

PRE-PROOFS VERSION

The Myths of Data-Driven Campaigning

Jessica Baldwin-Philippi

Assistant Professor, Fordham University

Just as social media campaigning has been valorized since the Obama 2008 campaign, data-driven campaigning has gained vaunted status in the aftermath of both the 2012 and 2016 elections. The analytics teams that supported a variety of the Obama campaigns' work in both 2008 and 2012 have been mythologized, and in the months since Donald Trump's election, his data operation—led by the UK-based firm Cambridge Analytica and Brad Parscale--has been valorized by journalists (Grassegger & Krogerus, 2017; Green & Issenberg, 2016; Halpern, 2017), while the digital efforts of the Clinton campaign have been blamed for her loss (Gold, 2017; Wagner, 2016). This piece examines the gap between the rhetoric of data driven campaigning and the realities of on-the-ground practices, especially as they relate to the position of the Trump campaign compared to others. All but the most sophisticated digital and data-driven strategies are imprecise and not nearly as novel as the journalistic feature stories claim. Although there is usually a gap between the practices of presidential and down-ballot races, similarities can be seen between the Trump campaign and practices of congressional and state-wide candidates. This essay assesses the way the Trump campaign compared to other presidential and down-ballot campaigns in its use of analytics—namely email and social media platforms.

Data driven campaigning involves two main features: targeting, or deciding which messages go to what potential voters at what time during the campaign, and testing, or empirically measuring how well messages perform against one another and using that information to drive content production and further targeting. While these practices can

PRE-PROOFS VERSION

be undertaken in any part of the campaign, it is the field teams—those behind mobilizing supporters, making contact with likely supporters via phone and in-person canvassing, and getting out the vote—that have a deeper history of data-driven work. Since the late 1990s, academics and practitioners have experimented with using public voting records and census data to successfully target voters for informational canvassing and get out the vote efforts, and had lesser successes using “lifestyle” data pulled from consumer data like credit card purchases and magazine subscriptions (Hersh, 2015). As the use of such data has largely been institutionalized at the level of the party (Kreiss, 2012a, 2016), both the DNC and RNC currently have relatively similar capacity to do this type of targeting. The wider differences, as well as much of the public praise of data-driven campaigning, can be seen in the range of ways communications and digital teams use data to create and test their digital messages.

Email analytics: A widening gap

Although email is a rather old technology, it is the site of an enormous amount of content creation, and the vast majority of analytics based message testing (Baldwin-Philippi, 2016; Hestres, 2017; Issenberg, 2012; Karpf, 2016). Most campaigns use customer management systems (CRMs) that can organize email lists and facilitate targeting and testing of messages using information a campaign often gets when individuals sign up to receive email, such as zip code and areas of political interest. CRM systems can also collect data on whether an email was opened, if it was shared in other social media spaces, and/or what links were clicked, allowing campaigns to test how well certain messages worked, and with what populations they were successful. Although

PRE-PROOFS VERSION

these systems require little effort or infrastructure, the ways campaigns put them to use and the attention given to and rigorous use of analytics varies wildly.

Prior to 2016, the exact amount of email a contemporary campaign generated was difficult to know due to researchers' inability to access every targeted version of an email, a conservative estimate is that campaigns and parties send out hundreds of emails per account. In 2012, testing messages against each other—from subject line to full text of an email—was central to presidential-level campaigns, though there was a wide gap between the Obama and Romney campaigns in doing so. *ProPublica* counted 1703 Obama campaign emails, 446 Romney campaign emails during the general election (ProPublica, 2012).¹ While the Obama campaign noted that a central component of its strategy was to “test every single thing,” (Cramer, 2012) the RNC's 2012 autopsy highlighted message testing as an area where the DNC and Obama campaign significantly outpaced Romney and the RNC (Republican National Committee, 2013).

In 2016, the gap between emails sent by the Trump campaign and the Clinton campaign was even wider. Notably Trump campaign sent many fewer emails, signaling much less targeting and testing. In May 2016, during the final throes of the primary, the Clinton campaign sent out 658 emails, the Sanders campaign sent out 272, and the Trump campaign sent out 21 (Kaye, 2016). While the Trump campaign's numbers increased over time, they did not come close to matching his opponent. From the summer through late October, the Clinton campaign sent out 6 to 7 times the amount of emails as the

¹ Because this number was gathered via submissions from individuals rather than tracking emails, the actual numbers are surely higher for each campaign, but they point to a discrepancy in the amount of targeting and testing that would be possible for each campaign

PRE-PROOFS VERSION

Trump campaign.² This level persisted until the last two weeks of the race, when the Trump campaign substantially increased its email campaign, sending out an amount of email that was only 13% lower than Clinton's. It is clear that in throughout the summer the Trump campaign was neither targeting nor testing its emails. As the number ramped up, we cannot be sure whether there was both testing and targeting going on, but at the very least, targeted messages were being deployed.

