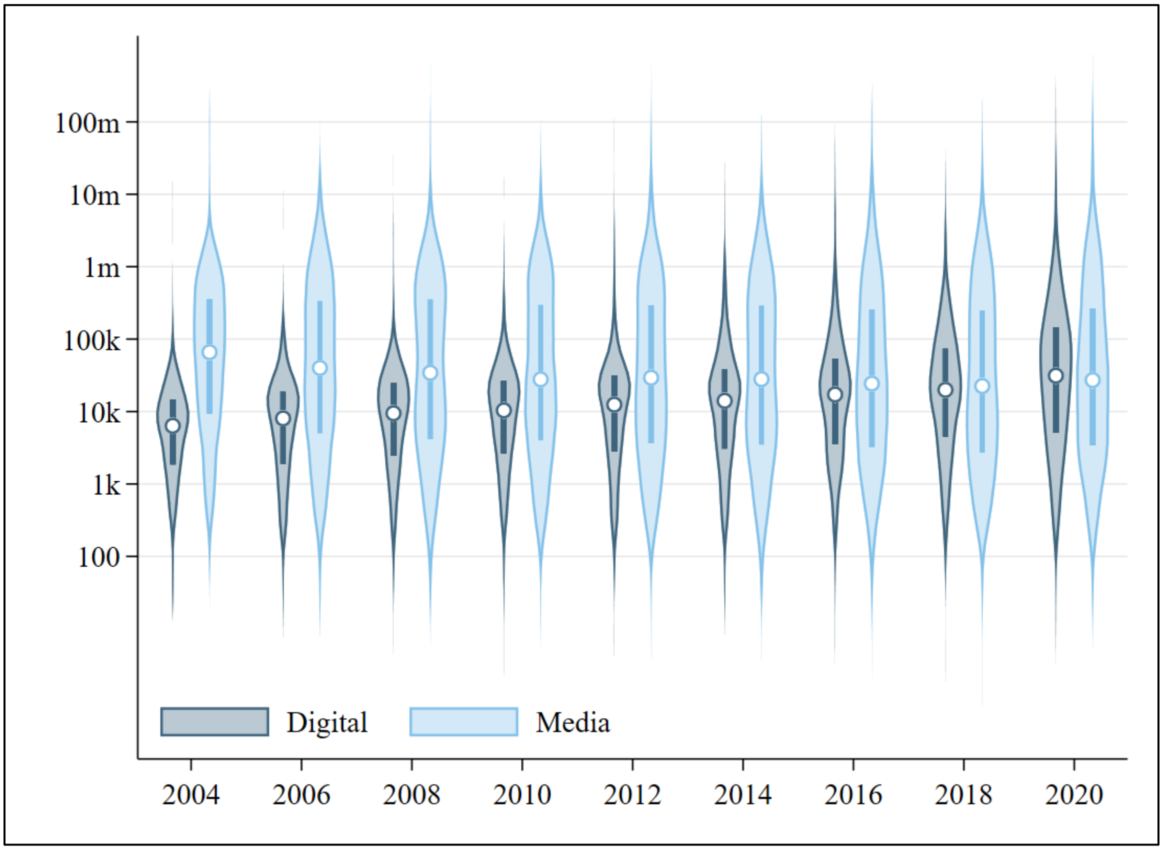

Figure 3: Distribution of spending on digital and media by cycle (log scale)

Note: major party committees with positive spending in both categories

Before we explore these spending patterns in greater detail, there are two features of FEC reports that could affect the accuracy of our estimates. First, in some cases it is difficult to differentiate digital advertising from spending on traditional media such as

15 See Table A5 for descriptive statistics.

television or radio. Descriptions of expenditures such as “ads”, “media buy”, or “advertising” are classified by our model as media whether they appear online or offline.[16] This could result in an underestimate of digital spending. Our estimates could also reflect changing practices as campaigns get better at differentiating digital from traditional media when reporting expenditures to the FEC.

We examine this problem more closely by looking at payments made directly to platform companies such as Facebook and Google. Because we know that 100% of advertising on Facebook or Google is online, we can calculate how much digital spending on platforms is missed by our model due to ambiguous or under-specified descriptions in the underlying data. In 2020, for example, 14% of payments to platform companies used broad terms such as “advertising” and was classified as media by our model. If we assume the same proportion of digital advertising is similarly classified in the rest of our data, we can estimate the additional online spending in each cycle our model might miss due to vague or imprecise expenditure descriptions.[17] As shown in Figure 4, although the overall trend remains unchanged, our re-estimate using the share of platform payments classified as media results in more than $800 million in digital advertising in 2016. Overall, our revised estimate suggests total digital spending (including digital services and consulting) grew from $1.1 billion in 2016 to $2.6 billion in 2020.

[16] Table A6 includes the most common media and digital descriptions.
[17] The appendix describes the estimation procedure.

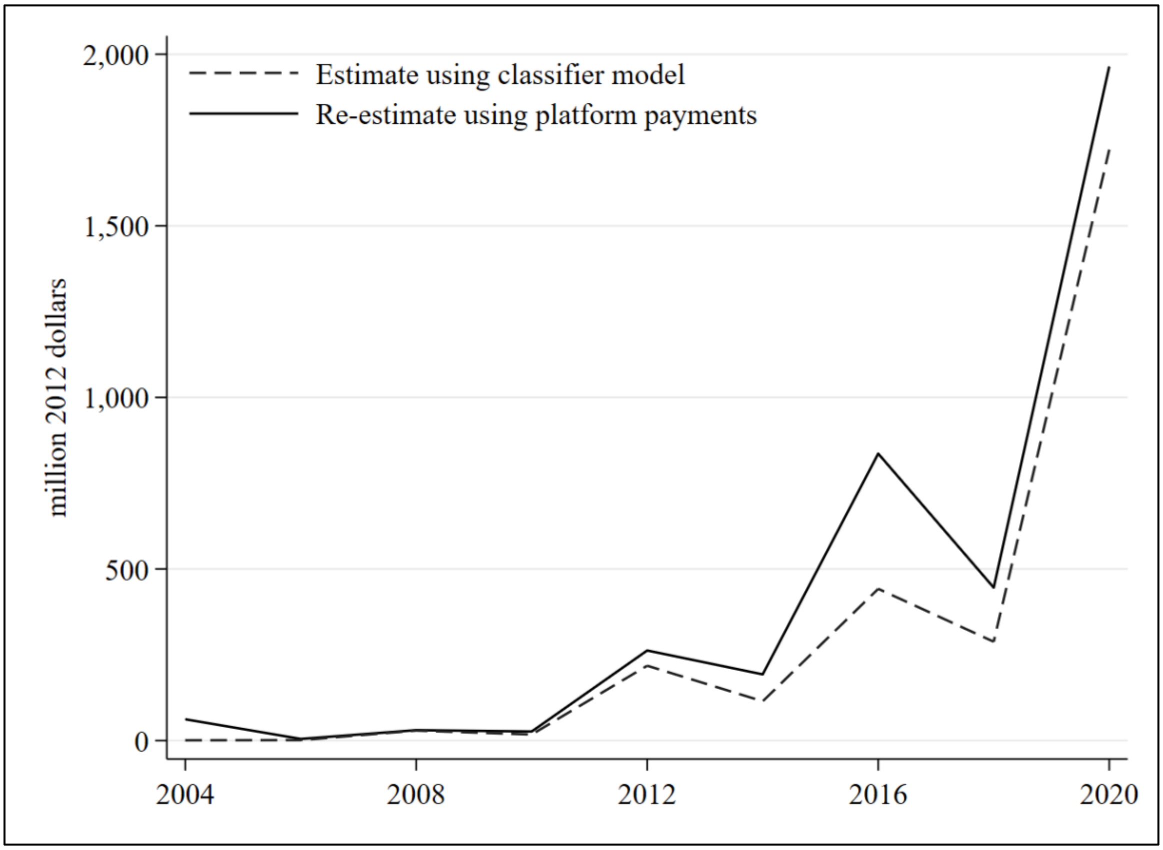

Figure 4: Digital advertising expenditures, 2004-2020

A second potential limitation of using FEC data is that federal regulations only require political entities to report election advertising if it occurs 30 days before a primary and 60 days before a general election. Dark money groups sometimes exploit this loophole in order to make large political expenditures without disclosing them to the FEC.[18] To see if this results in an undercount of television spending in our data, we compare our estimates

[18] FEC rules define election communication as "any broadcast, cable or satellite communication that refers to a clearly identified federal candidate, is publicly distributed within 30 days of a primary or 60 days of a general election and is targeted to the relevant electorate" (https://www.fec.gov/help-candidates-and-committees/other-filers/making-electioneering-communications/, accessed September 30, 2022. For a discussion of how dark money groups exploit this loophole, see https://www.opensecrets.org/dark-money/process (accessed September 29, 2022).

of media expenditures derived from FEC reports with tracking data collected by the Wesleyan Media Project on all political ads in House and Senate races that aired during the 2018 elections (Fowler et al. 2020b). For purposes of comparison, we exclude expenditures in our data that mention other types of media such as radio, newspapers, or billboards. We also differentiate between candidate, PAC, and party committee spending to match the sponsorship of ads appearing in the Wesleyan Media Project data.

Figure 5 displays weekly advertising totals for the 2018 election cycle, comparing data sources and committee types. Our estimate of media spending closely tracks the weekly ad spending from the Wesleyan Media Project. In the case of PACs, for example, we do not find evidence of large advertising buys outside of the 30- or 60-day reporting period. We further explore the similarity between data sources using a regression of weekly advertising expenditures reported to the FEC on the estimated cost of television ads from the Wesleyan Media Project. Our baseline model suggests that each dollar of advertising purchased on television is associated with $1.28 in media expenditures reported to the FEC. This difference likely reflects the markups consultants charge for the production and placement of ads (Martin and Peskowitz 2018). When we include an interaction term of committee types, each dollar of ad spending on television is associated with $1.09 in candidate spending and $1.60 in PAC spending as reported to the FEC. This difference is also consistent with previous research that finds higher markups paid to consultants by PACs compared to candidates (Ibid.).[19]

[19] We find almost no difference in ad spending FEC expenditures reported by party committees. See the appendix for descriptive statistics and coefficient estimates.

