

Article

Big Data under Obama and Trump: The Data-Fueled U.S. Presidency

Barbara Trish

Political Science Department, Grinnell College, 50112 Grinnell, USA; E-Mail: trish@grinnell.edu

Submitted: 2 May 2018 | Accepted: 28 September 2018 | Published: 21 November 2018

Abstract

The much-heralded use of data, analytics, and evidence-based decisions marks the U.S. presidency, wherein many processes and decisions are structured by the analysis of data. An approach with historical precedent, reliance on data was prominent under Obama, and is even under Trump, despite signals to the contrary. This article examines three cases from the Obama era: microtargeting in electoral campaigns, performance management in government, and signature drone strikes employed by the national security apparatus. It also reflects on the early Trump administration. The processes described are highly dependent on data, technically big data in two instances. The article examines the cases both on their own terms and in the context of a critical lens that directs attention to the political economy of the data. The analysis helps unpack the allure of data and analytics as well as the challenges in structuring an environment with a measured approach to data and big data, which would examine both their potential and drawbacks.

Keywords

analytics; big data; data; drone strikes; evidence-based; microtargeting; Obama; performance management; president; Trump

Issue

This article is part of the issue “Big Data Applications in Governance and Policy”, edited by Sarah Giest (Leiden University, The Netherlands) and Reuben Ng (National University of Singapore, Singapore).

© 2018 by the author; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

“Evidence-based” is the 21st century coin of the realm, with broad, seemingly unbounded applicability. Practices relying on evidence, their counterpart “data-based decisions”, along with the tracking, data, and analytics that fuel all have infused the public realm, private lives and all areas between. Presidential politics in the U.S. is no exception.

Reliance on data, in some cases technically “big data”, marks the contemporary presidency to the extent that it has become the default approach, part of its institutional DNA. In some measure, this mirrors the broad progression of thought and practice that extends beyond the narrow scope of the U.S. presidency, from politics and governing generally, to commerce, sports, and all variety of enterprises that value efficiency, either as means of deploying resources or as an ultimate goal. But for U.S. presidents and their administrations, reliance on data has attendant advantages, quite apart from the operations and decisions that unfold—and these serve to deflect atten-

tion from problems, including difficult ethical challenges, that accompany the data-driven presidency. Because of this, the presidency needs the rare combination of expertise and detachment to yield effective decision making, in order to weigh the traps and biases associated with this world of data against the advantages.

This article focuses primarily on the role of data in the Obama administration, in structuring processes and decision-making, with cases drawn from three different domains—presidential selection, internal governance, and tactical national security decision-making. Early indications for the Trump presidency suggest consistency in the role of data, despite some signals to the contrary. Taken together, these cases demonstrate how reliance on data is buttressed by recurring calculations that emphasize efficiency, and in some cases a private sector that provides the data, at an extreme even extracting data from individuals without compensation. In other words, the political economy of data in the presidency helps explain its modern allure—and recognizing it can also inform prudent action in the future.

2. Epistemological and Political Roots

The turn to data in the U.S. federal executive is consistent with early twentieth century ideas, both epistemological and ideological. Emphasizing data has roots in logical positivism, which envisioned knowledge as the result of empirical evidence. In the U.S., the ideological counterpart to positivism was turn-of-the-century Progressivism, valuing information and expertise as tools to disrupt the stranglehold of moneyed interests and patronage-fueled political parties on political power (Greider, 1992; Orren & Skowronek, 2017). But as the century advanced, information and expertise in the domain of the executive became vehicles not for disruption, but for leadership by the president in the U.S. separated system, with a massively-expanded state apparatus. In other words, data and expertise became instrumentally valuable in politics.

Like the political world, the academy pivoted to embrace empirical evidence. About when Progressive political sentiment took hold, reformers in the American academy moved to establish a distinctive approach to the study of politics and government, shedding “the legalist and theoretical way in which political life was studied in the European academy” (Susser, 1992, p. 4). The result was a new discipline of political science, at times with some academics inserting themselves into the rough and tumble of politics. But the more pervasive quality of this new approach—which was fully formed by the 1960s—was its social-scientific orientation, emulating the scientific model and placing a premium on empirical, especially quantitative, evidence.

Against this backdrop, contemporary applications of data in the U.S. presidency—that is, reliance on evidence-driven practices and data—are not fundamentally new. Indeed, well before “analytics” and “big data” emerged, there was a strong element of data-based politics in the U.S., extending over the nation’s entire history and with remarkable scope. From census data informing allocation of congressional representation starting at the founding, to data-fueled economic projections mandated by the New Deal, and even to Ronald Reagan-as-president using polling data to refine his rhetoric, the historical examples are abundant. However, politics and governing in the U.S. has reached a critical juncture, with reliance on data so pervasive that it’s difficult to imagine anything but; it has become the default choice, the go-to solution for decisions, management and administration. With this in mind, the following section describes how data are employed in three different domains of the contemporary U.S. presidency, beginning with campaign politics, marked by what can justifiably be called big data.

3. Data and Big Data under Obama

3.1. *The Data Science of Campaigns*

While the road to the White House had been paved with data and analytics for some time, the near obsession

of Barack Obama’s two campaigns with evidence-based practices represents a difference in kind. What’s more, the campaigns’ successes impelled the wide diffusion of the data-centric campaign model. Even Donald Trump, who conveyed skepticism about data—at times eschewing it—subscribed to fundamentals of a data-driven campaign model.

