Divide to Rule: Deconcentration as Coalition Manipulation

by

Zach Warner

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Political Science)

at the

University of Wisconsin–Madison

2018

Date of final oral examination: 2 April 2018

The dissertation is approved by the following members of the Final Oral Committee:

John Ahlquist, Associate Professor, School of Global Policy and Strategy, University of California, San Diego
Rikhil Bhavnani, Associate Professor of Political Science
Scott Gehlbach, Professor of Political Science
Noam Lupu, Associate Professor of Political Science, Vanderbilt University
Nils Ringe, Professor of Political Science

To Naomi

Acknowledgments

My mother's father, James Turek, grew up in Green Bay, Wisconsin. His family owned a small cattle auction business that supplied the local meatpacking industry. In 1953, after serving in the Korean War, he enrolled at St. Norbert College in De Pere. His ambition was to teach political science. But after his father fell ill, he was forced to abandon his studies and take over the family business. He was never able to complete the degree.

This dissertation is evidence of my singularly good fortune: unlike my grandfather, I have been able to lean on an exceptional support network throughout graduate school. Although mine is the only name that appears on this dissertation, it has been a team project from beginning to end, and I would not have been able to finish it without the sacrifices of a great many people. I have incurred too many debts to fully enumerate them here, but certain individuals' contributions deserve special mention.

I could not have asked for a better dissertation committee. John Ahlquist has been an outstanding chair, spending countless hours discussing my ideas, pushing me to sharpen them, and helping me to communicate them more effectively—despite, frequently, an eight-hour time difference between us. It is also to him that I owe much of my technical training, and subsequently my understanding about what social science research is and what it can tell us about the world. His mentorship

has undoubtedly made me a better scholar. Scott Gehlbach has also gone above and beyond as my *de jure* chair. From my first days at Wisconsin, Scott has been a constant source of insight. He is among the very best teachers I have ever had, not only in the classroom but also in one-on-one settings, where he showed me how to keep stripping away layers of the problem posed by deconcentration until I got to the core. Noam Lupu reminded me to keep my eye on the big picture when I got bogged down in identification strategies or coding problems. Collaborating with him has been a practical education of the highest order. Working as a research assistant for Rikhil Bhavnani gave me my first experience as a data hound, and he continues to push me to find new and better ways to test my theoretical claims. Nils Ringe has been a source of constant support while encouraging me to think more deeply about the institutions in the deconcentration story.

I have also drawn heavily on Wisconsin's Africanist community. Michael Schatzberg and Scott Straus were among the first to help me acclimate to graduate school and have always been a source of excellent advice. I am grateful for the opportunity to draw on Crawford Young's seemingly endless knowledge—on more than one occasion I heard him mention he "knew a bit about" a topic, only to later find out he'd written a book about it. I also benefited from many conversations with Africanist graduate students past and present, particularly Kate Carter, Barry Driscoll, Kathleen Klaus, Evan Morier, Nat Olin, Jeff Paller, Taylor Price, and Charlie Taylor. And although we didn't overlap at Wisconsin, Brandon Kendhammer was among the first Madisonians I spoke to back in 2012; his generosity was as exemplary then as it is today.

My work at the Latin American Public Opinion Project (LAPOP) has also given me the opportunity to engage in the fantastic community at Vanderbilt. I am grateful to Liz Zechmeister for this opportunity, and for her continued support

and understanding as my contribution to the Project has fluctuated with frequent moves between Nashville, Madison, and London. I also thank the entire LAPOP team, with whom it has been a joy to collaborate. In particular, Mollie Cohen and Mariana Rodríguez have saved my bacon on more than one occasion.

More broadly, Larry Bartels, Guy Grossman, Charles Hankla, Yoi Herrera, Melanie Manion, Ellie Powell, Byron Shafer, Erica Simmons, Tariq Thachil, and Aili Tripp provided valuable feedback at various stages of this project. Alex Tahk helped me work through tricky methodological problems time and again. While an interloper at Nuffield College, I benefited from helpful conversations with Ben Ansell, Ray Duch, Andy Eggers, Bryn Rosenfeld, and Luis Schiumerini. Audiences at the annual conferences of the Society for International and Organizational Economics and the American Political Science Association, as well as seminar participants at Wisconsin, Vanderbilt, and Nuffield, all contributed important insights. Financial support from the University of Wisconsin-Madison and the United States Department of Education's Foreign Language and Area Studies Fellowship program facilitated fieldwork in Nigeria and in London. Finally, I thank the Lagos State Government's Ministry of Economic Planning and Budget, the National Bureau of Statistics, and my hosts in Nigeria: Sanuth Jafar, Tara Vaughan, and Tomi Vaughan.

My knowledge of and love for Nigerian politics is entirely due to Olufemi Vaughan. I am continually grateful that Femi responded to an email from an unknown undergraduate in search of a research job back in 2011. His knowledge of Nigerian history and politics is only rivaled by his kindness and his generosity, and I hope that my claims about Nigeria in this dissertation do justice to the training he has given me.

Of course, my friends and family have been a constant source of strength. I was lucky to arrive at Madison as part of a large and talented cohort. Beyond those

already mentioned, I thank Annie Anderson, Sarah Bouchat, José Luis Enriquez Chiñas, Rachel Jacobs, Clarence Moore, Anna Oltman, Camilla Reuterswärd, Degi Uvsh, Jack Van Thomme, Logan Vidal, and Bri Wolf. Adam Auerbach, Sanja Badanjak, Nick Barnes, Jack Edelson, Brad Jones, Ning Leng, Ryan Powers, and Ben Toff provided much-needed guidance over the years. Kyle Marquardt helped me navigate the transition to dissertator, and later, job marketeer. Hannah Chapman and Sam Vortherms were always available for advice. Emily Sellars helped me take my first steps in both R and econometrics, and continues to be generous with her time and her knowledge. I've learned a lot from Evan Crawford, Steph Dillon, and Rachel Schwartz—mostly at trivia. And although his football loyalties are suspect, Ben Power got me through many weekend mornings at Cooper's.

I am also grateful to have such a strong support network in London, having spent so much time there over the past decade. I thank Padmini and Nilupa Sugunasingha, who have put up with me and my dissertation as much as anyone. The Sugunasingha-Pathinayake clan is far too vast to list here, but I am grateful to them all for the many ways in which they've helped over the years, and I'm lucky to call them family. The Royal Holloway crew have provided much-needed breaks from the solitude of writing a dissertation. In particular, I thank Sam Ault, Danny Ribeiro, and Matt Stavrou—may the "wing crawl" tradition never end.

The Schultz family deserves special mention. Laura, Jeff, Sam, and Becky were the first to take me in when I arrived in Madison six years ago, and have supported me ever since. I am grateful for the countless favors they have done for me, and for all of the Badgers games to which they've taken me, but I am most grateful that they let me into their lives over these years. It's been a privilege to share Madison with them, and I will miss them dearly. *Once you've said Wisconsin… you've said it all!*

I am grateful to my mother, my sister, and my step-father for more than I can describe here. This dissertation would not be possible were it not for their sacrifices.

Finally, I am thankful for the endless love and support that I have received from my wife, Naomi. She has been an unwavering source of inspiration for twelve years, and I remain in awe of her courage, her curiosity, and her compassion. I look forward to our next adventure together.

Table of Contents

List of Tables

List of Figures

Abstract

Deconcentration—the creation of new subnational units of government—has quietly become commonplace, with nearly 400 episodes across 126 countries over the period 1960-2010. Recent studies have argued that deconcentration is the product of local and national elites jointly pursuing patronage and electoral gain. Yet this perspective leaves unresolved why patterns of contestation over deconcentration fail to map onto dominant cleavages such as partisanship or ethnicity. Nor does it clarify the constraints that determine when and where attempts to deconcentrate succeed.

This dissertation proposes a new explanation centered on elite bargaining over political coalitions—groups of agents who choose and implement a joint course of action, such as parties or factions. I argue that politicians pursue deconcentration because it allows them to manipulate the structure of these coalitions. Creating new subnational units of government empowers new politicians, altering the potential coalitions that can emerge from elite bargaining. Thus, politicians form preferences over deconcentration according to how they believe it will impact their position in the coalition structure. Deconcentration is in turn constrained by the institutions structuring coalition bargaining.

I test the empirical implications of this theory in three ways. First, I introduce new data on deconcentration worldwide over the period 1960-2010, and use supervised machine learning models to demonstrate that the observed variation

in deconcentration matches my theoretical expectations. Second, I show that sudden leadership deaths destabilize coalition structures and thereby substantially increase the probability of deconcentration. Finally, I draw on primary documents held at the National Archives of the United Kingdom to trace the coalition-manipulation mechanism through a close case comparison of three statehood movements in Nigeria over the period from 1946 to 1966.

This study contributes to a number of areas of scholarly interest. In particular, this dissertation documents the global surge in deconcentration over the last sixty years, and its implications for relationships between citizens and states the world over. It also sheds new light on debates over selectorate theory, the endogenous evolution of representative institutions and party systems, and why self-interested political actors trade away their own power.

1 | The deconcentration puzzle

In 1963, Nigeria carved a new Mid-West Region out of the old Western Region. While this reform was unsurprising to contemporaries—local demands for autonomy had been voiced since 1946, when the British colonial government reorganized Nigeria into a tripartite federation—the identity of its architects was unexpected. The Northern People's Congress (NPC) and the National Council of Nigeria and the Cameroons (NCNC) were dominant within the Northern and Eastern Regions, respectively, and together formed a national coalition government. Even if voters in the Mid-West rewarded the NPC and NCNC by electing a sympathetic regional government, its support would be immaterial for their legislative agenda. And this prospect looked uncertain at best: the Mid-West would be composed of districts where the opposition party, the Action Group (AG), had been making enormous gains in recent years (Brand 1965; Vickers 2000, 238). Should the Mid-West return an AG government, it would enjoy substantial powers, including control over the Region's education, health, industrial development, and agriculture policies (Mackintosh 1962), as well as the ability to block the creation of more regions.

Initial confusion over why the dominant NPC-NCNC alliance would risk empowering an opponent was quickly resolved. While the AG leadership opposed mid-West separatism, a loud minority faction were willing to defect from the party in pursuit of their new region. By backing Mid-West statehood, the NPC and NCNC

forced the AG to oppose it. This opposition sparked a crisis that broke the party in half, and by the time the 1964 federal elections arrived, the AG could no longer compete as a national power. As Suberu (2001, 83) explains, the NPC-NCNC plan had been to use the creation of the Mid-West to "annihilate the AG and decimate its base in the Western Region." It was a roaring success.

Military ruler Sani Abacha's decision to create six new states in 1996 was borne of a very different strategic logic. By the early 1990s, Nigeria's four regions had splintered into 30 states, and in contrast to the NPC's hegemony, Abacha's regime was creaking under the strain of dwindling popular support, economic crisis, and international outcry over his regime's use of detention and torture of political opponents. Abacha's state creation came as a surprise to Nigerians, as it fractured existing patronage networks—which had helped him personally embezzle upwards of $5 billion (Enweremadu 2013)—and did not open new rifts among the opposition, as the creation of the Mid-West Region had in 1963. Nor was it simple majority-rigging, as no elections were planned or held during Abacha's reign.

Again the motive behind creating subnational units appeared opaque to contemporaries, and again this motive was coalitional realignment. State creation allowed Abacha to fold disgruntled minorities into his regime, winning himself much-needed elite civilian allies. The new Gombe State would quiet growing discontent in the North-East, Abacha's base, while Nasarawa State would empower groups in the diverse "Middle Belt" that were largely disenfranchised in the military government. Most importantly, Ekiti State would provide Abacha a foothold in the old Western Region, and draw the powerful Yoruba elite closer to his regime. As in 1963, state creation appears to have paid dividends: these new allies voiced "unflinching support" for Abacha (Kraxberger 2004, 428), likely prolonging his rule until he died suddenly in 1998.

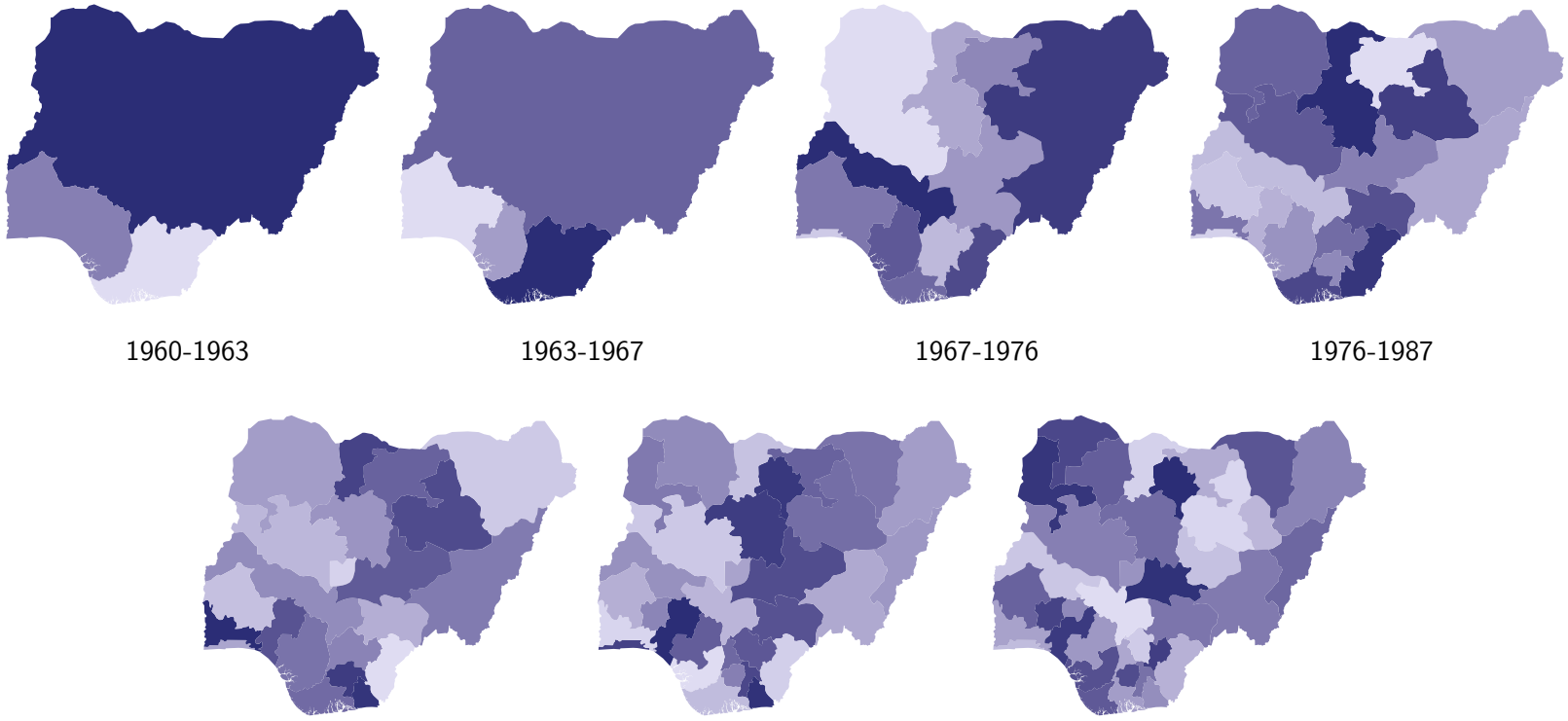


Figure 1.1: Nigeria's top-level administrative units since independence. Each panel is a map of Nigeria's top-level administrative units (regions until 1967 and states thereafter) following an episode of deconcentration.

Deconcentration—the creation of new subnational units of government—is not uniquely Nigerian. Grossman and Lewis (2014, 198) define "administrative unit proliferation" (AUP) as a special case of deconcentration, and document at least 25 episodes in sub-Saharan Africa since 1990. Malesky (2009) describes "gerrymandering, Vietnamese style" as the creation of new provinces to generate a national coalition in favor of economic reform, and Arjomand (2009) traces the rise of Iran's New Right through the strategic placement of hardliners in executive positions in newly-created provinces.[1] Nor is deconcentration limited to nondemocracies. Stewart and Weingast (1992) demonstrate how in the United States (US), Democratic Senators helped create new Western states expected to vote Republican, in the service of a larger Western caucus. Over the period 1960-2010, nearly 400 cases of deconcentration have occurred in 126 countries worldwide, as depicted in Figure 1.2. Fully 83% of all people live in countries that have deconcentrated in the last 60 years.[2]

Yet explanations for deconcentration are scarce. Scholars have devoted significant attention to tracing the origins of federalism, but almost none to understanding where subnational units come from more generally. Equally, studies of decentralization are essentially silent on the causes of deconcentration, focusing on why governments transfer new powers to existing units instead of why governments transfer existing powers to new units. Further, among the few studies that have explicitly theorized deconcentration, all have struggled with the puzzling patterns of contestation over unit creation. Scholars still lack a framework for making sense of battles over deconcentration that run orthogonal to partisanship and other natural fault lines, as in Nigeria and during the American Reconstruction. Nor are we able to explain how rulers intent on deconcentration can overcome the opposition of

1. See also Okafor (1987) on "jurisdictional partitioning," Swianiewicz (2010) on "territorial fragmentation," and Grossman, Pierskalla, and Dean (2017) on "government fragmentation."
2. I describe data sources and coding rules in Chapter 3.

powerful elites in national and subnational government.

Without a better understanding of the causes of deconcentration, we cannot understand how rulers organize governments spatially. This choice is of vast consequence to people all over the world. For many, local and regional governments are the "face" of the state with which they most frequently interact (Ostrom 1990). This relationship is particularly pronounced in the developing world, where structural adjustment programs in the 1980s and democratization assistance in the 1990s commonly included provisions for expanding and strengthening subnational government (Manor 1999). The structure of subnational government also has intrinsic meaning for citizens, providing a vehicle through which to express their national, regional, and local identities (Hooghe et al. 2016). Further, in democracies,

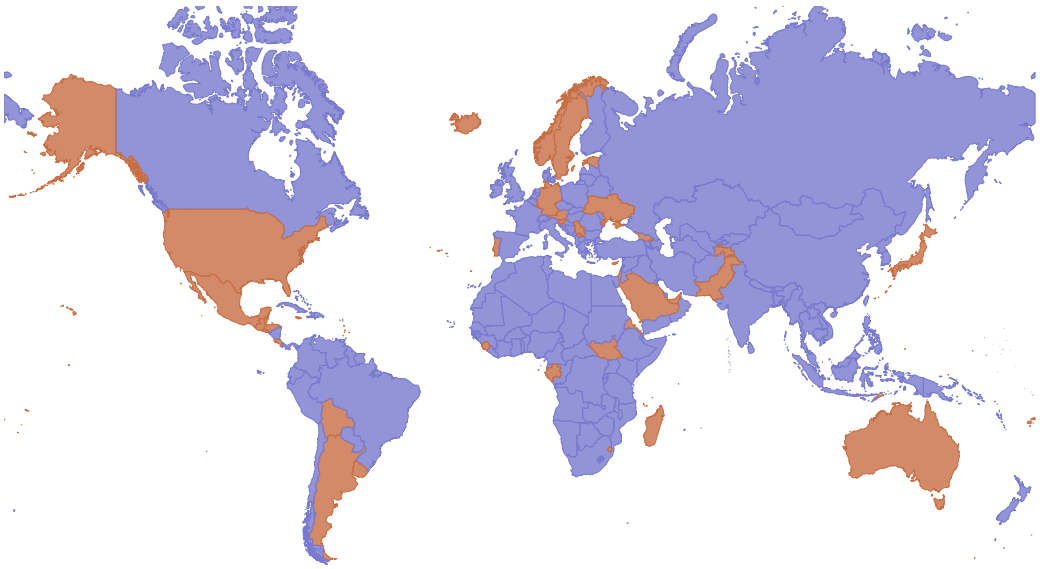


Figure 1.2: Worldwide deconcentration, 1960-2010. Countries that created new top-level administrative units or new layers of government between 1960 (or after independence) and 2010 are plotted in blue, while those that did neither are plotted in orange. Data sources and coding rules are provided in Chapter 3.

subnational boundaries typically coincide with electoral constituencies, so that the quality of representation depends on the process producing such boundaries (Bhavnani 2015). In short, understanding how governments engage with the governed requires an understanding of the politics of deconcentration.