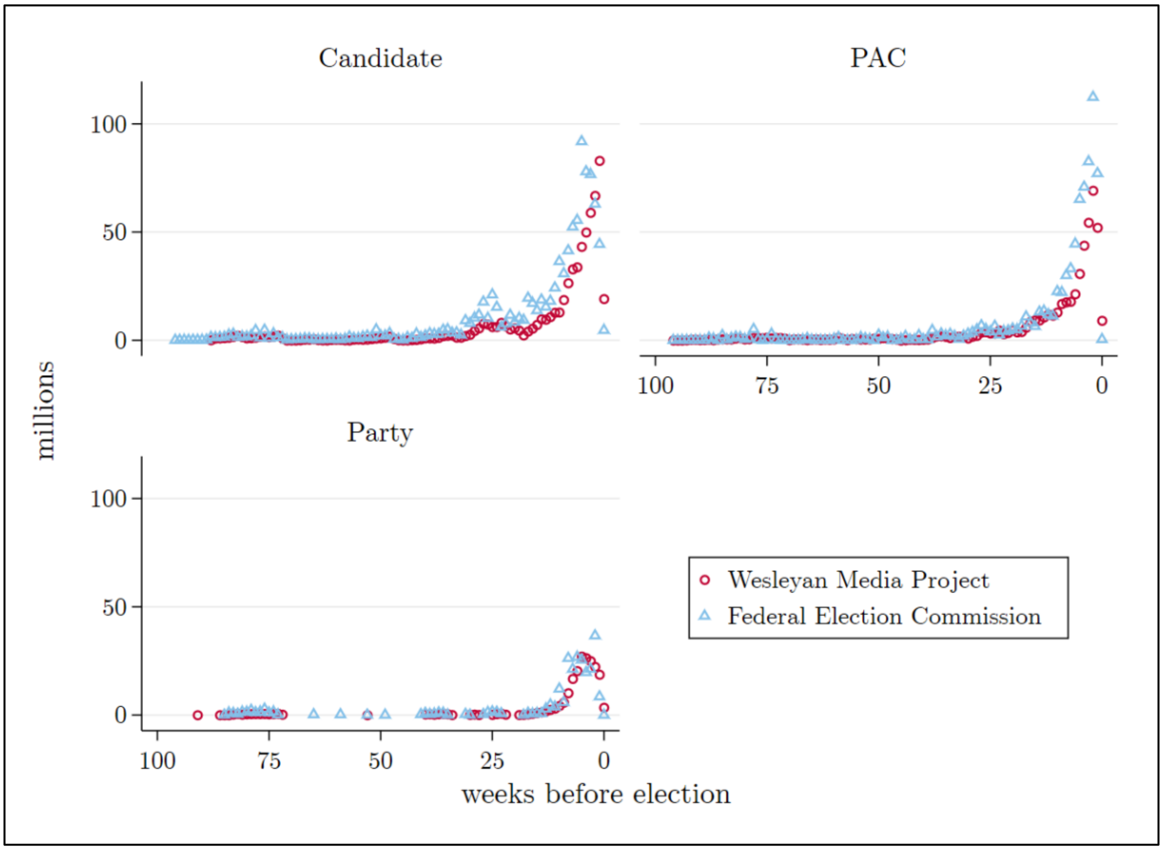

Figure 5: Weekly media spending by committee type and data source

Despite the inherent limitations of our data, we believe our estimates offer the most comprehensive picture to date of political spending in U.S. campaigns. When we compare spending reported to the FEC with ad tracking data, for instance, we do not find evidence of unreported media buys by dark money groups. Although some of the expenditures classified in our data as media likely appeared online, we can use direct spending on platforms such as Facebook and Google to estimate how much additional digital advertising our model might miss due to ambiguous and unstandardized descriptions of expenditures. Even after accounting for this likelihood, our estimates still point to a profound growth in digital spending and a pronounced shift in how campaigns allocate resources.

Partisan Patterns of Digital Spending

RQ2 asks whether there is a partisan pattern to the growth of digital spending. Figure 6 compares the mean share of spending allocated to digital ads and services by party from 2004 to 2020.[20] Consistent with previous work on digital campaign innovations (Kreiss 2012, 2016), Democratic committees and campaigns allocated more money on average than their Republican counterparts. This partisan gap has grown considerably since 2016.

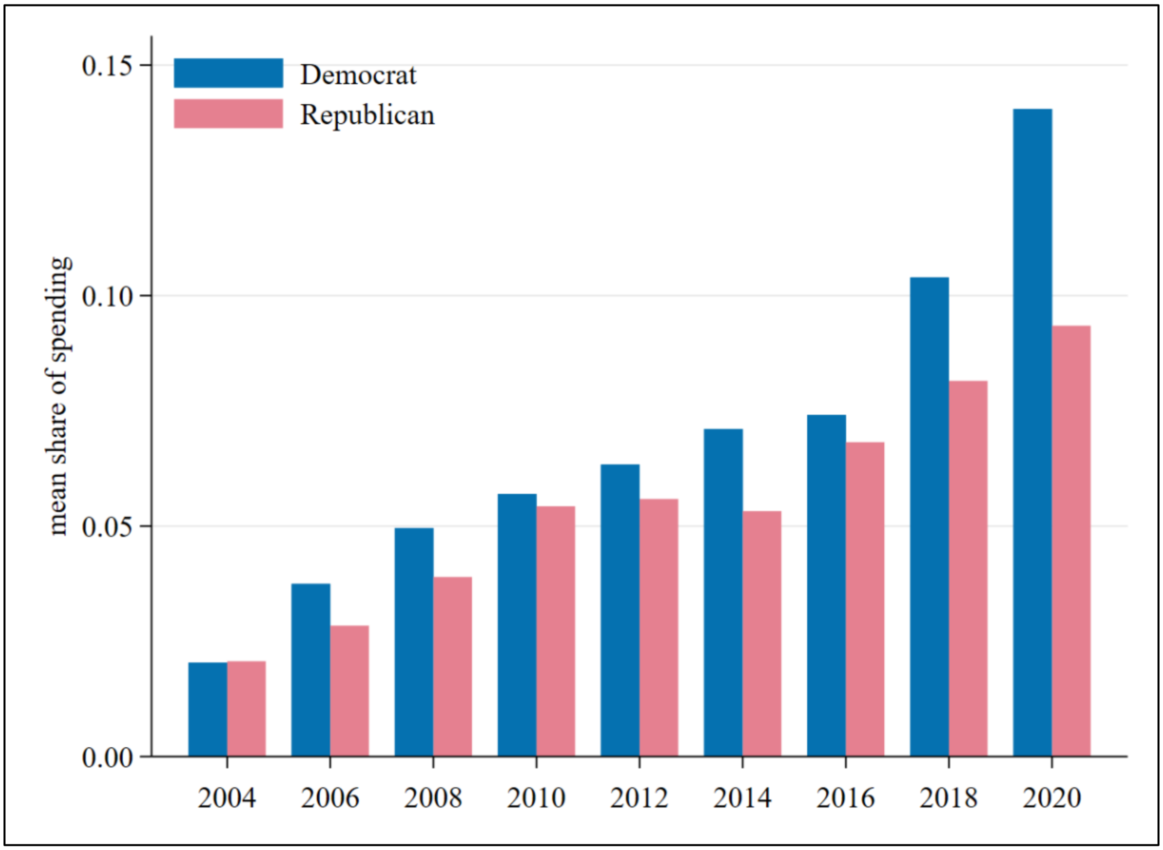

Figure 6: Spending allocated to digital ads and services by party, 2004-2020
Note: operating expenditures only

[20] Because independent expenditure committees typically limit spending to media (online and offline), Figure 6 only includes committees with operating expenditures.

Recent work by Fowler et al. (2021) uses the within-candidate difference in spending to evaluate how campaigns allocate resources between digital advertising and traditional media. Figure 7 adopts a similar approach and displays density plots for both major parties for each cycle since 2004. Positive values indicate committees that spent more on television and cable, negative values show committees that spend more on digital than traditional media. Figure 7 shows there has been a shift from the early 2000s, when committees typically spent more on traditional media, to the 2010s when committees begin allocating more resources to digital ads and services. Figure 7 also points to small but persistent differences with Republican committees tending to spend more on traditional media and Democratic committees spending more on digital.

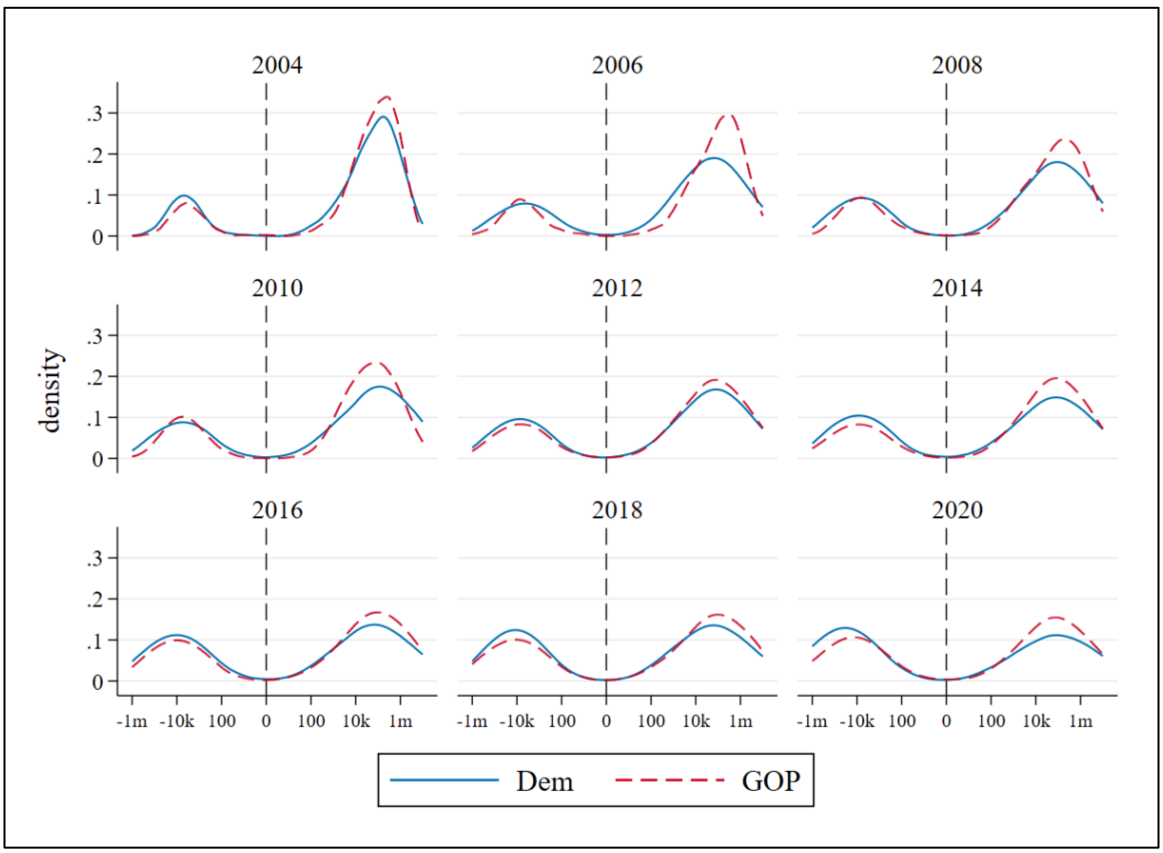

Figure 7: Within-committee difference between media and digital (log scale)
Note: committees with positive spending on digital and media

Figure 8 shows how these partisan differences stack up across different types of committees.[21] With some exceptions, Democratic candidates, party committees, and PACs allocate a larger share of spending to digital ads and services than Republicans. Figure 8 also shows that presidential and Senate campaigns devote a larger share of resources to digital advertising than House campaigns. This is surprising given research showing digital spending is greater among candidates with fewer resources (Fowler et al. 2021). Figure 8 also calls into question some of the received narratives about the role of digital strategy in presidential elections. Despite the innovations of the first Obama campaign, for example, Republican presidential candidates in 2008 allocated a greater share of resources to digital ads and services than Democrats.[22]

[21] Figures for presidential, House, and Senate campaigns include spending by joint fundraising committees and PACs associated with a candidate.

[22] This difference is not statistically significant. Independent group t-test: Democrats (M=.029, SD=.014) - Republicans (M=.076, SD=.039): $t(49)=-1.144$, $p=0.258$.