The predominant narrative of both Obama’s 2008 nomination and general election wins emphasized that the campaign’s data-driven operation successfully mobilized voters, especially new ones, to the polls. In 2008, a ground game flush with money fueled a sophisticated data-rich field operation, enhanced by online capacity which included new platforms to engage voters. The 2012 addition to the narrative emphasized that the re-election campaign was metric-driven and informed by the insights of social scientific research. Both campaigns fundamentally ran on data, not unprecedented in approach, but certainly in scope.

The data at the heart of mobilization efforts are voter lists, used by campaigns to identify potential supporters and mobilize them—through direct contact—to the polls. These basic lists are longstanding, in fact the byproduct of the early Progressive Era introduction of voter registration. Ironically, the information collected by this turn-of-the-nineteenth-century reform, meant to weaken political parties, became the raw material for the mobilization efforts of the parties and their candidates. Before the advent of polling, these lists provided a rare portrait of relative party strength among the electorate as well as measures—like party affiliation and demographics collected by the state—that could inform mobilization efforts (Hersh, 2015, p. 49).

Obama’s voter contact data were simply an advanced version of those early lists, but digitally enhanced and readily operational through a user interface. The campaign used “VoteBuilder”, the Democratic Party’s proprietary data, accessed through a user interface purchased from a left-leaning for-profit, NGP-VAN. VoteBuilder lists offer—at their core—the same type of information that in a prior era a party agent might have retrieved from official voting records, namely voting history and demographics of the registered vote. Now these lists are augmented to include additional individual-level information about the voter, drawn from a number of sources, including commercial firms, as well as parties and campaigns themselves which glean information from field staff and volunteer interactions with voters.

By 2008, the use of data like these was common, not just for presidential campaigns. So too were forays in microtargeting, procedures to further augment data by means of statistical analyses, a process that had been evolving over the prior decade. Microtargeting techniques produce synthetic measures of voter characteristics—“model scores”—by means of predictive analytics, integrating the results of large-n survey data with the augmented voter file. The model scores serve as criteria for a particular voter contact effort, tar-

geted to specific individuals. Indeed, the “micro” aspect of this enterprise involves data analysis at the *individual* level, with targeted contact efforts similarly aimed. It is, in effect, an algorithmic decision-making process, though the practice began well before the term was applied to it.

Obama in 2008 engaged in a virtually uninterrupted process of modeling and refining the data, and then modeling again (Issenberg, 2012a). Progressive data giant Catalist, another for-profit firm partnering with the Obama effort, extended the data beyond the traditional cache of registered voters to include unregistered voters, not represented in traditional voter files. The Catalist data offering, according to the firm’s own accounts, numbered 265 million cases, reflecting approximately the universe of voting-age adults in the U.S.

Conventional wisdom holds that the prowess of Obama’s resource-flush 2008 campaign contributed heavily to his win. But by 2012, the reelection efforts would be enhanced significantly by an evidence-based understanding of the effectiveness of voter mobilization techniques, drawing heavily from experimental research with ties to the social-scientific community. The 2012 campaign, gripped by a culture of experimentation, employed evidence-informed programs (IEPs) not only in direct voter contact protocols used in the field, but also in digital and fundraising campaigns (Issenberg, 2012b).

The 2012 campaign was structured to give data and analytics a strong voice. At the Chicago Obama headquarters, a team of fifty analytics professionals worked out of the “the cave”, with a direct line of reporting to the Chief Innovation and Integration Officer, who reported directly to the campaign manager. Data and technology departments constituted an estimated 30–40% of headquarters staff, and “[a]nalytics was the breakout star of 2012” (Engage, 2012). The data-focused structures and practices were complemented by the attitudes and norms of personnel, including senior staff with a willingness “to listen to numbers people rather than consultants acting on old-fashioned political intuition” (TechPresident, as cited in Engage, 2012.)

These 2008 and 2012 campaigns represented state-of-the-art data operations—in fact “big data”, at least in the sense that they integrated data from a variety of sources, augmenting them repeatedly with newly-acquired information. The data even approached an $N = \text{all}$ (Mayer-Schöenberger & Cukier, 2013) quality, in Catalist’s case approximating the universe of the voting age population. But not all data that fuel the presidency are big in this sense. In fact, data collected under management protocols are rather conventional, despite being part of a monumental data-collection enterprise.

3.2. Tracking the Executive with Performance Management

Performance management in the abstract focuses on how efficiently and effectively the executive branch administers the programs of the federal government. It’s

data-driven management, which—much like the use of data in campaign politics—has become systematic and elaborate over time, especially since the 1990s. The data at the heart of performance management as practiced in the Obama years were collected by agencies, permitting judgment of the extent to which the outcomes of their activities met the goals of the programs and of the administration.

Most observers trace the development of performance management to the Bill Clinton era, though the impetus to employ management to consolidate the president’s leadership of the executive branch came earlier, during the Richard Nixon administration. Nixon, in an effort to harness the discretion of the federal agencies and to ensure agreement across the administration with his policy priorities, layered on management responsibilities to the existing budget office, creating the Office of Management and Budget (OMB) in 1971. While OMB offered an institutional arm to the president for management, it wasn’t until later, in 1993, that performance management as a systematic approach was codified by Congress in the 1993 Government Performance and Results Act (GPRA), requiring federal agencies to engage in strategic planning every five years and undertake annual performance reviews. And then in 2010, Congress passed the GPRA Modernization Act (GPRAMA), which revised specific expectations for performance management, including movement from annual to quarterly reviews/reporting, making performance management an ongoing process. This world is data-heavy, requiring that agencies establish goals and track progress toward achieving them. The initial statute instructed agencies to develop “quantifiable and measurable program targets” as well as “outcome measures”, metrics by which the real-world success of the programs would be judged (Harris, 2015, pp. 105–106).