Beyond sheer numbers, the Trump campaign's email efforts were riddled with errors that showed a lack of understanding and care for the analytic outcomes of their messages. Most notably, purchasing lists and emailing people who had not opted into receiving Trump emails, and emailing from multiple, new accounts made the campaign's emails more likely to end up in their receivers' spam filters. This happened for an astoundingly high 60% of the addresses for the campaign's first fundraising email, and 8% of all other emails, a rate that is quite high by industry standards, where a normal rate is closer to ~1% (Kaye, 2016; Miller, 2016). The Trump campaign had lower open rates than the Clinton campaign, averaging 12% to Clinton's 18% as of October 2016, a number that had been much higher before the size of Trump's list increased.

Overall, these practices are more on par with a high-level congressional or mid-level senate campaign than a presidential one, and the Trump campaign's use of an old technology, email, fell well short of standards of a presidential campaign. In down-ballot campaigns, there are many valid reasons for practices like these. The relatively small labor pool of staffers devoted to digital, as well as the small payoff in terms of total gain for the campaign. For instance, if a presidential list of 10 million people experiences a

² July saw 212 Clinton emails, versus 35 from the Trump campaign; and in the two weeks after labor day the Clinton camp sent out 537 to Trump's 82 (Landsman, 2016a, 2016b).

PRE-PROOFS VERSION

0.5% increase in donations because of testing, that amounts to an extra 500,000 donations, but returns diminish for smaller campaigns. The reasons for this happening at the presidential level are likely simple: creating and testing many versions of email takes many hours of labor. The Trump campaign's digital operation was relatively thinly staffed, as the "over 100" number that's been publicly discussed also included non-digital help, such as "elderly call center volunteers" (Green & Issenberg, 2016).

Social media: Targeted ads and bullhorns

In congressional, senate, and gubernatorial races, campaigns' facility with targeting digital ads far surpasses any of its other communications efforts, and the Trump campaign was no different. As early as 2010, even congressional campaigns were much more familiar with targeting ads in Facebook and looking at the ad analytics Facebook provides in order to assess whether an online ad was succeeding, and making the appropriate changes from these findings. Campaign staffers who did not A/B test any e-mail or web content were familiar with targeting Facebook ads by location, age, education, gender, and other demographic features (Baldwin-Philippi, 2015) due to their knowledge of general targeting practices used in direct mail and web ads. In the intervening years, Facebook has added to its ad platform to make microtargeting easier, allowing campaigns to target more precise interests that are based on keywords or categories that users have entered, geographical data, and algorithmically-created lifestyle profiles.

The Trump campaign's greatest triumph in data-driven strategy was its use of microtargeted Facebook ads, taking the form of sidebar and in-feed ads, targeted by demographics and other Facebook-created interest categories, as well as "dark posts"

PRE-PROOFS VERSION

(nonpublic posts that the campaign can make visible to whomever it wants) and promoted posts. The campaign used positive ads targeted to Trump supporters as well as negative attack ads for swing voters or lean-Clinton voters in an effort to reduce her turnout using many of Facebook's built in targeting tools. This practice is what the majority of news stories on the Trump campaign's digital strategy focused on and what the campaign's digital consultant, Cambridge Analytica and campaign staffers themselves have stated was the revolutionary practice that won them the election (Grassegger & Krogerus, 2017; Green & Issenberg, 2016).

The amount of money the Trump campaign poured into Facebook ads—the majority of Parscale's \$90 million budget, which was about three times the size of Clintons's digital ad buy (Lapowsky, 2016)—shows dedication to using the platform and its ability to target, and implies a significant number of ads were created and run. Although the specifics are relatively unknowable due to the organization's reticence to go on the record and conflicting claims,³ these practices largely mirrored what the Obama 2012 campaign did while also taking advantage of some more precise features. Those include “lookalike audiences” that would find people similar to known supporters (or undecideds) and target them accordingly, dark posts, and a new feature of Facebook's: advice from Facebook staffers dedicated to helping campaigns (Kreiss & McGregor, forthcoming). While many have touted Cambridge Analytica's “psychometric” targeting methods (in which psychological profiles are added to interest and demographic profiles) as the next step in increasingly-micro targeting practices (Grassegger & Krogerus, 2017;

³ The campaign claimed to have created and ran 175,000 variations of ads in one day on Facebook (Lapowsky, 2016), but the skeleton staff of the Trump campaign, with only a dozen embeds from CA joining Parscale's digital team, makes it somewhat dubious. The sheer amount of labor needed to create that many variants would have crippled what was already an incredibly thin staff.