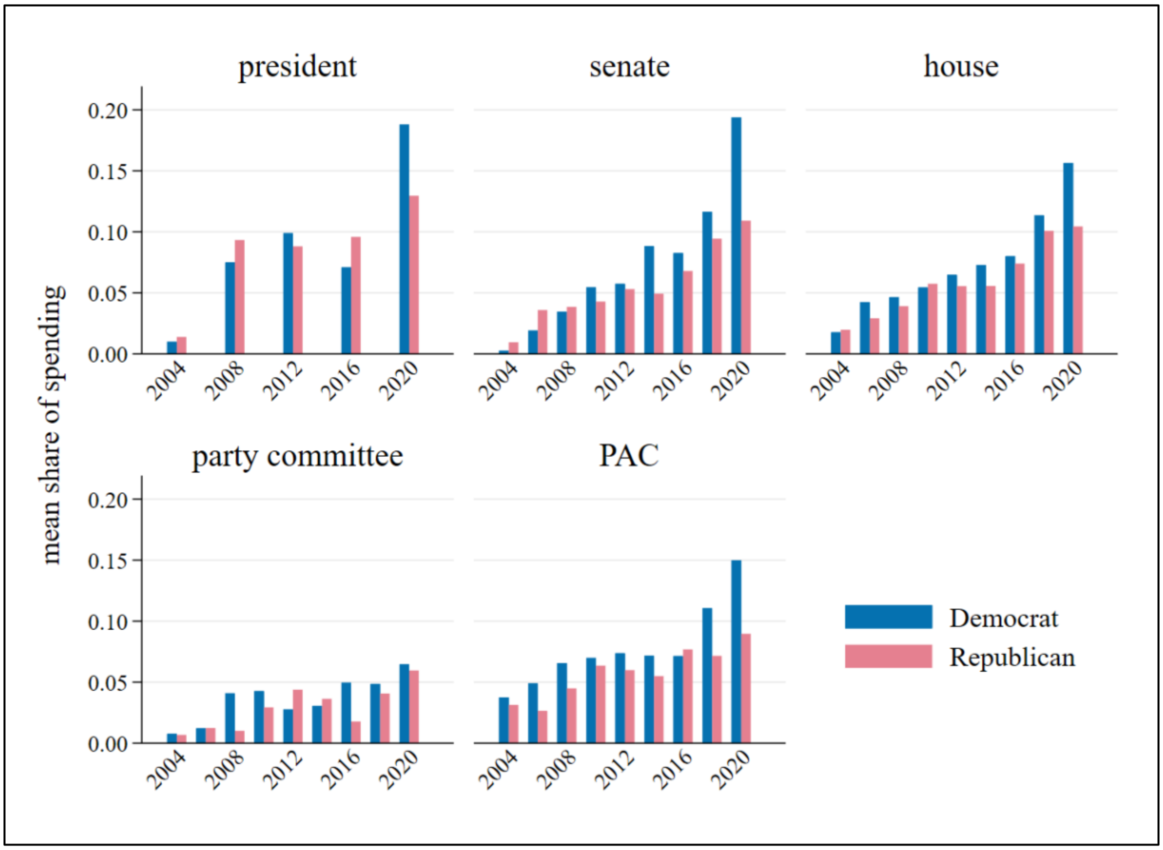

Figure 8: Spending allocated to digital ads and services by committee type
Note: operating expenditures only

Because presidential candidates use a variety of spending vehicles to execute an online strategy, one should look beyond the official campaign committee to get an accurate picture of which party enjoys a digital advantage (Williams and Gulati 2018; Ridout et al. 2021). Figure 9 shows this by comparing different sources of digital spending on behalf of the major party presidential nominees since 2004.[23] Looking at the 2016 race, for example, the Clinton campaign used a joint fundraising committee to buy most of its digital ads. By

[23] Figure 9 includes spending by each nominee's campaign committee, ten joint fundraising committees (McCain-Palin Victory 2008, Obama Victory Fund, Romney Victory, Hillary Victory Fund, Trump Victory, Trump Make America Great Again, and Biden Victory Fund) and independent expenditures by outside groups.

contrast, most of Trump's digital spending went through his main campaign committee. In 2020, the pattern was reversed: Biden's digital spending came from his campaign whereas President Trump relied on a joint fundraising committee. Including outside groups, Democrats outspent Republicans in both 2016 and 2020.[24] Figure 9 also shows that digital spending on behalf of major party presidential nominees increased markedly in 2020.

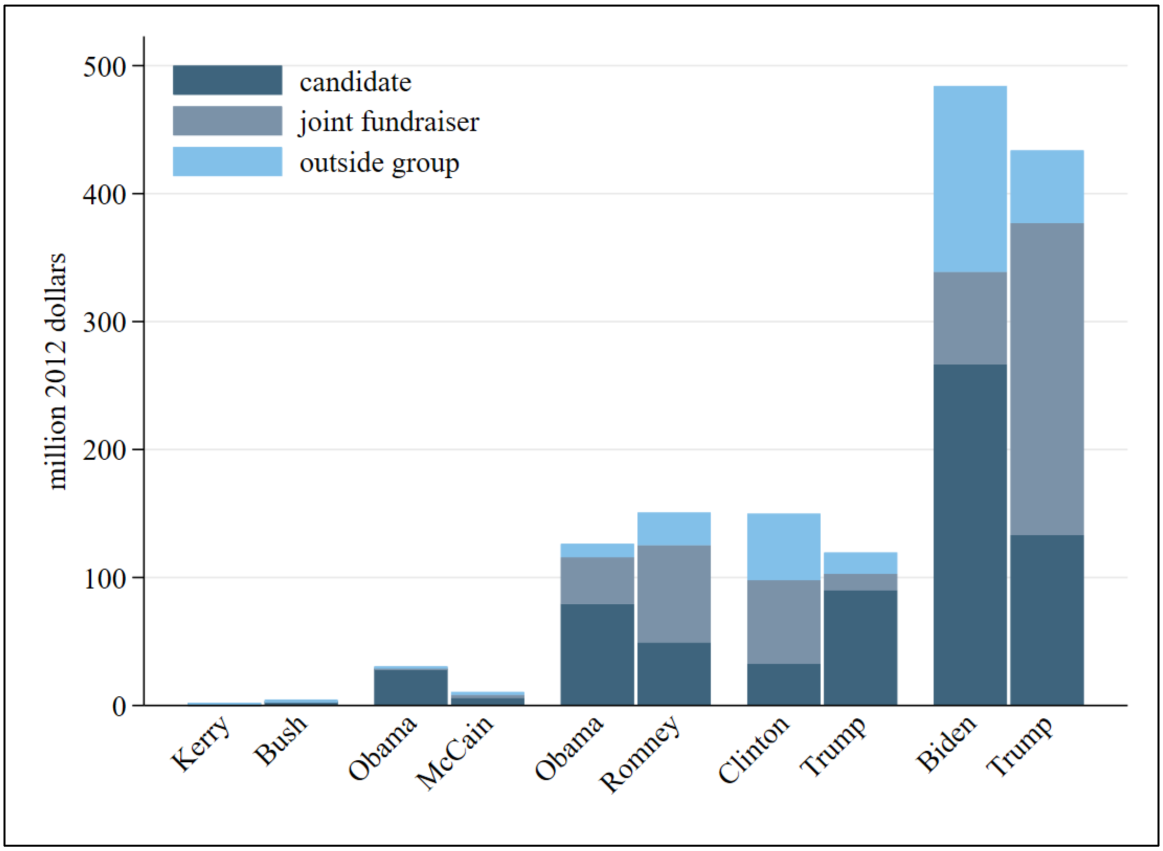

Figure 9: Digital spending on behalf of major party presidential nominees

Another issue to consider when comparing digital spending by presidential campaigns is that platform ad libraries only capture a portion of the resources allocated to

[24] Outside groups include independent expenditures reported to the FEC as supporting/opposing Biden or Trump.

online advertising. Although most online political spending now takes place on Facebook and Google, digital ad expenditures also include payments to other platforms, SMS (text) advertising, and the commissions and fees consultants charge for producing and placing online ads. Table 3 compares totals for digital advertising reported to the FEC with spending figures derived from the Facebook and Google ad libraries. Excluding outside groups, the Biden campaign reported $366 million in digital advertising to the FEC. This includes $81 million in direct payments to Facebook and Google and $286 million in digital ads executed through consulting firms.[25] The Trump campaign reported $334 million in digital advertising to the FEC, all of it paid to a single company: American Made Media Consultants. Based on their analysis of data from the Facebook and Google ad libraries, Ridout et al. (2021) estimate $213 and $276 million in digital ad spending by the Biden and Trump campaigns, respectively. This implies an additional $212 million in spending, or 30% of the total digital advertising budget, allocated to SMS advertising, other platforms, and consultant markups and fees. However, as we examine in the next section, the influence of Facebook and Google in modern campaigns reaches beyond the dollar value of ads appearing on their platforms.

Table 3: Digital advertising payments by committee and source

	Trump	Biden
Campaign committee*	143.7	284.3
Joint fundraising committee†	190.6	82.1
Total reported to FEC (A)	334.3	366.4
Total in Facebook and Google Ad libraries‡ (B)	276.0	212.6
Other Digital Spending§ (A-B)	58.3	153.8

*Donald J Trump for President; Biden for President; †Trump Make America Great Again; Biden Victory Fund; ‡As reported in Ridout et al. 2021, p. 476; §Includes SMS advertising, other platforms, and consultant markups and fees

[25] Of the spending executed through consultants, three firms make up 98% of Biden's digital spending: GMMB, Bully Pulpit, and InfoGroup. See Table A7 for details.

The Network of Digital Consultants

RQ3 asks how platform companies like Facebook and Google fit in the partisan market for digital ads and services. We build on previous research by using a network analysis to visualize this market as a set of relationships between consulting firms with shared clients (Nyhan and Montgomery 2015; Barrett 2021). Figure 10 displays undirected networks between the top fifty firms providing digital ads to Democratic and Republican campaigns, respectively, during the last four presidential election cycles.[26] Each node in the network is a consulting firm or platform; each line or edge is a connection between firms that provide digital ads to the same candidate or committee. Blue nodes (circles) are firms exclusively serving Democratic clients, red nodes (squares) are firms serving Republicans, and green nodes (triangles) are firms that serve clients in both parties.[27] Larger nodes indicate firms with higher total revenues for digital ads and services. Labels indicate the positions of Facebook and Google in each network.

[26] In 2008, there were 42 firms providing digital ads to Republican committees.
[27] Figure 10 excludes firms with no clients in common.