In its rhetoric, the Obama administration embraced performance management. It built onto the efforts of the Bush administration (Jochum, 2009), which had itself prioritized performance management and had devised and touted its Performance Assessment Rating Tool (PART), a quantitative assessment of goals and performance used by over 200 federal programs, estimated to account for 20% of the federal budget. Jeff Zeints, the acting director of OMB at the start of the Obama administration, described GPRA and Bush’s PART as important starting points for the new administration. Ratcheting up the hype, Obama’s performance management was spearheaded by “performance guru” Shelley Metzenbaum, who was responsible for developing www.performance.gov, a tool to both articulate the administration’s approach and to provide access to copious reports and reviews filed under the program.

A wide variety of data is collected under performance management protocols, with each agency establishing annual performance goals for its mission areas, identifying metrics for assessing the goals, and then reporting the actual performance. Much of the data is straightfor-

ward, the product of counting and tracking. For example, the U.S. Department of Agriculture (USDA) reports the number of wetland acres (in millions) restored by the Conservation Reserve Program (CRP), a measure compiled from CRP contracts. But other measures entail estimation procedures, requiring technical expertise and skill, with results that may be marked by irregularities. Consider USDA's metric for enabling access to healthcare facilities for rural areas. Measured as the percentage of people who are provided access to new and/or improved essential community facilities, this metric involves estimation of a geographic service area for each facility and the population within it, both difficult to assess (USDA, 2013). Messiness aside, this is not "big data" by any conventional definition, though monumental in scope, since every agency collects measures to assess a large number of goals. To get a sense of the magnitude, consider the 2013 performance management reports for Commerce and USDA alone, each listing over forty metrics, while Housing and Urban Development (HUD) reported more than fifty. Twenty-four agencies are subject to this data-driven management process.

Data collected by the General Accountability Office (GAO) offer a valuable window to the experiences of agency managers in performance management, and the 2013 Federal Managers Survey reveals a mixed picture about the implementation of performance management. Administered in late 2012 and early 2013, the survey sampled a population of some 148,000 managers and supervisors in agencies that undertake performance management as specified by the two statutes, producing approximately 3,000 usable responses. The survey found that a full 48.8% had never even heard of the GPRAMA. Just under one-fifth (18.8%) of managers suggested that performance data were not easily accessible to them, while close to one-third (31.2%) felt the data were not easily accessible to their employees. When asked whether data were formatted in a manner easy to use, 27.7% reported they were not, and 30.2% expressed that they did not have sufficient analytic tools to collect, analyze and use the data (GAO, 2013). These indicators can be taken as glass half-full or half-empty, but regardless suggest that the reality of data-based performance management under Obama might not have met the promise of the accompanying build-up.

That said, performance management is transparent, with not only the data publicly available, but also the perception of the practitioners revealed systematically. Furthermore, the data are quite traditional, metrics with which to judge performance and, presumably, instruct decision makers. The following example offers a dramatic contrast, drawn from the realm of national security, involving the specification of targets for military attack.

3.3. High-Stakes, Covert Signature Strikes

To base decisions about whom to attack on evidence is unsurprising. But the Obama administration employed a

controversial technique to target individuals or groups of individuals for drone strikes when they bore the characteristics—the "signature"—of those likely to be engaged in terrorist activity. In contrast to "personality strikes", in which the U.S. targets known terrorists (Zenko, 2012a), signature strikes are based on patterns of behavior indicative of terrorists, even if the individual target is not known to be a terrorist. President George W. Bush was the first to authorize such signature strikes (Zenko, 2013, p. 12), though this came at the end of his tenure.

These strikes are shrouded in secrecy, extreme even compared to the typical opacity of national security. Micah Zenko notes that signature strikes have not been acknowledged officially. "[N]o U.S. Government official has ever acknowledged the practice of signature strikes". Nor has any official "described the practice, justified it, or explained how it is consistent with the...laws of war" (M. Zenko, personal communication, 26 April 2018). Even more so, information about precisely what data and analytical tools inform the targeting is sketchy, with what is known owing largely to Edward Snowden's June 2013 leaks of National Security Agency (NSA) data.

Snowden revealed that the NSA mines metadata, essentially the trail that follows digital and cell-phone communication (Hu, 2017, p. 235). Even absent the content of the communications, these metadata allow the analyst, most likely relying on a combination of machine learning and network analytical techniques, to identify potential terrorists by their patterns of connections to others, including to known terrorists. Journalist Glenn Greenwald emphasizes that the validity of the data is not confirmed by traditional techniques, like engaging "operatives or informants on the ground" (Scahill & Greenwald, 2014). This threat of imperfection is captured in the oft-repeated comment, attributed to an unnamed State Department official: "[T]he C.I.A. sees 'three guys doing jumping jacks'...[and] thinks it is a terrorist training camp, [adding that those] loading a truck with fertilizer could be bombmakers—but they might also be farmers" (Becker & Shane, 2012).