In this dissertation I provide a novel explanation. I argue that deconcentration is the result of politicians' attempts to manipulate political coalitions. The key feature of deconcentration is that it invests new politicians with formal authority derived from holding office. This authority enables them to enter the process of bargaining over political coalitions within and across levels of government. Including these newly-empowered politicians in coalition bargaining can result in outcomes that substantially alter the extant coalition structure. Thus, politicians' preferences over deconcentration are induced by their preferences over the coalition structures they expect to emerge with and without deconcentration. Whether deconcentration occurs depends both on these induced preferences and the institutions that structure coalition bargaining.

In demonstrating the origins of deconcentration in coalition manipulation, this study contributes to a number of areas of scholarly interest. To my knowledge, this dissertation is the first attempt to document and explain deconcentration as a global phenomenon. By moving beyond individual case studies, this broader perspective helps shed light on deconcentration's "hard cases" which do not fit existing explanations. I also extend selectorate theory by showing how winning coalitions and selectorates can be negotiated simultaneously by politicians bargaining over coalition structures (Bueno de Mesquita et al. 2003). Further, my model of deconcentration provides a rigorous account of how parties and institutions of government can evolve endogenously, reconciling perspectives in which party systems are generated by electoral institutions, and those in which the causal arrow is

reversed. Finally, this dissertation improves our understanding of a fundamental question in political economy: why political actors give away powers to others. I show how some politicians are willing to trade away power in order to improve their relative coalitional position. Focusing on the simultaneous negotiation of coalition structures and institutions helps make sense of such disparate problems as power-sharing in nondemocratic regimes, post-war disarmament, court-packing, and the expansion of the right to vote.

1.1 What is deconcentration?

1.1.1 Definitions

I define deconcentration as the creation of new subnational units of government. Typically, this process occurs in one of two forms. First, rulers can create entirely new "layers" of governing units. Second, rulers can create new governing units at existing layers. A layer (also called a "level" or "tier") is a group of units with approximately equal rights and responsibilities vis-à-vis the other units of government, and which do not overlap spatially. A unit is any unique jurisdiction within the national territory. I use "intergovernmental" to refer to relations between layers of government. Where the context is unambiguous, I use "state creation" as a synonym for deconcentration.

An example helps clarify these distinctions. In 1961, one year removed from French colonial rule, Côte d'Ivoire underwent a complete reorganization of subnational government. Ivoirian President Félix Houphouët-Boigny divided the country into three layers: the center, four primary units known as *départements* and, below them, 102 local governments or *sous-préfectures*. Two early episodes of unit creation in the mid-1960s added both *départements* and *sous-préfectures* to

Table 1.1: Deconcentration in Côte d'Ivoire

	Year								
	1961	1963	1967	1969	1972	1975	1977	1978	1985
(Central gov't)	1	1	1	1	1	1	1	1	1
Regions	0	0	0	**6**	6	6	6	6	6
Departments	4	**6**	6	**24**	24	**28**	28	**32**	**49**
Subdivisions	102	102	**108**	108	**127**	127	**162**	162	162

Each cell provides a count of the number of units. Names of layers reflect (translated) conventions at the time of Houphouët's death in 1993. Changes across years are bolded. See Chapter 3 for data sources and coding rules.

the map, but kept the layers of government unchanged. In 1969, however, Houphouët created an entirely new layer of government, transferring the name *départements* to these units while renaming the top-level units "regions." Côte d'Ivoire continually subdivided *départements* and *sous-préfectures* over the next quarter-century, so that by 1985, the country's administrative divisions had been radically transformed from those of 1961. Table 1.1 summarizes the evolution of the Ivoirian state through these distinct episodes of deconcentration.

This definition of layers differs subtly from those of previous studies. Layers are often defined as sets of non-overlapping jurisdictions that cover the entire territorial extent of the country, leaving no "ungoverned" space, such that units at one layer are completely contained within a single unit at the next level up the hierarchy (e.g., Treisman 2007). For instance, in the US, all counties belong to exactly one state, and each state's territory is completely subdivided into counties; similarly, the nation as a whole is completely divided into non-overlapping states.[3] I argue that this definition obscures much of the nuance of deconcentration. Countries

3. As the capital territory, the District of Columbia (DC) is excluded from statehood. While such cases constitute exceptions in other categorizations, as defined here, DC exists on par with the other American states as top-level units. Similarly, I count Nigeria's Federal Capital Territory of Abuja alongside its 36 states.

frequently grant special autonomous status to cities and regions due to natural geographic features, ethnic clustering, or divergent historical trajectories (Hale 2004; Walter 2006b, 2006a). In the United Kingdom (UK), growing emphasis on disparate regional identities in the late twentieth century led to the creation of special representative assemblies in Scotland, Northern Ireland, and Wales. Without an equivalent assembly, England was essentially left as "ungoverned" or residual territory at the new regional level. A definition of deconcentration that emphasizes completeness of each layer would miss these reforms, which together represent one of the more radical changes to the structure of British government since the Acts of Union in 1800. Our definition of deconcentration should allow for such asymmetries.

Empirically, layer creation is much rarer than unit creation: of the 371 cases of deconcentration I identify worldwide over the period 1960-2010, only 23 are coded as layer creation (see Chapter 3). Given this imbalance, it might seem intuitive to narrow my definition of deconcentration to just unit creation. However, layer creation is a superset of unit creation, since by definition there must be at least one new unit at the new layer of government. To ignore layer creation would be to ignore episodes of unit creation. Nor is it clear that there are any differences *in general* between creating some number of units at an existing layer versus creating the same number of units at a new layer. Both are manifestations of the same general phenomenon.

Three further clarifications are helpful. First, an *episode* or *case* of deconcentration refers to a specific edict, law, or judicial decision that formally creates at least one new administrative unit. Nigerian military ruler Yakubu Gowon's television address on 27 May 1967, which informed the nation of regime's decision to create 12 new states, represents a single case. Second, deconcentration is a general phenomenon that encapsulates more specific processes, chief among them administrative unit

proliferation. Grossman and Lewis (2014, 198) describe AUP as "a political process resulting in a large number of local governments splitting into two or more units over a relatively short period." AUP therefore captures only unit creation, not layer creation, and by imposing constraints of both scope ("large number") and time ("short period"), AUP sets aside smaller episodes of deconcentration. Finally, deconcentration as defined here is a process, not a state of being.[4] The outcome of interest in this study is whether this process occurs, and not its particular effect on the map. As Chapter 3 discusses in detail, my measure of deconcentration is a binary indicator, not on ordinal measures of how many administrative units are created (as with, e.g., "government fragmentation" as defined in Grossman, Pierskalla, and Dean 2017).

1.1.2 Related concepts

Deconcentration thus defined is theoretically and empirically distinct from decentralization. Most importantly, deconcentration is more permanent than decentralization, as it draws new lines on the map, requires new government offices, and creates (or formalizes) new political identities. It also commonly requires primary legislation or a constitutional amendment. In contrast, decentralization is limited to the renegotiation of the distribution of rights and responsibilities among political actors, and can be implemented through bureaucratic initiative, (non-)enforcement of existing law, and budgetary procedures. Thus, while deconcentration only occurred in 4.6% of country-years between 1960 and 2010, practically every country de- and recentralizes to some extent every year. Further, unlike decentralization, deconcentration does not *in general* involve transfers of power. Clearly, new units acquire powers where they previously did not exist, yet this differs qualitatively

4. A country with any subnational governmental units is said to be "deconcentrated."

from the systematic transfer of power of downward which defines decentralization. Given these stark differences, I argue that deconcentration and decentralization are, in general, unrelated.

This characterization departs significantly from how scholars typically relate these concepts, mapped in Figure 1.3. Deconcentration is generally understood as but one form of the broader phenomenon of decentralization. Specifically, recent studies have adopted the tripartite division of decentralization into its fiscal, political, and administrative variants (Norris 2008; Schneider 2003).[5] Fiscal decentralization

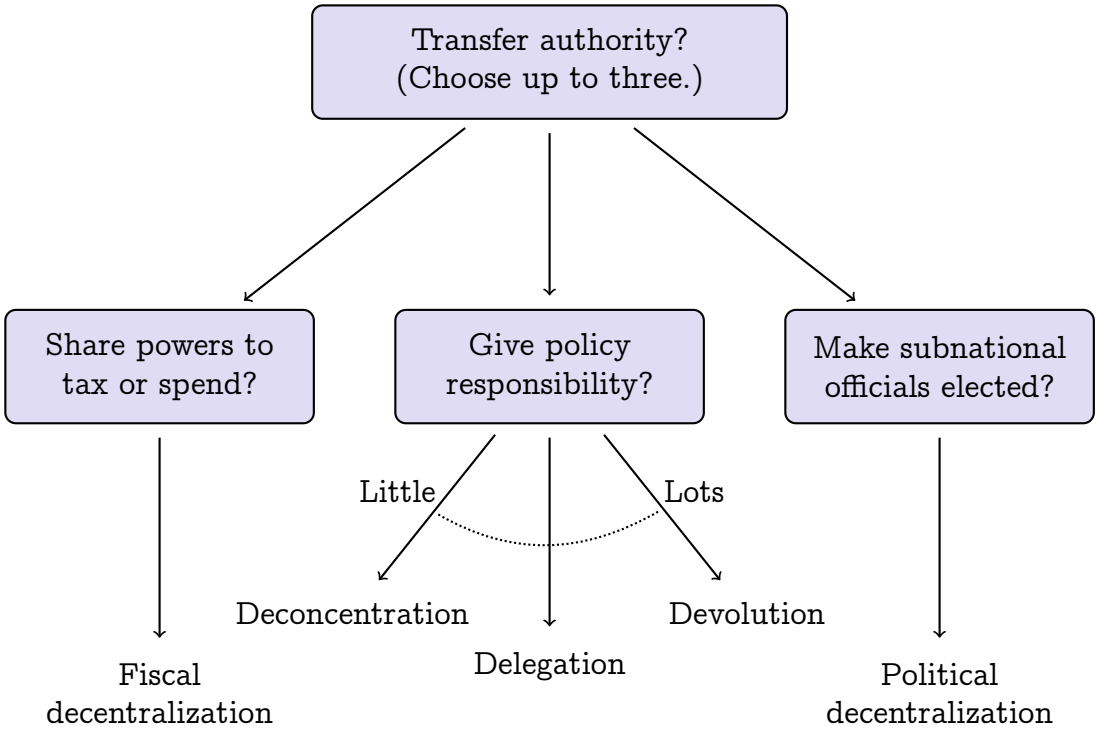


Figure 1.3: Deconcentration and decentralization in the literature. Scholars generally consider deconcentration a weak form of administrative decentralization. Dotted lines indicate a range of outcomes.

5. There are certainly other means of classifying institutional reforms that fall under the umbrella of decentralization. Hooghe et al. (2016), Manor (1999), and Treisman (2007) offer notable alternatives. However, the Schneider (2003) tripartite classification remains influential even among these alternative conceptual maps.

refers to the transfer of state prerogatives to borrow, tax, and spend. In contrast, political decentralization increases local political actors' autonomy vis-à-vis their national counterparts, for instance by mandating the election of local officials. Administrative decentralization in turn refers to a range of reforms that grant more policy and personnel authority to local units. The reforms that fall under the administrative umbrella are usually laid along a spectrum according to the extent to which they transfer real power. At one end is devolution, in which local governments are made largely autonomous, as with the constituent nations of the United Kingdom. Less radical change is typically described as delegation, in which local actors remain accountable to a central authority but acquire policymaking power. Last is deconcentration, which nominally has the least impact on the balance of power between levels of government.

Yet understanding deconcentration as a weak form of administrative decentralization is problematic for both theoretical and empirical reasons. Theoretically, there is little reason to believe that creating more subnational units of government *ipso facto* transfers power downward. Empirically, many scholars have documented a connection between deconcentration and (re-)centralization—the transfer of rights and responsibilities from subnational governments (back) to the center. For instance, deconcentration in Uganda has diminished the administrative capacity and fiscal independence of the average district (Grossman and Lewis 2014). With more units, each individual district has less intergovernmental bargaining power, less technical capacity, and less information about its own citizens. All of these factors also make districts more dependent on patronage from the center (Lewis 2014). Similar mechanisms have linked deconcentration to centralization throughout sub-Saharan Africa, in Argentina and Brazil, and elsewhere (e.g., Dickovick 2011; Eaton and Dickovick 2004; Khemani 2010). An influential World Bank study argues (Manor

1999, 5, emphasis in original), for example:

> When deconcentration occurs in isolation, or when it occurs together with fiscal decentralization but *without simultaneous democratization*—that is, when agents of higher levels of government move into lower level arenas but remain accountable only to persons higher up in the system—it enables central authority to penetrate more effectively into those arenas without increasing the influence of organized interests at those levels. The central government is not giving up any authority. It is simply relocating its officers at different levels or points in the national territory. In such circumstances, it tends in practice to constitute centralization, since it enhances the leverage of those at the apex of the system.

Because deconcentration has no clear effect on the balance of power across levels of government, such studies conclude that deconcentration "hardly warrants consideration" and focus exclusively on decentralization (Manor 1999, 6). I instead set aside the question of how the intergovernmental balance of power shifts and study deconcentration alone. Thus, I remain agnostic on whether deconcentration occurs contemporaneously with decentralization or any other such reforms. Figure 1.4 provides the conceptual map underpinning this dissertation.

1.2 Lessons from previous theories

Because scholars have typically mapped deconcentration into the decentralization framework in Figure 1.3, very few theories attempt to explain deconcentration itself. For example, neither of the two most prominent recent studies of the "architecture of government" mention deconcentration by any name (Filippov, Ordeshook, and Shvetsova 2004; Treisman 2007). Instead, its causes are largely inferred from broader theories about the organization of government, scattered across the federalism and decentralization literatures. I organize these disparate arguments into three broad explanations for deconcentration: elite bargains made to counter military threats,

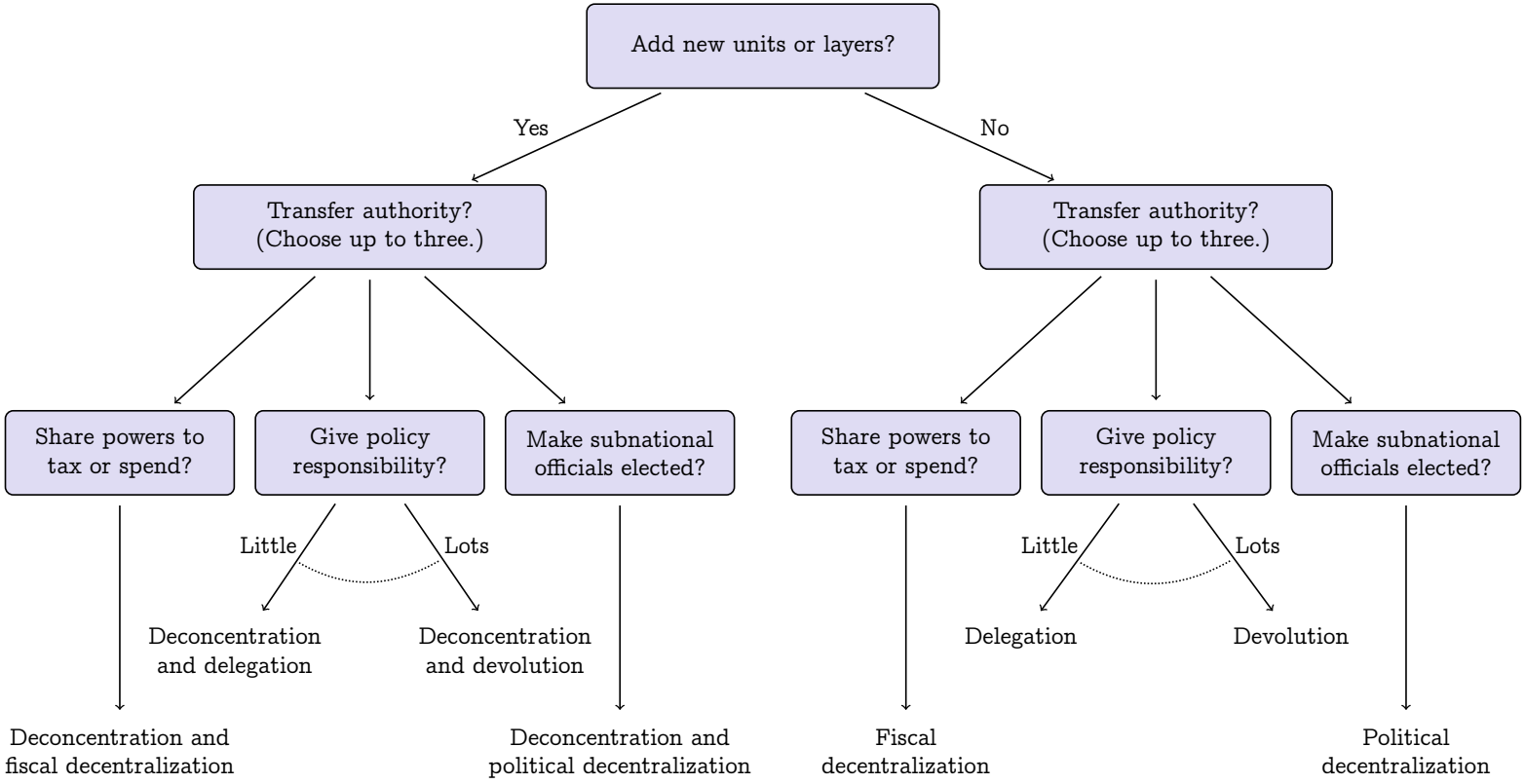


Figure 1.4: Deconcentration is orthogonal to decentralization. Deconcentration and decentralization are independent processes. The dotted line indicates a range of outcomes.

reforms to induce normatively positive outcomes, and institutional manipulation to produce electoral or partisan advantages. Throughout this dissertation, I refer to these theories as the *external threat*, *outcomes*, and *majority-rigging* theories.

1.2.1 The external threat theory

Modern theories of the structure of government trace their roots to early studies of federalism. Federalism is typically defined following Riker (1964, 11):

> A constitution is federal if (1) two levels of government rule the same land and people, (2) each level has at least one area of action in which it is autonomous, and (3) there is some guarantee (even though merely a statement in the constitution) of the autonomy of each government in its own sphere.

Federalization, then, is the process of establishing a federal system by: making a single-tiered government into a multi-tiered one with properties (2) and (3); creating a new "area of action" in which extant levels of government each have distinct authority; or by creating institutional guarantees for each level's autonomy in existing areas of action. Only the first of these involves deconcentration.