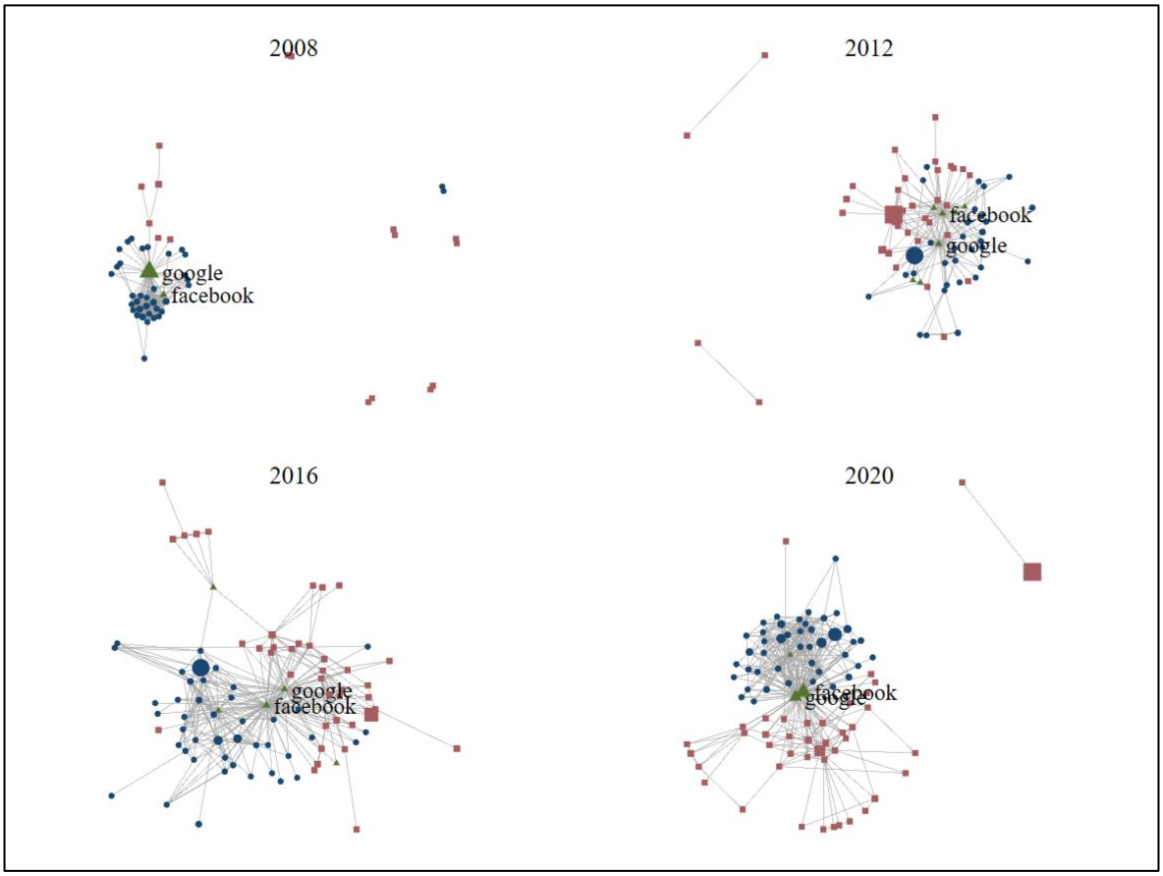

Figure 10: Network visualization of top fifty digital firms by party

Figure 10 shows how the market for digital ads and services has expanded over time. In 2008, there was a small but organized network of firms serving Democratic clients, with far fewer Republican firms by comparison. In 2012, two partisan clusters begin to emerge. The largest nodes in the network are two firms, Bully Pulpit and Targeted Victory, that handled digital strategy for the Obama and Romney campaigns, respectively. By 2016, Facebook and Google appear more centrally positioned in the network. By 2020, Democratic firms appear as a dense network of connections compared to the more widely dispersed array of Republican firms.[28] Nodes for Facebook and Google are prominent at

[28] Table A11 includes separate network measures for Democratic and Republican firms.

the center of the network. The pair of nodes set apart from the central 2020 network represent American Made Media Consultants and Parscale Strategies, the two firms that handled digital strategy for the Trump reelection campaign. The absence of any connections to other Republican firms may reflect Donald Trump’s outsider status in the party, even after his election as President. Research on “never Trumpers” found that political consultants are among Trump’s most vocal critics within the Republican establishment (Saldin and Teles 2020).

Table 4 points further to the developing market for digital campaign services using several network measures. Edges are the number of connections between firms. Density measures the proportion of possible connections. Increasing numbers of edges and higher density scores indicate a more complex network. Table 4 also includes three measures of network centrality: degree, betweenness, and closeness. These are, respectively, the count of ties for each node, the frequency a node is the shortest path connecting all other pairs, and the proximity of each node to all other nodes. Each measure tries to capture a different kind of network influence (Freeman 1978). Nodes with high degree centrality are influential by virtue of their many connections. Nodes with high betweenness serve as a bridge or broker between parts of the network. Nodes with high closeness are important because their actions diffuse rapidly through a network. Table 4 reports measures of centralization, or how unequally distributed centrality measures are in each network (Freeman 1978). Centralization equals 1 when a single node has the maximum possible score and equals zero when all nodes have the same centrality score. Higher centralization means a small number of firms are focal points in the network.

Table 4: Network measures by cycle

	2008	2012	2016	2020
edges	260	184	326	409
density	0.068	0.043	0.071	0.086
degree	0.366	0.445	0.658	0.628
betweenness	0.126	0.273	0.424	0.365
closeness	0.133	0.245	0.408	0.455

Measures of centrality also show the prominence of Facebook and Google in the network for digital advertising. In 2008 and 2012, Google ranked first in all three centrality measures. In 2016 and 2020, Facebook ranked highest, followed by Google. Figure 11 represents these trends graphically.

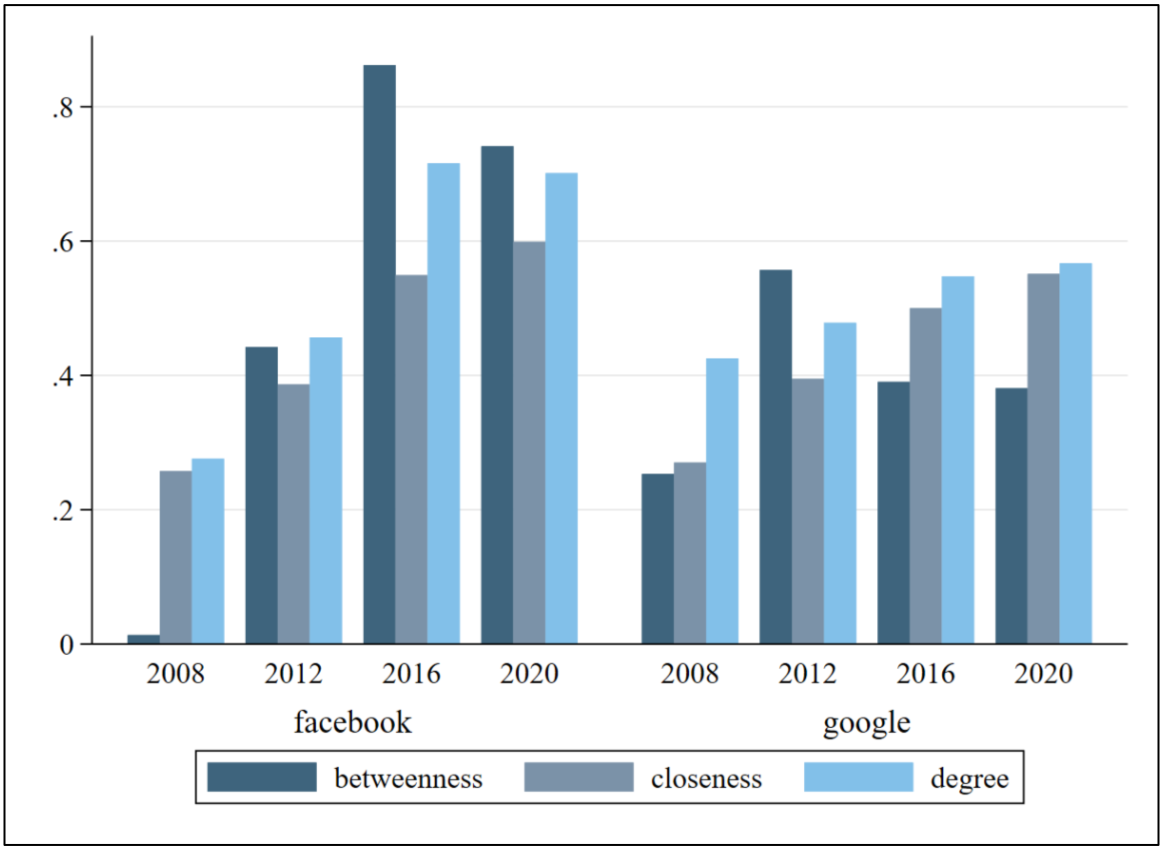

Figure 11: Network centrality of Facebook and Google, 2008-2020

The centrality of Facebook and Google raises interesting questions about how two ostensibly nonpartisan platform companies influence the conduct of partisan political campaigns (Barrett 2021). Although a full answer is beyond the scope of this article, network analysis offers insight into Facebook's growing influence in U.S. elections. Over the course of the 2020 cycle, for example, Facebook instituted new rules for buyers of political ads[29], announced it would exempt political ads from fact-checking[30], and instituted a ban on new political ads one week before Election Day.[31] Measures of network centrality help to illuminate the reach and significance of these actions. Because of Facebook's high degree centrality, the decision to ban political ads had widespread consequences for the market. Facebook's high betweenness centrality means the company operates as a kind of gatekeeper, influencing practices regarding misinformation among firms from both parties. Finally, high closeness centrality means that Facebook's verification requirements for ad purchases diffused quickly through the market.

Characteristics of digital firms

RQ4 asks whether firms providing digital ads and services differ from other types of political professionals. In their study of media consultants, Martin and Peskowitz (2018) argued that ideological similarity between consultants and their clients aligns the financial interests of the former with the electoral interests of the latter. Candidates seek out like-minded consultants to make sure these professionals are motivated by politics and not just by profit. Aligning the political and pecuniary incentives of consultants creates distinct partisan markets for professional advice. The network analysis showed this relationship extends to digital consultants as well, most of whom serve clients from one party or another—platform companies like Facebook and Google being important exceptions.

The ideological alignment between candidates and consultants extends beyond party labels alone. Using Bonica's (2014) common space estimates of ideology derived

---

[29] "Facebook Tightens Rules on Verifying Political Advertisers." *The New York Times*, August 28, 2019. https://www.nytimes.com/2019/08/28/technology/facebook-election-advertising-disinformation.html?smid=url-share. Accessed February 12, 2022.

[30] "Facebook Says It Won't Back Down from Allowing Lies in Political Ads," *The New York Times*, January 9, 2020. https://www.nytimes.com/2020/01/09/technology/facebook-political-ads-lies.html. Accessed February 12, 2022.

[31] "Facebook Widens Ban on Political Ads as Alarm Rises Over Election." *The New York Times*, October 7, 2020. https://www.nytimes.com/2020/10/07/technology/facebook-political-ads-ban.html. Accessed February 12, 2022.

from campaign contributions (CF scores), Martin and Peskowitz find that ideological proximity is a strong predictor of a consultant working for a specific candidate from the same party. Their finding suggests that consulting firms occupy distinct ideological niches within each party rather than draw clients uniformly from a set of Democratic or Republican candidates.

Building on this insight, figure 12 compares the mean CF scores of candidates and committees served by digital consultants compared to specialists in media, fundraising, general consulting, and polling.[32] As expected, the distribution of average client ideology is bimodal for all specialties, reflecting the partisan organization of professional advice. In addition, Figure 12 shows differences in client ideology across specialties. The average CF scores of clients served by digital firms are between one quarter and one third of a standard deviation to the left and right of other specialties.[33] Consistent with the idea that consultants occupy a distinct ideological niche, firms specializing in digital advertising serve candidates and committees who are on average more ideologically extreme than the rest of their party.

[32] Firm specialization is defined as more than 50% of spending classified in a professional category. Firms with less than $50k in payments for professional services are excluded.
[33] See Table A11 for descriptive statistics and Table A12 for results of one-way analysis of variance.