Making decisions based on incomplete evidence is not new to the world of military tactics. Zenko (2012b) recounts an anecdote conveyed by General Colin Powell about his early history in Vietnam:

I recall a phrase we used in the field, MAM, for military-age male. If a [helicopter] spotted a peasant in black pajamas who looked remotely suspicious, a possible MAM, the pilot would circle and fire in front of him. If he moved, his movement was judged evidence of hostile intent, and the next burst was not in front but at him.

The pilot in Powell's account makes a judgment based on the available evidence, flawed as it might be. The drone strike likely reflects similar judgment—but with split-second processing, the application of algorithms to data, perhaps even real-time geo-location data. No-

tably missing in this endgame is direct human judgment. Hu (2017, p. 231) asserts that the absence of human judgment distinguishes this big-data approach from the “small-data” methods of the past, which relied on human perception and human decision making. The other significant element is that the algorithmic process yields a quantitative measure of likelihood that a person is a terrorist—or that a targeted geographic space would encompass terrorists.

Critics find especially concerning that the decision to kill is based on a likelihood generated by an algorithm. But perhaps more problematic is the high civilian death toll associated with drone strikes. Data from New America Foundation (NAF), Long War Journal (LWJ), and The Bureau of Investigative Journalism (TBIJ), mostly from 2004–2012, estimate over 400 drone strikes in Pakistan, Yemen and Somalia, with approximately 12% (401) citizens among the 3,430 killed (Zenko, 2013, p. 13.)

Advances in technology—drones, the widespread use of mobile technology, as well as the ability of the NSA and CIA to track and analyze the exhaust—have opened the door for signature strikes. And while the details of the data and analytics that undergird them escape public scrutiny, it’s clear that this is an executive branch big-data enterprise, not just in terms of volume, but also in the substitution of machine judgment for human judgment.

4. *Plus Ça Change...in Early Trump?*

Donald Trump, despite being an unconventional candidate and president, over the two years of his administration has signaled that in many respects he follows in the footsteps of his predecessor regarding the use of data. This orthodoxy, however, is especially notable given that the president has at times vocally eschewed evidence-based practices. Plenty of time remains for the Trump approach to data to take shape, but at this juncture it looks like rhetoric does not always mesh with actions.

Candidate Trump expressed disdain for campaign data, calling it “overrated” (Vogel & Samuelsohn, 2016b), but then assembled a rather conventional voter data and mobilization effort, admittedly smaller and flying under the radar more so than his predecessor’s (Vogel & Samuelsohn, 2016a). By early 2016, the operation was staffed by two former Republican National Committee (RNC) operatives, low key in orientation, but with experience working with the RNC’s Voter Vault, the counterpart to the Democratic National Committee’s VoteBuilder, which had fueled Obama’s and—eventually—Hillary Clinton’s campaigns. Largely undetected, the Trump campaign assembled “Project Alamo”, an ambitious digital database that aided in online and offline targeting, strategic decisions and voter mobilization—as well as a dose of voter *demobilization*, attempting to limit the Hillary Clinton vote (Green & Issenberg, 2016).

And then there was Cambridge Analytica. Trump turned to the U.K.-based data and analytics firm, which

was later revealed to have misappropriated Facebook data for the purpose of its “psychographic modeling” activities. The New York Times, working with London’s Observer, reported that Cambridge Analytica, with close ties to central figures in the Trump orbit like Steve Bannon and the Mercers, acquired personal information on Facebook users by means of an academic, who claimed the data were for the purpose of academic research (Rosenberg, Confessore, & Cadwalladr, 2018).

While the tangled web of the Facebook data breach and possible connection to Russian collusion remained unresolved by late-2018, finance reports confirm that Cambridge Analytica was a player in the Trump data operation. Federal Election Commission (FEC) records from 2016 compiled by the Center for Responsive Politics (CRP) show disbursements of \$5.9 million from the Trump campaign to the data firm. Notably, these disbursements were dwarfed by the \$87.8 million paid to Giles-Parscale, the San Antonio digital marketing firm that was responsible for Project Alamo.¹

Trump’s embrace of data in the campaign phase is replicated in management of the administration, and there’s even sign of the same Obama-era promotional voice. However, the ends to which Trump’s performance management are directed are distinctively-Trump: to limit the reach of the federal government. Of course, this same goal is advanced by the record-number of key appointed positions in the executive branch unfilled well into the term (Kruzel, 2018) and the marginal shrinking—through attrition—of the size of the civilian work force (Jacobson, 2018). Not surprisingly, President Trump’s approach to performance management, while similar in its practices to Obama’s, aspires to a business model, envisioning the citizens as customers and holding federal employees accountable. Margaret Weichert, Deputy Director for Management at OMB, sees a central role for data in this enterprise, with “drivers” of the agenda being information technology, data accountability/transparency, as well as a modern workforce (Clark, 2018, p. 16). But if not for prototypically Trump-like messages signaling disdain for career bureaucrats, advocating streamlined processes to remove poor performers, and pushing back at unions (Katz, 2018, p. 7), this technology-driven emphasis might well have come from Trump’s predecessor.

