Abstract

The idea of data-driven campaigning has gained increased prominence in public discourse. Journalists, policymakers, and scholars have played an important role in highlighting the potential threat posed by this activity. This chapter outlines the nature of current debate and argues that there is a need to reconsider the claim that parties are engaging in extensive, sophisticated, and concerning practice. Problematizing the dominance of the US case, and highlighting a lack of empirical insight, this chapter poses three questions that guide the remainder of the book: What is data-driven campaigning? How does data-driven campaigning practice vary? What explains different data-driven campaigning practices? Defining data-driven campaigning as comprising data, analytics, technology, and personnel, the chapter makes the case for an in-depth analysis of each of these components. Introducing an empirical focus on 18 parties in five countries, and the use of over 300 interviews, this chapter also outlines the structure of the book.

data-driven campaigning; US dominance; democratic threat; sensationalism; definition; interview; case

Introduction

Election campaigns represent a key moment when political parties around the globe reach out to the electorate to gain support. The leaflets, posters, doorstep conversations, and messages that parties disseminate have long been seen as an essential and laudable part of democratic practice. And yet, over the past 10 years, such virtuous depictions of campaigning have been challenged by a new, more sinister account. Political parties are seen to have amassed vast databases of highly personal information. They are seen to be using complex analytical techniques to profile and gather unprecedented insight into our personal lives. And they are perceived to be deploying these techniques to manipulate voters and elections (Rubinstein 2014, 879). Such practices paint a particular image of modern campaigning and the role that data plays in elections today that has raised a range of democratic anxieties. Journalists have accordingly highlighted privacy concerns (Cadwalladr 2017a) and data breaches (Murphy 2019; Jones & Cinelli 2017; Scally 2019). Scholars have documented threats such as voter suppression, manipulation, exclusion, deception, privacy violations, and a fragmentation of the public sphere (Bennett 2016; Jamieson 2018; Moore 2018; Taylor 2021; Zuiderveen Borgesius et al. 2018). And policymakers and civil society organizations have sought to mitigate risks associated with data-driven practices such as micro-targeting (Cicilline 2020; European Commission 2021; IDEA 2018; Kofi Annan Foundation 2020).

Within this book we suggest that the world of data-driven campaigning (DDC) is more complex and diverse than these (often sensationalized) accounts suggest. At present, however, the breadth and nature of data-driven campaign activity are not fully appreciated because of a focus on US practices and the risks of DDC. What is needed, therefore, is greater empirical understanding of how data is used in campaigns around the world, with more attention paid to whether (and why) usage varies internationally. Only with this insight is it possible to engage in a more empirically grounded discussion of the impact of DDC on democracy.

At present, when thinking about modern campaigning, US practices are commonly evoked to proclaim the routinization of DDC. There is extensive evidence available to support this idea. Senior officials in US presidential campaigns from both the Republican and Democratic parties have spoken about the importance of data and analytics. Catherine Tarsney, analytics director at the Democratic National Committee (DNC), for example, has discussed the need to build up a "360-degree view of voters by routinely incorporating data from new sources" (Data Council 2019). Elsewhere, the annual Reed Awards (2023), handed out by the US-based Campaigns and Elections magazine, now recognizes and celebrates the best data analytics solution, the best use of machine learning in online fundraising, the most sophisticated targeting in direct mail fundraising, and the best application of AI (artificial intelligence) technology to optimize targeting. There are also numerous examples of companies that offer data and analytics services to support US campaigns, ranging from the Republican group Data Trust (n.d.)—who advertise themselves as "the leading provider or voter and electoral data to Republican and conservative campaigns, parties, and advocacy organizations" (n.p.)—to fundraising platforms such as the Democrat-supporting Act Blue, which describes its process of "constantly A/B testing our contribution forms" to maximize donations (Act

Blue n.d., n.p.). US parties are also reported to be investing in data personnel and systems. The DNC, for example, hired a chief technology officer to oversee a staff of 65 employees, and increased investment in a new data system and new data points by purchasing 65 million cell phone numbers in 2020 (Ryan-Mosley 2020). Data is therefore a well-established component of campaigning in the US context.

Looking beyond the United States, there is some evidence that these practices are found elsewhere. In Australia, for example, the former national secretary and campaign director for the Australian Labor Party, Noah Carroll, argued that "[t]he interconnectivity of field work, data, analytics, research and messaging is the clear systems requirement of current and future campaigning" (quoted in Bramston 2016, n.p.). There is also evidence of international parties purchasing data lists for campaign purposes, with Canadian parties working with Environics Analytics, which breaks down the population into lifestyle clusters (Delacourt 2012), and UK parties hiring Experian and Data8 to access, analyze, and store data (Information Commissioners' Office 2018). Moreover, LinkedIn shows an international job market for data analysts and statisticians within party campaigns, with vacancies advertised in Australia, Canada, the United Kingdom, and Germany. Such examples suggest that data has become a central component of modern campaigns and is playing a key role in guiding electoral strategy.