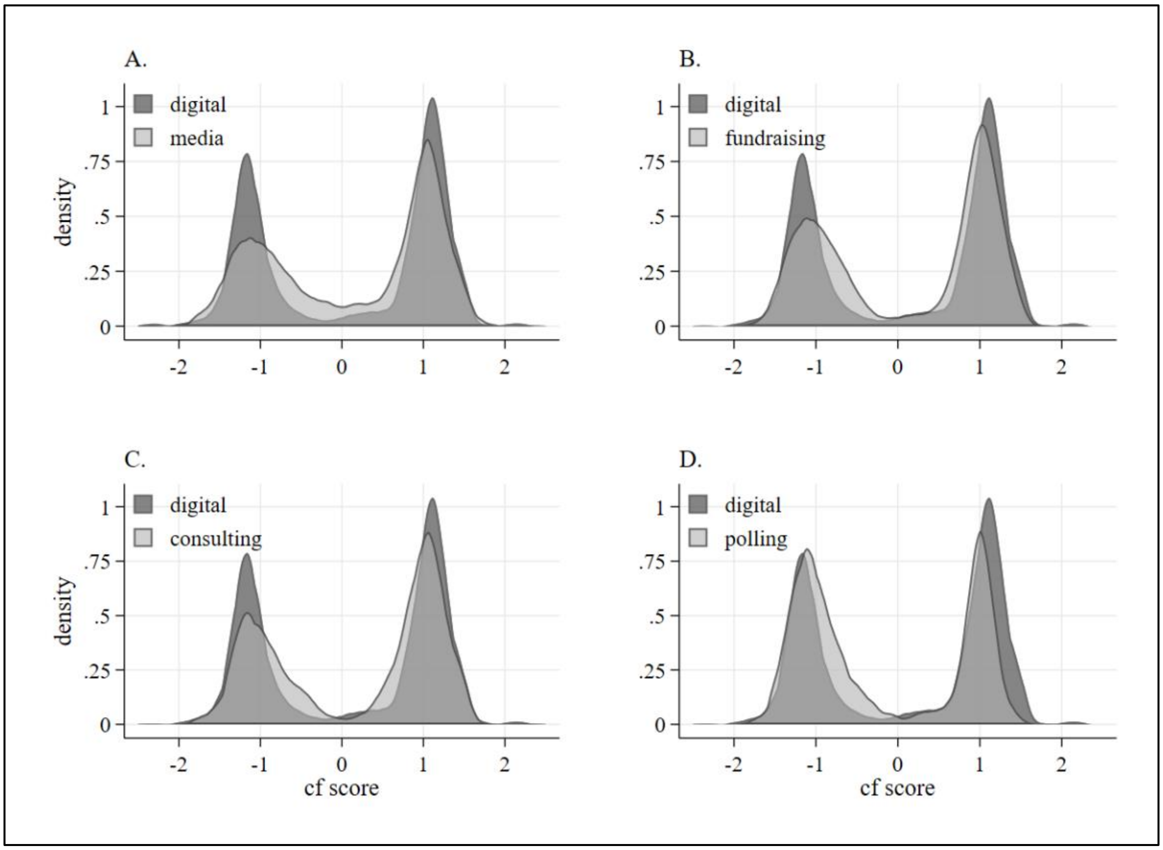

Figure 12: Distribution of client ideology by firm specialty
Note: Firms with at least $50k in payments for professional services

Research on online campaigning suggests at least two explanations for the pattern observed in figure 12 (Fowler et al. 2021). The lower unit cost of Facebook compared to traditional media makes advertising online more attractive to less well-funded candidates. If candidates at the ideological extremes are also under-resourced, perhaps because they lack connections to the partisan network of consultants and fundraisers, they may be more likely to utilize digital advertising than more mainstream candidates. This same research also finds that Facebook increases the ideological extremism and variability of advertisements relative to the same candidates' messages on television. This suggests that ideologically extreme clients might seek out assistance from digital firms to better reach voters attracted to the polarized and partisan political discourse they find online.

Discussion

This paper explored the evolution of spending on digital ads and services in U.S. federal elections. Using a unique dataset derived from FEC expenditure reports, the paper makes four key observations. First, there has been a profound growth in resources allocated to digital campaigning with the 2020 election being a watershed cycle in this development (RQ1). Although more money still goes to advertising on television, the spending gap between traditional and digital media is narrowing and many candidates spend more on digital media than radio or television compared to previous cycles. The extraordinary growth of digital advertising points to a pronounced change in how campaigns allocate resources in federal elections (Limbocker and You 2020).

Second, there are persistent differences between the two major parties when it comes to digital campaigns (RQ2). Democrats allocate a larger share of spending on digital ads and services than Republicans, a pattern that may reflect differences in the early adoption of digital tools (Kreiss 2012, 2016). This partisan gap appears to be growing, although evidence is mixed at the presidential level and depends on whether spending takes place through candidate committees or joint fundraising committees. We suspect this distinction is of little strategic importance because campaigns often use the same consultants regardless of which fundraising vehicle they choose to execute their digital strategy.

Third, platform companies play a central role in the market for digital ads and services by virtue of their service to candidates and committees in both parties (RQ3). Measures of network centrality point to the potential power of platform companies like Facebook to influence the behavior of other digital firms, create rules that diffuse quickly across the political spectrum, and serve as a gatekeeper for the use and practice of digital advertising. Our network analysis also shows a clear partisan difference in the market for digital ads. Firms specializing in digital political communication emerged earlier among Democratic firms and the network is characterized by a denser connection of shared clients compared to firms serving Republican clients.

Fourth, firms specializing in digital ads and services occupy a more extreme ideological niche within the parties compared to other types of consulting firms (RQ4). This finding is consistent with previous work on the ideological alignment between candidates and consultants (Martin and Peskowitz 2018) and fits with research that candidate advertising is more ideologically extreme on Facebook than on television (Fowler et al. 2021).

The paper is limited in several ways. As a descriptive analysis, the paper does not consider what caused the growth of digital spending or examine the consequences of digital spending for the electoral competitiveness of candidates. Nevertheless, descriptive inference plays an essential role in social scientific explanation and our data provide a valuable starting point for future research on the role of digital media in modern campaigns. Understanding whether candidates substitute lower cost digital advertising for traditional media or if more extreme candidates seek out digital consultants because of political affinity, affordability, or a mixture of the two requires basic data regarding how much candidates spend on digital advertising versus other types of media. We hope our work, along with the underlying data on which it is based, will enable scholars to explore the causes and consequences of the spending patterns we identify.

Another limitation of the study is the data itself. Although we have high confidence in the accuracy of our model, the use of ambiguous and unstandardized descriptions means that payments to vendors reported simply as “advertising” could be for digital or traditional media. As shown in figure 4, estimates of digital spending using the share of platform expenditures classified as media does result in significantly higher levels for 2016 than we derived using our classification model alone. Although the revised estimates do not change the overall pattern of digital spending, scholars should keep this limitation in mind when interpreting data derived from FEC reports.

A second data issue is the lack of transparency in campaign finance reporting. Although we did not find evidence of unreported media spending when comparing FEC expenditure reports with tracking data collected by the Wesleyan Media Project, transparency concerns extend beyond the loopholes exploited by dark money groups. A complaint filed with the FEC by the Campaign Legal Center alleges the Trump campaign "disguised nearly $170 million in spending" by using American Made Media Consultants and Parscale Strategies to hide "payments to Trump family members or senior campaign staff" (Campaign Legal Center 2020, 2). The case points to a broader limitation of data derived from FEC reports. Current law only requires that political entities disclose who they paid and for what purpose; it does not require the recipients of these payments to disclose how they spend the money. This raises the possibility that payments to consultants and other vendors could be used for purposes other than what is reported to the FEC and in ways that contravene campaign finance rules (Babwah Brennen and Perault 2021).

Conclusion

Elections in the United States are a multi-billion-dollar business. Increasingly, this business takes place online. Our study provides the most comprehensive picture to date of the extraordinary growth in spending on digital ads and services in U.S. federal elections. Our method of classifying expenditures provides insights that, until now, had been unavailable to scholars because of the challenges associated with cleaning and coding millions of FEC records. Our findings have implications for scholarship on how campaigns allocate resources, whether candidates advertise online or offline, and the growing prominence of platform companies in the conduct of U.S. elections.

Our study suggests three lines of future research. First, campaign spending data points to the enduring influence of political professionals in U.S. elections (Sheingate 2016). Despite mixed evidence regarding the persuasive effects of advertising (Sides et al. 2022; Kalla and Broockman 2018), media and digital spending together account for almost 60% of all political expenditures—more than $8.4 billion in 2020. By comparison, campaigns spend comparatively little on direct voter contact such as canvassing, phone, or mail ($312 million in 2020) despite evidence of its effects on turnout (Enos and Fowler 2018). One possible reason why campaigns continue to allocate the lion's share of resources to advertising is because consultants earn commissions and fees for the production and placement of ads.[34] The extraordinary growth in online advertising suggests digital campaigning has created new sources of revenue for an industry that continues to play a crucial mediating role in political communication. Future research combining our dataset with platform ad libraries could explore consultant markups in digital advertising more systematically.

Second, Facebook and Google play an unusual role in the market for digital advertising because they serve both parties in what is usually a partisan network of professional advice (Barrett 2021). In this role, platform companies do more than broadcast ads to users: they provide strategy and data analytics to campaigns (Kreiss and McGregor 2018, 2019). Although Facebook ended its policy of providing on-site support staff, anecdotal evidence suggests the company still provides specialized consulting services to political campaigns.[35] The ability of platform companies to serve both parties is an important departure from the structure of the political consulting industry with firms aligned to the partisan affiliation of their clients. Our data suggests that an increasing number of candidates and committees pay platform companies directly for their services and future work should explore whether this reduces the reliance on traditional consultants.

Third, our method of classifying expenditures makes it easier to study the ongoing development of digital campaigning in the United States and abroad. One advantage of our model over keyword searches is that it can be more easily replicated and applied to future election cycles, enabling scholars to explore whether online political advertising continues its remarkable growth or if the political parties converge in their patterns of spending. The longitudinal nature of our data can also shed light on whether or when campaigns substitute digital advertising for television. Our study also illustrates the advantages of natural language processing to classify campaign expenditures, opening possibilities for

[34] Because field campaigns rely extensively on volunteers, they provide far fewer opportunities for consultants to generate income (Nielson 2012).
[35] "Facebook, Google Still Offering 'Embed'-Like Consulting to Campaigns, Tech Transparency Project," https://www.techtransparencyproject.org/articles/facebook-google-still-offering-embed-like-consulting-campaigns, accessed March 16, 2022.

comparative work in countries such as the United Kingdom or Brazil where public databases of political spending are available (Dommett et al. 2022).

Finally, the limitations of publicly available data raise normative considerations about transparency in U.S. elections. Nearly half a century ago, Congress created a system of campaign finance rules that requires candidates and committees to submit detailed spending reports to the FEC. These rules offer us important protections against political slush funds, vote buying, and other practices that undermine the integrity of elections. Changing technologies and practices, including the rise of digital media, underscore the need for updates and refinements to an outdated set of reporting requirements. Standardized descriptions of expenditures that accurately reflect the reality of modern campaigns would make it easier to track political spending, but public databases are meaningless if campaigns can conceal their spending through shell companies or create political entities that exploit loopholes in the law. Without improvements and enhancements in our ability to follow the money, we have a great deal to lose—as scholars and citizens alike.

References

Babwah Brennan, J.S., and Perault, M. (2021). *Breaking blackout black boxes: Roadblocks to analyzing platform political ad bans*. Duke Center on Science & Technology Policy.

Baldwin-Philippi, J. (2015). *Using technology, building democracy: Digital campaigning and the construction of citizenship*. Oxford University Press.

Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, 34(4), 627-633. https://doi.org/10.1080/10584609.2017.1372999

Barrett, B. (2021). Commercial Companies in Party Networks: Digital Advertising Firms in US Elections from 2006-2016. *Political Communication*, 39(2), 147-165. https://doi.org/10.1080/10584609.2021.1978021

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2), 367-386. https://doi.org/10.1111/ajps.12062

Dommett, K., Power, S., MacIntyre, A., and Barclay, A. (2022). *Regulating the business of election campaigns: Financial transparency in the influence ecosystem in the United Kingdom*. International Institute for Democracy and Electoral Assistance.