Similarities between Obama and Trump regarding use of data extend to drone strikes as well. The day after inauguration, Trump authorized the use of strikes in Syria though departing from Obama national security processes, which reserved the strike capacity for the Pentagon. Under Trump, the CIA both collects the intelligence that fuel the targeting and carries out the strikes. The turf maneuvering between the Pentagon and CIA could have real ramifications, since they reportedly employ different standards of algorithmic certainty, with the CIA’s “near certainty” decision-rule more demanding than the Pentagon’s “reasonable certainty” (Lubold

¹ For more information see www.opensecrets.org/pres16/expenditures?id=n00023864

& Harris, 2017). At the same time, empowering the CIA to conduct the strikes removes the process even more so from the scrutiny of congress and the courts, with CIA activities, relative to the Pentagon's, shielded more from view. Put differently, the data and analytics may remain the same, but process differences could have a real impact.

Still, in other areas that move beyond the cases explored in this article, Trump has taken aim at data. He famously banned the Center for Disease Control (CDC) from using terms like "evidence-based" and "science-based" (Sun & Eilperin, 2017), removed data from the website of the Environmental Protection Agency, and disbanded advisory councils that might challenge his own beliefs about the climate and the economy. Furthermore, Trump purged from the web the White House Visitor Log, citing national security risks; the machine-searchable list of visitors to the White House and the Eisenhower Executive Office Building (EEOB) had been made readily available under the Obama administration in the interest of transparency and probably not much of a national security threat. Not surprisingly, given signals on both sides of the question, something of a debate still wages about whether Trump carries on a "war on data".

5. Theoretical Insight Regarding Data and the Presidency

This article's description of data and evidence-based approaches used in the U.S. presidency is necessarily incomplete, dependent on a handful of cases, focusing on one presidency with some insight into another. Indeed, the portrait of the prominence of data in the presidency is if anything modulated by these cases. Performance management is a bit of a sleeper, and the data-driven campaign model, since showcased by Obama, has been diffused widely, across parties, down the ballot and even to campaigns in different international settings, so much so that even it is a little *passé*.

But Obama did deploy data in far more contexts than described in this article. Technology reporter Nancy Scola dubbed him the "big data president", with some eighty-five big data projects ongoing in his time in the executive (Scola, 2013). Under Obama's watch, the National Institutes of Health (NIH), worked to facilitate delivery of healthcare targeted to a patient's unique genetic makeup, with the goal of collecting data from one million volunteers (NIH, n.d.). The Justice Department (2016), in conjunction with the Whitehouse and law enforcement agencies, released data on police actions, "to increase transparency and accountability and build trust with...communities". And as documented—along with other projects—by the Executive Office of the President (2012), "Mission-oriented Resilient Clouds" would detect and respond to security threats in cloud computing.

The allure of these programs—as well as that of the cases described in this article—is readily apparent, given their stated ambitions. And precisely because of natural

allure we should proceed warily, to remember that a critical lens would caution against ignoring the underlying assumptions and power relationships that undergird the processes related to data. With this in mind, this article returns to the three cases, introducing a focus on the political economy of the data.

5.1. A Presidential Data "Revolution"

At first blush, the data-heavy model of campaign politics, performance management and drone strikes under Obama all entail a conceit that the practices are democratizing. After all, a microtargeting process resulting in direct voter contact—campaign personnel reaching out on the phone and at the door—is a decided departure from the mass media model of campaigns that had become prominent over the final decades of the twentieth century. The democratic nature of performance management in the executive is a little different, but it entails the ability of the electorate, as mediated by representatives in Congress, to hold the vast unelected bureaucratic state accountable. Even signature drone strikes, obscured from the view of the public and most elected officials, arguably have an attendant democratic sense. Protecting the U.S. military from ground combat, as uncertain as that is up against non-state opponents in the fight against terrorism, is democratically significant in that military personnel are disproportionately drawn from lower economic classes.

In each of these cases, the data and analytics operations exploited new technologies. Granted microtargeting was around long before Obama (Malchow, 2003), and even the basic architecture of the data employed by Obama was already in place (Hersh, 2015; Kreiss, 2016). But the extent of processing power and the servers, especially in the Catalist world of data mining and modeling with essentially a universe of cases, were fundamentally new and characteristics of big data. In contrast, the technology of data collection in performance management was not cutting edge, but the dissemination of data was. The Obama Administration prioritized the distribution of data, with its performance.gov portal, along with the heralded data.gov portal, which in late October 2015 offered some 189,000 data sets and as of late 2018 over 300,000 data sets to the public. As for signature drone strikes under Obama, it wasn't so much the new tool of drones, because unmanned aerial vehicles have a long history. It was the ability to equip the operation with digital and mobile data, abundant on the ground, then mined, integrated and analyzed to inform algorithmic decisions.

These three data applications, beyond holding democratic allure and the draw of new technology, have the added appeal of secondary instrumental benefits. Regarding drones strikes, a variety of polling data shows that the American public is not particularly critical of them, and presumably successful strikes, not putting American personnel directly at risk, secures stronger public support. A similar byproduct accompanies perfor-

mance management. Early applications of the practice did result in an upswing of trust of government (Kamarck, 2013). And political campaigns realize multi-faceted instrumental benefits from data and metric-driven efforts. The metrics are used to motivate volunteers and staff and to hold them accountable. Even more so, they serve as concrete and persuasive evidence for donors of the impact and promise of the campaign in the absence of more definitive measures like election outcomes.

It may only be a slight exaggeration to suggest these uses of data in the presidency hold some apparent revolutionary potential. Not only do the data contribute to a desired outcome, but they purport to even change usual power dynamics—offering voice to those not typically heard, as well as a new role for or protection of the average American. But this same revolutionary potential makes it easy to glance away, ignoring inherent biases and threats associated with data in general and more specifically in the presidency.