PRE-PROOFS VERSION

Lapowsky, 2016), Matt Oczkowski, Cambridge Analytica's head of product, admitted to a post-election group at Google that "we actually didn't do any psychographics with the Trump campaign" ("Google Post-Election Review," 2016).

These microtargeting efforts in Facebook were well planned and executed, took place at a large scale, were likely instrumental in fundraising, and according to the campaign, got undecided voters who thought poorly of Trump to stay home rather than vote for Clinton. That said, even when down-ballot races were still stumbling head-first into figuring out best practices, the familiarity with ad buys and targeting—grown out of direct mail, and website and Youtube ads—is something that is comfortable for staffers in a way that the other social media analytics, such as impressions, reach, and engagement are not. Thus, while the campaign's Facebook ad strategy was productive was ultimately similar in quality and more extensive in quantity than that of the Clinton campaign, both of which were largely an extension of the Obama 2012 campaign.

The more potentially revolutionary use of social media fell completely outside the realm of data or analytics, with its practice of using Twitter to bolster outside voices, including news organizations, activists groups, and individuals. What the Trump campaign did was effectively break norms for how candidates (and elected officials) engage in social media production (Lawrence & Boydston, 2017; Wells et al., 2016). Authorship, often blurred by the fact that social media staffers run the accounts of candidates and officials alike, was further erased by the Trump campaign and presidency. While many candidates and officials sign some tweets or Facebook posts to signal that they, not their staffers, authored a post, the Trump campaign did not. In Twitter, the blended voice of the candidate/campaign/administration was also purposefully

PRE-PROOFS VERSION

interrupted by retweeting major pundits and little-known supporters alike, with seemingly little attention given to researching who the entity being retweeted was (often to scandalous results). This practice does not seem to be data-driven, as Trump's original tweets garner far higher engagement metrics than those he retweets. Additionally, some have hypothesized that it was an important feature for establishing Trump's authenticity, a trait increasingly associated with a "good" social media presence (Marwick & boyd, 2011; McGregor, 2017). It remains to be seen whether these moves constitute the beginning point for changing of norms, or remain outliers, but one thing is certain: these practices are novel insofar as they privilege voices outside of the candidate's controlled message, and they lack any payoff in terms of identifiable increases in reach or engagement.

Data-rich and data-less

Contemporary political campaigns are operating in a data rich environment. Every email and ad, every word of copy or image used, can provide data point for who can be persuaded to take a particular action at any given time. Often however, campaigns do not take advantage of these opportunities. Even at the presidential level, the labor it takes to craft and test messages is significant. While campaigns have generally been inching forward in testing and analytics since 2010, the Trump campaign was a mixed bag, severely behind the Clinton campaign and even the 2012 Romney campaign in testing email, while it was relatively comparable in its use of Facebook ad analytics.

So, why does what campaigns are doing—and, more to the point, aren't doing—matter? The myths of data-driven campaigning are deeply tied to the field's literature on the democratic implications of microtargeting, most of which are dystopian concerns

PRE-PROOFS VERSION

about the ability for citizens to be informed and their relegation to filter bubbles (Howard, 2006; Kreiss, 2012b; Pariser, 2011; Tufekci, 2012). Increasingly, these fears are being revealed to be just that—concerns of theoretical impact rather than actual impact, with recent studies showing the following: the “micro” part of targeting is no more effective than using public records (Hersh, 2015), and we are not stuck in filter bubbles, instead audiences duplicate and overlap frequently (Fletcher & Nielsen, 2017). Coupled with the knowledge that campaigns often do not execute analytic-based campaigning tactics as fully or rigorously as possible, the concern that we are being driven away from democracy by data is overstated. While campaigns certainly try to control citizens’ actions and information, and often succeed, this system does not have the strict control that many of these concerns rely on.

PRE-PROOFS VERSION

Reference:

- Baldwin-Philippi, J. (2015). *Using Technology, Building Democracy: Digital campaigning and the construction of citizenship*. New York: Oxford University Press.
- Baldwin-Philippi, J. (2016). The Cult(ure) of Analytics in 2014. In J. A. Hendricks & D. Schill (Eds.), *Communication and Midterm Elections: Media, Message, and Mobilization*. New York: Palgrave Macmillan.
- Cramer, R. (2012, November 9). Messina: Obama Won On The Small Stuff. *BuzzFeed*. Retrieved from <https://www.buzzfeed.com/rubycramer/messina-obama-won-on-the-small-stuff-4xvn>
- Fletcher, R., & Nielsen, R. K. (2017). Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *Journal of Communication*. <https://doi.org/10.1111/jcom.12315>
- Gold, D. (2017, February 9). “Data-Driven” Campaigns Are Killing the Democratic Party. *POLITICO Magazine*. Retrieved from <http://politi.co/2krhUuC>
- Google Post-Election Review. (2016, December 13). C-SPAN. Retrieved from <https://www.c-span.org/video/?420077-1/google-hosts-postelection-review&start=6905>
- Grassegger, H., & Krogerus, M. (2017, January 28). The Data That Turned the World Upside Down. *Motherboard (Vice)*. Retrieved from https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