And yet, while there is some evidence of campaigners' increased investment in data, there are also signs that the data practices of many parties are not as extensive, sophisticated, or concerning as often depicted. In many parties outside the United States, campaigners are highly curtailed regarding the type of data they can collect and store. In Germany, for example, there has been coverage of the way in which data protection law fundamentally restricts the type of personal information that parties are able to collect (Dachwitz 2017; Jaursch 2020; Kolany-Raiser & Radtke 2018), and of the strong cultural norms that reduce public acceptance of data collection and targeting (YouGov 2017: Völlinger 2017). In that regard the Berlin Data Commissioner (2021, n.p.) concluded in their annual report that:

parties can certainly use modern digital technologies for party work and election campaigns if this is done in accordance with data protection regulations. In Germany, however, it is not permissible to create profiles of voters, as is the case in the US, for example. By and large, the parties adhere to this. However, less comprehensive data about supporters and voters also needs protection. Therefore, political parties must exercise due diligence and consistently implement data minimization and anonymization.

Similarly, there are signs that in many countries, parties do not possess reliable systems to facilitate the use of data. In Canada, for example, a post-2021 election review published by the New Democratic Party reported that two of the party's data systems—CallHub and Dandelion—"crashed when they were needed most," leading to calls for tools to be capacity tested ahead of Election Day (National Democratic Party 2021, 7). In the United Kingdom, even the most highly resourced parties reported limitations with their databases, with the Labour Party's post-2019 General Election review noting that "vital systems and platforms were frequently unreliable, slow, hard to use, glitch-ridden, or tied up by complicated access restrictions," with the party's database, Contact Creator, in particular, reported to "not have the capacity to cope with high levels of data input" (Labour Together 2019, n.p.).

Meanwhile, in terms of data analytics and targeting, there is evidence that many international parties are restricted by data-protection rules. In Germany, for example, the Social Democrats published a campaigning fairness code which explains that the party will only use voter targeting which is exclusively within the framework of the high European and German data-protection standards (SPD-Parteivorstand 2021). There is also evidence that some parties lack the resources to invest in developing and updating sophisticated models (Kefford et al. 2022). In the United Kingdom, for example, the Liberal Democrats' post-2019 General Election analysis suggested that the party did not invest in ongoing and continually updated modeling, but rather relied on a single multilevel regression and poststratification (MRP) model conducted in June, meaning that their strategy was based on data that quickly went out of date (Kearns & Alexander 2020, 6). Such contrasting pictures of campaign practice suggest there are important questions about the way in which data is being used in modern campaigning. It indicates that far from there being a single manifestation of DDC, data is used in a variety of ways. This has important implications for any

attempt to understand the use of DDC and suggests that concerning practices are not an inherent feature of DDC, but may be present to different degrees. At present, however, three things are lacking: first, a framework for understanding what we mean when we refer to DDC; second, empirical insight into the variety of ways in which data is presently being used in campaigns internationally; and third, an explanation for that variation that pays attention to the contexts in which DDC occurs. Only by addressing these three gaps is it possible to fully appreciate the role of data in modern campaigns and consider its impact on democracy.

Contribution

In this book we provide the first internationally comparative study of DDC. We do so by exploring DDC across five advanced democracies: Australia, Canada, Germany, the United States, and the United Kingdom. Studying practices within 18 parties, we offer unprecedented insight into the reality of modern campaigns. Seeking to move beyond the prevailing emphasis on extensive data collection and concerning data analytics practices, we clarify what we mean by the term DDC, unpacking the spectrum of ways in which parties can utilize data, conduct analysis, deploy technology, and staff data-driven campaigns.

Adopting this approach, we combine conceptual mapping with empirical analysis to show not only that data can be gathered and utilized in a variety of different ways by political parties and campaigns, but also that specific parties in different countries vary in their precise engagement with data. Mapping variations, we also demonstrate that the vast majority of parties' data-driven practices are mundane, predictable, and removed from the hyperbolic accounts that dominate popular commentary, but that there are differences in how certain parties behave.

These conclusions are drawn from fine-grained qualitative research, and in many ways the book is unashamedly descriptive. As scholars who draw on a variety of methods in our research, we believe that any effort to understand DDC must begin with the provision of detailed "thick" descriptions of contemporary practice in order to generate new theories and accurate diagnoses of problematic practice. We therefore utilize qualitative methods to contextualize DDC practice with a view to advancing this body of scholarship.