5.2. *The Political Economy of Presidential Data*

Three common threads are woven through the cases examined, and they expose concerns that deserve to be addressed in both decisions to turn to data and evaluation of success. The first of these is the uncritical acceptance of efficiency as a goal, most directly borne out in performance management, with its emphasis on outcomes and the use of metrics to judge success.

The approach to performance management manifest in government used by of Obama—Trump too—and predecessors was adopted initially from the private sector in the 1970s. The “New Public Management” aspired to “create market like conditions within the government...to run them ‘more like a business’” (Muller, 2018, p. 51). But despite numerous shortfalls of this market application in governing—like metrics distorting incentives and representing overly simplistic conceptions of what motivates personnel—the practice was well-entrenched by the late 1990s (Muller, 2018, p. 55) and it continues today. Muller (2018) finds that metrics that drive management often operate perversely, drawing attention to only things that can be measured and even stifling innovation. There is, however, even a more fundamental concern with this business model, in that it poses efficiency as a preeminent goal.

Efficiency is a common metric employed in the private sector, but its adoption in politics and governing may be at the expense of other things valued. Microtargeting in campaigns, for example, is premised on the efficient deployment of resources to mobilize and persuade enough voters to win an election. But the flip side of targeting voters is that some are ignored, deemed either lost causes or even certain supporters, neither warranting attention by the campaign. And while this may effectively carry a candidate across the line in a given election, it represents a narrow, short-term focus that may not contribute to building an electorate that will support the

party in the future. Sociologist Robert Merton called this “the imperious *immediacy* of interests” (Merton, as cited in Muller, 2018, p. 170, emphasis added), wherein individuals look only as far as the short-term consequences of their action.

Zeynap Tufekci’s (2012) problem with the efficiency in campaigns is a little different, namely that they *will succeed* in efficiently engineering the electorate. Tufekci is first concerned that the “scalpel” of microtargeting is deployed in private, not subject to public scrutiny. But the bigger problem is that it just may be effective, especially for well-financed campaigns with the resources to devote to data and persuasive techniques. Even if not effective—even if the data which guide the appeals are flawed and replete with errors, as any staffer or volunteer who has worked with these data knows—that campaigns are treating the electorate as a target of their engineering efforts is itself a cause for worry.

Data enterprises that posit efficiency as a goal is a first thread that runs through the cases. A second thread is that the data-based presidency is inexorably tied to a private sector that both supports and benefits from it. The interface and sometimes the data that the campaigns use are held in private hands. NGP-VAN and Catalist, the left’s go-to data interface and source of mined data respectively, along with thousands of other paid vendors, constitute the for-hire network of data professionals, many of whom move back and forth between the campaigns and the party apparatus from election to non-election seasons (Kreiss, 2016). It’s notable that the combination of the Democratic data and the privately-held NGP-VAN interface, according to Kreiss (2016), serves as a “robust piece of infrastructure that the party’s technology ecosystem convenes around”. In other words, the private data actually structure the party organization. And in a related fashion, the dependence of a political party and its campaigns on a small number of private firms cannot help but affect where power rests within the party organization, not necessarily with the voter or the party elites, but with vendors.

Performance management, like the data and analytics in campaigns, is subject to a revolving door of sorts regarding leadership. While not universal, a common pattern is that top personnel responsible for performance management, and OMB directors as well, are drawn from the private sector—or at least from those with experience in the private sector. It’s also the case that many of these management leaders return then to the private sector after service. Admittedly, the career of Shelley Metzenbaum, President Obama’s “performance guru”, was more entangled with academia and other governmental positions than the private sector. But Trump performance management leader Margaret Weichert demonstrates a clear trajectory into government from the private sector.

The undercurrent of values and practices that inject a market-based influence into data in the presidency extends to a third dimension as a well: the transaction

marking the exchange of data. This is evident in the data used by campaigns and in signature strikes, with the first involving an implicit transaction with the state and the second with individuals, granted, many outside of the boundaries of the U.S.

The campaign data originate in lists of registered voters compiled by the U.S. states. As a condition of voting, individuals provide data to the state, but then private firms like NGP-VAN offer user-friendly tools for working with the data to parties and campaigns. Or in the case of Catalist, the data are augmented through integration with other sources, using algorithmic processes to add synthetic measures (Hersh, 2015). In the abstract, these data and tools, the likes of which are used by presidential candidates, represent an implicit transaction, one in which data collected by the state and made available at little cost is collected by businesses, then sold in a repackaged form to political organizations. Of course, compensating the intermediaries for the value added to the data seems only right. Yet it introduces the question of whether there is just compensation for the original data provider.

For Phil Howard (2018) the answer is “no”, at least with reference to the big social media players and political mining firms like Cambridge Analytica, which extract data with ease. Howard is concerned that the citizens have no effective control over their data, which will be used for political purposes. Among the mix of Howard’s recommendation to put some degree of control back in the hands of the public is that individuals should be able to donate their data to “the civic groups, political parties, or medical researchers they want to support” allowing them to leverage their own data for political purposes.