This "coming-together" federalism is typically the outcome of bargaining among political elites from the regions of the would-be federation, often the leaders of existing states, who share mutual interests such as a common external military threat (Deutsch et al. 1957; Riker 1964; Stepan 1999; Wheare 1946). In nineteenth-century Western Europe, these incentives were pre-emptive, with Italian and German statebuilders looking to expand the territory and military capacity of their would-be federations to fend off rivals. In both cases, the incentives to federalize were conditioned by the regions' governing capacity; federalism emerged in Germany because its units could credibly govern themselves, but unitary government emerged

in Italy because they could not (Ziblatt 2004, 2006). In the developing world, internal security threats have motivated federalization more commonly than have external ones (Boone 2003). Elites in sub-Saharan Africa, for instance, struck federal bargains to project state power outward from the center and establish control over sparsely-populated peripheral territories (Dickovick 2014; Herbst 2000). The common thread connecting these theories of federalization is the focus on elite bargains made to address military threats.

Explanations grounded in this security dilemma do not travel well to deconcentration. The defense-enhancing effects of Rikerian federalization do not seem likely to hold for deconcentration, since creating new units of government does not bring new territory under a country's control or provide added strength against a foreign military aggressor.[6] Nor does the threat of civil war appear to provide much leverage. Empirically, many of the countries that have experienced deconcentration in the last half-century—e.g., Canada, Denmark, and New Zealand—seem less susceptible to internal security challenges than other countries that have not experienced deconcentration, such as the Republic of Congo or Madagascar (see Figure 1.2).

Nevertheless, studies of federalism underscore the importance of elite-level bargaining to the reorganization of government.[7] Elites' mutual self-interest in reorganizing the state is a minimal condition for creating new political units; without some level of elite buy-in, activists' efforts to create more subnational units (or grant existing units more authority) are dead on arrival. To understand where this buy-in

6. This also makes explanations for what Stepan (1999) refers to as "putting-together federalism"—when a dominant power conquers neighboring countries and incorporates them into a multinational state, such as when Georgia, Azerbaijan, and Armenia were forcibly folded into the Soviet Union—inappropriate for understanding deconcentration.
7. This approach stands in contrast to modernization theories, which view decentralization and deconcentration as the natural results of socio-economic development (e.g., Arzaghi and Henderson 2005).

comes from, scholars have turned to elites' shared interest in political and economic reform.

1.2.2 The outcomes theory

A long tradition in political economy has linked federalization, decentralization, and deconcentration to a range of normatively positive outcomes. Most prominent in this vein are arguments that these reforms can make public policy more economically efficient by ameliorating information asymmetries or facilitating competition (Alesina and La Ferrara 2005; Crémer and Palfrey 1999; Hayek 1944; Montesquieu 1989 [1748]; Oates 1972; Persson and Tabellini 1996; Tiebout 1956; Treisman 2007). Federalization and decentralization are also thought to promote: better democratic representation by making lines of accountability clearer (Hamilton, Jay, and Madison 1892 [1787-1788]; Tocqueville 1964 [1835]; Weingast 1995); interethnic peace and stability by protecting minority groups (Lijphart 1984); and economic growth by fostering local investment (Roland 2000). Implicitly or explicitly, these studies suggest that elites restructure government with the specific aim of inducing these positive outcomes, whether for their direct benefit or due to the electoral incentives tied to realizing these outcomes.

Four problems limit the usefulness of these arguments for understanding deconcentration. First, almost all of them depend on the assumption that the context for reform is a liberal democracy. Many scholars make this requirement explicit (Beramendi 2007; Beramendi and León 2015; Diaz-Cayeros 2006; Filippov, Ordeshook, and Shvetsova 2004; Hale 2004; Watts 1998; Wibbels 2005), while others impose minimal conditions that are hallmarks of liberal democracies, such as functioning constitutional courts (Rodden 2005).[8] Crucially, the mechanisms

8. Treisman (2007, 25) also notes that "some might argue that [subnational elections are] implied by the

linking federalization and decentralization to the positive outcomes are themselves the institutions by which scholars recognize democracies: electoral accountability, non-partisan bureaucracies, effective judicial systems, and others. For instance, Tiebout (1956) provides a classic argument in which federalization produces more efficient public policy. Citizens locate themselves in districts where the median preference (for spending on some public good) is closest to their own, voters elect representatives that best represent their district's preferences, and then bureaucrats implement each district's chosen policy. Thus, citizens "voting with their feet" only works with free movement, free elections, and a functioning bureaucracy.

Yet deconcentration frequently occurs in nondemocracies. A quick glance at Figure 1.2 suggests that established democracies are no more likely (and perhaps less likely) to deconcentrate than nondemocracies.[9] This interpretation is substantiated by a number of case studies. For instance, it does not seem likely that the examples highlighted above—Iran, Uganda, Reconstruction-era US, and Vietnam—meet the minimal conditions of the democratic exemplars assumed in the outcomes school. And while Nigeria's Mid-West Region was created under the country's first civilian government, all of the following six episodes of Nigerian state creation occurred under military rule. In short, if deconcentration were caused by elites jointly pursuing normatively good outcomes, then we would not expect to observe it where the mechanisms that generate those outcomes do not exist—and yet we regularly do.[10]

Second, these arguments also depend on subnational units having the capacity

Riker and Dahl definitions" of federalism.

9. I test the relationship between deconcentration and democracy more rigorously in Chapter 3.
10. This pattern also makes it unlikely that deconcentration can be explained by external pressures as arising from international financial institutions (IFIs). Although the World Bank (1997) notes that deconcentration and decentralization seem "a logical continuation" of the donor-backed third wave of democratization, if deconcentration were largely the result of pressure from such institutions, then it should be observed most frequently in countries dependent on these IFIs. This does not appear to hold generally. Nor does there appear to be a spike in deconcentration during the years when IFI pressure to deconcentrate was strongest (see Chapter 3).

to produce these positive effects. For instance, federalization or decentralization to promote interethnic peace requires subnational governments to have the capacity to constrain the center (and each other) from impinging on minority rights (Horowitz 1985). This seems a poor description of the typical subnational unit in a country experiencing deconcentration. Countries with very weak subnational units, such as Uganda and the Philippines, have been among world leaders in deconcentration—while those with the strongest subnational units, including the US and Australia, have not deconcentrated at all. Were deconcentration motivated by elites' desire for better economic and social outcomes, then it would not be observed where subnational units are so weak that they cannot produce them. Here again, the data do not match theoretical expectations.

Third, even if these democracy and capacity conditions hold, there is not much empirical evidence that either federalization or decentralization actually produces these desirable effects. For example, gains in economic efficiency and growth appear to be contingent on a number of intervening variables such as inequality and the structure of domestic capital (Beramendi 2012; Besley and Coate 2003; Treisman 2007). Further, decentralization has also been linked to a number of negative outcomes. For instance, federations in which subnational units have considerable spending and taxing authority suffer from fiscal irresponsibility, policy stasis, and consequently the inability to respond to macroeconomic crises (Rodden 2002). Decentralization may also contribute to ethnic fragmentation and civil conflict (Bakke and Wibbels 2006). And when territorially overlapping levels of government share the authority to tax, they can end up "overfishing" their shared revenue base, decreasing economic efficiency (Berry 2008). Given these ambiguities and contradictions, it is unsurprising that scholars can find no empirical relationships linking decentralization to normatively positive outcomes *in general* (Treisman

2007). Of course, it is possible that rulers *believe* that deconcentration will yield better outcomes despite this mixed evidence. Yet this suggests a tautological causal story, i.e., "deconcentration is the result of rulers believing it will cause positive economic effects, which we know they believe because we observe deconcentration."

Finally, the outcomes theory struggles to explain the process by which interests coalesce around whatever outcome is being pursued. The explosion of scholarly research in the past few decades in the tradition of the "new political economy" has thoroughly demonstrated that institutional change almost always has distributional implications (Harriss, Hunter, and Lewis 1995; Knight 1992; North 1990). This insight, which follows directly from the Coase (1937, 1960) Theorem, suggests that self-interested actors will contest institutional change if they fear losing out. Thus, federalization and decentralization are likely to be contested by at least some elites, whether in the center or in subnational units. The structure and resolution of these conflicts is therefore a central concern to outcome-centered theories, yet studies in this vein typically assume away or black-box such complications.

In short, theoretical inconsistency and empirical incongruities indicate that deconcentration is not likely driven by elites seeking to induce normatively positive outcomes. But by shining a light on the enormous distributive implications of federalization and decentralization, these studies point to the importance of mechanisms for resolving intra-elite conflict. If the external threat theory teaches us that the reorganization of government depends on elites' mutual interests, the outcomes theory indicates that deconcentration also depends on overcoming the opposition of those who stand to lose out from reform.

1.2.3 The majority-rigging theory

Recent scholarship adopts this conflict-centered approach. The central claim of these studies is that administrative unit proliferation is the result of a temporary alliance struck between citizens, activists, and rulers to enhance their respective prospects in future distributive conflicts (Grossman and Lewis 2014). I discuss each group in turn.

Citizens demand local autonomy because creating new units formally recognizes minority groups and increases ethnic homogeneity within districts (Kasara 2006; Pierskalla 2016). It also requires infrastructure, which creates jobs and spurs investment in community facilities (O'Dwyer 2006). Once established, new units are guaranteed either some measure of national resources or control over local taxation and spending, allowing for more locally targeted public goods provision (Fitrani, Hofman, and Kaiser 2005; Hassan and Sheely 2017; Kimura 2013).

Activists, typically marginalized elites and civil society leaders, also lobby the center for new provinces. Besides benefiting from their share of the boom to be expected from increased local spending, activists are also then able to parlay their role in deconcentration into elected office, putting patronage resources at their disposal (Green 2010). Their career prospects are also propelled forward by residents' greater access to government services (Hassan 2016), the credit for which they are able to claim.

Finally, rulers use AUP to extend their hold on power. Decentralization can present a lifeline for a governing party that expects to lose a looming election at the national level, but to retain power in regional strongholds (O'Neill 2005). Beyond preserving the ruling party's dominion over part of the country, this form of emergency decentralization can also weaken a rival incoming government. More

broadly, rulers can remain in power at the center through deconcentration by expanding patronage networks, appeasing vocal minorities, and rigging electoral majorities (Ayee 2012; Falleti 2005; Hassan 2014; Hassan and Sheely 2017; Kimura 2013; Mawdsley 2002; O'Dwyer 2006; Pierskalla 2016; Riedl and Dickovick 2014). For instance, Green (2010) argues that Uganda's ruling party builds electoral alliances by creating new districts, establishing formal channels for patronage, and then transferring state resources to these districts. Policy preferences are a final motivation: in Vietnam, elites in favor of economic reform created new provinces to build support for relaxed economic controls (Malesky 2009).

This theoretical framework goes a long way in explaining the empirical variation in deconcentration, yet it leaves a number of questions unresolved. For one, it is unclear why citizens, activists, and rulers are not satisfied with less costly and more immediate means of achieving their respective goals. Local demands for patronage can be satisfied directly, without the need for new political units. Activists can be folded into the existing state apparatus or dismissed on a case-by-case basis. And last-minute electoral majorities can be rigged at the ballot box instead of imperfectly engineered years in advance. Since the relatively mundane goals of patronage distribution and self-preservation can be met through less permanent means, it is unclear why deconcentration is needed at all.

Moreover, this framework suggests that citizens and activists will always want new units for themselves (though possibly not for others), and that rulers will always want to manipulate the structure of government to suit their governing agendas. Yet none of these studies offers a compelling explanation for why deconcentration movements are successful in one context but fail in another. Just as grievance theories of civil war are largely unable to explain why some demands are more pressing or their proponents better mobilized (Fearon and Laitin 2003), studies of

AUP do not distinguish why some local activists are more effective than others. Since demand for new units should be near-universal, the empirical variation in AUP and deconcentration more generally is likely to be a function of the *constraints* on those who seek new subnational units. The majority-rigging theory provides little guidance on what those constraints are or when they bind.

Answering these questions may require understanding the preferences of politicians holding subnational office or representing subnational constituencies in national assemblies, such as district administrators, governors, and senators.[11] In many countries, these actors have formal voice in any renegotiation of the structure of the state, through constitutional or legislative channels. Even where such channels do not exist, ties of partisanship and patronage commonly bind rulers to the will of these subordinates. Few scholars have focused on these politicians' role in deconcentration, yet their capacity to constrain rulers has been noted anecdotally in a variety of contexts (Ayee 2012; Grossman and Lewis 2014; Malesky 2009). With both opportunity and motive, these politicians seem well-placed to kill deconcentration movements, and indeed their opposition to unit creation has frequently been decisive (e.g., Suberu 2001).[12] Understanding deconcentration requires a better sense of the factors driving these politicians' support for and opposition to deconcentration.

11. This is a relatively narrow definition, since it excludes bureaucrats, civil society activists, and elites who do not hold office. Yet I argue that this focus is appropriate because the rules and procedures governing deconcentration typically exclude these actors; while they can mobilize support for or against deconcentration, their role is essentially informal, and their influence on the outcome highly uncertain. In contrast, as explained below politicians generally place formal constraints on rulers seeking to create new units or layers.
12. See also the close case comparison of three Nigerian statehood movements, two of which were ultimately unsuccessful, in Chapter 5.

1.2.4 Unanswered questions

This puzzle can be drawn out explicitly in two parts. First, scholars lack a clear understanding of how rulers overcome or bypass opposition to deconcentration among politicians, both within their governing coalition and without.

Politicians representing subnational constituencies seem well-placed to block rulers' attempts to deconcentrate because every polity has mechanisms to distribute power vertically. Liberal democracies extend shared decision-making powers downward, so that local, state, and regional authorities provide active checks on the ability of the center to restructure government without their consent. Federal governments, and particularly confederations,[13] make deconcentration essentially impossible without buy-in from a significant portion of subnational elites (Lijphart 1984). For example, Article IV, Section 3 of the US Constitution formalizes the required consent from Congress, giving every Senator and Representative a voice in state creation.

Such constraints may be lacking in less established democracies, but can be reinforced or replaced by party institutions and particularly clientelist ties. In these contexts, elites at the center sit atop vast partisan networks on which they rely for funding, voter turnout, and lobbying efforts (Ichino 2008; Kitschelt and Wilkinson 2007; Stokes et al. 2014). Deconcentration drives a wedge between rulers and these local officials, and so while the latter may have no formal authority to stop such reform, they do wield a considerable weapon in the ability to cut the party out from underneath the center (Boone 2003). Even in the least democratic contexts, rulers must distribute power vertically among subordinates within their coalitions (Svolik 2012). Deconcentration necessarily affects this balance, potentially upsetting

13. Confederation are typically defined as voluntary unions of sovereign states, wherein each state has a unilateral veto over any policy implemented at the center, comparable to a very loose federation.

the selectorate on which the ruler depends (Bueno de Mesquita et al. 2003). In the extreme, nondemocratic rulers intent on deconcentration may face rebellion or secession if they override elites' objections (Zanker, Simons, and Mehler 2015). Thus, while it is possible to imagine a context in which rulers are able to deconcentrate at will, without the consent of subnational actors, this would seem a poor reflection of reality. Identifying whose support rulers need—and how they get it—is crucial for understanding the empirical variation in deconcentration.

Scholars still lack a consistent theoretical framework for answering these questions. For instance, Malesky (2009) suggests that reformist parliamentarians in Vietnam were able to capitalize on the short-sightedness of those opposed to reform. While this explanation may be plausible in certain cases, it depends on an exceptional asymmetry (whether in information or time horizons) between those fighting for and against deconcentration. This asymmetry seems unlikely to hold more broadly, particularly in countries that have experienced AUP, and where the distributional and institutional consequences of deconcentration are well-known and highly politicized (Suberu 2001).

Others have suggested that subnational elites are aware of the negative consequences of deconcentration, but can be bought with one-off side payments from the center (Grossman and Lewis 2014).[14] Yet this explanation misses the importance of credible commitments. One-off side payments can be recaptured or reneged upon in the future, and so are far less valuable than institutionalized decision-making powers and venues for bargaining with the center. Politicians are keenly aware of this tradeoff, as has been documented in constitutional deliberations in a variety of settings (Negretto 2013; North and Weingast 1989; TNA, CO 879/159).

14. They name both direct and indirect mechanisms for such transfers, including outright bribery and through reducing campaign costs by making districts more homogeneous.

Individual transfers alone are unlikely to sufficiently motivate politicians to give in on deconcentration.

Thus, rulers' attempts to deconcentrate seem likely to run into formidable opposition. We still need a framework that can answer: *Whose support do rulers need to deconcentrate? How are rulers able to mobilize this support?*

Answering these questions depends on understanding subnational politicians' preferences. Despite calls for greater emphasis on these actors' interests (Grossman and Lewis 2014; Pierskalla 2016), scholars still do not understand what explains variation in their support for deconcentration. This presents the second part of the puzzle: deconcentration depends on some measure of support from subnational elites, and yet for this group, deconcentration seems almost certainly a disaster.

There are a number of mechanisms by which these elites are likely to be worse off after deconcentration. First, when new administrative units are carved out of existing ones, many politicians are left with "rump" provinces. Such elites may end up with control over less territory and a smaller population, decreasing their national prominence and stature within their party. Second, deconcentration may expand the pool of political rivals, making each individual's potential climb to national leadership less certain. Third, deconcentration can create more claimants on national common-pool resources, such as those derived from Nigeria's petrochemical exports (Suberu 2001). Every new unit of government has the potential to reduce the average transfer to existing districts, even those whose borders are not directly affected by deconcentration. As subnational governments tend to be dependent on the center for resources (Wibbels 2005), any decrease in such funding can significantly hinder subnational elites' ability to pursue policy goals, provide public goods, create jobs for constituents, sustain clientelist networks, steal for personal consumption, or otherwise use state funds to pursue power and influence. Fourth, to the extent that districts

must coordinate to implement policy—for instance, in adopting painful economic reforms (Wibbels 2005)—these goals may become more difficult to achieve. The incentive to free ride on others' fiscal prudence and shift the negative externalities of bad policy choices to others increases with the number of districts (Rodden 2005). Finally, as the group of subnational elites expands, the collective action problem of resisting central-government incursion on their authority may become more difficult to overcome. Coordination against such encroachments may be more difficult to engineer not only because the cost of organizing could rise, but also because the "focal points" that signify boundaries on the central government's authority may fade (Weingast 1995). As these lines blur, subnational politicians may find it harder to determine if their peers are likely to mobilize, discouraging collective action and diminishing their collective ability to resist the center in the future (Lewis 2014). Given all of these reasons to think deconcentration would hurt these elites,[15] it is hard to imagine them ever supporting it.

Subnational politicians' opposition to deconcentration is indeed common and vigorous—but not unanimous. Often, at least some faction of governors or district chiefs are vocal in their support for new government units. And contrary to scholarly expectations, these conflicts do not fall onto easily interpretable fault lines, such as partisanship, ethnic animosity, or even direct material benefit. For example, deconcentration in Vietnam split the dominant ruling party according to districts' relationship to state-owned enterprises (Malesky 2009). In 1864, five Democratic Senators voted against admitting Nevada to the United States because Nevadans

15. In some contexts, there may be offsetting electoral incentives for subnational politicians to accept deconcentration. For instance, it may be that individual politicians favor deconcentration because their electoral prospects improve if opposition strongholds are carved out of their districts. However, it is unclear why these more favorable districts could not be engineered through simple redistricting, as is common even in advanced democracies (Engstrom 2006; Stephanopoulos and McGhee 2015), rather than via deconcentration.

were widely expected to return Republican Senators (as indeed they did for the following 30 years), but four Democrats voted in favor of the bill. Republicans were similarly divided, voting 18-7 (in favor). Deconcentration has fractured parties and generated unlikely alliances in Kenya, Nigeria, and elsewhere (Cheeseman, Lynch, and Willis 2016; McCarty, Poole, and Rosenthal 1998; Suberu 2001).