Adopting this empirically grounded qualitative approach, our analysis is well-placed to explain different practices, and we do so by introducing an original theoretical framework that shows the relevance of systemic, regulatory, and party-level factors for any attempt to understand variations in DDC. Moving beyond the prevailing tendency to cite the importance of a country's data and privacy regulation in shaping DDC activity, we explain the significance of factors such as the electoral system, campaign regulation, and party ideology for DDC practice. Adopting this approach, we reveal why actors in the same jurisdiction use data differently, and how country context makes a difference in what is possible. Together, these insights help us to understand not only what is happening in political campaigning, but also how a variety of contextual factors can inform how and why data is used. These insights are likely to be highly significant for regulators or other actors seeking to respond to perceived challenges presented by DDC, revealing the very different factors that could be altered to change data practices.

In tracing the components of DDC and seeking to map and understand its contemporary practice, we argue that it is important not to overstate the novelty of this phenomenon. Data, as one form of information, has long been used to shape strategic decision-making. Whether gathering feedback from customers, conducting focus groups to test messages, or simply recording the demographics of people spoken to, the process of gathering and looking at patterns in information or data is a long-standing component of much political and commercial activity. As a result, in studying DDC, we engage with the rise of new tools and capacities, but we also spell out the many long-standing practices which remain core to data collection and analytics today. We show that rather than representing a radical new activity, over recent years DDC has evolved. Understanding this heritage is vital in the context of growing concern about the democratic implications and consequences of these activities, as it suggests that the use of data can be a valued component of democratic contact and should not, therefore, be universally condemned.

Cumulatively, these insights lead us to contend that:

- DDC is not a uniform practice, but can appear in a range of different forms;
- DDC is not inherently problematic, but can be used in ways that may be more or less acceptable to citizens;
- DDC is not new, but is the latest evolution of a long-standing practice of gathering and analyzing data in efforts to secure electoral success;
- Systemic, regulatory, and party-level factors affect the form of DDC.

What Do We Actually Know?

In making these contributions, we build on an emergent academic literature on DDC that, while expanding in recent years, so far lacks a detailed, comparative analysis of practices in different countries. The role of data in politics has long been recognized, with studies of, for example, parties' use of polling to gather insights into voters in the 1960s (Abrams 1963) and voter-segmentation techniques (Phillips et al. 2010, 311; Webber 2006). However, the most recent work on political campaigning suggests that there has been a substantive shift in attitudes toward, and use of, data.

Diagnosing the emergence of a new fourth era of campaigning (Röemmele & Gibson 2020), scholars have directed attention to "an organizational and strategic dependency on digital technology and 'big data'; a reliance on networked communication, the individualized micro-targeting of campaign messages, and the internationalization of the campaign sphere" (Röemmele & Gibson 2020, 595). While scholars often fail to define or characterize this activity in precise terms (Dommett et al. 2023), it has been widely suggested that there has been a "big data revolution" (Hersh & Schaffner 2013, 520), and a growing reliance on microtargeting for "every aspect of modern elections" (Rubinstein 2014, 883). Beyond these broad depictions, however, there have so far been few attempts to pinpoint the precise indicators of DDC or to explore the degree to which similar practices are found in different parts of the world.

Reviewing scholarly depictions of DDC, what emerges is therefore a relatively homogenous account of the role of data in modern campaigns that mirrors the popular focus on US practices (described above). Hersh (2015, 24), for example, offers a detailed study of data use in the United States that explains:

In an effort to win an election, campaigns seek to mobilize supporters and persuade undecided voters. In order to contact these voters and transmit mobilizing or persuasive messages, campaigns must predict which voters will be responsive to their appeals, and they must decide which voters should get which kinds of appeals. To make these decisions, campaigns gather data and form impressions about the voters.

He goes on to detail how data can derive from a range of sources, but "in more recent years, campaign organizations develop statistical models that generate a score for each voter, which estimates the probability that they support a particular party or candidate or that they will be likely to vote" (Hersh 2015, 28). This depiction of data-driven targeting and modeling has been widely replicated. Indeed, Chester and Montgomery (2017, 3–4) describe how campaigns "can now take advantage of a growing infrastructure of specialty firms offering more extensive resources for data mining and targeting voters." This includes:

data about individuals from a wide variety of online and offline sources, including first-party data from a customer's own record, such as the use of a supermarket loyalty card, or their activities captured on a website, mobile phone, or wearable device; second-party data, information collected about a person by another company, such as an online publisher, and sold to others; and third-party data drawn from thousands of sources, comprising demographic, financial, and other data-broker information, including race, ethnicity, and presence of children.