PRE-PROOFS VERSION

- Green, J., & Issenberg, S. (2016, October 27). Inside the Trump Bunker, With Days to Go. *Bloomberg BusinessWeek*. Retrieved from <https://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go>
- Halpern, S. (2017, June 8). How He Used Facebook to Win. *New York Review of Books*, 64(10). Retrieved from <http://www.nybooks.com/articles/2017/06/08/how-trump-used-facebook-to-win/>
- Hersh, E. (2015). *Hacking the Electorate: How Campaigns Perceive Voters*. New York: Cambridge University Press.
- Hestres, L. E. (2017). Tools Beyond Control: Social Media and the Work of Advocacy Organizations. *Social Media + Society*, 3(2). <https://doi.org/10.1177/2056305117714237>
- Howard, P. (2006). *New media and the managed citizen*. New York: Cambridge University Press.
- Issenberg, S. (2012). *The victory lab: The secret science of winning campaigns*. New York: Crown Books.
- Karpf, D. (2016). *Analytic Activism: Digital Listening and the New Political Strategy*. New York: Oxford University Press.
- Kaye, K. (2016, June 23). Trump's First Fundraising Email Had a 60% Spam Rate. Retrieved November 1, 2016, from <http://adage.com/article/campaign-trail/trump-s-fundraising-email-a-60-spam-rate/304673/>
- Kreiss, D. (2012a). *Taking our country back: The crafting of networked politics from Howard Dean to Barack Obama*. New York: Oxford University Press.

PRE-PROOFS VERSION

Kreiss, D. (2012b). Yes We Can (Profile you): A brief primer on campaigns and political data. *Stanford Law Review Online*, 64, 70–74.

Kreiss, D. (2016). *Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy*. New York: Oxford University Press.

Kreiss, D., & McGregor, S. (Forthcoming). Active Intermediaries, Not Distribution Channels: How Technology Firms Shape. *Political Communication*.

Landsman, J. (2016a, July 25). Nugget: Email Wrap-up - Post Republican Convention. Retrieved from <http://blog.edatasource.com/nugget-email-wrap-up-post-republican-convention/>

Landsman, J. (2016b, September 19). Presidential Campaign Emails: Post Labor Day Check-in. Retrieved July 31, 2017, from <http://www.edatasource.com/presidential-campaign-emails-post-labor-day-check-in/>

Lapowsky, I. (2016, November 15). This Is How Facebook Actually Won Trump the Presidency. Retrieved July 31, 2017, from <https://www.wired.com/2016/11/facebook-won-trump-election-not-just-fake-news/>

Lawrence, R. G., & Boydston, A. E. (2017). What We Should Really Be Asking About Media Attention to Trump. *Political Communication*, 34(1), 150–153. <https://doi.org/10.1080/10584609.2016.1262700>

Marwick, A. E., & boyd, danah. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114–133. <https://doi.org/10.1177/1461444810365313>

PRE-PROOFS VERSION

- McGregor, S. C. (2017). Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. *New Media & Society*, 1461444816686103. <https://doi.org/10.1177/1461444816686103>
- Miller, S. (2016, July 8). Growing Pains for Trump's Email Effort. *Campaigns and Elections*. Retrieved from <https://www.campaignsandelections.com/campaign-insider/growing-pains-for-trump-s-email-effort>
- Pariser, E. (2011). *Filter Bubble: How the new Personalized Web is changing what we read and how we think*. New York: Penguin Books.
- ProPublica. (2012). *Message Machine*. Retrieved from <http://projects.propublica.org/emails/>
- Republican National Committee. (2013). *Growth and Opportunity Project*.
- Tufekci, Z. (2012, November 16). Beware the Big Data Campaign. *The New York Times*. Retrieved from <https://www.nytimes.com/2012/11/17/opinion/beware-the-big-data-campaign.html>
- Wagner, J. (2016, November 9). Clinton's data-driven campaign relied heavily on an algorithm named Ada. What didn't she see? *Washington Post*. Retrieved from <https://www.washingtonpost.com/news/post-politics/wp/2016/11/09/clintons-data-driven-campaign-relied-heavily-on-an-algorithm-named-ada-what-didnt-she-see/>
- Wells, C., Shah, D. V., Pevehouse, J. C., Yang, J., Pelled, A., Boehm, F., ... Schmidt, J. L. (2016). How Trump Drove Coverage to the Nomination: Hybrid Media Campaigning. *Political Communication*, 33(4), 669–676. <https://doi.org/10.1080/10584609.2016.1224416>

PRE-PROOFS VERSION