In each case, scholars have offered context-specific explanations for this phenomenon. Yet none of these accounts offers a reason why conflict over deconcentration generally fails to map onto well-established fault lines. We still lack a framework that can answer: *What explains whether politicians support or oppose deconcentration?*

Deconcentration has become common around the world, yet scholars do not have an adequate understanding of the causal processes that produce it. We still are not able to explain how rulers mobilize allies to push through deconcentration, nor even the basic patterns of contestation over such reforms. How can rulers deconcentrate if so many actors have a stake in stopping them? Why do subnational politicians support deconcentration when it is almost sure to make them worse off, while others break with their co-partisans to oppose deconcentration? Whose support is necessary for deconcentration, and when do they give it?

1.3 Overview of the argument

I argue that elites pursue deconcentration to manipulate political coalitions. Creating new units of government empowers new politicians, in turn changing the outcome of coalition bargaining within and across levels of government. Because politicians are primarily concerned with their welfare under alternative coalition structures, their support for or opposition to deconcentration is a function of how it affects their

position in the structure of political coalitions. Two institutions determine whether these patterns in support and opposition then translate into successful reform: those determining who has a (formal) say in deconcentration and how governments form. I refer to these rules as the voting and majoritarianism rules, respectively.

I define a coalition as a group of agents who agree to behave cooperatively by choosing and implementing a joint course of action (Ray 2007). This broad definition includes political "teams" ranging from narrow factions within larger organizations (e.g., caucuses in the US House of Representatives) all the way up to groups of parties that together form a coalition government. A "coalition structure" refers to the set of all coalitions, and a "coalition partition" (or just "partition") describes all coalitions and their memberships.

The key feature of deconcentration is that it creates new political offices and so empowers new politicians. I refer to this as the "empowerment effect." Holding office allows these politicians to enter the process of bargaining over political coalitions that occurs within and across levels of government. Crucially, this formal role authorizes them to take part in selecting the national executive and governing coalition, which they could not previously access even if they were already leading figures in economic or social realms. This empowerment effect distinguishes deconcentration from other reforms—redistricting, decentralization, and so on—that alter the intergovernmental balance of power. Only deconcentration lets new politicians take part in executive selection.

Coalition bargaining is complex, and so coalition structures are extremely sensitive to the entry of new politicians. Thus, no matter how few units are created, or how weak they may be, adding new politicians can change the outcome of coalition bargaining in complicated, often radical ways. For instance, Zairian dictator Mobutu Sese-Seko continually rotated allies in and out of key positions so as to cut down rivals

from within his regime who had grown too powerful, allowing his ruling coalition to expand and contract according to political expedient (Young and Turner 1985). Even where new politicians' entry does not directly affect the ruling coalition, it can generate ripples throughout the coalition structure, destabilizing opposition groups. This is precisely what the creation of the Mid-West Region in Nigeria accomplished: by empowering a breakaway faction within the opposition AG, deconcentration facilitated the self-destruction of a major non-governing coalition. As these examples indicate, there is no direct mapping from a new politician's ideology, interests, or power and the likely result on the extant coalition structure.[16]

Politicians' preferences over deconcentration are induced by their preferences over alternative coalition structures—they support or oppose deconcentration according to their expectations about how such changes to the coalition structure affect their individual welfare. But because these changes are so complex, and so dependent on the precise set of people bargaining, politicians' preferences over deconcentration are unlikely to map onto neat divisions such as partisanship or ethnicity. Instead, they are induced by expectations about factors such as whether they will be in the ruling coalition and their relative position within it. So while politicians' motivations may be very similar, the effect of deconcentration across individuals may be radically different. It is the complexity of coalitional bargaining that generates the inscrutable alliances observed in the US, Nigeria, and elsewhere.

Two institutional details determine whether deconcentration proposals are ultimately successful. The first is the voting rule, which establishes who must vote for a deconcentration proposal for it to pass. The second is the majoritarianism

16. This feature of the theory also distinguishes it from majority-rigging arguments. My model allows rulers to use deconcentration to trim their coalitions down closer to minimum-winning size, or to divide opposition parties without building up their own majorities. Further, by focusing on coalitions more broadly than legislative majorities, I am able to account for deconcentration in non-electoral contexts.

rule, which describes the (super-)majority required to form a governing coalition. Both of these rules have recognizable real-world analogues: the voting rule captures provisions such as Article IV, Section 3 of the US Constitution, which specifies the procedure for admitting new states to the Union, while the majoritarianism rule can be seen in the thresholds required to form governments in many parliamentary democracies.

This theory of deconcentration helps address the two main questions scholars have yet to answer. First, to be successful, deconcentration proposals require support from at least enough politicians to satisfy the voting and majoritarianism rules—at least enough politicians to overcome procedural hurdles and to form a ruling coalition. Second, whether these rules are satisfied is a function of how deconcentration proposals affect coalition structures. Politicians support deconcentration if they think they would be better off under the new coalition structure and oppose it otherwise. The complexity of patterns in support for deconcentration reflects the underlying complexity of coalition bargaining.

1.4 Contributions

This dissertation contributes to a number of areas of scholarly interest. Most directly, it is the first attempt to document and explain the global surge in deconcentration over the last half-century. By identifying deconcentration's empowerment effect and demonstrating its importance for coalitional outcomes, I provide a new way of thinking about deconcentration as its own phenomenon, distinct from decentralization and federalization. Understanding deconcentration as coalition manipulation sheds new light on the forces that structure subnational government around the world.

This dissertation also asks new questions of selectorate theory (Bueno de

Mesquita et al. 2003), in which all polities are characterized by a two-dimensional institutional space consisting of the winning coalition (those who govern) and the selectorate (those who choose who governs). Prominent studies in this vein have been largely devoted to studying rulers' attempts to manipulate the size and shape of the winning coalition, and the economic and socio-political effects thereof. However, relatively little attention is paid to studying how politicians can manipulate the selectorate, commonly the broader political elite from which the winning coalition is drawn. I show how both of these institutions can emerge endogenously in a single game of coalition bargaining. Politicians bargain over whom to empower into the selectorate through deconcentration, which then determines the shape of the winning coalition.

My analysis also provides insight into the strategic origins of parties and electoral institutions in both democracies and nondemocracies. For decades, scholars have documented parties' origins in politicians' strategic response to extant electoral institutions including presidentialism and majoritarianism (Cox 1997; Duverger 1954). More recently, however, such institutions have themselves been shown to be determined by partisan competition (Boix 1999; Robinson and Torvik 2015). The explosion in scholarly study of "democratic" institutions in nondemocratic regimes has only deepened this divide (Boix and Svolik 2013; Brancati 2014; Gehlbach, Sonin, and Svolik 2016). The wealth of evidence on each side of this debate suggests that both forces are likely at play. Yet comparatively few studies have provided a rigorous explanation for how parties and representative institutions evolve endogenously. My theory highlights coalition bargaining as one mechanism that jointly produces new partisan alignments and new institutional arrangements.

More broadly, my theory provides an answer to why strategic actors are sometimes willing to give away their power. I show that the changing structure of

coalitional memberships can motivate politicians to prefer deconcentration, even if it diminishes their power relative to other actors. This finding has broad implications for theories wherein rulers and other elites are faced with the option to trade away their own authority. Among the applications of this framework are post-conflict disarmament, demobilization, and reintegration (Humphreys and Weinstein 2007); the transition from military to civilian rule (Acemoğlu, Ticchi, and Vindigni 2010); and the emergence of representative institutions in nondemocracies (Svolik 2012). My findings suggest that in each case, the powerful can find it in their own interest to give power away if they expect to be better off under the new coalition structure induced through deconcentration.

1.5 Layout of the dissertation

I develop the theory of deconcentration as coalition manipulation more fully in Chapter 2. After describing the empowerment effect, I show how deconcentration can impact coalition bargaining, and describe the role of institutions in structuring bargaining outcomes. I then derive four empirical implications relating the institutional and coalitional context to the probability of deconcentration. Although the exposition does not rely on a mathematical model, throughout the chapter, I provide brief formal examples to help build intuition.

Chapter 3 presents the first tests of the theory. After introducing new data on deconcentration worldwide over the 1960-2010 period, I use observational data to examine the predictive power of the coalition manipulation model. The majority of this chapter is devoted to a large supervised machine learning task in which I compare the informational content of the key variables in my theory against a set of controls and predictors derived from alternative explanations. Results from this exercise

indicate that the coalition bargaining theory provides much greater explanatory power than do alternative accounts, and generates predictions approximately as accurate as those of predictive models studied elsewhere in the discipline.

In Chapter 4, I turn to establishing the causal link between deconcentration and coalition manipulation. In line with the predictions of Chapter 2, exogenous shocks to the set of politicians bargaining over coalitions significantly increase the probability of deconcentration. After presenting the main results, I study a variety of robustness checks and threats to inference. I then demonstrate, using data on cabinet size and composition, that the mechanism linking sudden leadership deaths to deconcentration holds. Together, these results isolate coalitional considerations as the driving force behind the creation of new subnational units. More than just being a useful predictive tool, my theoretical model uncovers the causal origins of deconcentration.

I trace this causal mechanism more closely in Chapter 5. Using new data from the National Archives of the United Kingdom, I conduct a case comparison of the three most prominent statehood movements in late-colonial and early-independent Nigeria: those advocating the proposed Mid-West, Calabar-Ogoja-Rivers, and Middle Belt regions. Among these, only the Mid-West was successful in becoming an independent region before a coup ended civilian government and sparked a chain of events culminating in the Civil War of 1967-1970. This matched comparison allows me to study variation across movements under fixed conditions, as well as institutional variation over time, as key political actors struggled to alter coalitions in their favor. Together, these cases demonstrate the causal role of coalitional considerations in dictating deconcentration in Nigeria's First Republic.

Finally, Chapter 6 reviews the argument and evidence, and discusses their implications. I demonstrate the utility of these findings for theories of regime

transition, institutional change, and political economy. This chapter also suggests ways to extend this theory of deconcentration in order to better understand how governments organize spatially more broadly.

2 | The coalition manipulation theory

> Everybody can see what is happening in Scotland and in Wales. What has England got out of devolution? Not a sausage.
> —Mayor of London Boris Johnson to a Parliamentary Select Committee (3 March 2014)

This chapter introduces the theory of deconcentration as coalition manipulation. I proceed in six steps. I begin by describing the empowerment effect by which deconcentration invests new politicians with institutionally-derived authority. I then show that this authority enables politicians to enter coalition bargaining. Next, I develop a small bargaining model to demonstrate how adding new political actors can radically change coalition structures. I then show how these different bargaining outcomes can provoke politicians' preferences over deconcentration, before describing the institutions that structure outcomes and ultimately determine whether deconcentration occurs. Finally, I close by deriving the hypotheses that guide the empirical analysis in Chapters 3-5.

2.1 The model

2.1.1 Deconcentration empowers new politicians

Deconcentration invests new politicians with power. In even the most centralized polities, subnational administrative units require political officers to govern them. Deconcentration necessarily generates executive (and sometimes legislative) vacancies in new units which, in the aggregate, requires some offices to be filled with candidates

who did not previously hold such positions of formal authority. Thus, creating new government units increases the pool of officials who hold executive or legislative authority at any level of government within a country (whom I broadly refer to as *politicians*).

The key feature of deconcentration's "empowerment effect" is that it grants formal, institutional voice to new actors. Scholars may be concerned that deconcentration only makes formal pre-existing, informal power relationships—in other words, that it simply reflects the set of politicians rather than changes it. However, this argument is difficult to square with the simple observation that far more demands for deconcentration are denied than are granted. All over the world, activists mobilize for new government units and fail. In the US, more than a dozen proposals for new states have broken down at various stages of achieving recognition, while DC and Puerto Rican statehood remains a contentious issue. In Indonesia, lobbying for new districts has become relatively universal, with only a selection of the proposals accepted (Fitrani, Hofman, and Kaiser 2005). And in Nigeria, hundreds of unsuccessful statehood applications outweigh the 33 new states that have been created. Forty-five such movements made applications at the 1994-1995 Constitutional Conference, only six of which were successful, as presented in Table 2.1. It is unlikely that activists and politicians would spend considerable effort—often involving years of lobbying and enormous sums paid to key actors—to secure deconcentration if it merely reflected existing power relationships.

Even if it were just formalizing existing relationships, deconcentration would still effect an important change in giving new politicians formal access to the leadership selection process. For instance, in many political systems based on the Westminster model, the Prime Minister is chosen through a vote among parliamentarians (as in the UK). Elsewhere, party officials have veto powers over

Table 2.1: Statehood movements at Nigeria's 1994-1995 Constitutional Conference

State (1994-1995)	Proposed state	State (1994-1995)	Proposed state
Abia	Aba	Kano	Tiga, Gari, Tigari
Adamawa and Taraba	Sardauna	Katsina	Karadua
Akwa Ibom	Itai, Atlantic	Kogi	Okura, Okun
Anambra	Ezu	Kwara	Oya, Yoruba, Ekiti
Bauchi	*Gombe*, Katagum	Niger and Kebbi	Kainji, Ndaduma, Nupe
Benue	Apa, Katsina-Ala	Ogun	Ijebu-Remo
Cross-River	Ogoja	Ondo	*Ekiti*
Delta	Anioma, Toru-Ebe	Osun	Oduduwa
Edo	Afemesa	Oyo	New Oyo, Oke-Ogun, Ibadan
Enugu and Abia	*Ebonyi*	Plateau	*Nasarawa*
Imo	Njaba	Rivers	*Bayelsa*, Niger Delta, Orashi, Ogoni, Rivers East, Port Harcourt, New Rivers, Oloibiri
Jigawa	Hadejia, Lautai, Bayajida	Sokoto	*Zamfara*, New Sokoto, Sakkwato
Kaduna	Gurara	Taraba	Mambila

"State (1994-1995)" refers to the state from which the proposed state would be excised. Proposals in italics were successful, and became new states in 1996. Data are from Suberu (2001).

presidential candidates, whether preceding elections (Nigeria) or through normal legislative processes (China). Even in contexts where the executive is directly elected, politicians still hold enormous sway over the composition of the cabinet, and can thereby impose their will on the broader shape of the ruling coalition. In short, across regime types, one constant is that powerbrokers need to hold office to have a formal say in executive selection. Even if deconcentration just reflects existing power relationships, the act of reflection itself has important consequences for whose voice matters in deciding who will govern.

More abstractly, even in the classical view of institutions as epiphenomenal (Riker 1980), the key feature of institutions as "congealed tastes" is that they are *congealed*. While deconcentration likely grants offices to actors who already enjoy some public stature, it also embeds them in new sites of formal authority, crystallizing their political resources into permanent structures. These structures then constrain and stabilize political outcomes in a way that informal power relationships cannot. "Structure-induced equilibria," which depend not only on the players' preferences but also on the procedural rules underpinning bargaining, are stickier than "preference-induced equilibria" precisely because formal institutions are more difficult to change than their informal counterparts (Shepsle 1989). Thus, even under the view of deconcentration as capturing existing power relationships, the act of formalizing new sites of political power is important in and of itself because it alters the set of potential bargaining equilibria.[1]

To be clear, empowering new politicians is not the only effect of deconcen-

1. This understanding resonates with much recent theorizing within the "institutional turn in comparative authoritarianism" (Pepinsky 2014). Scholars have identified a diverse range of institutions that nondemocratic rulers use to structure political order, including parties, elections, constitutional courts, and "free" media (Brancati 2014). These institutions, while endogenous to elite competition, place constraints on feasible short-run outcomes and so are important parameters in their own right (Gandhi and Lust-Okar 2009; Gehlbach, Sonin, and Svolik 2016).

tration. Creating new subnational governments has other important implications for intergovernmental relations, particularly as it affects resource distribution and the scope for collective action among subnational elites, as discussed in Chapter 1. However, these other effects offer limited power for explaining deconcentration. Such factors often influence subnational politicians evenly, and so cannot explain variation in support for deconcentration among them. For example, state creation in Nigeria creates more claimants on common pool resources. We might expect deconcentration to generate uniform opposition among Nigerian governors, who would stand to lose funding, and yet they frequently advise and support statehood movements. Further, while the empowerment effect is unique to deconcentration, these other effects also obtain for alternative means of reorganizing the architecture of government, including federalization, decentralization, and redistricting. A theory of deconcentration built around its effect on electoral constituencies would be less convincing, for example, because the same theory could equally explain gerrymandering, and would struggle to distinguish the two. Thus, focusing on the empowerment effect helps make sense of variation in deconcentration in ways that studying other mechanisms cannot.

2.1.2 Politicians bargain over coalitions

Empowering new politicians changes who is able to engage in bargaining over coalitions. Following deconcentration, actors who were previously civilians or activists enter formal political office, where they ally with other politicians to fulfill the duties of their new office. Empowerment via deconcentration is the necessary condition for participation in formal coalitional bargaining. It is also sufficient, since holding political office without engaging in coalitional politics is virtually "unthinkable" (Schattschneider 1942, 1).

Building and maintaining coalitions is fundamental to holding political office. Teamwork allows politicians to achieve mutually favored outcomes, generating "coalitional surplus" which is divided among coalition members. This coalitional surplus can take many forms. In democracies, parties allow politicians to access a range of important benefits, including a share of party-controlled resources, the opportunity to have a voice in legislating, and an electoral "brand" (Aldrich 1995; Cox 1987; Lupu 2016). Parties in nondemocracies typically offer careers and patronage opportunities that are otherwise unobtainable for most politicians, while also providing important monitoring and power-sharing mechanisms for rulers (Brancati 2014; Svolik 2012). Further, coalitional surplus is not limited to parties. Factions form within parties to influence policy positions and to direct pork-barrel spending (Ceron 2012; Persico, Pueblita, and Silverman 2011). Legislative member organizations generate social networks that facilitate the exchange of policy-relevant information (Ringe and Victor 2013). Coalitions of parties emerge to form governments, giving party leaders greater influence on policy outcomes and private benefits such as ministerial posts (Laver 1998). In sum, coalition structures themselves are important determinants of "who gets what, when, how" (Lasswell 1936).

Precisely how coalitional surplus is distributed is the subject of a long-standing scholarly debate. Canonical models of coalitional bargaining allow coalition members to credibly agree to the division of the group "pie" in advance (von Neumann and Morgenstern 1944). Yet this approach assumes that coalitional contracts are universally enforceable, and ignores the negotiation of these contracts entirely. In contrast, models of non-cooperative bargaining are built around individuals making coalition proposals, and other players voting to accept or reject these proposals, in the absence of credible commitments. Participants vote according to their expected utility under each outcome, conditional on their expectations about how everyone

else is likely to vote. Despite the absence of enforceable contracts, coalitions emerge in equilibrium even when players are strictly self-interested and cannot commit to sharing surplus, as each individual understands that others share the same incentives to act as-if cooperatively. This theoretical framework has come to dominate scholarly thinking on coalition formation over the last 40 years (Baron and Ferejohn 1989; Chatterjee et al. 1993; Ray 2007; Ray and Vohra 2014; Rubinstein 1982; Ståhl 1977).

My theory of deconcentration is built around non-cooperative bargaining because agreements to form *political* coalitions are notoriously unenforceable (compared to economic ones). Negotiations are often conducted outside formal channels even in rigidly institutionalized and transparent political systems, through what Golder, Golder, and Siegel (2012, 428) call "secretive backroom dealings." Even widely-publicized contracts (such as the agreement between the UK's Conservatives and Liberal Democrats struck in May 2010) carry no mechanism for sanctioning coalition members who renege on their commitments. Typically, the only option available to those not receiving their share of the coalitional surplus is to exit the agreement and receive none of it.[2] This fluidity is only exaggerated in less democratic contexts; Adolf Hitler and Indira Gandhi both famously reneged on coalitional agreements immediately after assuming power. Moreover, since coalitional contracts are known to be unenforceable, politicians cannot be induced to accept payoffs that depend on side payments after bargaining finishes. Politicians' attempts to manipulate coalition structures must be independent of such transfers if they are to be successful.