In line with such accounts, it is common to see descriptions of political databases that "hold records on almost 200 million eligible American voters" wherein:

[e]ach record contains hundreds if not thousands of fields derived from voter rolls, donor and response data, campaign web data, and consumer and other data obtained from data brokers, all of which is combined into a giant assemblage made possible by fast computers, speedy network connections, cheap data storage, and ample financial and technical resources. (Rubinstein 2014, 879)

While offering a range of valuable insights into the dynamics of modern campaigning, this literature has several shortcomings that we argue warrants a new approach. Particularly notable is the lack of comparative international analyses of DDC. While a small number have emerged in other contexts—with studies of practice in Canada (Bennett 2016; Munroe & Munroe 2018), Germany (Kruschinski & Haller 2017), and Australia (Kefford 2021)—for the most part our understanding of DDC reflects norms and practices found in US presidential campaigns, and is rarely comparative (cf. Kefford et al. 2022).

Although it is of course valuable to understand the dynamics of US campaigning—particularly given the influence that these practices can have on campaign activity elsewhere (Vaccari 2013), we argue that US practices should not be equated with general practice across advanced democracies because they are not simply diffused around the world through processes of "modernization," "imposition," and "imitation" (Pasquino 2005, 4; t'Veld 2017, 3). Rather, we argue that there are important systemic, regulatory, and party-level factors which mean that DDC practices found in the United States will not be replicated elsewhere (Kruschinski & Haller 2017).

Whether thinking about variations in data-protection law, party finance, or even electoral systems, the boundaries of legally and socially acceptable behavior are not consistent across advanced democracies. Indeed, even within the United States, there are reasons to think that this depiction may not accurately describe all DDC activity. The focus on well-resourced campaigns supported by expert data professionals can, for example, overlook the data practices of grassroots activists and lower-order (or down-ballot) campaigns that may not have the resources to cultivate large-scale data collection and analytics operations (Kefford et al. 2022). It also overlooks the potential for state-by-state variation where different local regulations and practices can cause DDC practices to vary. At a very basic level, therefore, we need to interrogate what we mean by DDC and how and why this practice might look different in different contexts and circumstances.

We also note that while there has been a growing interest in DDC, there have been relatively few empirical studies of the actual practice of this activity. Due to a range of challenges in securing access to observe campaigns in practice, and widespread reticence (especially among conservative parties) to give interviews about campaign activity (Dommett & Power 2021), many studies have based their depiction of DDC on claims made by companies selling DDC services. These assertions about campaign capacity can be easily located. Taking just one example, the C|T Group, an international consultancy company led by Lynton Crosby—an Australian political strategist who has advised on election campaigns around the world—claims it can gather "reliable, high-quality data to shape and influence behaviour in the desired direction by targeting the motivations of key actors and utilising identified pressure points to achieve the desired outcome" (C|T Group n.d., n.p.). Such sources can provide some insight into the objectives and goals of campaigns, but social scientists need to be wary of such assertions, which are often self-serving. As the Cambridge Analytica scandal has demonstrated, companies often make significant claims about their capacities that are not reflected in actual practices (ICO 2020).

For these reasons, there is a need to study the actual practice of DDC cross-nationally to ensure that our understanding reflects the real rather than potential use of these techniques. Pursuing study at this level, it becomes possible to understand how parties' use of data varies in different countries, or even varies between or within parties in the same context. It is also possible to ask whether we can observe one form of DDC in all countries, or whether there are different types of this practice in different contexts. Furthermore, consideration can also be given to why DDC appears more sophisticated in certain contexts, and more basic in others.

To generate these insights, this book reacts to the existing literature by posing three interlinked questions. First, we ask at the most basic level: What is DDC? Offering a definition that distinguishes four components of this practice—data, analytics, technology, and personnel—we outline how activity can differ at each level. Second, and operationalizing these frameworks, we ask: How does DDC practice vary? Presenting empirical data gathered from our five case study countries, we show exactly how parties are using data in different countries, highlighting variations and similarities exposed through our data collection. Finally, we ask: What explains different DDC practices? In doing so, we consider the relevance of systemic, regulatory, and party-level factors as drivers of campaign practice. By posing these questions, we offer unprecedented insight into modern campaigning, and demonstrate through detailed descriptive accounts of the mechanics of DDC cross-nationally that the drivers and potential responses to DDC are multifaceted.

In writing this book, we rely not only on our own academic analysis; we also integrate the voices of leading campaign professionals. At the end of each of our four substantive chapters (on data, analysis, technology, and personnel), we ask practitioners from across our countries to reflect on the logic of prevailing narratives around DDC, providing more direct insight into the way that campaign professionals understand the dynamics of the modern campaign. This book accordingly provides an important juncture from much previous work, helping scholars, practitioners, and those concerned about these practices to better understand the subtle nuances and influences upon parties' use of data.

What Is DDC?

