**Appendix**

**Data**

This paper relies on a dataset of 840 unique digital political advertisements (737 pro-Obama and 103 pro-Romney), collected from a Moat search for the brands “Obama” and “Romney” on November 6, 2012.[[1]](#footnote-1) Moat defines a “unique” creative based on the content and size in kilobytes (KB). Ad size could vary, for instance, based on the dimension of the ad associated with a particular location (top vs. side) on a website. As this is a new data source, we want to provide a more thorough discussion of the nature, limitations, and alternatives to the Moat sample.

Founder Jonah Goodhart explains Moat as follows: “It’s basically like Google, except Google basically skips the ads and just grabs the content. We skip the content, and just grab the ads.” Like most search engines, the specific search algorithm is proprietary, somewhat limiting our ability to evaluate the extent and nature of the coverage error between our sample and the full population of online display ads in 2012. Nonetheless, there are several reasons to conclude that the sample offers a reasonable—if imperfect—collection of display ads by which to explore the campaign communication strategies in online display ads in presidential elections.

First, the Moat algorithm does a more complete job of capturing ads on the most popular 50,000 websites compared to those on less popular websites. This parallels the type of omission found in the CMAG database of TV ads, which has traditionally covered ads in the top 100 of the 210 US media markets (covering 75% of the population). In the past, CMAG has tracked ads on 53 different broadcast and cable networks, but there are 1,774 broadcast television stations and 6,100 cable systems. As Moat relied on a combination of machine learning, logo detection and crowd-sourced human identification in classifying ads, we might think they would be more likely to miss ads that do not use the campaign logo. However, we can think of no reason that ads on less popular websites or without logos would be systematically more targeted to core supporters.

We were also able to compare the nature of our sample to one generated by a more comprehensive for-pay service, Moat Pro, which launched after research was underway. We were granted temporary access to Moat Pro to conduct a quality check of the sample. The reported match estimate is conservative because we scored as a match only those ads with identical text. By our definition, many unmatched ads had a near match. For example, “Join Michelle and tell Barack you’re in.” could end with “Are you in?” or “Get started”; we did not consider these a match. Overall, Moat Pro identifies 1405 unique Obama ads and 157 unique Romney ads, but the definition of “unique” differs across samples and are not directly comparable. Reassuringly, we find that our sample captures more than 55% of the unique ads listed in Moat Pro. More importantly, we found no systematic pattern in the content of ads not included in the Moat sample.

As a robustness check on the limitation that we do not have the complete list of every website where the ad appeared, as Moat only reports the last two websites on which an ad was shown, we reran our analyses on the subset of 319 ads (38% of the sample; 269 Obama, 50 Romney) that appeared on two or fewer websites. We find that our general trends hold: while Obama placed more ads on left-leaning sites (26.4% Obama, 4% Romney) and Romney placed more ads on right-leaning sites (12.6% Obama, 78% Romney), each also placed a significant number of ads on politically neutral sites (29.4% Obama, 12% Romney) and non-political sites (47.9% Obama, 16% Romney). The only change is that in this subsample, Romney placed a somewhat smaller proportion of ads on neutral and non-political sites.

Given our imperfect sample, it is also reasonable to consider alternative approaches. We could, for instance, conduct qualitative interviews with campaign strategists. However, campaign staffers might have incentives to conceal or misrepresent campaign communication strategies.[[2]](#footnote-2) Another alternative is to seek the full population of advertising produced directly from the campaign. Although rare, there have been cases where scholars have been able to work with the campaigns to gather detailed information about their communication strategies (e.g., Shaw 2006). Unfortunately, that does not (yet) appear to be possible for online advertising. According to Romney’s Director of Advertising Analytics, it would be “impossible” to pull together a comprehensive list of digital ads for the Romney campaign because the campaign did not maintain records of all the online ads placed, in part because of the frantic pace in which various creatives were put together, especially towards the end of the campaign. The campaign would produce many different versions of a creative, but final placement and use were often determined by a third-party intermediary. Campaigns would simply specify placement factors such as specific websites, keywords on website, or audience characteristics.

In sum, our sample is not comprehensive, but it captures a narrow and important slice of online communication strategy and offers a rare opportunity to look at the strategies behind contextual online display ads in presidential elections.

**Coding Details**

The following characteristics were directly recorded from Moat for each ad: kilobyte size, sponsor, date last seen, and last two websites on which the ad was shown. A screenshot of the ad was captured and the written content of the ad was recorded verbatim. The content of the ads was then coded using a set of questions modeled after the Wisconsin-Wesleyan Advertising Project.[[3]](#footnote-3) This content analysis captured the type of actions urged by the ad, the characteristics of the individuals in the ad, the issue content of the ad, and some 50 other characteristics. A complete codebook is available online on the Dataverse Network as part of the dataset. Ads were initially coded by one of the authors before hypotheses were developed. A sample of 100 ads was independently coded by an undergraduate research assistant for calculating reliability statistics. When possible, we calculate two reliability statistics—percent agreement and Cohen’s kappa—for each variable. Percent agreement is thought of as a liberal reliability measure, and kappa is thought of as a conservative reliability measure. Together, the two provide adequate information on the inter-coder reliability of our variables, which is generally high. The following variables were used in the presented analysis.