Besides through contracting, credibility might also arise naturally. In selectorate theory, politicians are assumed to have innate affinities to one another due to cultural or economic ties (e.g., Bueno de Mesquita et al. 2003, 62). These affinities

2. Or as Laver and Shepsle (1996, 12) suggest, "nailing the SOB next time around."

are common knowledge, so coalitional proposals are naturally credible if the proposed division of coalitional surplus matches them. However, credibility via affinity makes little sense once the possibility of deconcentration is introduced, since it flows entirely from shared affinities *among the current set of politicians bargaining*. If membership in the selectorate is itself endogenous to coalition bargaining, then offers are no longer credible because future bargaining might include new politicians with greater affinity to coalition formateurs. Thus, like contracts, affinities cannot be relied upon to generate credible coalitional promises.

2.1.3 Empowering politicians can change coalition structures

One mechanism for manipulating coalitions that does not rely on either credible contracts or innate affinities is adding new players to the bargaining. The key insight from formal models of coalition formation is that, in general, *equilibrium coalition structures depend on the identities of the players*. Adding even one player to a bargaining game can generate enormous effects that ripple through the extant coalition partition, changing a number of coalitional alignments.

This sensitivity is one of the central insights of Acemoğlu, Egorov, and Sonin (2008, "AES").[3] In the AES model, imposing even a very weak bargaining protocol (similar to that of Baron and Ferejohn 1989) yields a unique equilibrium coalition structure. But if the player set fed into the bargaining changes, then the collection of potential equilibrium coalitions changes, and the payoffs to each player from each of these potential outcomes changes. Further, the number of ways to partition players into coalitions grows with the Bell (1938) numbers, an exponential sequence.[4] Thus, adding new players changes *every* potential equilibrium payoff for *every* player, and

3. See their Proposition 2.
4. For instance, four players can organize into 15 different coalition structures, five into 52, and six into 203. With 20 players bargaining there are more than 50 trillion potential coalition structures.

adds many new potential payoffs. Even the smallest deconcentration "corresponds to a 'large' shock and may change the nature of the ruling coalition dramatically" (Acemoğlu, Egorov, and Sonin 2008, 1003).

A simple model helps illustrate this point. Consider a bargaining game in which three politicians are trying to form a government to administer their collective territory. Each wants to rule so that she can control the distribution of an infinitely divisible natural resource of fixed size (normalized to one). To achieve such control, she may join a (single) team called a coalition. Assume that successfully forming a ruling coalition requires 75% of the total *power* controlled by the three politicians. Here, power is just an exogenously-given individual endowment that fully encapsulates each politician's identity. I remain agnostic on the source of these endowments: they can be financial (e.g., support from major donors) social (the loyalty of an ethnic group), military (close ties to armed groups), or otherwise (such as support of an expatriate community, familial ties to other powerbrokers, or friendships with foreign dignitaries). Broadly, these power endowments correspond to the ability to push political outcomes closer to individuals' ideal points.

In each bargaining round, an agenda-setter is chosen at random with equal probability. She then proposes any *feasible* governing coalition, one that satisfies the 75% threshold, which can be a singleton if anyone is individually powerful enough to rule alone. Everyone named in the proposed coalition then votes sequentially.[5] If the would-be rulers unanimously consent—no one can be forced to join a coalition—then the proposal passes, the governing coalition forms, the game ends, and the resource is distributed. If at any stage a politician votes against a coalition that would include

5. Sequential voting rules out weakly dominated strategies and uninteresting equilibria that arise from coordination failures such as "I vote no because I think you will too, and you think likewise" (Ray and Vohra 2014, 255). This bargaining protocol also echoes that of Acemoğlu, Egorov, and Sonin (2008) and Baron and Ferejohn (1989).

her, the proposal fails and a new agenda-setter is chosen at random from among those who have not already been picked. Assume that if each politician has taken a turn as the agenda-setter and no governing coalition has formed, then the grand coalition of all politicians forms, the game ends, and the resource is distributed among everyone.

Index the politicians $i = 1 \dots N$, where N in this example is just three, and each politician's power endowment is given by γ_i. Denote a coalition S, the set of all possible coalitions $\mathcal{S}$, a coalition structure π, and the set of all possible coalition structures $\prod$. The power of any coalition S is simply the sum of its members' endowments such that $\gamma_S = \sum_{i \in S} \gamma_i$. Allow the governing or ruling coalition to be represented by $R \in \pi$ and the set of all ruling coalitions by $\mathcal{R}$. I refer to the j^{th} coalition structure as π^j, and the coalition that rules it as R^j.

Preferences in this model are a function of idiosyncratic endowments and coalitional memberships. For example, we can define a utility function

$$u_i(R) = \begin{cases} \frac{\gamma_i}{\gamma_R} & \text{if } i \in R, \\ 0 & \text{else.} \end{cases} \tag{2.1}$$

Thus, each politician prefers to be part of the ruling coalition, and to be more powerful within it. This particular utility function is not necessary to obtain the results below, but its simplicity helps focus attention on the key mechanism by which deconcentration can upset a coalition structure.

In order to characterize the equilibrium of this bargaining game, we need to fix the identities of the politicians by assigning them endowments. Suppose without loss of generality that these are given by $\{\gamma_1, \gamma_2, \gamma_3\} = \{6, 7, 8\}$. These endowments can be alternatively expressed as a proportion of the total power in the polity, given

by $\{\widetilde{\gamma}_1, \widetilde{\gamma}_2, \widetilde{\gamma}_3\} = \{0.29, 0.33, 0.38\}$. (I use γ for clarity of exposition, since $\widetilde{\gamma}$ varies as a function of the player set.) For these politicians, the set of all possible coalition partitions and their associated governing coalitions are given by

$$\prod = \left\{ \begin{array}{l} \pi^1 = \{\{1,2,3\}\}, \\ \pi^2 = \{\{1,2\},\{3\}\}, \\ \pi^3 = \{\{1,3\},\{2\}\}, \\ \pi^4 = \{\{2,3\},\{1\}\}, \\ \pi^5 = \{\{1\},\{2\},\{3\}\} \end{array} \right\} \text{ and } \mathcal{R} = \left\{ \begin{array}{l} R^1 = \{1,2,3\}, \\ R^2 = \emptyset, \\ R^3 = \emptyset, \\ R^4 = \emptyset, \\ R^5 = \emptyset \end{array} \right\}. \tag{2.2}$$

Trivially, no one- or two-player coalitions are powerful enough to rule. Since this polity requires a 75% majority to govern, any ruling coalition must possess $\gamma_R \geq .75\gamma_N$. Only the grand coalition satisfies this constraint, so as the only feasible ruling coalition, it is the one that forms in equilibrium. Denote this equilibrium coalition structure π^* and define a payoff function $\varphi(\cdot)$ from the utility function in Equation 2.1. Payoffs are then given by $\varphi(\pi^*[N=3]) = \{\frac{6}{6+7+8}, \frac{7}{6+7+8}, \frac{8}{6+7+8}\}$, or just $\varphi(\pi^*[N=3]) = \{0.29, 0.33, 0.38\}$.

Contrast this outcome with the ruling coalition that emerges from an environment identical in every respect, save the set of politicians. Assume now that four politicians with endowments $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4\} = \{6, 7, 8, 3\}$ are bargaining, so that their relative endowments are given by $\{\widetilde{\gamma}_1, \widetilde{\gamma}_2, \widetilde{\gamma}_3, \widetilde{\gamma}_4\} = \{0.25, 0.29, 0.33, 0.13\}$. With more politicians, the solution space expands considerably to 15 potential coalition structures. Omitting those in which there are no feasible ruling coalitions,

the potential partitions are

$$\prod = \left\{ \begin{array}{l} \pi^1 = \{\{1,2,3,4\}\}, \\ \pi^2 = \{\{1,2,3\},\{4\}\}, \\ \pi^3 = \{\{2,3,4\},\{1\}\} \end{array} \right\} \text{ and } \mathcal{R} = \left\{ \begin{array}{l} R^1 = \{1,2,3,4\}, \\ R^2 = \{1,2,3\}, \\ R^3 = \{2,3,4\} \end{array} \right\}. \tag{2.3}$$

Three salient features of this environment allow us to pin down which coalition structure emerges in equilibrium.[6] First, since a majority of power is required to govern, there are no disjoint candidate ruling coalitions (i.e., each element in $\mathcal{R}$ must have at least one politician in common with another element in $\mathcal{R}$). Second, as noted above, conditional on being a member of the ruling coalition, each politician's payoff is highest when her allies' power is minimized. Finally, since no one can be forced to join a coalition, any politician can reject less attractive coalitional memberships and wait for an offer for which she will receive a higher payoff. Together, it follows that *at least two players must share an ideal coalition structure, and they can veto all other proposals, ensuring their most favored outcome obtains.*

In this example, politicians two and three must be included in any ruling coalition—no coalition can meet the 75% threshold if either is excluded. They are therefore able to hold out for their most-preferred coalitional ally, the weakest politician who lets them clear this threshold, and veto any other proposals. Here, that ally is the fourth politician. The equilibrium partition is therefore π^3, so that the politicians receive payoffs $\varphi(\pi^*[N = 4]) = \{0.00, 0.39, 0.44, 0.17\}$.

This stylized environment demonstrates that adding a politician can substantially change the outcome of coalition bargaining, even when everything else is

6. This assumes a weak genericity assumption holds—that two coalitions with exactly the same total power are identical. See Assumption 2 and Theorem 3 in Acemoğlu, Egorov, and Sonin (2008). As AES note, this assumption is without loss of generality.

unchanged. Just by adding a fourth politician—even a relatively weak one—the identities of the rulers changed, and the polity as a whole turned consensus into divided government. This result does not depend on a complicated model: here, the resource size is fixed, there are no opposition coalitions, all information is common knowledge, there is no discounting, and the bargaining protocol need only include sequential voting. Nor does it depend on the particular parameters chosen in this example, as AES demonstrate. In general, the identities of those who govern is very sensitive to the identities of those involved in coalition bargaining.

2.1.4 Politicians form expectations about coalition structure change

We can now answer one part of the deconcentration puzzle. Since empowering new politicians can have such a strong effect on which coalitions form, politicians' preferences over deconcentration are induced by their preferences over alternative coalition structures. Their support or opposition to unit creation depends on how it will affect their position in the coalition structure.

Continuing with the running example, suppose that, before any coalition bargaining occurs, the three politicians consider ceding some of their territory to a *candidate politician* (or just "candidate") and folding him into coalitional negotiations. This stage of *deconcentration bargaining* occurs immediately before coalition bargaining, and with a similar protocol. One politician is chosen at random with equal probability, who can then propose deconcentration or pass (i.e., she cannot propose the existing player set). Assume that, unlike coalition proposals, deconcentration proposals require only a simple majority to succeed. If she proposes to add the candidate, the three politicians vote on the proposal. If the

vote succeeds, the candidate is added and all four proceed to coalition bargaining. If the agenda-setter passes, or if politicians reject her proposal, then a new proposer is chosen at random, and the cycle continues. If all politicians have had a turn as agenda-setter and no deconcentration proposal has succeeded, then coalitional bargaining begins with just the existing set of three politicians. The three- and four-player bargaining environments introduced above can then be described as subgames of the bigger deconcentration game, presented in Figure 2.1.

During the deconcentration bargaining phase of the game, politicians "look

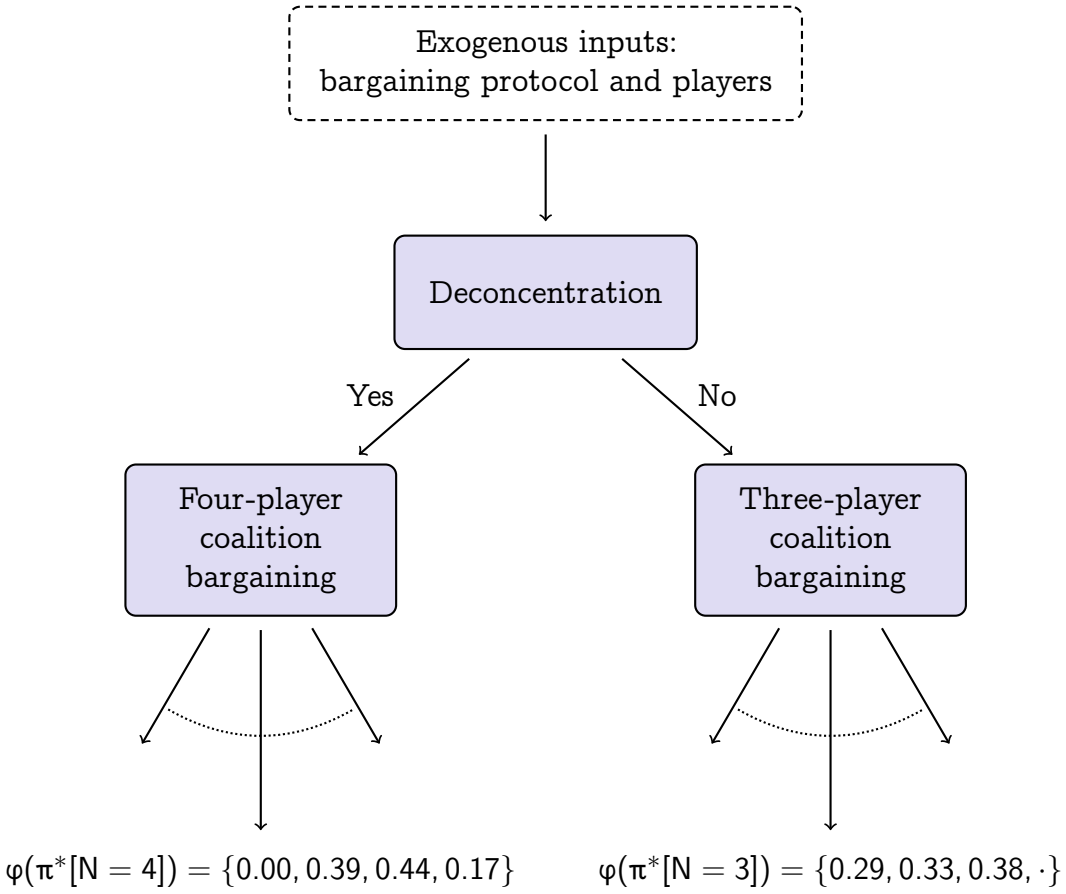


Figure 2.1: The deconcentration game. Dotted lines indicate a range of outcomes. The candidate politician does not receive any payoff (zero or otherwise) if deconcentration does not occur. The unique equilibrium in this game is $\pi^*[N = 4]$, in which deconcentration occurs.

down the game tree" to see how empowering a candidate will change the coalition structure, and what their subsequent payoffs will be. Votes in the deconcentration bargaining stage follow directly from politicians' expectations about the coalitions that emerge from each subgame. In this example, the first politician is worse off after the candidate politician is empowered, and so we would expect her to vote against any deconcentration proposals. However, since the second and third politicians receive higher payoffs from the coalition structure that emerges from four-player bargaining, they will propose and vote for the candidate to be added. And since a simple majority is sufficient for a proposal to succeed in this polity, deconcentration occurs in the unique equilibrium of this game.

This analysis assumes that politicians are aware of the candidates' identities. Preferences over deconcentration are only provoked because politicians can back out its effect on their respective positions in the coalition structure, which requires some knowledge of the new politicians joining coalitional bargaining. However, in general, this assumption seems reasonable. An "identity" is formally defined as a power endowment, so politicians need only know the rough shape of a player's political resources to guess the likely change in the coalition partition caused by his entry. These features are publicly known in most contexts, especially since most advocates of deconcentration mobilize around the creation of *specific* new subnational units. For instance, returning to the US example, Western Democrats voted in favor of Nevada's statehood with the knowledge that whomever Nevadans voted into office would support pro-Western policies. Stewart and Weingast (1992, 13) explain:

> The economic development of the west—defined as the states and territories west of the Mississippi—hinged on the adoption of generous land and public works policies to speed settlement of the region and exploitation of its resources. Because the addition of two more western senators would improve the prospects for generous land, public works,

> and railway policies, western legislators—even Democrats—were more likely to support the admission of Nevada.

Thus, Western Democrats favored deconcentration because they knew the preferences and power endowments of their prospective colleagues: pro-West, with two votes to add to the Senate's then-48 seats.

This example also highlights two important features of the theory. First, deconcentration can impact co-partisans differently. While empowering the candidate increases payoffs for the second and third politicians, the first politician is made worse off. A real-world analogue has already been noted in Mobutu's strategic promotion of weak allies to cut out stronger ones who posed a potential threat (Snyder 1992). Where the context is far more complex than the stylized environment studied here, winners and losers are likely to be found both inside and outside of the ruling coalition. These disparities help clarify why deconcentration frequently produces strange fault lines. Second, politicians can prefer deconcentration even when it will make themselves weaker. The second and third politicians ceded territory to the candidate they empowered, and correspondingly saw their share of the total power in the polity drop from 33% and 38% to 29% and 33%, respectively. Nevertheless, deconcentration improves their payoffs because it makes them stronger within the ruling coalition. The prospect of a better position in the coalition structure can motivate politicians to give power away.

2.1.5 Institutional rules determine feasible outcomes

We can now answer the second half of the deconcentration puzzle. The institutions determining whether deconcentration occurs are those determining whether deconcentration proposals are successful in the first stage of the game, and the minimal size

of a ruling coalition in the second stage. I call these the voting and majoritarianism rules, respectively.

First, the voting rule establishes *whose consent is required for adding new players to the bargaining environment.* Formally, the voting rule specifies particular politicians, or proportions of politicians, who must vote in favor of a deconcentration proposal for the vote to pass. This rule is a close cousin of the voting rule in the Baron and Ferejohn (1989) model of legislative bargaining, as well as the q rule in Austen-Smith and Banks (1999, 62). In the foregoing example, I assumed a voting rule of a simple majority, where each politician's vote counted equally.

The voting rule need not be so simple. For instance, vote shares could be unequally weighted, as with publicly-traded companies' shareholder votes. It may also differentiate types of actors. We can imagine a bargaining context where the players N consist of multiple types, for instance, senators and governors (so that $N = \{N_s, N_g\}$). A voting rule in this context could specify that deconcentration requires the consent of at least half of the senators and two-thirds of the governors. Further, the voting rule could identify *particular* actors whose consent is necessary for deconcentration, equivalent to a list of veto players. For most polities, this will include at a minimum whoever wields ultimate national executive authority (Tsebelis 2002). Thus, the voting rule can become more complex not only in requiring a larger proportion of players to vote for deconcentration, but also in requiring a greater number of specific actors' votes.

The voting rule also need not be explicit. Deconcentration bargaining frequently occurs in informal contexts lacking clear rules about whose vote matters. All of Nigeria's deconcentration since 1967 has occurred under military rule, which suggests a voting rule equivalent to "sufficient support among the top brass so as

to not provoke a coup." Starker still, Friedrich and Brzezinski (1956) suggest that from the late 1920s until 1953, the Soviet Union's voting rule was simply Stalin's approval. Thus, while many voting rules are codified, like the US Constitution's New States Clause, this need not be so. The voting rule is instead best understood as being defined by a bundle of formal and informal mechanisms establishing whose consent is necessary for deconcentration.