Within this book we define DDC as a mode of campaigning that seeks to use data to develop and deliver campaign interventions with the goal of producing behavioral or attitudinal change in democratic citizens. We see DDC as composed of four central components: data, analytics, technology, and personnel. In offering this definition, we depart from much existing scholarship. Many academics have tended to avoid defining this phenomenon, with a recent systematic review (Dommett et al. 2023) showing only a handful that have outlined the traits of DDC (for example, Baldwin-Philippi 2019; Munroe & Munroe 2018; Kefford 2021). This tendency has allowed certain implicit and untested assumptions to become endemic and leaves many unanswered questions about the boundaries of this activity.

In our previous work, we have asserted the need to recognize variations in what data is being used, who is using data, and how data is being mobilized as part of a campaign (Dommett 2019; Dommett et al. 2021; Kefford et al. 2022; Kruschinski & Bene 2022). And we have demonstrated the potential for parties to simultaneously exhibit data analytics and targeting practices that are highly complex in some communication channels and simple and mundane in others (Kefford 2021; Kruschinski & Haller 2017). Spotlighting these different possibilities, we argue that there is limited utility in characterizing one set of (the most sophisticated) practices as indicative of DDC. Instead, we argue that there is a spectrum of types of data and mechanisms for data collection, as well as a range of different analytical techniques that can be deployed by different personnel to engage in DDC.

Our definition reflects this argument, meaning that we do not reify particular forms of DDC, but rather focus on the different ways in which data can feature within, and be used by, different organizations. By considering our four key elements of DDC—data, analytics, technology, and personnel—in turn, we distill the range of different possible practices that can be observed in each area of campaign activity. This approach allows us to move attention beyond US presidential campaigns and the handful of international instances in which US-style practices appear to be evident, to offer a more encompassing picture of the way data is being used by political parties across advanced democracies.

Why Do We Need to Understand More about DDC?

DDC is, if we believe much of the commentary, a serious threat to the effective functioning of liberal democracy. Though DDC is only one component of a far wider debate about the capacity of technology to revitalize or undermine democratic practice, numerous scholars and policymakers have raised concerns about the impact of data-driven practices (ICO 2020, US Senate 2018). These prognoses are significant because they are not confined to academic discussion (Harker 2020, 157; Jamieson 2013; Nadler, Crain, &

Donovan 2018, 34; Rubinstein 2014, 886; t'Veld 2017, 3), but have begun to inform wide-ranging proposals for democratic reform. In countries around the globe, proposals have begun to be made by policymakers for improved data-protection law, increased transparency, and regulatory oversight of data-driven practices (Kuehn & Salter 2020, 2600–2601). These proposals have the potential to dramatically shape what constitutes acceptable (and legal) practice, and yet they are based on limited empirical evidence. In providing more insight into DDC, this book seeks to facilitate more sophisticated debate around the democratic implications of these practices, and appropriate responses to this activity. It also provide a template for thinking about the study and analysis of new technologies in politics more generally. As exemplified by new debates around the use of AI in politics (Kapoor & Narayanan 2023; Robins-Early 2023), there is a tendency to focus on the negative potential and to rush to regulate concerning practices, but our analysis suggests the need for a more empirically grounded and nuanced approach (Jungherr & Schroeder 2023).

First, at the most basic level, we argue that in order to appreciate the threat that DDC poses to democracy, there is a need to have a clear conception of the problem DDC poses. As Nielsen (2020, n.p.) has argued, "whenever we deal with any large public issue that requires a societal response, we've got to get the problem right otherwise our responses will be at best ineffective and at worst counterproductive." At present, the dominance of the US case has elevated concerns around particular practices which are deemed problematic. This includes voter-suppression activity (Kim 2018) and political redlining (Harker 2020, 155–156; Judge & Pal 2021; Kreiss 2012). However, it is not clear the extent to which these practices are found elsewhere, or whether other "problems" may be perceived in different contexts. With decades of scholarship showing variations in public attitudes toward privacy and democratic expectations, it is by no means to be expected that citizens (or policymakers) in different countries will perceive DDC practices in the same way (Kozyreva et al., 2021). In one context, it may therefore be deemed completely unacceptable to purchase information about citizens without their consent, while in another it may be relatively unproblematic. These possibilities make it vital to more fully appreciate what is happening in different countries in order to facilitate more informed discussion of the type and extent of democratic threat posed by DDC. Only when equipped with such knowledge can we determine which potential responses are most likely to produce the changes sought by democratic citizens in different contexts.

Second, and related to the point above, we draw attention to the current tendency for policymakers to propose data and privacy regulation as a means of curtailing problematic DDC practices. While an important type of response, we argue that this is not the only means of influencing the nature of data-driven campaigns. Indeed, our analysis shows that variation in DDC can be a product of a complex interplay of systemic, regulatory, and party-level factors. As such, efforts to shape DDC can usefully recognize the influence of a range of different factors, making it informative to examine how particular contextual factors affect the way DDC is manifest if seeking reform.