Enos, R.D. and Fowler, A. (2018). Aggregate effects of large-scale campaigns on voter turnout. *Political Science Research and Methods*, 6(4), 733-751. https://doi.org/10.1017/psrm.2016.21

Fowler, E.F., Franz, M. and Ridout, T.N. (2021). *Political advertising in the United States*. Routledge.

Fowler, E.F., Franz, M.M. and Ridout, T.N. (2020a). Online political advertising in the United States. *Social Media and Democracy: The State of the Field, Prospects for Reform*, pp.111-138.

Fowler, E.F., Franz, M.M. and Ridout, T.N., and Baum, L.M. (2020b). "Political Advertising in 2018." Version 1.0 [dataset]. Middletown, CT: The Wesleyan Media Project, Department of Government at Wesleyan University.

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. https://doi.org/10.1017/S0003055420000696

Freeman, L.C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3), 215-239.

Gulati, J., and Williams, C.B. (2021). Digital Media Expenditures in Presidential Campaigns, 2008–2020. *The Internet and the 2020 Campaign*.

Hersh, E.D. (2015). *Hacking the electorate: How campaigns perceive voters*. Cambridge University Press.

Kalla, J.L. and Broockman, D.E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review*, 112(1), 148-166. https://doi.org/10.1017/S0003055417000363

Koger, G., Masket, S. and Noel, H. (2009). Partisan webs: Information exchange and party networks. *British Journal of Political Science*, 39(3), 633-653. https://doi.org/10.1017/S0007123409000659

Kolodny, R. and Logan, A. (1998). Political consultants and the extension of party goals. *PS: Political Science & Politics*, 31(2), 155-159.

Kreiss, D. and McGregor, S.C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 US presidential cycle. *Political Communication*, 35(2), 155-177. https://doi.org/10.1080/10584609.2017.1364814

Kreiss, D. and McGregor, S.C. (2019). The "arbiters of what our voters see": Facebook and Google's struggle with policy, process, and enforcement around political advertising. *Political Communication*, 36(4), 499-522. https://doi.org/10.1080/10584609.2019.1619639

Kreiss, D. and Saffer, A.J. (2017). Networks and innovation in the production of communication: Explaining innovations in US electoral campaigning from 2004 to 2012. *Journal of Communication*, 67(4), 521-544. https://doi.org/10.1111/jcom.12302

Kreiss, D. (2012). *Taking our country back: The crafting of networked politics from Howard Dean to Barack Obama*. Oxford University Press.

Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. Oxford University Press.

Limbocker, S. and You, H.Y. (2020). Campaign styles: Persistency in campaign resource allocation. *Electoral Studies*, 65, 102-140. https://doi.org/10.1016/j.electstud.2020.102140

Martin, G.J. and Peskowitz, Z. (2015). Parties and Electoral Performance in the Market for Political Consultants. *Legislative Studies Quarterly*, 40(3), 441-470. https://doi.org/10.1111/lsq.12083

Martin, G.J. and Peskowitz, Z. (2018). Agency Problems in Political Campaigns: Media Buying and Consulting. *American Political Science Review*, 112(2), 231-248. https://doi.org/10.1017/S0003055417000594

Massoglia, A. (2021). Liberal "dark money" groups' revenue soared ahead of 2020 elections. https://www.opensecrets.org/news/2021/12/liberal-dark-money-groups-revenue-soared-ahead-of-2020-elections. Accessed February 17, 2022.

Nielsen, R.K. (2012). *Ground Wars: Personalized Communication in Political Campaigns*. Princeton University Press.

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. https://doi.org/10.1111/ajps.12143

Ridout, T.N., Fowler, E.F. and Franz, M.M. (2021). Spending fast and furious: Political advertising in 2020. In *The Forum* 18(4), 465-492. https://doi.org/10.1515/for-2020-2109

Saldin, R.P. and Teles, S.M. (2020). *Never trump: The revolt of the conservative elites*. Oxford University Press.

Sheingate, A.D. (2016). *Building a business of politics: The rise of political consulting and the transformation of American democracy*. Oxford University Press.

Sides, J., Vavreck, L. and Warshaw, C. (2022). The effect of television advertising in United States elections. *American Political Science Review*, 116(2), 702-718. https://doi.org/10.1017/S000305542100112X

Williams, C.B. and Gulati, G.J. (2018). Digital advertising expenditures in the 2016 presidential election. *Social Science Computer Review*, 36(4), 406-421. https://doi.org/10.1177/0894439317726751

Williams, C.B., Gulati, J. and Zeglen, M. (2020). Following the money: Uses and limitations of FEC campaign finance data. *Interest Groups & Advocacy*, 9(3), 317-329. https://doi.org/10.1057/s41309-020-00096-8

Appendix: Digital Advertising in U.S. Federal Election, 2004-2020[1]

Our dataset is based on all operating and independent expenditures for each of the nine election cycles between 2004 and 2020, more 13.7 million records in total.[2] Each record includes a text description of the purpose of the expenditure as reported to the FEC. For transactions with missing purpose descriptions, we use and clean the text included in the memo field or category description.[3] In the case of transactions where the purpose is listed as in-kind expenditures or reimbursement, we use the memo field or category description if available. Expenditures with no usable description are coded as "other".

To train and test the dataset, we developed a coding scheme based on categories used by OpenSecrets, with some important differences.[4] Following work by Limbacker and You (2020), we create separate categories for polling and consulting. In addition, we distinguish digital expenditures from other forms of media. If a description fit in multiple categories, we used the following tiebreaker: Consulting > Digital > Media > Legal > Polling > Fundraising > Field > Travel > Administrative. For example, we classified an expenditure for "digital consulting" as "consulting".[5]

A major challenge working with FEC data is that campaigns and committees use a variety of terms to describe expenditures that fall within the same category of spending. Our model addresses this issue in two ways. First, we select an initial set of keywords based on our definition of each category and add additional terms that appear frequently in our training and testing data. A full list of keywords is in table A1. Second, we use the

---

[1] For replication files, see Sheingate, Scharf, and Delahanty (2022).
[2] For a description of the operating expenditure data, see https://www.fec.gov/campaign-finance-data/operating-expenditures-file-description/. For independent expenditures see https://www.fec.gov/campaign-finance-data/independent-expenditures-file-description/.
[3] For a list of FEC category descriptions, see https://www.fec.gov/campaign-finance-data/disbursement-category-code-descriptions/.).
[4] For OpenSecrets categories, see https://www.opensecrets.org/campaign-expenditures/methodology
[5] As described in the text, we recode expenditures for digital consulting as digital and distinguish it from spending on digital ads and services.

Datamuse API, a word-finding query engine, to identify synonyms for our keywords in each category.[6]

Table A1: Keywords used for query expansion

Media (34)	
	ad, ad air, ad buy, ad placement, ad production, advert, advertising, advertising buy, advertising tv, air, buy, and media, buy, communications, media, media buy, media placement, media production, media supporting, news, newspaper, placement, production, production costs, radio, radio ad, television, television advertising, television production, tv, tv ad, tv ads, tv advertisement, tv advertising, art,
Digital (28)	
	advertising – internet, advertising – online, adwords, and digital, data, data software, digital, digital ad, digital advertising, digital media, email, facebook, google, google ad words, internet, internet advertising, internet communications, microtargeting, online, online advertising, social, social media, software, twitter, web, web ad, web advertising, website
Polling (6)	
	opinion, polling expenses, research, survey, surveys, polling
Legal (11)	
	accounting, attorney, audit, compliance, election law, law, lawyer, lawyers, legal fees, professional services, legal
Field (31)	
	ballot, ballots, bank polling, banking telephone, banks polling, calls, canvass, canvassing, canvassing services, canvass-related, contact, contact calls, contact mail, contact phone, contact phones, door hangers,

[6] For documentation, see http://www.datamuse.com/api/. We only selected keywords for eight of our categories, excluding travel. This increases the likelihood of correct assignment.

Table A1: Keywords used for query expansion

	door to, field, gotv, gotv phone calls, list rental, outbound voter, petition, phone bank, phone banking, phone banks, phone calls, polling phone, signature, telephone voter, voter contact
Consulting (19)	
	advising, advisor, analyst, analytics, campaign consulting, consult, consultant, consultants, consultation, consulting, consulting, counsel, counseling, purchase consulting, retainer, strategic, strategic, strategy, win bonus
Fundraising (24)	
	act blue, actblue, actblue processing fee, caging, contribution, direct mail, dissemination, donations, donor, fundraiser, fundraising, mail costs, mail design, mail direct, mail list, mail production, mailer, services direct, solicitation, telemarket, telemarketing, telemarketing, winred, winred processing fees
Administrative (32)	
	admin, air, airline, beverage, cater, collateral, costs, director, event, food, gas, insurance, management, manager, paper, payroll, phone, postage, postcards, printing, refreshments, rental, salary, shipping, staff, staffing, supplies, taxes, telecom, travel, wage, administrative

Our classification model employs a traditional natural language processing approach.[7] First, the model splits each text description into individual words (tokenization) and removes stopwords such as ‘a’ or ‘the’. Next, the model transforms the remaining text into a vector by combining the term frequency and inverse document frequency of all terms into a composite weight for each observation in the training dataset (Manning et al. 2008).[8]

[7] Code, training, and testing data is available at https://github.com/sheingate/Campaign_Classifier

[8] Term frequency is the occurrence of term *t* in document *d*. Document frequency is the number of documents *d* that contain term *t*.

The composite weight for term *t* in document *d* where N is the total number of documents (in our case, purpose descriptions) is calculated by the formula,

$$\mathrm{tf}-\mathrm{idf}_{t,d} = \mathrm{tf}_{t,d} \ \mathrm{x} \ \log\frac{N}{d\mathrm{f}_t}$$

We slightly modify the tf-idf weighting by assigning greater value to keywords and their synonyms (and lower value to frequently occurring words that are neither keywords nor synonyms).[9] We assign additional weights to keywords and synonyms by repeating these terms several times when they appear in a purpose description.

We train and test our model using the scikit-learn SGDClassifier, a linear classifier that allows classification into multiple categories by combining several binary classifiers in a "one versus all" scheme (Pedregosa et al. 2011).[10] The SGDClassifier employs stochastic gradient descent to optimize the fit of a linear classifier such as a support vector machine. The classifier relies on the following packages: NLTK (Bird et al., 2009); NumPy (Van der Walt et al., 2011); Pandas (McKinney 2010); and Python 3+ (Oliphant 2007; Millman and Aivazis 2011). Because the purpose descriptions contain spelling errors, we are unable to use more advanced language processing techniques to train and test our model such as BERT or word2vec. For training, a separate model for each category assigns a positive value for descriptions predicted to belong in that category and a negative value otherwise (e.g. media/not media; digital/not digital). For testing, the classifier calculates a confidence score for each binary model and assigns the description to the category with the highest score.

To get a better picture of the accuracy of our model, we ran the classifier on the entire set of unique descriptions (N=620,292) and then examined a subset that together sum up to 95% of the total spending in each of the nine categories. Table A2 reports our results. As indicated, the model is 96% accurate classifying the top expenditures in each category.

[9] For predefined keywords, the term frequency is multiplied by 25. For synonyms, term frequency is multiplied by 5.
[10] For more detail, see https://scikit-learn.org/stable/modules/sgd.html.

Table A2 also indicates that expenditures are highly concentrated in a relatively small number of unique descriptions. According to the support column in table A2, 3.25 percent of observations, or just over 20,000 unique descriptions, account for 95% of all spending between 2004 and 2020. Given the accuracy of our model, coupled with our ability to check and correct the top expenditures in each category, we have very high confidence that our data can capture trends in spending on digital advertising and other types of campaign expenditures.