The second key institutional detail is the majoritarianism rule, which defines the minimal (super-)majority required to form a governing coalition. In the example above, I fixed the majoritarianism rule at $.75\gamma_N$. More broadly, this rule is likely to reflect some proportion $(0.5, 1]$ of the total power required to govern. For example, in the classical majoritarianism rule of Riker (1962), a simple $50\% + 1$ rule is assumed. In the US Senate, the practical requirement for governing the chamber is the 60% required to overcome an opposition filibuster. Elsewhere, highly fractured social landscapes may give rise to consociational rules that require near-consensus governing coalitions (Lijphart 1984). As with the voting rule, the majoritarianism rule need not be codified: in nondemocracies, it may just represent a majority of guns or factories. Whether formal or informal, this rule is an important determinant of deconcentration because it constrains the solution space, determining which coalition partitions are feasible in equilibrium.

2.1.6 Summing up

Figure 2.2 presents the full theory, stated generally. Politicians' payoffs are a function of their relative position within the coalition structure, which emerges from non-cooperative coalitional bargaining. Prior to this bargaining, politicians decide whether to expand the player set through deconcentration. Their preferences over

deconcentration are induced by their preferences over alternative coalition structures. The set of feasible outcomes is structured by the voting rule, which defines who must approve of a deconcentration proposal for it to pass, and the majoritarianism rule, which defines the minimum power required for a coalition to govern. Like the set of players N, the set of candidates C, and the distribution of power endowments γ, both of these rules (represented in Figure 2.2 as δ and α, respectively) are assumed to be exogenous.

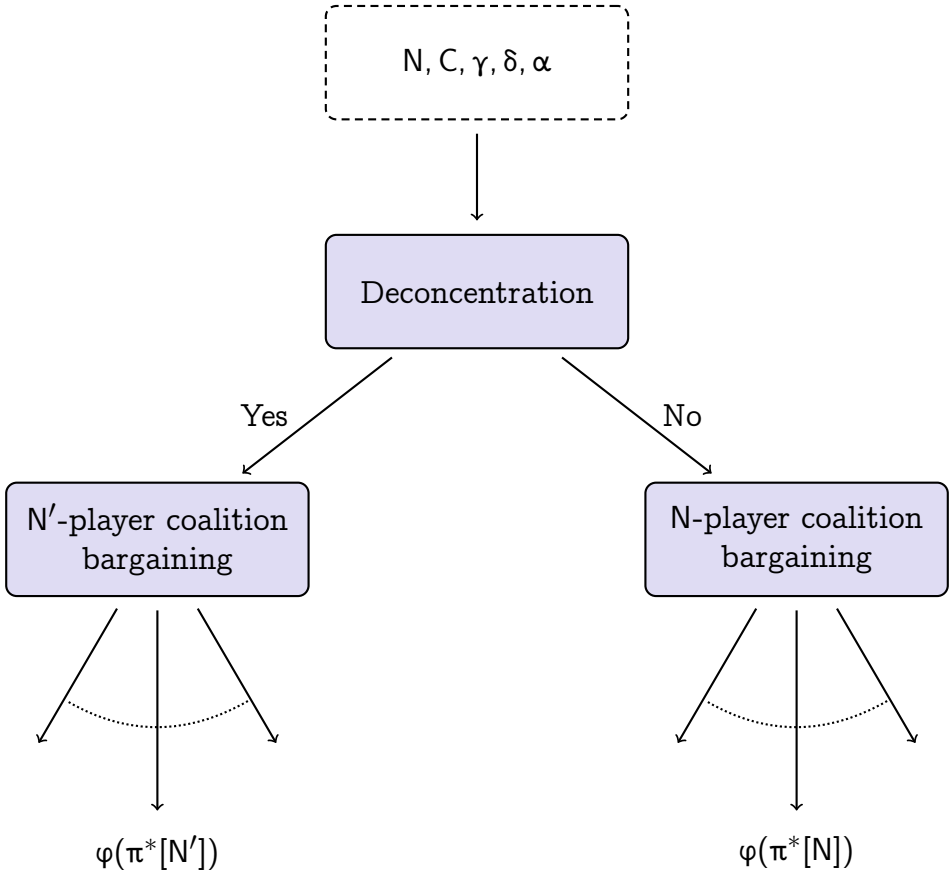


Figure 2.2: A map of the theory. Dotted lines indicate a range of outcomes. The politicians N, candidates C, endowments γ, voting rule δ, and majoritarianism rule α are exogenously given. Payoffs given by $\varphi(\cdot)$ are a function of the optimal coalition structure π^* under alternative player sets.

Scholars may be concerned by the assumption of exogenous voting and majoritarianism rules. In reality, political actors can and do manipulate both. In 2013, Senate Democrats exercised the "nuclear option," effectively lowering the majoritarian threshold for overcoming filibusters of many presidential appointments, which Senate Republicans extended in 2017. Stalin's purges can be seen as pruning the voting rule until his consent was both necessary and sufficient for major institutional changes such as deconcentration. It is not difficult to imagine these rules being similarly endogenous elsewhere.

However, while these institutions may be endogenous in the macro-historical sense, they remain exogenous in shorter-term bargaining because renegotiating them is prohibitively costly. In the Ostrom (1990) typology, these rules are constitutional, determining who has a say in leadership selection. The transaction cost of renegotiating such rules typically exceeds any short-term gains to be won by their revision (Greif and Laitin 2004; Shepsle 1989). Individual battles over deconcentration may change the set of players making collective decisions, but they rarely involve renegotiation of the constitutional rules over how such decisions are made. The voting and majoritarian rules therefore act as fixed constraints in the short term, and so can be treated as exogenous for answering the questions with which this dissertation is concerned.

2.2 Empirical implications

The coalition manipulation theory suggests four empirical implications. First, I hypothesize that deconcentration is increasing in the number of existing subnational units. Intuitively, as more players are bargaining over coalitions, the set of feasible coalition structures grows exponentially. Since deconcentration adds new players to

coalition bargaining, the solution space becomes increasingly dominated by coalition structures that result from deconcentration.

For instance, three politicians can organize into five distinct coalition partitions, but four can organize into fifteen. Thus, "no deconcentration" outcomes account for $\frac{B_N}{B_N+B_{N+C}} = \frac{5}{5+15} = 25\%$ of the solution space. When four politicians are bargaining over adding a fifth, this ratio decreases to $\frac{15}{15+52} = 22\%$. The exponential growth of the Bell (1938) numbers means that the proportion of the solution space for "no deconcentration" outcomes monotonically declines toward zero as the set of players bargaining increases, as depicted in Figure 2.3. Since the probability of deconcentration is just $1 - \Pr(\text{No deconcentration})$, deconcentration becomes increasingly likely as N grows, *ceteris paribus*.

These calculations assume that the probability of a given coalition partition being optimal is uniformly distributed. This is a reasonable assumption because there is no direct relationship between politicians' or candidates' identities and their individual effects on the extant coalition structure, as demonstrated above. Because equilibrium coalition structures depend on the identities of all players bargaining, it is not possible to characterize the likelihood of a given coalition being optimal *in general*, even with knowledge of its total power endowment or those of its players (Acemoğlu, Egorov, and Sonin 2008). Assigning each coalition the same probability of being optimal is the most reasonable prior belief in this context.

Trivially, deconcentration should also be increasing in the number of candidate politicians who may be added to coalition bargaining. If $C' > C$, and again we assume all coalition structures are equally likely to be optimal, then it follows that $\frac{B_N}{B_N+B_{N+C}} > \frac{B_N}{B_N+B_{N+C'}}$. Here the denominator grows at the same rate as the Bell (1938) numbers while the numerator remains fixed, shrinking the number of "no deconcentration" solutions much faster than when C is fixed and N grows. Thus,

increasing the number of candidate subnational units should substantially increase the probability of deconcentration.

These predictions are unique to my model. In particular, the external threat, outcomes, and majority-rigging theories all depend on new subnational units of government providing value that *decreases* with the number of existing units. For instance, if deconcentration's main effect is to create more homogenous subnational

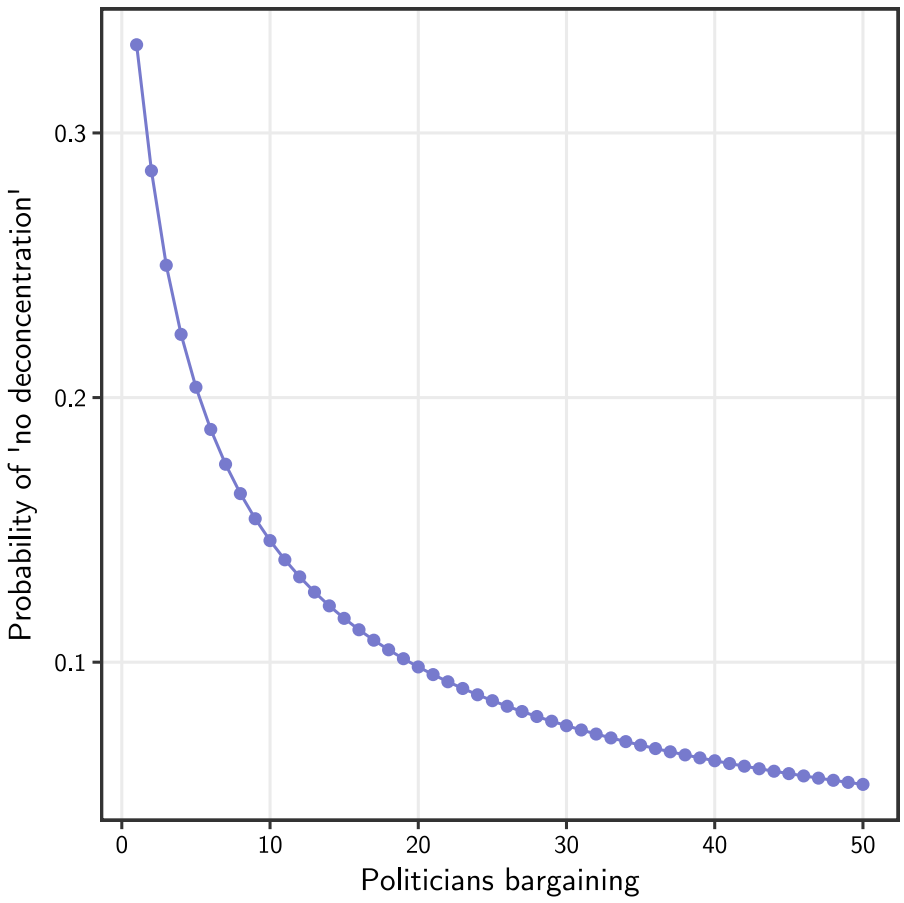


Figure 2.3: Deconcentration is increasing in the number of existing units. Dots represent $\frac{B_N}{B_N+B_{N+1}}$, the proportion of "no deconcentration" outcomes in the solution space, where N refers to the number of players bargaining and B_N is the Bell number for a set of size N, and the probability of a given coalition partition arising in equilibrium is uniformly distributed.

units so that public spending can more closely match districts' median preferences, then deconcentration will be less valuable where more districts already exist. Citizens will have already "voted with their feet" and the gains to deconcentration will already be priced in. Similarly, if new units are just vehicles for rigging majorities, then more extant units implies smaller constituencies and smaller electoral gains from each district rigged, both in votes and in legislative seats. Because these predictions stand in stark contrast to those of my model, the relationship between the number of existing units and deconcentration should be particularly useful for evaluating my theory.

At the same time, deconcentration and the number of units cannot spiral upward together forever, so something must impose an upper bound on this relationship. However, it is straightforward to impose a fixed cost on deconcentration that serves this role. Deconcentration can only serve to increase individual players' share of a fixed resource, normalized to size one in the simplified models above. As more players are added, ruling coalition sizes typically grow, and so the average gain from deconcentration in capturing more of that resource shrinks (but remains positive), reflecting the law of diminishing marginal returns. If an arbitrarily small fixed cost is added to the model, then eventually the benefits from deconcentration will fail to exceed this cost, *ceteris paribus*. Once this threshold is reached, politicians will have a hard time finding deconcentration proposals that improve their coalitional position.

Second, I hypothesize that deconcentration is decreasing in the stringency of the voting rule. Stringency in this context is an increase in either scope or specificity (or both) of veto players. When a greater proportion of politicians must consent to deconcentration for it to happen, or when a greater number of specific actors must consent, the probability of finding a deconcentration proposal to satisfy the voting

rule drops. In the extreme, a rule that requires unanimous consent—everyone is a veto player—will effectively eliminate deconcentration, since at least one member of the ruling coalition must be worse off under deconcentration.[7] The permanent members of the United Nations Security Council provide a real-world example. It is hard to imagine either of the major blocs creating new seats for opponents—e.g., Russia accepting Germany's addition, or the US accepting Iran's. Since each permanent member has a unilateral veto over deconcentration, my model predicts that the Council is very unlikely to add permanent members under current institutional arrangements. With more veto players comes greater coalitional stability.

Third, I also hypothesize that deconcentration will be decreasing in the strictness of the majoritarianism rule—that is, decreasing in the supermajority required to govern. The intuition here is more subtle. In the simple model above, politicians outside the ruling coalition can only "win" from deconcentration by being brought into the fold. In contrast, the "losers" from deconcentration were exclusively drawn from inside the extant ruling coalition, as their share of the winnings may be diminished by adding new allies, or worse, they could be kicked out of the ruling coalition entirely. All else equal, a larger majoritarian threshold will generate ruling coalitions that include more of the politicians in the polity—shrinking the number of potential winners and increasing the number of potential losers. To take an extreme case, a polity where the grand coalition among all politicians is required to form a government is very unlikely to experience deconcentration because the addition of any new actors would diminish the equilibrium payoffs for everyone. Under government by consensus, no one can win from deconcentration, and everyone stands to lose. Moreover, if the voting rule names specific veto players, then the probability

7. For example, it is apparent from Equation 2.1 that anyone being removed from the ruling coalition is automatically worse off, and if no one is removed, then everyone in the ruling coalition's payoff decreases since their share of the coalition's total power decreases.

that one of these politicians loses out from deconcentration increases for higher majoritarian thresholds. Through these two mechanisms, greater majoritarianism makes deconcentration less likely.

Existing accounts are generally silent on the role of these institutions in determining deconcentration, but they imply opposite predictions to those of my theory. For instance, in the external threat theory, deconcentration should correlate with stronger subnational units, enshrined in stricter voting and majoritarianism rules, as they provide protections from the tyranny of the center in new federations. Similarly, in the majority-rigging theory, if deconcentration is used to buy elections, then stricter voting rules or greater supermajorities should require more deconcentration because rulers must trade more units for more votes.

Finally, I hypothesize that deconcentration is more likely during periods of coalitional transition. With a fixed set of politicians, preferences, and rules, deconcentration is unlikely to occur once coalition structures reach equilibrium, because every feasible coalition structure (including those with new players added) is considered at the beginning of the game. Because deconcentration is decided endogenously, no further deconcentration is feasible under the same conditions. However, a sudden shift in the player set can knock the coalition structure out of equilibrium. After such a shock, bargaining begins anew, with politicians scrambling to push coalitional bargaining toward a more favorable new outcome. During these periods of renegotiation, entirely new coalition structures become possible, and those that had previously been off equilibrium path may become optimal under the new player set. Many of these new options will include deconcentration.[8]

8. Re-equilibration may frequently occur without deconcentration. My theory just implies a higher rate of deconcentration during such periods of coalitional realignment than when coalitions are already in equilibrium.

2.3 Conclusion

In this chapter I have outlined a novel theory of the causes of deconcentration. My central claim is that deconcentration is the result of politicians trying to manipulate political coalitions. This theory is clearly functionalist: deconcentration is simply a useful tool for achieving a desired outcome, namely, a change in the coalition structure.

This perspective answers both parts of the puzzle identified in Chapter 1: unexplained variation in political elites' support for deconcentration, and the conditions under which their support is sufficient for reform. Politicians' preferences over deconcentration are induced by their preferences over coalitional outcomes, so deconcentration will occur if the actors who prefer the coalition structure after deconcentration over the extant structure are sufficient to satisfy the voting and majoritarianism rules. Proponents of deconcentration need not hope that they can sneak reform past a short-sighted opposition. They need not cobble together contracts for post-deconcentration side payments where there are no mechanisms for enforcing them, nor must they rely on innate affinities to make their offers credible. Instead, they can make deconcentration offers that imply coalitional outcomes that are better for the necessary actors, defined by voting and majoritarianism rules.

My theoretical framework suggests four direct empirical implications. I hypothesize that deconcentration is more likely where there are more subnational units, fewer veto players, lower majoritarian thresholds, and coalitional instability. These four empirical implications provide substantial leverage with which to evaluate the coalition manipulation theory: three of these hypotheses cut directly against alternative explanations, while the fourth is a novel prediction of my model. The rest of this dissertation is concerned with evaluating the evidence for these claims.

3 | Forecasting deconcentration

A good theory of deconcentration should at least correspond to its observed variation across time and space. My first test of the coalition bargaining theory is therefore whether its empirical implications match the global incidence of deconcentration. This chapter sets out to examine these patterns, and in particular, to determine whether the variables identified in Chapter 2 are able to predict deconcentration. If models relying on coalition bargaining variables can predict deconcentration with reasonable accuracy, then it is strong evidence that the theory is a useful one.

I begin by introducing new data on the global occurrence of deconcentration over the 1960-2010 period. I then describe the variables used to test the hypotheses described in Chapter 2, as well as those derived from alternative explanations. After presenting results from a series of logistic regressions, I impute missing data and pool estimates. Both sets of models provide suggestive support for my theoretical framework. However, this approach does not allow me to adjudicate which predictors are most useful for understanding variation in deconcentration. Nor is falling back on standard model fit statistics feasible: with half a million possible covariate combinations (assuming no interactions or non-linear relationships), it is impossible to manually test every possible model.

The majority of this chapter is devoted to addressing this model selection problem using a large machine learning task. Machine learning (ML) refers broadly

to statistical techniques that identify patterns in data. The goal of *supervised* ML is to predict an outcome using a series of inputs (Hastie, Tibshirani, and Friedman 2009); where the outcome is binary, it is referred to as a "classification task." Political scientists are increasingly aware of the pitfalls of basing model choice on individual covariates' statistical significance (Gill 1999; Ward and Ahlquist 2018; Ward, Greenhill, and Bakke 2010), but have largely ignored supervised ML tools developed specifically for these tasks. Recently, however, scholars have turned to supervised ML to evaluate broad empirical claims where the evidence was mixed (Hainmueller and Hazlett 2014; Hill and Jones 2014).[1] The deconcentration literature, where theoretical guidance and empirical evidence are both relatively scarce, provides fertile ground for deploying ML methods.

Since my outcome of interest is binary, my goal is to build models that best separate observations into the "no deconcentration" and "deconcentration" classes. By holding some data back from the model-building process, I can rigorously evaluate the predictive performance of alternative models using as-if new data. This procedure allows me to determine which models are most useful for understanding the incidence of deconcentration. Among these, I then determine the predictors to which they give the greatest weight, based on a series of "variable importance" metrics discussed in detail below.

After outlining the methodology behind this task in detail, I provide brief introductions for each of the sixteen models I use, broadly organized into five categories: generalized linear models (GLMs) and extensions; variations on discriminant analysis; multivariate adaptive regression splines (MARS); decision trees and forests; and neural networks.[2] All of these models are implemented in R via

1. Earlier uses of ML models in political science focused primarily on methodological rather than empirical contributions (Beck and Jackman 1998; Beck, King, and Zeng 2000).
2. More commonly, these models are grouped into three categories: kernel methods, tree-based ensemble

the caret package with relatively low computational costs (Kuhn 2008), making it straightforward to investigate many models, and decreasing my reliance on specific implementations. However, these sixteen represent just a small fraction of the available algorithms (Fernández-Delgado, Cernadas, and Barro 2014); estimating all of them would be prohibitively costly and difficult to interpret.[3] I select these sixteen for computational stability, diversity of underlying approach, and predictive performance.[4]

This ML task reveals broad support for the coalition manipulation theory. Among the models that best predict deconcentration, the most informative variables relate to coalitional bargaining. In addition, the direction of these relationships are as hypothesized in Chapter 2: deconcentration is more likely where more politicians are bargaining over coalitions, where fewer veto players can block deconcentration, where a smaller majority coalition can rule, and during periods of coalitional transition. Finally, the predictive performance of these models is comparable to those studied elsewhere in the discipline, e.g., for predicting US Supreme Court decisions and local violence during civil war (Blair, Blattman, and Hartman 2017; Montgomery, Hollenbach, and Ward 2012). In contrast, alternative explanations fare poorly, providing relatively little predictive power and often cutting against the directions expected by these theories. Taken together, these findings indicate that the coalition bargaining framework is the best theory for understanding global variation in deconcentration.