For these reasons, we argue that while it is reassuring to see action taken that is designed to protect the democratic system, there is a danger that without a clear understanding of precisely how data is being used and what is driving these practices, regulatory interventions will either tackle the "wrong" problem, or tackle the "right" problem but in ways which do not produce the desired result. As such, our book not only is important for academic understanding, but also has serious implications for democratic practice.

Our Cases

Our analysis of DDC by political parties focuses on practices in five advanced democracies: Australia, Canada, Germany, the United Kingdom, and the United States. These cases, while similar in many ways, exhibit significant variation on key systemic and regulatory dimensions that underpin our theoretical framework. These differences allow us to explore why DDC looks different in alternative country contexts, and yet we also seek to explain variations within each of our cases. As Vaccari has argued, "[e]ven within the same political context, campaign techniques are adopted in different ways by different political actors" (2013, 11). For this reason, we study a range of different parties within each of our case studies, considering how party-level factors exert an influence on the practice of DDC.

As outlined in Table 1.1, our five cases vary across many of the key variables commonly associated with campaigning, especially that of political parties. In terms of systemic variables, this includes variations in the electoral systems (mixed; majoritarian), systems of government (federal; unitary), party systems (multi-party; two-party), and hybrid media systems (high, mixed, low). Further, our five case study countries offer a balanced selection of weak, medium, or strong party, and campaign, data, and privacy, as well as media, regulations. While there were an infinite number of variables we could have presented to demonstrate the variation among our cases, we argue that the differences highlighted in Table 1.1—and discussed in more detail in Chapter 2—provide sufficient justification to draw generalizations from these five cases to other advanced democracies. We, of course, would have liked to extend this analysis to explore cases that varied even further, but ultimately, we chose cases that would provide sufficient generalizability while ensuring that we were able to gather the required empirical insights to fully understand the manifestation of DDC in each context.

Table 1.1 Key Systemic and Regulatory Variables

	Systemic Variables				Regulatory Variables			
	Electoral System	**System of Government**	**Party System[a]**	**Hybrid Media System[b]**	**Party Regulation**	**Campaign Regulation**	**Data and Privacy Regulation**	**Media Regulation**
Australia	Mixed: Alternative Vote and Single Transferable Vote	Federal	Multi-party	Mixed	Weak	Weak	Weak	Weak
United Kingdom	Majoritarian: Plurality	Unitary[c]	Multi-party	Mixed	Medium	Medium	Medium	Medium
United States	Majoritarian: Plurality	Federal	Two-party	High	Weak	Weak	Weak	Weak
Germany	Mixed: Plurality and Party list	Federal	Multi-party	Low	Strong	Strong	Strong	Strong
Canada	Majoritarian: Plurality	Federal	Multi-party	Mixed	Medium	Medium	Medium	Medium

[a] While there are numerous ways to classify party systems, for our purposes here we have simply used a two-party/multi-party distinction to highlight the number of "relevant" parties in the party system (see Sartori 2005, for more on this).

[b] To provide a sense of the level of hybridity in each of the media systems, we have considered the ways that old and new media are combined, which provides opportunities for grassroots participation and that new media and old media shape each other and are integrated with one other so that new media logic emerges. To assess this variable, we distinguish between systems as displaying "high," "mixed," or "low" levels of hybridity. In particular, we draw on the work of Mattoni and Ceccobelli (2018) here to outline what we consider to be of most interest for DDC.

[c] The UK here is classified as a unitary state, but due to devolution of certain powers to Scotland and Wales is technically an asymmetrically decentralised unitary state.

Across the five country case studies, our analyses detail the practices of 18 political parties. As outlined in Table 1.2, within each country we include the major parties, as well as a selection of minor parties. The explanation for this relates squarely to issues of access. Within many of our case study countries, we were unable to secure sufficient numbers of interviews, or to identify documents for analysis to ensure we were able to verify our insights for each party in the party system. In such instances, we chose to exclude these parties rather than risk misrepresenting practice. The 18 parties covered in this book provide a cross section of major and minor parties and, reflecting our theoretical framework (outlined in detail in the next chapter), capture a variety of differences related to party resources (high, medium, and low), structure (hierarchical, stratarchical, federated), ideology (social democrats, conservative, greens, liberal, left, right), and attitudes toward campaigning (enthusiasm, mixed, reticence). These cases accordingly allow us to generalize about political parties across advanced democracies and not only to point to the variations we identify, but also to theorize about the drivers of variation by using our framework.