This transactional sense of data emerges in Evgeny Morozov’s (2017) analysis of artificial intelligence (AI) as well. Morozov asserts that the compensation received by individuals for the data that fuel AI research and development—compensation that is nothing more than access to a social network—is modest when considering the price that government and individuals will pay for products created by AI. Signature strikes invoke an element of this same transactional logic, with a troubling addition. The data that fuel the strikes represent the digital exhaust of users on the ground, snatched up by surveillance operations. In this, access to internet and mobile technology is the compensation for the user, which admittedly may be of substantial value. But that the data are then deployed to target for the purpose of killing individuals with only some stated degree of statistical certainty, and that this practice captures innocent bystanders as well, has an element of perversity to it. Admittedly, national security and covert operations are not the same as AI enterprises, and it’s absurd to suggest that those being surveilled should be better compensated for their data. But this transactional calculus regarding signature strikes, just like those implied or described by Kreiss (2016), Howard (2018) and Morozov (2017), at a minimum, points to the merits of looking well beyond

the effectiveness of the data and processes as measured by numbers of terrorist killed.

5.3. Moving Forward

This article has suggested that the world of data—including big data—is borne out in the U.S. presidency, in some cases accompanied by the buzz that this is fundamentally new, even to the point of revolutionary, potentially disruptive of traditional power arrangements. But the subtle irony is that viewed through a critical lens, those traditional power arrangements may prevail, in some cases enhanced by the perceived revolutionary potential of the data and data-related processes used in the presidency.

But one need not focus on the political economy of data to identify ways in which the popular understanding of data and the potential they hold are entangled with political and ethical concerns. Consider the alarm generated by looking closely at algorithms, challenging the conceit that they are immune from prejudice. Cathy O’Neil (2016) demonstrates how algorithmic-informed decisions can reinforce the existing biases of society, that policing tools using predictive modeling carry the appearance of objectivity but can be “tools of math destruction”, perpetuating existing traditional class biases. What’s more, data journalism outlet ProPublica offers a telling rejoinder to statisticians who judged as fair the algorithms used in sentencing recommendations, finding that when applied to actual people, the algorithms systematically overestimate the threat of recidivism for African American defendants and underestimate it for White defendants (Caplan, Donovan, Hanson, & Matthews, 2018).

Ethical concerns come into the picture typically in ways subtler than the life/death calculations marking signature drone strikes, and because of this they slip by undetected, especially when safety procedures don’t intervene successfully. Cambridge Analytica’s misuse of data was facilitated by Facebook permitting academic Aleksandr Kogan to harvest its user data, despite Kogan’s research proposal being rejected by his university’s ethics board (Weaver, 2018). Even with a definitive say by an Institutional Review Board, these bodies tend to gravitate toward a legalistic review of proposals, and—furthermore—are frequently ill-equipped to tease out the ethics of big data (Metcalfe, Keller, & Boyd, n.d.). It goes without saying that many of the decisions regarding data escape scrutiny by experts tasked with reviewing ethics, especially in the realm of politics and government.

This is all to say that from a number of perspectives—whether viewing data deliberately through a critical lens or simply examining the current areas of concern regarding data and big data—it’s clear that the U.S. presidency faces substantial data-related challenges. Despite the occasional utterance of Trump that calls expertise into question, the U.S. political system continues to value

information and expertise, both of which contribute to the political capital of actors and institutions. And the norms of science still prevail, indeed even with some new role for social-scientific applications in the political world. In short, there is no reason to believe that “data-driven” is a passing phase. But it’s time to contemplate what a measured approach to data would look like. To be concrete, the goal should be to deploy data in effective and ethical ways, all the while alert to the biases and shortcomings that underlay their collection and use. This is no small task, especially in an environment that routinely prioritizes quick action over deliberation, particularly on matters that may require extraordinary technical expertise, though sound and detached political judgment as well.

Acknowledgments

The author thanks colleagues Michael Guenther (History), Janet Seiz (Economics) and Eliza Willis (Political Science) for sharing their insight.

Conflict of Interests

The author declares no conflict of interests.