Table A2: Accuracy of Classifier in top 95% of spending categories

	precision	recall	f1-score	support
administrative	0.96	0.96	0.96	8,462
travel	0.99	0.98	0.98	2,409
fundraising	0.95	0.95	0.95	2,305
consulting	0.99	0.96	0.98	2,141
field	0.91	0.96	0.94	2,026
digital	0.95	0.98	0.96	1,519
media	0.94	0.94	0.94	635
legal	1.00	0.90	0.95	433
polling	0.96	0.94	0.95	214
macro avg	0.96	0.95	0.96	20,144
weighted avg	0.96	0.96	0.96	20,144
accuracy			0.96	20,144

Preparing and cleaning the dataset

We carry out additional preparation and cleaning of the dataset prior to analysis. Using a unique identifier for each committee registered with the FEC, we merge the expenditure data with the committee and candidate master files for each cycle between 2004 and 2020. The committee files include a unique candidate identifier for the principal campaign committees of candidates for federal office (House, Senate, and President), as well as the name of the committee, the committee designation under FEC rules (e.g. joint fundraiser or leadership PAC), and a code for whether it is associated with a candidate, PAC, or political party. We use the candidate identifier to merge the data with the FEC candidate master files for each cycle. The candidate files include the candidate's name, party affiliation, district and state, office sought, and candidate status (incumbent, challenger, or open seat).

We clean each vendor name to correct for small variations in spelling or naming conventions.[11] We then create a party affiliation for each candidate and committee using several sources. The committee and candidate master files assign a party label to 40% of committees in the dataset. To fill in missing information, we merge the data with the committee bulk files from OpenSecrets for the 2004-2020 cycles. The OpenSecrets data provides the missing party affiliation for 14.5% of the committees in our data. Finally, we impute the party affiliation from FEC data on independent expenditures using the amount spent by committees for and against candidates from each of the two major parties. This provides party affiliation for an additional 2.8% of the committees. Overall, 48.7% of the committees in the dataset are affiliated with a major party, 6.8% with a third party, and 44.5% do not have a party affiliation. Table A3 explains the procedure for imputing party.

Table A3: Imputation of Party Affiliation from Independent Expenditures

Step 1: Calculate a mean party support score for committees based on the following measures:

Party Support 1: Equals 1 if Democratic share of total expenditures in support of all candidates is greater than .5. Equals 2 if Republican share of support is greater than .5.

Party Support 2: Equals 1 if 10% of independent expenditures supports Democrats and 0 support for Republicans. Equals 2 if 10% of independent expenditures supports Republicans and 0 support for Democrats.

[11] We use the OpenRefine application to cluster and clean vendor names using key collision and nearest neighbor algorithms. See https://github.com/OpenRefine/OpenRefine/wiki

Table A3: Imputation of Party Affiliation from Independent Expenditures

Party Support 3: Equals 1 if committees spend any money supporting Democrats and no money either supporting or opposing Republicans. Equals 2 if committees spend any money supporting Republicans and no money either supporting or opposing Democrats
Party Support 4: Equals 1 if committees spend more in support of Democrats than in support of Republicans and spend more supporting Democrats than opposing Democrats. Equals 2 if committees spend more in support of Republicans than in support of Democrats and spend more supporting Republicans than opposing Republicans
Step 2: Calculate a mean party opposition score using the following measures
Party Oppose 1: Equals 1 if committees spend 100% of independent expenditures opposing Republicans. Equals 2 if committees spend 100% of independent expenditures opposing Democrats.
Party Oppose 2: Equals 1 if committees spend more opposing Republicans than opposing Democrats and spend more supporting Democrats than supporting Republicans. Equals 2 if committees spend more opposing Democrats than opposing Republicans and spend more supporting Republicans than supporting Democrats
Party Oppose 3: Equals 1 if committees spend more combined on opposing Republicans and supporting Democrats than combined on supporting Republicans and opposing Democrats. Equals 2 if committees spend more combined on opposing Democrats and supporting Republicans than combined on supporting Democrats and opposing Republicans.
Step 3: Impute party, using the mean party support score (the average of the four measures).
For committees that do not have a mean support score, use the mean party opposition score (the average of the three measures).

A final difficulty when working with the FEC data is the risk of double-counting operating expenditures for itemized expenses paid for with a credit card or through a vendor.[12] The FEC data includes a field indicating whether an entry is part of an itemized expense. Using a unique identifier for each transaction and a back reference for itemized expenses, we construct an indicator variable for the parent expense. We then remove the "parent" transactions so that our spending totals only include the itemized amounts.[13] Similarly, we remove duplicate transactions from the data files for independent expenditures.[14]

Index of Political Expenditures (Figure 2)

Table A4 reports total inflation-adjusted expenditures by category and cycle. The deflator is drawn from Johnston and Williamson (2020). The total category includes uncoded expenditures, which constitute 0.8% of spending in our dataset. For ease of presentation, table A4 reports presidential cycles only.

Table A4: Inflation-adjusted campaign spending by category and cycle (in millions)

	2004	2008	2012	2016	2020
media	1,084.16	1,618.07	2,490.67	2,459.14	5,209.38
digital	43.42	124.23	426.09	662.31	2,440.85
administrative	1,386.15	1,552.63	1,409.00	1,556.07	2,436.34
fundraising	269.26	457.08	576.96	495.98	960.58
consulting	207.74	265.29	418.11	483.15	759.46
field	130.18	128.07	110.50	134.31	274.93
travel	156.05	262.64	176.63	256.77	194.64

[12] For instance, a committee may list a $5,000 payment to "Chase Bank" and then itemize five, separate $1,000 payments for "airline tickets" in the same report. In this example, a simple sum of transactions is $10,000 but the actual campaign outlay is $5,000.

[13] In addition, we created an indicator variable using a simple text search for "in kind" and removed these transactions before calculating spending totals.

[14] The bulk files for independent expenditures from the FEC contain transactions from both original and amended reports. We match amended records to their original transactions according to their file number and then exclude the original transaction from our calculations.

Table A4: Inflation-adjusted campaign spending by category and cycle (in millions)

polling	41.12	93.48	136.71	135.65	182.61
legal	30.61	51.55	60.89	79.00	112.13
total	3,641.95	4,575.26	5,813.08	6,274.16	12,580.47
deflator	84.84	94.42	100.00	105.74	113.65

Comparison of Media and Digital Spending (Figure 3)

Figure 3 compares spending on digital and media by major party committees since 2004. Looking only at committees with positive spending in both categories, spending on traditional media far exceeds spending on digital; however, this gap has narrowed considerably in recent years. Table A5 show the median committee spends over $30,000 a year on digital ads and services.

Table A5: Descriptive Statistics (million 2012 dollars)

	N	Mean	SD	p50	p10	p90
Digital						
2004	957	43.57	509.50	6.36	0.58	30.14
2006	1290	39.56	511.77	8.05	0.47	39.29
2008	1493	80.25	969.71	9.48	0.64	50.29
2010	1894	49.96	438.22	10.40	0.62	59.48
2012	1946	214.09	2902.32	12.47	0.54	83.08
2014	1796	112.87	857.92	14.08	0.64	122.61
2016	1898	331.71	3093.89	17.18	0.95	199.44
2018	2326	191.34	1125.91	19.95	0.91	251.13
2020	2476	922.62	9497.83	31.28	0.96	600.77
Media						
2004	957	1054.67	8964.35	66.07	1.48	985.25
2006	1290	675.59	3290.75	40.08	0.85	1376.83
2008	1493	1078.35	12063.56	34.34	0.80	1242.03
2010	1894	577.77	2694.54	27.87	0.72	1196.91

Table A5: Descriptive Statistics (million 2012 dollars)

2012	1946	1266.93	11918.18	29.30	0.70	1235.51
2014	1796	661.08	3439.20	28.08	0.68	1160.12
2016	1898	1275.66	9538.46	24.33	0.74	1300.42
2018	2326	836.84	5383.58	22.49	0.60	1205.67
2020	2476	2072.88	17611.59	27.19	0.79	1640.90

Digital estimates using payments to platform companies (Figure 4)

One of the main challenges of working with FEC data is that campaigns often use ambiguous or under-specified descriptions of expenditures. This makes it difficult in some cases to differentiate between online and traditional advertising. To illustrate, table A6 reports the top ten expenditure descriptions for the digital and media categories and the share of total spending in each category since 2004. Whereas digital spending tends to be clearly demarcated in FEC reports, descriptions of media spending tend to be more general. More than half of media spending in our data is described as “media” or “advertising” without reference to a specific format such as radio or television.

Table A6: Descriptions of Digital and Media Spending

Digital		
Description	Millions	Share of Category
Digital Advertising	816	19.6%
Online Advertising	502	12.1%
On Line Advertising	136	3.3%
Digital Consulting Online Advertising	96	2.3%
Data Services	83	2.0%
Software	77	1.9%
Digital Advocacy	72	1.7%
Internet Advertising	69	1.7%
Digital List Rental Services	56	1.3%
Digital Advertising Estimate	46	1.1%

Table A6: Descriptions of Digital and Media Spending

All Other	2209	53.1%
	Media	
Description	Millions	Share of Category
Media Buy	4928	27.3%
Media	2208	12.2%
Media Placement	903	5.0%
TV Advertising	712	3.9%
Placed Media	660	3.7%
Advertising	646	3.6%
TV Media Placement	632	3.5%
Media Buy Estimate	408	2.3%
Television Advertising	321	1.8%
Media Production	318	1.8%
All Other	6331	35.0%

We can estimate how much digital spending we might miss because of ambiguous descriptions using data on direct payments to platform companies such as Facebook and Google. We identify payments to eight platform companies using a text search of vendor for each transaction (Facebook, Google, Twitter, Snapchat, Instagram, YouTube, AOL, and Yahoo). Our model classified almost 17% of platform expenditures as media because of ambiguous descriptions such as “advertising” or “ad buy”. We recode these as digital because they are paid to a platform company. In addition, we use the share of payments to platform companies classified as media for each cycle to estimate digital spending we might miss because of ambiguous or under-specified expenditure descriptions.

Estimated digital spending in year t is represented by the formula:

$$\text{Estimate}_t = f(x)\text{Digital}_t$$

Where $f(x)$ is a function of platform expenditures for media and digital in year t.

For example, in 2020,

81% of platform expenditures are classified as digital advertising

14% of platform expenditures are classified as media

Total platform expenditures in 2020 is represented by the equation,

$$(1) \qquad .81(\text{Platform}_{2020}) + .14(\text{Platform}_{2020}) = .95(\text{Platform}_{2020})$$

Rewriting equation 1 as a function of platform expenditures classified as digital (PD),

$$\text{PD}_{2020} = .81(\text{Platform}_{2020})$$

$$\text{Platform}_{2020} = \frac{\text{PD}_{2020}}{.81}$$

$$(2) \qquad .81\left(\frac{\text{PD}_{2020}}{.81}\right) + .14\left(\frac{\text{PD}_{2020}}{.81}\right) = .95\left(\frac{\text{PD}_{2020}}{.81}\right)$$

Using the share of platform expenditures classified as digital and media, we can now estimate total digital spending in 2020.

$$\text{Digital}_{2020} + \frac{.14}{.81}\text{Digital}_{2020} = \frac{.95}{.81}\text{Digital}_{2020}$$

This can be simplified as

$$(3) \qquad \text{Digital}_{2020} = 1.18\left(\text{Digital}_{2020}\right)$$

In other words, for every \$100 of expenditures correctly classified as digital by our model, we estimate an additional \$18 of digital advertising was classified as media because of ambiguous expenditure descriptions.