Table 1.2

Party Variables and Parties Included in Analysis

	Party Variables			
	Resources	Structure	Ideology	Attitudes to DDC
Australia				
Australian Labor Party	Medium	Stratarchical	Social Democrats	Enthusiasm
Liberal Party	Medium	Stratarchical	Christian Democrats/Conservatives	Enthusiasm
Australian Greens	Low	Federated	Greens	Mixed
United Kingdomi				
Labour Party	Medium	Stratarchical	Social Democrats	Enthusiasm
Conservative Party	Medium	Hierarchical	Christian Democrats/Conservatives	Enthusiasm
Liberal Democrats	Low	Stratarchical	Liberals	Enthusiasm
Green Party of England and Wales	Low	Federated	Greens	Mixed
United States				
Democratic Party	High	Stratarchical	Social Democrats	Enthusiasm

Republican Party	High	Stratarchical	Christian Democrats/Conservatives	Enthusiasm
Germany				
Social Democratic Party (SPD)	Medium	Stratarchical	Social Democrats	Enthusiasm
Christian Democratic Union (CDU)/ Christian Social Union (CSU) (together: Union parties)	Medium	Stratarchical	Christian Democrats/Conservatives	Enthusiasm
The Greens	Medium	Federated	Greens	Mixed
Free Democratic Party (FDP)	Medium	Stratarchical	Liberals	Enthusiasm
The Left	Low	Stratarchical	Democratic socialism/Left-wing populism	Reticence
Alternative for Germany (AfD)	Medium	Stratarchical	Nationalism/Right-wing populism	Mixed
Canada				
Liberal Party of Canada	Medium	Stratarchical	Liberals	Enthusiasm
Conservative Party	Medium	Hierarchical	Christian Democrats/Conservatives	Enthusiasm
New Democratic Party	Medium	Federated	Social Democrats	Enthusiasm

a Resource was calculated based on Political Party Database Data (2022). Due to inconsistencies in other possible resource metrics (i.e., a lack of available data on total party income for our cases), we looked at data on campaign spending for Australia, Canada, and Germany. For the United Kingdom we used data from national spend from the database. US data was gathered from the FEC website. The most recent available data was utilized, and we converted all national currencies into US dollars to ensure comparability. In reporting this data, we distinguish between low (less than $10 million spent), medium (between $10m–$100m), and highly ($100m+) resourced parties).

b To assess this variable, we applied the framework provided by Bolleyer (2012, 320), in distinguishing between parties as being hierarchical, stratarchical, or federated. This framework suggests that while hierarchical parties see power held centrally,

in stratarchical parties power is held across levels, whereas in federated parties it is held regionally. An assessment was made across each of the variables specified by Bolleyer in reaching a classification. For Canada, there is debate about the best way to characterize the party structure (see Carty 2004; Coletto et al. 2011). Likewise, for the German parties, there is some debate, however, we have used Bolleyer's assessment of their structure here as a way of highlighting organizational differences.

c This coding was based on the Political Party Database Data from 2020. The United States is not included in this database, hence entries for this country were produced by the authors.

d Attitudes toward DDC vary across parties and there is no preexisting measure to monitor variations. Accordingly, we drew on our existing research and the secondary literature to differentiate between "Enthusiasm," "Mixed," or "Reticence."

The focus of this book is on DDC by political parties. However, our view is that it is essential to recognize that data is used by a wide range of organizations and individuals. Whether thinking about governments, universities, civil society organizations, or businesses, data is a valuable currency. This is because, in reflecting the well-known adage that "information is power," data provides actors with insights about the world that can be collected and mobilized to deliver a range of outcomes. Whether helping to identify an audience for a desired product, to enable targeted messaging to "nudge" a recipient into a desired action, or to test audience reception to specific initiatives, data is integral to the way a whole range of actors work today. Parties are therefore by no means unique in collecting and analyzing data, meaning that the findings of this book resonate far beyond these organizations. Although we focus on these actors as key institutions within liberal democracy (Bartolini & Mair 1990), we recognize the potential to expand this study to reflect on the activities of other institutional types.

Our Empirical Data

DDC often reaches the public consciousness via the sales pitches of campaign professionals and companies (such as Cambridge Analytica)—painting a distorted picture of the capacities and uses of these techniques. In contrast, this book promotes the voices of other actors within the system, offering unique insight into the experiences and views of those who engage in DDC. We do so in two ways. First, we draw on 329 interviewsii that we conducted with party operatives, campaign consultants, pollsters, and data brokers to contextualize developments. Undertaken between 2017 and 2022, we spoke primarily to senior party officials working within campaign headquarters, as well as grassroots campaigners (most often at a regional level) and to actors within external campaign organizations who were supporting party campaigns.iii These interviews focused on the use of DDC, how usage had developed over time, how DDC was viewed, and the conditions under which it was (and was not) deployed. All interviews were recorded and transcribed where consent was given, or interview notes were taken and approved. This formed a corpus of documents that were then coded and analysed by the researchers to identify recurring themes and ideas. This interview data offers unprecedented insight into how we understand DDC, as well as its variation both within and between different parties around the globe. We use these interviews to understand not only what is happening, but also why these practices have come about.