---

methods, and multi-layer perceptron/deep learning algorithms (Bagnall and Cawley 2017). The classification adopted here allows me to describe these models as they are more commonly understood by political scientists.

3. These models take approximately 16 hours to run on a mid-2012 MacBook Air with 8Gb of memory and a 2 GHz Intel Core i7 processor.
4. I also limit my analysis to "soft-type classifiers," which output each observation's predicted probability of deconcentration, to facilitate model comparison. ("Hard-type classifiers" simply return the predicted outcomes with no associated class probabilities.)

3.1 Data

3.1.1 Deconcentration

Deconcentration data are sparse. Grossman and Lewis (2014) note that over half of African countries have increased the number of local government units by 20% since 1990, and Grossman, Pierskalla, and Dean (2017) examine government fragmentation in middle-income and poor sub-Saharan African countries over the period 1960-2012. Case studies provide snapshots of individual episodes in countries such as India (Mawdsley 2002) and Indonesia (Kasara 2006). Studies of federalism and regional authority occasionally provide sufficient information from which cases of layer creation can be inferred (Hooghe et al. 2016). Yet combining these data yields an incomplete picture of global deconcentration. Limited coverage, divergent definitions, inconsistent coding rules, and non-random selection (usually on the dependent variable) limit the inferences that can be drawn from the extant data.

To address this shortcoming, I provide new data on deconcentration worldwide since 1960. Like others studying the structure of subnational government, I draw on the Statoids project (Law 2016).[5] Statoids tracks changes to administrative divisions over time by processing official updates to organizations that define statistical and boundary names. These organizations include: the International Organization for Standardization; United States's Federal Information Processing Standard; the European Union's Nomenclature of Territorial Units for Statistics; the United Kingdom's Office for National Statistics; the Canadian Standard Geographic Classification; and the United Nations's M.49 coding system. Where these organi-

5. I refer to Statoids using the present tense. However, the creator and maintainer of the site, Gwillim Law, passed away in August 2016 after a brief illness. While the site remains accessible, it is unclear if it has been or will continue to be updated. To guard against the potential disappearance of these change histories, I updated my original data (coded in 2013) in October 2016 and captured a record of each country page as it appeared then. These records are available on request.

zations disagree on the timing and scope of changes to administrative divisions, Statoids reconciles the differences by drawing on primary sources and, occasionally, expert knowledge.

Statoids provides these data freely available online as "change histories," or written descriptions of updates to the structure of subnational government. Figure 3.1 provides a typical example of what these change histories look like. Since the data are in text format, highly irregular, and unstructured, I code these changes manually. Using the change history and occasional snapshot tables, I work forward and backward through time for each country, gathering complete time series for 189 sovereign states. The resulting data describe the number of government units and layers, and the years in which changes to either are made. From these data I determine whether a country experienced deconcentration in a given year, which I

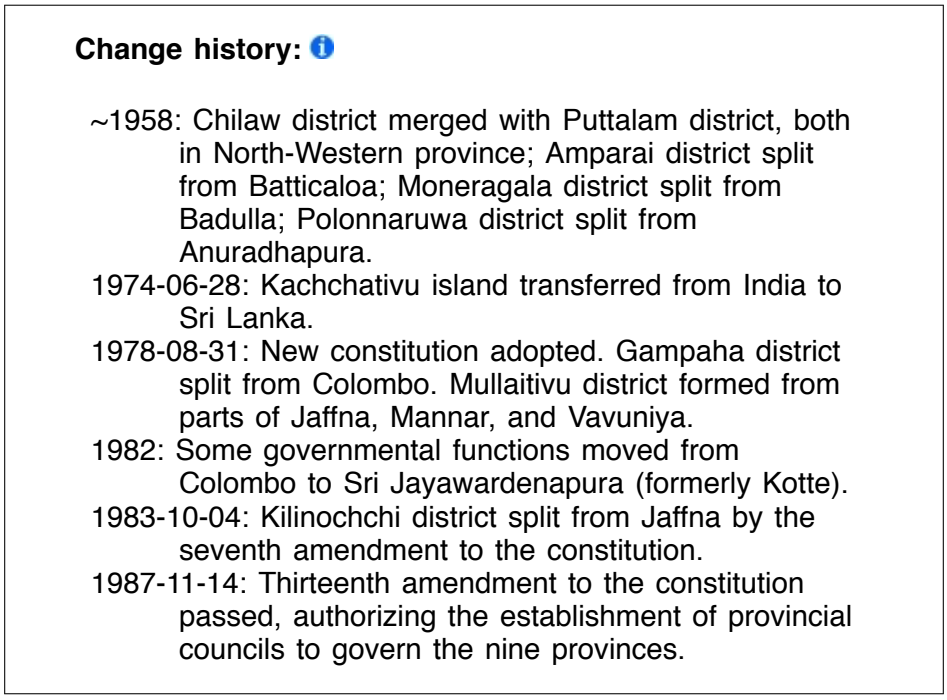


Figure 3.1: Statoids change history for Sri Lanka. An example of the text from which the deconcentration data are coded.

code as 1 if either unit creation or layer creation occurred and 0 otherwise.

Unit creation includes simple changes such as the excision of Kilinochchi district from Jaffna in Sri Lanka in 1983 (see Figure 3.1). It also includes a total restructuring which results in more, smaller top-level units, such as when Nigeria transformed four regions into twelve states in 1967. However, cases where units simply change status—as when Mexico's Baja California Sur and Quintana Roo territories became states in 1974—are not coded as unit creation. In addition, due to data constraints, I am only able to study unit creation for *primary* administrative subdivisions. These are the largest subnational units with real authority, typically states, regions, provinces, districts, or federal entities. By "real authority" I mean that I exclude purely statistical units used as a bureaucratic convenience (e.g., in the US, Federal Reserve Districts) and cultural regions without formal definition ("New England"). I also allow the definition of "top-level" to vary within countries where no one layer extends over the entire territory; for instance, in 2010, Russia's included 46 oblasts, 21 republics, nine krais, four autonomous okrugs, two cities of federal significance, and an autonomous oblast.

I code layer creation as occurring any time a new level of government is added to the map. Such cases include, for instance, when Côte d'Ivoire created regions in 1969, as in Table 1.1. However, layer creation does not include cases where existing layers are merely refashioned, like when Burundi's provinces became *arrondissements* in 1962. Again I exclude layers that lack any "real authority"; purely statistical or bureaucratic layers are ignored.

For both of these types of deconcentration, I use the year of the *decision* to deconcentrate, not the year the change was implemented. This choice is motivated by two considerations. First, the decision to deconcentrate is much more directly observable, and so measuring by date of decision is more reliable than measuring

by date of implementation. Second, any definition of "implemented" must vary considerably across and within countries, given the great variety in the structure of subnational units and the responsibilities devolved to them. Allowing the definition to flex this much across space and time would limit the validity of the measure and limit any inferences that could be drawn from it; focusing on the decision to deconcentrate avoids much of this variation. (Changes for which I cannot pin down a year are not included in the data.)

These coding rules yield 371 distinct episodes of deconcentration, accounting for 4.6% of the 8,089 country-years in the data. Among these, the vast majority are unit creation (348 cases) as opposed to layer creation (23 cases). I record at least one episode of deconcentration in fully 126 countries. The global distribution of deconcentration is plotted in Figure 3.2, which provides a heatmap by region and

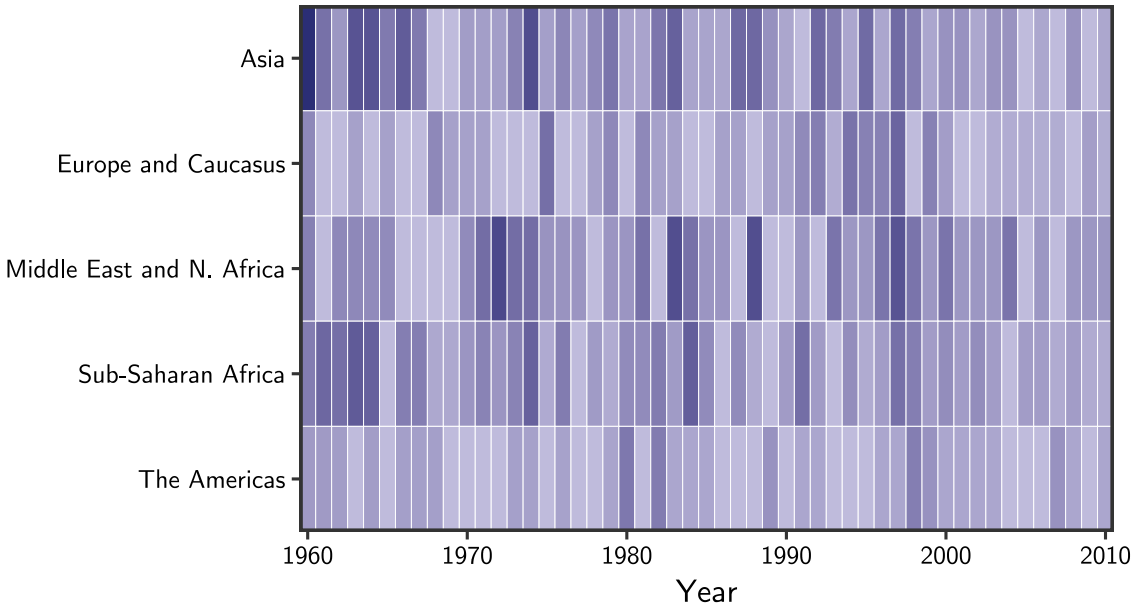


Figure 3.2: The global prevalence of deconcentration. Each rectangle is shaded according to the proportion of countries in that region and year that deconcentrated. Darker blue indicates more deconcentration.

year.[6] The visual evidence suggests some clustering during the era of decolonization in the 1960s and a tapering-off in recent years. However, overall, the picture is one of global prevalence and relative continuity over decades. These patterns indicate that my results are unlikely be driven by the characteristics of just one region or epoch.

3.1.2 Predictors

I merge these data with predictors provided by a variety of sources. From among the near-limitless set of variables that could be studied, I focus on those identified by my theory and by alternative explanations.

First, I hypothesize in Chapter 2 that deconcentration is more likely in countries where more players are bargaining over coalitions (from the theoretical model, where N is larger). I measure this concept in two ways. The first is a proxy for the number of politicians: the number of top-level administrative units, as coded from Statoids. The second is a count of the number of cabinet-level ministers, available from the Cross-National Time-Series Data Archive (CNTS; Banks and Wilson 2013). This measure is effectively a proxy for the size of the ruling coalition, which in turn will generally be larger where N is larger: because ruling coalitions typically require a majority of seats, guns, or dollars to remain in power, ruling coalitions will typically (but not always) grow with the size of the set of politicians. Deconcentration should be more likely for larger values of either measure.

Second, I hypothesize that deconcentration should be decreasing in the strictness of the voting rule. To test this relationship, I study a measure of veto players provided by the Database of Political Institutions (DPI; Beck et al. 2001), as well as judicial independence provided by the Political Constraints dataset

6. These regional classifications are available from the World Bank. I merge North America with Latin America and the Caribbean, and South Asia with East Asia and the Pacific, due to the small country samples in each.

(POLCON; Henisz 2002). Veto player data are an ordinal scale running from 1 to 18, where the index is incremented by one for opposition control of legislative chambers, for the separation of legislative and executive leadership, and for other checks on centralized power. Judicial independence is a binary indicator, coded as 1 if the judiciary exerts institutionalized constraints on executive authority and 0 otherwise. Both variables measure the dispersion of power away from the national executive, and are therefore likely to co-vary with, if not directly capture, the dilution of executive power over deconcentration.

Third, I also hypothesize that deconcentration should be decreasing in the degree of majoritarianism required to govern. I test this hypothesis using an ordinal measure of legislative electoral independence and an indicator for whether local government is elected, both from DPI. The legislative independence index runs from 1 ("no legislature") to 7 ("largest party got less than 75%" in a competitive election), while local government is coded as 1 if either the executive or legislative branch (or both) of the lowest level of government is elected, and 0 otherwise.

There are two reasons to expect that these variables serve as decent measures of majoritarianism. By definition, elected bodies draw from larger selectorates than do their unelected counterparts. Controlling elected legislatures or local governments therefore requires the governing coalition to meet a higher majoritarian threshold than does controlling appointed but otherwise equivalent bodies (Bueno de Mesquita et al. 2003). Further, each governing body that is elected separately increases the effective supermajority required to rule. If the electorates choosing the representatives in each chamber are all identical, then the degree of supermajority is simply the highest ruling threshold for any assembly. But if the electorates are different, then the effective supermajority required to govern grows by the amount by which the electorates differ. I therefore expect both legislative electoral independence and local

elections to be correlated with less deconcentration.

Finally, I hypothesize that deconcentration is more likely during periods of coalitional transition. I measure this in two ways. First, I code a country as "in transition" using the regime transition variable from Polity (Marshall, Jaggers, and Gurr 2011), which is coded as 1 for years in which the revised combined Polity score shifts by three or more points (on a -10 to 10 scale). I also code transition as 1 for the two years following independence, in line with the prediction outlined in Chapter 2. Second, to capture coalitional transitions that might not be reflected in institutional changes such as democratization or independence, I study mean party age, available from DPI. Recent transformations to the extant coalition structure should, on average, be reflected in younger parties.

Scholars may be concerned that many of these predictors are correlated with, or are constituent components of, democracy more broadly. These measures might just be picking up the same general effect of democratic institutions, and not the more specific features of the coalition bargaining environment identified in my theory. However, this issue is unlikely to be a problem for the ML task. As discussed below, many of the models I study impose shrinkage regularizations, in which estimates of correlated predictors are constrained. Intuitively, if a series of variables are all providing the same information, then all of their estimates will be "shrunk" toward zero and their importance severely downweighted, since they are individually less informative. If legislative electoral independence, local elections, veto players, and judicial independence are all picking up the same general variation in democracy—and not specific institutions related to coalition bargaining—then the ML models will adjust for this multicollinearity and ignore these variables. If instead the models rely heavily on these covariates, it is because they provide more specific information about the coalition bargaining environment that helps predict

deconcentration.

To provide a baseline by which to evaluate the utility of my theory, I also compare it to alternative explanations introduced in Chapters 1 and 2. The first of these alternative accounts, the Rikerian *external threat* theory, understands deconcentration as a response to perceived military threats. I measure threat perception using (logged) military spending per capita from the World Bank (2017), as well as an indicator for whether a country is engaged in an international war, from the Center for Systemic Peace (CSP; Marshall 2015). Both are expected to be positively correlated to deconcentration. More indirectly, scholars have noted that both civil and international wars are less likely in rich democracies (e.g., Benoit 1996; Hegre 2014). Wealthier, more democratic countries should therefore be less likely to deconcentrate because they do not experience military threats of the same magnitude as do poorer, less democratic countries. I test these correlations using logged gross domestic product (GDP) per capita in constant 2010 US dollars and "revised combined" Polity scores (Marshall, Jaggers, and Gurr 2011; World Bank 2017), both of which I expect to be negatively associated with deconcentration.

The second set of alternative explanations for deconcentration, which I refer to collectively as the *outcomes theory*, are those that view deconcentration as reform undertaken to induce normatively positive outcomes such as better democratic representation, faster economic growth, or better provision of public services. I test three hypotheses from the outcomes theory. As noted in Chapter 1, all of these arguments depend on democratic institutions, so in contrast to the military threat theories, this perspective would expect a positive correlation between democracy and deconcentration. Moreover, the gains to economic growth and public services from deconcentration are both expected to accrue as a result of ideological and economic diversity among citizens: new subnational units can only generate growth and

Table 3.1: Predicted relationships

Coalition manipulation		Alternative explanations		
Variable	Sign	Variable	Theory	Sign
Number of units	+	Military spending	ET	+
Cabinet members	+	International war	ET	+
Veto players	−	GDP	ET	−
Jud. independence	−	Democracy	ET/O	−/+
Leg. independence	−	Ethnic diversity	O/MR	+
Local elections	−	Inequality	O/MR	+
Transition	+	Exec. election	MR	+
Mean party age	−	Leg. election	MR	+
		Population	Control	+
		Geographic size	Control	+
		Ruggedness	Control	+

"ET" refers to the external threat theory, "O" to the outcomes theory, "MR" to the majority-rigging theory, and "Control" to control variables. Signs indicate the predicted effect on the probability of deconcentration of increasing each variable (or changing binary variables from 0 to 1).

efficiency if "citizens voting with their feet" makes the average district substantially more homogenous. Deconcentration should therefore be more likely in countries that have greater ethnic diversity and economic inequality. I test ethnic diversity using ethnolinguistic fractionalization from Fearon (2003). Because income data are scarce—available for only 10-15% of the observations in my data—I follow Blaydes and Kayser (2011) in measuring inequality with per capita caloric consumption data from the Food and Agriculture Organization of the United Nations (FAO 2017). The outcomes theory predicts that all three variables—democracy, ethnic diversity, and inequality—are positively correlated with the probability of deconcentration.

I also test the *majority-rigging* argument, that deconcentration is used to distribute patronage and manipulate elections in weak democracies. Since this set of explanations centers on the utility of deconcentration for electoral competition, we should observe more deconcentration in country-years where elections occur. I

therefore also include indicator variables for legislative and executive elections, which I take from DPI, in all of the models. Less directly, the mechanisms underlying these explanations typically center on ethnic contestation over scarce resources (e.g., Grossman and Lewis 2014), and so predict that deconcentration is more likely where ethnic diversity and inequality are greater (as with the outcomes theory).

Finally, it will also be useful to evaluate the utility of the coalition manipulation theory relative to control variables that might be expected to influence deconcentration. It is plausible that larger, more populous countries have more subnational units of government, and are thus more likely to deconcentrate, so I include logged geographic size and logged population, both from the World Bank (2017), in the models. I also test whether deconcentration is a function of ruggedness of terrain, since more mountainous countries might have higher fixed costs for administering territory, and therefore might deconcentrate more (Nunn and Puga 2012). However, as with similar studies using variable importance to evaluate competing theories (e.g., Hill and Jones 2014), I do not include measures of spatial or temporal dependence in the ML task. While accounting for such dependence marginally improves the predictive performance of my models,[7] my main goal here is to compare the informational content of theoretically-relevant variables. Since modeling such dependence does not change the relative importance of the predictors drawn from theories of deconcentration, but imposes greater computational burdens, I exclude it from the main analysis. Pairwise correlations among all predictors are provided in Table 3.2.

7. Specifically, I include among the predictors the number of years since an episode of deconcentration, as well as an indicator for whether a neighboring country has experienced deconcentration in the previous three years, to capture temporal and spatial dependence, respectively. None of the results presented below are substantially changed by including these covariates.