References

- Becker, J., & Shane, S. (2012, May 29). Secret ‘kill list’ proves a test of Obama’s principles and will. *The New York Times*. Retrieved from www.nytimes.com
- Caplan, R., Donovan, J., Hanson, L., & Matthews, J. (2018). Algorithmic accountability: A primer. *Data & Society*. Advanced online publication. Retrieved from datasociety.net/output/algorithmic-accountability-a-primer
- Clark, C. (2018). Improving performance (in Trump’s government makeover). *Government Executive*. Retrieved from www.govexec.com
- Engage. (2012). Inside the cave. *Engage*. Retrieved from enga.ge
- Executive Office of the President. (2012). *Big data across the federal government*. Washington, DC: Executive Office of the President. Retrieved from obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_final_1.pdf
- General Accountability Office. (2013). *Managing for Results: 2013 federal managers’ survey on organizational performance and management issues* (GAO-13-519SP Report). Washington, DC: General Accountability Office.
- Green, J., & Issenberg, S. (2016, October 27). Inside the Trump bunker, with days to go. *Bloomberg*. Retrieved from www.bloomberg.com
- Greider, W. (1992). *Who will tell the people?* New York, NY: Simon & Schuster.
- Harris, S. (2015). Managing for social change: Improving labor department performance in a partisan era. *West Virginia Law Review*, 17, 100–158.
- Hersh, E. (2015). *Hacking the electorate: How campaigns perceive voters*. New York, NY: Cambridge University Press.
- Howard, P. (2018, July 16). Our data, ourselves. *Foreign Policy*. Retrieved from foreignpolicy.com/2018/07/16/our-data-ourselves-democracy-technology-algorithms
- Hu, M. (2017). Metadeath: How does metadata surveillance inform lethal consequences? In R. Miller (Ed.), *Privacy and power: A transatlantic dialogue in the shadow of the NSA-affair*. Cambridge: Cambridge University Press.
- Issenberg, S. (2012a). *The victory lab*. New York, NY: Crown Publishers.
- Issenberg, S. (2012b, May 22). The death of the hunch. *Slate*. Retrieved from www.slate.com
- Jacobson, L. (2018, January 22). Taking the measure of the federal workforce under Donald Trump. *Politifact*. Retrieved from www.politifact.com/truth-o-meter/article/2018/jan/22/taking-measure-federal-workforce
- Jochum, E. (2009, September 24). OMB will create new performance management framework for agencies. *Government Executive*. Retrieved from www.govexec.com
- Justice Department (2016). Growing number of communities are using data to improve policing and criminal justice. *Department of Justice Archives*. Retrieved from www.justice.gov/archives/opa/blog/growing-number-communities-are-using-data-improve-policing-and-criminal-justice
- Kamarck, E. (2013). Testimony before House Committee on oversight and government reform: Lessons for the future of government reform. *Governance Studies at Brookings*. Retrieved from www.brookings.edu
- Katz, E. (2018). Modernizing the workforce (in Trump’s government makeover). *Government Executive*. Retrieved from www.govexec.com
- Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. New York, NY: Oxford University Press.
- Kruzel, J. (2018, March 16). Why Trump appointments have lagged behind other presidents. *Politifact*. Retrieved from politifact.com
- Lubold, G., & Harris, S. (2017, March 13). Trump broadens CIA powers, allows deadly drone strikes. *Wall Street Journal*. Retrieved from www.wsj.com
- Malchow, H. (2003). *The new political targeting*. Washington, DC: Campaigns and Elections.
- Mayer-Schöenberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. Boston, MA: Houghton Mifflin Harcourt.
- Metcalfe, J., Keller, E., & boyd, d. (n.d.). Perspectives on big data, ethics, and society. *Council for Big Data, Ethics, and Society*. Retrieved from bdes.datasociety.net
- Morozov, E. (2017). *Do we have a right to our data? Data ownership and the inequality debate*. Public address

- at Grinnell College, USA.
- Muller, J. (2018). *The tyranny of metrics*. Princeton, NJ: Princeton University Press.
- National Institutes of Health. (n. d.). About the All of Us research program. *National Institutes of Health*. Retrieved from allofus.nih.gov/about/about-all-us-research-program
- O’Neil, C. (2016). *Weapons of math destruction*. New York, NY: Crown.
- Orren, K., & Skowronek, S. (2017). *The policy state: An American predicament*. Cambridge, MA: Harvard University Press.
- Rosenberg, M., Confessore, N., & Cadwalladr, C. (2018, March 17). How Trump consultants exploited the Facebook data of millions. *The New York Times*. Retrieved from www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html
- Scahill, J., & Greenwald, G. (2014, February 10). The NSA’s secret role in the U.S. assassination program. *The Intercept*. Retrieved from theintercept.com/2014/02/10/the-nsas-secret-role
- Scola, N. (2013, June 14). *Sizing up the executive branch: Fiscal year 2017*. Washington, DC: United States Office of Personnel Management. Retrieved from www.opm.gov/policy-data-oversight/data-analysis-documentation/federal-employment-reports/reports-publications/sizing-up-the-executive-branch-2016.pdf
- Sun, L., & Eilperin, J. (2017, December 15). CDC gets list of forbidden words: Fetus, transgender, diversity. *The Washington Post*. Retrieved from www.washingtonpost.com/national/health-science/cdc-gets-list-of-forbidden-words-fetus-transgender-diversity/2017/12/15/f503837a-e1cf-11e7-89e8-edec16379010_story.html?noredirect=on&utm_term=.5bc086fc9789
- Susser, B. (1992). *Approaches to the study of politics*. New York, NY: Macmillan.
- Tufekci, Z. (2012, November 16). Beware the smart campaign. *The New York Times*. Retrieved from www.nytimes.com/2012/11/17/opinion/beware-the-big-data-campaign.html
- USDA. (2013). *FY 2013 annual performance report*. Washington, DC: USDA. Retrieved from www.ocfo.usda.gov/docs/FY%202013%20Annual%20Performance%20Report.pdf
- Vogel, K., & Samuelsohn, D. (2016a, January 5). Trump quietly builds a data juggernaut. *Politico*. Retrieved from www.politico.com
- Vogel, K., & Samuelsohn, D. (2016b, June 28). Trump’s secret data reversal. *The New York Times*. Retrieved from www.nytimes.com/2012/11/17/opinion/beware-the-big-data-campaign.html
- Weaver, M. (2018). Cambridge University rejected Facebook study over ‘deceptive’ privacy standards. *The Guardian*. Retrieved from www.theguardian.com
- Zenko, M. (2012a). *Daniel Klaidman’s revelations*. New York, NY: Council on Foreign Relations.
- Zenko, M. (2012b). *Targeted killings and signature strikes*. New York, NY: Council on Foreign Relations.
- Zenko, M. (2013). *Reforming U.S. drone strike policies* (Special Report 65). New York, NY: Council on Foreign Relations.

About the Author



Barbara Trish is Professor of Political Science at Grinnell College (Grinnell, IA), where she also directs the Rosenfield Program in Public Affairs, International Relations, and Human Rights. Her analyses of U.S. politics have been published in scholarly journals, edited volumes and in the popular press.