In relying on interviews, it is, however, important to note that our ability to gain access to interviewees within different parties was not uniform. In many cases we found that parties simply lacked a large staff base, curtailing the available number of interviewees, but we also encountered overt unwillingness to engage. Evident most prominently in ideologically conservative parties, there were instances in which we were unable to secure interviews with current staff. In addition, we encountered the widespread use of non-disclosure agreements (often in major political parties), which meant that party staff and consultants were limited in their ability to be interviewed, or in what they could disclose. Encountering these varied issues, we used alternative sources of data to gather insight and to verify claims. Specifically, we collated a wide range of documentary evidence, including internal party reports, post-election analyses, regulatory reports, media coverage, and firsthand accounts of election campaigns. We also conducted reviews of party websites and social media archives, and we drew on analyses and reports from civil society groups and nongovernmental organizations. These documents were used to build up a rich picture of how data was talked about by different types of actors in different contexts. While most of these sources were publicly

available, we also gained access to a small number of internal party documents that were used to help us understand the actual practice of data use. Triangulating multiple data points, we used these sources to build up a rich picture of practice in hard to access campaigns. While our data collection was not, therefore, unproblematic, we have nevertheless compiled the most comprehensive account of DDC practice offered to date.

To augment our analyses of DDC, we also provide practitioner perspectives that allow us to give voice to those directly involved in DDC. While academic analysis can help us to unpick the common themes and ideas that run across our cases, we also want to give readers the chance to hear directly from practitioners themselves. For this reason, between chapters we provide opportunities for campaign professionals and regulators to offer their own perspective on the different aspects of DDC that we discuss. These perspectives are intended to provide tangible examples of the kind of practices we observe, offering a more extended glimpse into the perspectives we uncovered in our interviews. Moreover, they allow often behind-the-scenes practitioners to go on the record to communicate their own ideas about the use and significance of DDC in their own voice. Valuable not only for general readers, these interventions are intended to be of use to those teaching political communication, providing stimuli for students to consider the form and implications of this type of campaign activity.

Structure of the Book

The remainder of this book is structured as follows. In the next chapter we introduce our original theoretical framework to explain variations in DDC. Building on extant scholarship in political science and political communication that has utilized multilevel frameworks to understand the drivers of organizational change or political practices (such as Barnea & Rahat 2007; Esser & Strömbäck 2012; Gauja 2017), we develop a three-level framework which consists of systemic, regulatory, and party-level variables. We then move through the four central components of DDC that are especially significant for understanding variation. These are: data, analytics, technology, and personnel. Within each of these chapters we classify existing practice and then, presenting empirical evidence from our five cases, map variation in the form of each particular aspect of the data-driven campaign. In Chapter 7, we apply our theoretical framework to explain some of the variations outlined in previous chapters, showing the importance of considering systemic, regulatory, and party-level factors when seeking to understand and explain the existence of different data practices. In our final chapter, we discuss the conclusions that can be drawn from our analysis and consider the significance of our empirical evidence for debates about democracy.

End notes

i The parties within the United Kingdom were the national party. Interviews were not conducted within devolved parties (such as the Scottish Labour Party)

ii The interview breakdown is as follows: Australia 183 (82 Labor, 48 Greens, 53 Liberal and Liberal National parties); United Kingdom 49 (31 Labour, 5 Conservative, 6 Liberal Democrat, 7 Green); Canada 9 (3 Liberal, 4 NDP, 2 Conservative); Germany 51 (13 Christian Democratic Union, 14 Social Democratic Party, 7 The Greens, 6 The Left, 8 Free Democratic Party, 2 Alternative for Germany, 1 Free Voters Bavaria). We also completed 37 interviews with digital marketing firms, campaign consultants, and data brokers. It is important to note that while the number of interviews with Australian parties far exceeds the others, this is necessary as there is almost no information available in the public domain about campaign spending and associated matters in Australia as parties are not required to provide almost any information on these matters. Hence, interviews are the only methods available to access information about campaigns in any detail.

iii Interviewees from the parties were drawn from across the population of those with knowledge of these practices, including those who sit at different points in the wider campaign assemblage. This includes those in party headquarters, such as campaign directors, data analysts, field directors, and digital operatives. But it also includes those at the state and regional level, those organizing and running campaigns in individual seats, as well as campaign volunteers, which included member and non-member activists. While it would be cumbersome to outline the roles of each individual interviewed, they represent a cross section of those within these parties. These varying perspectives are important as they reveal that views of data are not always consistent within single organizations, a point we illustrate within the chapters that follow.