* **Ad Goal:** A classification of the primary goal of each ad is based on the judgment of two coders, categorized according to four mutually exclusive categories of appeals: mobilization/get-out-the-vote (GOTV), political persuasion (persuasion), donation for campaign (donation), recruitment for active role in campaign (recruitment). Don’t know (DK) and NA were also possible alternatives, but no ads were coded as such. In making this classification, coders were instructed to use all information available—the image, content, interactive links, etc. To create the most conservative conditions for our analyses, any ad asking for a donation request was coded "donation," even if it could have also fit in another category. Persuasion ads were defined as those that gave some argument for why the viewer should vote for one candidate or the other. These broad categories were then collapsed into persuasion or engagement, with engagement ads including "GOTV," "donation," or "recruitment." To be sure, these categories are admittedly artificial and more a reflection of political science literature on candidate strategy rather than insider knowledge of candidate motivation (e.g., Shaw 2006). However, the coding does at least allow for a clean testing of the conventional wisdom that online communications are focused on raising political donations. (93.6% coder agreement; Cohen's Kappa coefficient of .902).
* **Interactivity:** Coders answered “Does the ad direct the viewer to take any action (as opposed to merely providing information)? If action is urged, what is the action?” There were 16 possible responses: none, vote, donate, interact in social networks, learn more/find more information, buy a sticker, shop for candidate items, share your story, confirm polling place, confirm polling place and vote, register to vote, vote early, sign up/participate, and other. DK, and NA were possible responses, but zero ads were coded DK or NA. 96% of solicited interactive actions (718/744) were to sign up and/or participate, donate, or find out more information. (88.1% agreement; Cohen's Kappa coefficient of .847).
* **Issue Content:** Each ad was coded for mentions of 26 specific issues, reported in Table 2. Not shown are issue categories that no ads mentioned (media, religion/moral values, and Wall Street). (96.3% agreement; Cohen’s Kappa coefficient of .955).
* **Group targeting:** Coders identified if an ad was targeted to a subgroup of the electorate based on content and visuals. Categories included Business Owners, Blacks, Hispanics, Young People, College students, young, LGBT, Women, specific locations, and other (teachers, for instance, were occasionally targeted), and not targeted. (88.1%; Cohen’s Kappa coefficient of .706).
* **Website ideology:** In a separate coding exercise, the last two websites on which the ads were seen were coded for their ideological leaning. Websites were classified as either left-leaning, right-leaning, neutral (but still political), or nonpolitical. Classification was based on audience ideology estimates taken from Gentzkow and Shapiro (2011). For websites without audience measures, classification as left- or right-leaning was determined by candidate endorsement. For example, lasvegassun.com published an editorial on October 26, 2012 titled “Obama is our choice for president.” For websites without audience measures or candidate endorsements, an undergraduate coder reviewed the content of the website (on the dates of coding), with classifications validated by the authors. For instance, alternet.org has a clear left-leaning ideological appearance; as of July 2013 the site urged viewers to “Defeat the Republican majority.” Websites were coded nonpolitical if they had no political content, according to these three measures. To give some examples of the resulting classifications: nytimes.com, oprah.com, and colbertnation.com were all coded as left-leaning websites; orlandosentinel.com, drudgereport.com, and realclearpolitics.com were coded as right-leaning websites. Cnn.com, usatoday.com, and msn.com were coded as neutral websites; active.com, azlyrics.com, and ehow.com were coded as non-political websites.

**Measuring Issue Convergence**

To calculate issue convergence, we first consider the breakdown of issue mentions by each candidate in Table 2. Because ads can mention multiple issues, we divide the percentage of each candidate’s ads that mention each issue by the total issue mentions for each candidate, so that all percentages for each candidate sum to 100%. This gives us an idea of how much total attention was devoted to each issue by each side. From here, we can calculate an issue convergence coefficient proposed in Sigelman and Buell (2004),

 (1)

where PD and PR are percentages of the relative attention paid to a given issue (*i* in {1,…,*n*}) by Obama and Romney ads, respectively. As reported in Table 2 in the main article, the issue convergence coefficient for online campaigning in 2012 is 49.4.

**Appendix Works Cited**

Sigelman, Lee and Emmett H. Buell. “Avoidance or engagement? Issue convergence in US Presidential Campaigns.”*American Journal of Political Science*48, no. 4 (2004), 650-661.

Gentzkow, Matthew and Jesse M. Shapiro. “Ideological Segregation Online and Offline.”  *Quarterly Journal of Economics* 126*,* no. 4 (2011): 1799-1839.

Shaw, Daron R. *The Race to 270: The Electoral College and the Campaign Strategies of 2000 and 2004*. University of Chicago Press, 2008.

1. This set of ads included creatives from the campaigns as well as third-parties. Throughout our analysis we use the label “Obama ads” for pro-Obama from all sources and the label “Romney ads” for pro-Romney ads from all sources. We have replicated our results for the subset of ads produced by the campaign with similar conclusions. [↑](#footnote-ref-1)
2. Consider the common refrain from presidential candidates that they are speaking to *all* Americans. A recent survey found that 86% of Americans did not want tailored political advertising (Turow et al. 2012). [↑](#footnote-ref-2)
3. For more information about that project, including their coding fields, see <http://mediaproject.wesleyan.edu/wp-content/uploads/2015/05/WMP-2012-releasecodebook_v1.01.pdf> [↑](#footnote-ref-3)