We can now estimate total digital spending in year t using the formula,

$$\text{Digital Estimate}_t = \frac{\%M_t + \%D_t}{\%D_t}\left(\text{Digital}_t\right) \tag{4}$$

Where $\%D_t$ and $\%M_t$ are the share of digital advertising and media payments to platforms in year t

Table A7 reports total digital spending for each cycle using our model, the share of platform expenditures classified as digital advertising and media, and our re-estimate using the multiplier derived from equation 4.

Table A7: Re-Estimate of Digital Spending

cycle	%Digital Ads (Platforms)	%Media (Platforms)	Multiplier (Equation 4)	Digital Ads (Re-estimate)	Total Digital (Re-estimate)[a]
2004	0.01	0.62	70.96	62.35	104.78
2006	0.16	0.43	3.64	5.04	58.25
2008	0.90	0.08	1.09	30.45	125.44
2010	0.52	0.35	1.67	26.63	109.85
2012	0.78	0.17	1.22	262.59	470.38
2014	0.50	0.36	1.72	193.04	292.72
2016	0.47	0.43	1.92	836.26	1,056.35
2018	0.58	0.34	1.59	445.69	637.19
2020	0.81	0.14	1.18	1,964.48	2,683.22

[a]Includes digital services and consulting

Comparison of tracking data with FEC data (Figure 5)

We use ad tracking data from the 2018 Wesleyan Media Project (Fowler et al. 2020) to examine whether our estimates derived from FEC reports exclude advertising spending outside the 30- and 60-day windows. We compare all ad spending naming a House and Senate candidate (tracking data) with operating and independent expenditures classified in our model as media (FEC data). Operating expenditures include House and Senate candidate committees. Independent expenditures include PACs and party committees that report spending in support of or in opposition to a House or Senate candidate. Using a text search of expenditure descriptions, we exclude transactions that mention radio, print, or outdoor advertising. We do not include expenditures for media consultants (classified by our model as consulting where the term “media” appears in the expenditure description). The Wesleyan Media Project data codes ad sponsors as candidates, parties, coordinated between parties and candidates, or interest groups. Because coordinated spending is executed by party committees, we combine these ads with the party category to facilitate comparison with our FEC data. Table A8 includes descriptive statistics for the data sources by committee type.

Table A8: Descriptive Statistics of Media Spending by Source (in millions)

	N	Mean	SD	Min	Max
Candidates					
WMP	89	7.081	14.934	0.014	82.984
FEC	101	9.529	17.787	0.002	91.849
PACs					
WMP	97	5.158	11.740	0.001	69.186
FEC	101	7.479	18.783	0.000	112.322
Parties					
WMP	51	3.955	7.741	0.000	27.089
FEC	52	4.713	8.710	0.005	36.647

Note: WMP=Wesleyan Media Project; FEC=Federal Election Commission

Table A9 reports OLS regression coefficients of FEC expenditures on weekly ad spending collected by the Wesleyan Media Project. The results suggest that each dollar of ad spending on television is associated with $1.28 in FEC expenditures. Including an interaction term suggests that candidates and party committees report media spending that is only slightly higher than ad buys found in the tracking data but is significantly higher in the case of PACs. These results are broadly consistent with findings in the literature on consultant markups for media production and placement (Martin and Peskowitz 2018)

Table A9: Regression of FEC expenditures on Ad spending (WMP)

	(1)	(2)
Ad spending	1.279***	1.091***
	(0.142)	(0.221)
PAC		-3.567***
		(0.924)
Party Committee		-2.181*
		(0.984)
PAC x Ad spending		0.514**
		(0.195)
Party x Ad spending		-0.080
		(0.244)
Constant	1.104*	3.063**
	(0.516)	(0.945)
Observations	229	229
R-squared	0.839	0.874

Note: robust standard errors in parentheses (clustered by week)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Spending on behalf of major party nominees (Figure 9)

Presidential campaigns typically rely on one or two major firms to execute their digital strategy. In 2016 and 2020, the same digital firms worked for both the principal campaign committee and the joint fundraising committee. Table A10 shows total payments to the top firms providing digital ads and services to the Biden and Trump campaigns. The purpose description column is taken from the FEC data (after removing punctuation). Our data does not allow us to differentiate between consulting fees or advertising markups and the cost of placing ads on platforms such as Facebook and Google.

Table A10: Payments for digital advertising by the Biden and Trump campaigns[a]

Payee	Committee	Purpose Description	Total (mil.)

Table A10: Payments for digital advertising by the Biden and Trump campaigns[a]

	Biden Campaign		
GMMB	Biden for President	Digital Advertising	184.4
Bully Pulpit Interactive	Biden for President	Digital Advertising	41.3
Facebook	Biden Victory Fund	Digital Advertising	34.8
GMMB	Biden Victory Fund	Digital Advertising	29.9
Google	Biden for President	Digital Advertising	29.3
Infogroup	Biden for President	Digital Advertising	22.8
Google	Biden Victory Fund	Digital Advertising	14.5
Infogroup	Biden Victory Fund	Digital Advertising	2.5
Authentic Campaigns	Biden for President	Digital Advertising	2.1
Facebook	Biden for President	Digital Advertising	1.9
All other payments			2.9
	Trump Campaign		
American Made Media	Make America Great Again	Online Advertising	170.7
American Made Media	Trump for President	Online Advertising	96.6
American Made Media	Trump for President	SMS Advertising	21.8
American Made Media	Make America Great Again	SMS Advertising	10.4
American Made Media	Trump for President	Media Production Services	5.7
American Made Media	Trump for President	Online Ad. Subscriptions	2.1
American Made Media	Trump for President	Online Ad. Media Production	2.0
American Made Media	Make America Great Again	Online Ad. SMS Advertising	1.9
American Made Media	Trump for President	Placed Media Online Ad.	1.4

Table A10: Payments for digital advertising by the Biden and Trump campaigns[a]

American Made Media	Make America Great Again	Digital Media List Rental	1.2
All other digital advertising payments			20.7

[a]Includes payments by campaign and joint fundraising committees

Additional network measures (figure 10)

Table A11 reports separate centralization measures for the Democratic and Republican networks excluding platform companies. Consistent with the visualizations in figure 10, the network measures in table A11 confirm that the Republican network of digital firms is smaller and less densely connected than the Democratic network. By 2020, the density of ties between Democratic firms is more than twice that of Republican firms. Increases in the centralization measures suggest the emergence of a few connected firms in each network. The higher betweenness measure in the Republican network suggests there are one or two focal firms compared to a dense but more evenly distributed set of connections among Democratic firms. All network measures and visualizations use the nwcommands package in Stata (Grund 2015).

Table A11: Network measures by party

	2008	2012	2016	2020
Democrats				
edges	126	37	110	180
density	0.146	0.041	0.111	0.167
degree	0.359	0.107	0.240	0.257
betweenness	0.102	0.217	0.130	0.092
closeness	0.197	0.105	0.207	0.205
Republicans				
edges	8	32	72	83
density	0.011	0.035	0.070	0.074
degree	0.045	0.238	0.206	0.212

Table A11: Network measures by party

betweenness	0.003	0.155	0.159	0.165
closeness	0.024	0.103	0.178	0.210

Client ideology (Figure 12)

Table A12 reports descriptive statistics of client ideology firm specialization. We use Bonica's (2014) CF scores to place candidates and committees on a single ideological dimension. Firm specialization is defined as more than 50% of spending classified in a professional category. Firms with less than $50k in payments for professional services are excluded.

Table A12: Descriptive statistics of client ideology (CF Scores)

	N	Mean	SD	Min	Max
Democratic Firms					
Digital	476	-1.114	0.349	-4.395	1.265
Media	1515	-0.972	0.575	-4.848	1.780
Fundraising	2060	-0.976	0.356	-3.004	1.265
Consulting	1221	-0.997	0.351	-2.709	1.070
Polling	439	-0.999	0.332	-1.787	0.295
Total	5,711	-0.993	0.424	-4.848	1.780
Non-Digital	5,235	-0.982	0.428	-4.848	1.780
Republican firms					
Digital	647	1.090	0.257	-1.129	3.209
Media	2023	0.996	0.376	-2.313	4.834
Fundraising	2735	1.006	0.269	-1.531	4.496
Consulting	1806	1.015	0.304	-0.744	4.308
Polling	313	0.960	0.280	-1.473	1.487
Total	7,524	1.011	0.310	-2.313	4.834
Non-Digital	6,877	1.003	0.313	-2.313	4.834

A one-way analysis of variance indicates a statistically significant difference in the mean CF Scores of clients served by digital firms compared to the mean CF scores of clients served by other types of professional firms. Cell entries in Table A13 are differences in means (p-values in parentheses). Positive/negative cell entries indicate average CF Scores of the row category are to the right/left of the column category.

Table A13: Comparison of means (Bonferroni test)

	digital	media	fund	consult
Democrats				
media	0.141			
	(0.000)			
fund	0.137	-0.003		
	(0.000)	(1.000)		
consult	0.116	-0.025	-0.021	
	(0.000)	(1.000)	(1.000)	
polling	0.115	-0.026	-0.022	-0.001
	(0.000)	(1.000)	(1.000)	(1.000)
Republicans				
media	-0.095			
	(0.000)			
fund	-0.085	0.010		
	(0.000)	(1.000)		
consult	-0.076	0.019	0.010	
	(0.000)	(0.558)	(1.000)	
polling	-0.130	-0.035	-0.045	-0.054
	(0.000)	(0.599)	(0.150)	(0.040)

Note: p-values in parentheses

References

Bird, S., Klein, E. and Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. O'Reilly Media, Inc.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2), 367-386. https://doi.org/10.1111/ajps.12062

Fowler, E.F., Franz, M.M. and Ridout, T.N., and Baum, L.M. (2020). “Political Advertising in 2018.” Version 1.0 [dataset]. Middletown, CT: The Wesleyan Media Project, Department of Government at Wesleyan University.

Grund, T.U. (2015). "nwcommands. Network Analysis in Stata". https://github.com/ThomasGrund/nwcommands (accessed February 21, 2022)

Johnston, L., and Williamson, S.H. (2020). What Was the U.S. GDP Then? Available at http://www.measuringworth.org/usgdp/ (accessed July 28, 2020).

Limbocker, S. and You, H.Y. (2020). Campaign styles: Persistency in campaign resource allocation. *Electoral Studies*, 65, 102-140. https://doi.org/10.1016/j.electstud.2020.102140

Manning, C.D., Raghavan, P. and Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.

Martin, G.J. and Peskowitz, Z. (2018). Agency Problems in Political Campaigns: Media Buying and Consulting. *American Political Science Review*, 112(2), 231-248. https://doi.org/10.1017/S0003055417000594

McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*. 445(1), 51-56.

Millman, K.J. and Aivazis, M. (2011). Python for scientists and engineers. *Computing in Science & Engineering*, 13(2), 9-12. https://doi.org/10.1109/MCSE.2011.36

Oliphant, T.E. (2007). Python for scientific computing. *Computing in science & engineering*, 9(3) 10-20. https://doi.org/10.1109/MCSE.2007.58

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, pp.2825-2830.

Sheingate, A.D., 2016. *Building a business of politics: The rise of political consulting and the transformation of American democracy*. Oxford University Press.

Sheingate, A.; Scharf, J.; Delahanty, C. (2022): Replication Files for "Digital Advertising in U.S. Federal Elections, 2004-2020". https://doi.org/10.6084/m9.figshare.19372430.v3

Van Der Walt, S., Colbert, S.C. and Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2), 22-30. https://doi.org/10.1109/MCSE.2011.37