Table 3.2: Correlations among predictors of deconcentration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1) Number of units	1.00																		
(2) Cabinet members	0.14	1.00																	
(3) Veto players	0.12	−0.04	1.00																
(4) Jud. independence	0.14	−0.08	0.56	1.00															
(5) Leg. independence	0.14	−0.02	0.62	0.46	1.00														
(6) Local elections	0.18	−0.08	0.34	0.37	0.38	1.00													
(7) Transition	0.00	−0.08	−0.13	−0.09	−0.09	−0.10	1.00												
(8) Mean party age	0.19	−0.07	0.19	0.27	0.15	0.24	−0.11	1.00											
(9) Military spending	0.04	0.04	−0.17	−0.07	−0.26	−0.13	0.04	0.00	1.00										
(10) International war	0.02	0.06	−0.04	−0.07	−0.07	−0.04	0.05	0.02	0.11	1.00									
(11) GDP	0.11	−0.09	0.37	0.57	0.29	0.27	−0.14	0.48	0.09	−0.01	1.00								
(12) Democracy	0.17	−0.07	0.70	0.70	0.71	0.50	−0.03	0.31	−0.26	−0.04	0.48	1.00							
(13) Ethnic diversity	−0.10	0.09	−0.14	−0.20	−0.20	−0.19	0.03	−0.23	0.02	−0.01	−0.31	−0.20	1.00						
(14) Inequality	0.20	0.09	0.35	0.49	0.34	0.35	−0.14	0.36	0.06	−0.02	0.78	0.43	−0.28	1.00					
(15) Exec. election	0.03	0.00	−0.01	−0.02	0.06	0.04	0.12	−0.05	−0.01	−0.03	−0.08	0.06	0.00	−0.06	1.00				
(16) Leg. election	0.04	−0.01	0.10	0.12	0.12	0.09	0.07	0.06	−0.03	−0.02	0.09	0.17	−0.04	0.09	0.31	1.00			
(17) Population	0.39	0.48	0.12	0.01	0.10	0.14	0.05	0.08	−0.05	0.05	−0.10	0.13	−0.05	0.14	0.06	0.05	1.00		
(18) Geographic size	0.26	0.40	0.00	−0.04	−0.03	0.09	0.06	0.07	0.07	0.03	−0.14	−0.05	0.13	0.00	0.07	0.02	0.84	1.00	
(19) Ruggedness	−0.02	−0.19	−0.05	−0.01	0.02	0.13	0.01	0.02	0.09	0.03	−0.01	0.04	−0.15	0.05	−0.05	−0.02	−0.14	−0.27	1.00

Column numbers indicate predictors as labelled in the corresponding rows. Values are Pearson correlations among pairwise-complete variables.

3.2 Baseline relationships

I begin by regressing deconcentration on each of these predictors separately. Since deconcentration is binary, each model is a standard logistic regression with the country-year as the unit of analysis:

$$y_{i,t} \sim \mathcal{F}(\pi_{i,t}), \tag{3.1}$$

$$\pi_{i,t} = \text{logit}^{-1}(\alpha + \beta x_{i,t}), \tag{3.2}$$

where $y_{i,t}$ is deconcentration, $\mathcal{F}$ is the Bernoulli distribution, α is the intercept, β is the estimated effect of each predictor x, observations are indexed by countries $i \in \mathcal{I}$ and years $t \in \mathcal{T}$, and logit^{-1} is the inverse-logit function such that, e.g., $\text{logit}^{-1}(\theta) = \frac{1}{1+e^{-\theta}}$. I then estimate a saturated model where Equation 3.2 is instead

$$\pi_{i,t} = \text{logit}^{-1}\left(\alpha + \beta_1 x_{i,t,1} + \cdots + \beta_{19} x_{i,t,19}\right), \tag{3.3}$$

so that all predictors are included in the model simultaneously.

Results from these models are reported in the first two columns of Table 3.3. The initial evidence suggests support for deconcentration as coalition manipulation: all of the signs are in the expected direction, and all are statistically significant. For alternative approaches the evidence is more mixed. The two indirect proxies for external threats, GDP and democracy, are negative, but the direct measures in military spending and international war are indistinguishable from zero. For the outcomes theory, ethnic diversity appears a strong predictor, but inequality and democracy both have the "wrong" sign. Finally, theories of majority-rigging do not find much more support, as the effect of elections are also effectively nil. To assess the relative magnitude of these relationships, Figure 3.3 plots changes in the predicted probability of deconcentration from shifting each variable along

Table 3.3: Correlates of deconcentration

	Complete-case analysis		After imputation	
Variable	$\widehat{\beta}$ (ind.)	$\widehat{\beta}$ (full)	$\widehat{\beta}$ (ind.)	$\widehat{\beta}$ (full)
Number of units	0.02*	0.00*	0.02*	0.03*
Cabinet members	0.03*	0.00*	0.03*	0.04*
Veto players	−0.19*	−0.01	−0.19*	−0.18
Jud. independence	−0.51*	0.00	−0.51*	0.10
Leg. independence	−0.09*	−0.01	−0.09*	−0.13
Local elections	−0.40*	−0.01	−0.40*	−0.29
Transition	0.52*	−0.03	0.52*	−0.76
Mean party age	−0.01*	−0.00*	−0.01*	−0.02*
Military spending	0.07	−0.00	0.07	−0.09
International war	0.05	0.06	0.05	1.09
GDP	−0.24*	−0.01†	−0.24*	−0.27
Democracy	−0.03*	0.00*	−0.03*	0.08*
Ethnic diversity	0.72*	−0.02	0.72*	−0.83
Inequality	−0.51*	0.01	−0.51*	0.10
Exec. election	0.01	−0.00	0.01	−0.06
Leg. election	0.02	0.01	0.02	0.17
Population	0.22*	0.00	0.22*	0.02
Geographic size	0.15*	0.00	0.15*	0.10
Ruggedness	−0.07	0.00	−0.07	0.03

*$p < .05$, †$p < .10$. "Ind." refers to estimates from models with each covariate estimated separately, while "full" refers to estimates from the saturated model with all predictors included as covariates. Estimates using imputed data are pooled across 25 replicate datasets and combined as in Equations 3.4-3.5.

its interquartile range (or from shifting binary covariates from 0 to 1). Since the mean predicted probability of deconcentration across all models is approximately 10%, two percentage-point effect sizes are substantively large, representing 10-25% changes over the baseline. Table 3.3 also compares these results against estimates from the saturated model with all covariates. All effects are driven toward zero, though the covariates that remain statistically significant are all from my model, with the exception of democracy.

While it is tempting to interpret these results as suggesting that the variables

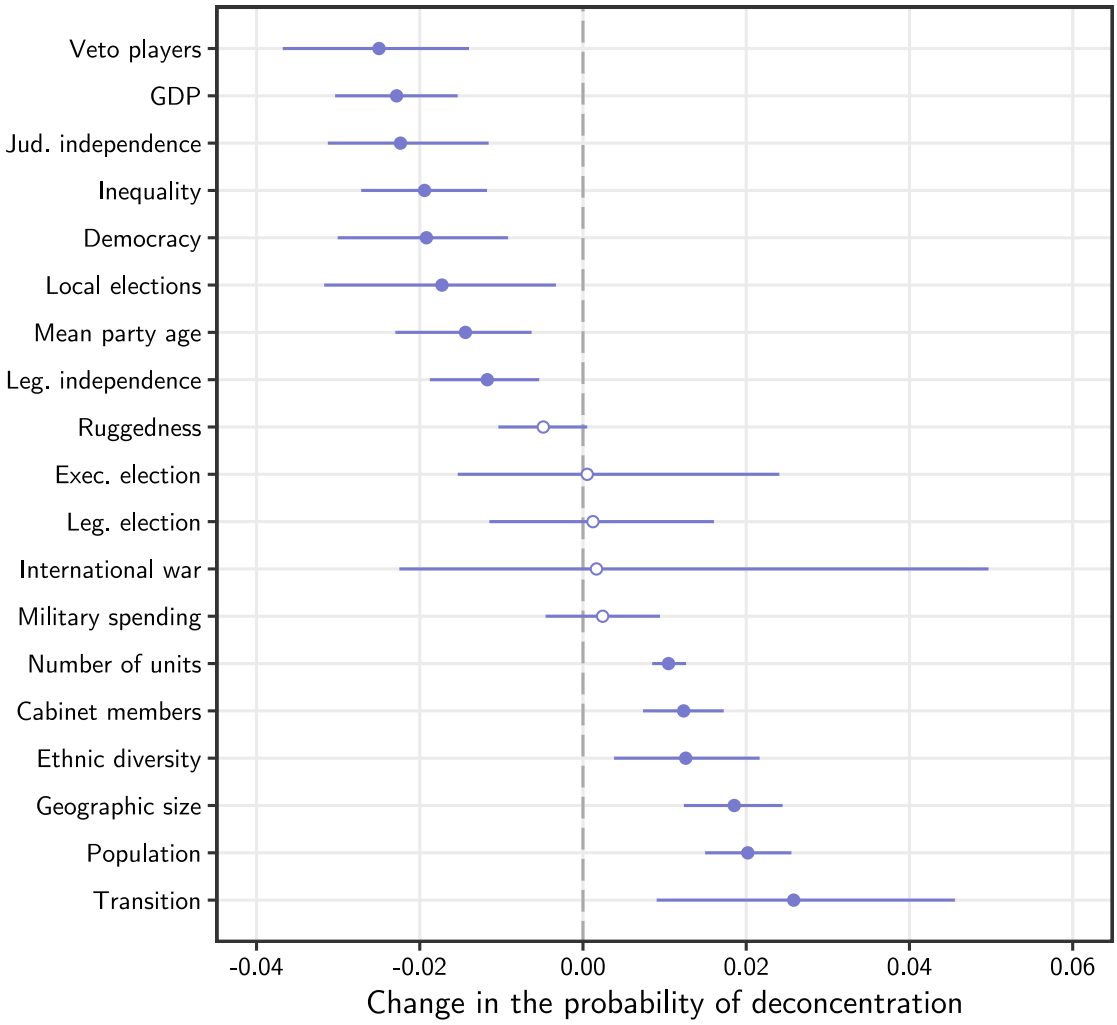


Figure 3.3: Bivariate correlates of deconcentration. Points plot the change in the predicted probability of deconcentration from varying each predictor from its 1st to 3rd quartile (or in the case of binary variables, from 0 to 1), with lines for 95% confidence intervals. Solid dots indicate estimates that are statistically significant at $p < .05$, with hollow dots for those that are not. Estimates and standard errors are constructed from 1,000 draws from $\mathcal{N}_{MV}(\widehat{\beta}, \mathbb{C}[\widehat{\beta}])$ for each of the baseline bivariate models.

from my theory have greater explanatory power, this conclusion is misguided. Beyond the inapplicability of statistical significance as a rubric for adjudicating models, it is also likely that these estimates reflect substantial data missingness. Listwise deletion (or complete-case analysis) produces inefficient estimates which are also biased if the data are missing not at random—as is true here. A better strategy is to impute the data, producing unbiased and more efficient estimates under reasonable assumptions about covariate distributions (Lall 2016).

I use conditional multiple imputation (MI), wherein the conditional distributions of partially-observed variables are iteratively modeled using the observed and imputed values for other variables (Kropko et al. 2014), implemented in R via mice (van Buuren and Groothuis-Oudshoorn 2011).[8] I generate $\mathcal{M} = 25$ imputed data replicates, each of which converges within 10 iterations (van Buuren et al. 2006), and ensure that the imputed data respect the scale constraints of each covariates (e.g., imputed ethnic fractionalization values are constrained to values between 0 and 1).[9]

Again I estimate the effect of each predictor separately and then in a saturated model. Each model, still given by Equations 3.1-3.2, is estimated on each of the imputed datasets and then the results are combined following Rubin (1987).[10] These

8. Conditional MI outperforms joint MI, in which the variables are modeled as a sample from a joint multivariate normal distribution, particularly for binary covariates and when using multinomial logit in the imputation model (Kropko et al. 2014).
9. Standard practice is to estimate just $\mathcal{M} = 5$ or 10 imputed data replicates. However, various studies have shown that larger values of $\mathcal{M}$ are required to capture Monte Carlo uncertainty across imputations, with a default value of the proportion of overall missingness suggested instead (Bodner 2008; Lall 2016; White, Royston, and Wood 2011). Here, the overall missingness rate is 23%, which I round up to 25. Convergence diagnostics are available on request.
10. Here the estimate and its variance are given by

$$\widehat{\beta} = \frac{1}{\mathcal{M}} \sum_{m=1}^{\mathcal{M}} \widehat{\beta}_m, \text{ where} \tag{3.4}$$

$$\mathbb{V}[\widehat{\beta}] = \frac{1}{\mathcal{M}} \sum_{m=1}^{\mathcal{M}} \mathbb{V}[\widehat{\beta}_m] + \left(1 + \frac{1}{\mathcal{M}}\right) \left[\frac{1}{\mathcal{M}-1} \sum_{m=1}^{\mathcal{M}} \left(\widehat{\beta}_m - \widehat{\beta}\right)^2\right]. \tag{3.5}$$

pooled estimates are reported in Table 3.3.

Three patterns stand out. First is the importance of imputation: while the individual models look the same as before imputation, the saturated models differ significantly. Second, the suggestive evidence in favor of the coalition bargaining theory is strengthened. All of the variables that achieve statistical significance in the saturated model with imputed data coincide with my theoretical expectations (save democracy), and although some variables related to coalition bargaining are indistinguishable from zero, the pooled results indicate that this framework provides much more leverage over deconcentration than do other approaches. Finally, the estimates vary considerably across models using imputed data, particularly those of Polity scores.

Although this approach allows me to overcome nonresponse bias, the estimates' variance across models still poses an important challenge. Even if I could manually search the model space by comparing all 2^{19} linear additive models that can be constructed with these predictors, standard fit statistics such as the Bayesian Information Criterion (BIC) break down when pooling across imputed data replicates. Nor would these models be able to address the problem of multicollinearity among variables closely related to democracy.

3.3 The learning task

I turn to supervised ML techniques to address these problems. My procedure is to build a series of models that best predict deconcentration, select the best-performing among them, and then study the variables on which these top performers rely. Before explaining the intuition behind each model, I describe the task in general, which is outlined formally in (Pseudo-)Algorithm 3.4. All analysis is conducted with

the caret package in R, and further details of each step are described in Kuhn and Johnson (2013).

First, I partition each of the imputed data replicates into a training and a validation sample, comprising 75% and 25% of each replicate dataset, respectively. Each replicate is split such that the marginal distributions of each predictor and the outcome are retained within each sample. Thus, each training sample consists of 6,324 observations with approximately 280 cases of deconcentration, while each validation sample is 2,108 observations with around 95 cases of deconcentration. I train and validate each model on each replicate independently, which helps retain some of the uncertainty associated with the imputation process.[11]

Second, I pre-process the predictors (known as "features" in the machine learning literature). I use the multivariate spatial sign, which projects the data onto a multidimensional sphere, akin to centering and scaling while reducing the influence of outliers (Kuhn 2008; Serneels, De Nolf, and Van Espen 2006).[12] The exact parameters of this projection are then used to pre-process values of the predictors in the validation sample. (Pre-processing the two samples together would pollute the training sample with information from the held-out data, breaking their independence and therefore my ability to evaluate model performance with a true validation sample.) In general, pre-processing contributes to better fit, and many models require it, so I run all of them on a common pre-processed sample to facilitate performance comparisons.

Third, each model run begins by subsampling each training data replicate

11. Alternative approaches that integrate multiple imputation and model selection more formally have been introduced in biostatistics (Schomaker and Heumann 2014). However, these approaches are computationally intractable for these data, and test runs fit to sub-samples of the data yield worse predictive performance than the classifiers considered here.
12. My results are robust to using the simpler centering and scaling approach, though in general the models' predictive performance degrades.

using the synthetic minority over-sampling technique (SMOTE; Chawla et al. 2002). In classification tasks where the outcome classes are unbalanced (as here), predictive performance can be improved by downsampling the majority class ("no deconcentration") or upsampling the minority class (deconcentration). SMOTE does both, but unlike other methods, constructs synthetic minority cases using nearest neighbors rather than simply resampling the minority class with replacement. Once again, this approach is motivated by predictive performance: SMOTE produces better model fit than does ignoring imbalance or using simpler subsampling techniques (He and Garcia 2009).[13]

Fourth, after subsampling, each model's hyperparameters (discussed below) are tuned on each replicate across a range of reasonable values initialized by caret. I set a tuning length of 10, so that a model with p hyperparameters tunes across a grid of 10^p hyperparameter combinations. For each point in this grid, repeated cross-validation with five folds and five repeats is used to generate a mean area under the curve (AUC) across receiver operating characteristics (ROCs), a widely-used measure of classification accuracy. The 10^p mean AUCs are then compared and the hyperparameter values that produce the best predictive performance (highest mean AUC) are selected and then used to fit the model on the full replicate training sample.[14] This tuning process accounts for most of the computational burden of the learning task, yet it is essential because within-model variation (across hyperparameters) dominates cross-model variation: predictive performance typically relies more on properly tuned models than it does on model choice (Bagnall and Cawley 2017).[15]

13. Again, this modeling choice does not affect the substantive results reported below, but it does improve predictive performance for these data across a variety of models.
14. My results are robust to using the Cohen (1960) Kappa measure of concordance to tune the models instead of AUC.
15. Where available, I use adaptive cross-validation, which resolves the tuning process more quickly by

Fifth, from this fitted model, I extract a measure of the information provided by each predictor, known in the ML literature as "variable importance." Since variable importance metrics differ across models, I discuss how each model operationalizes this measure below. By default, caret scales all variable importance measures to the range 0-100. Due to substantial differences across how variable importance is calculated for each model, I also study the importance *ranks* of covariates across models. Variables that provide more information about deconcentration will be more important, with values closer to 100 and higher ranks (e.g., 1^{st} instead of 5^{th}).

Finally, since this procedure outputs 16 variable importance summaries for each of the 25 imputed data replicates, I use predictive performance to adjudicate disagreements among models. For each replicate, I fit the trained models on the validation sample and compare performance using out-of-sample AUC, a standard metric for classification problems in political science (e.g., Ward, Greenhill, and Bakke 2010). I also study Brier (1950) scores for each of the models, effectively the mean squared forecast error, which is minimized when the predicted probabilities exactly match the observed outcomes (Merkle and Steyvers 2013).[16] Finally, I also compute a range of classification metrics which allow me to distinguish positive and negative predictive value, or the models' ability to predict deconcentration and no deconcentration, respectively.

Before introducing each of the models, I briefly state the ML task more formally. Allow the models to be indexed by $l \in \mathcal{L}$. Denote the imputed data replicate $m \in \mathcal{M}$ where D_m and V_m are the training and validation samples, respectively.

ignoring hyperparameter values unlikely to be optimal, typically without loss of performance. See Kuhn (2014) for details. My results are robust to using standard repeated cross-validation instead.

16. Specifically, the Brier score of a forecast is given by $BS = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2$, for observations $i \in N$, where $\hat{y}$ indicates a predicted class probability and y is the observed outcome class. Brier scores and other proper scoring rules are commonly utilized outside the discipline because they eliminate arbitrary classification thresholds from the model evaluation process (Gneiting and Raftery 2007; Merkle and Steyvers 2013; Montgomery, Hollenbach, and Ward 2012).

Inputs : $\mathcal{L}$ models; $\mathcal{M}$ imputed data replicates, each with $\mathcal{K}$ predictors, a response y, and observations $i \in N$

for Replicate $m = 1 \ldots \mathcal{M}$ **do**
  Split m into 75% training sample D_m and 25% validation sample V_m;
  Pre-process $\mathcal{K}$ in D_m using multivariate spatial sign;
  Apply pre-processing formula to $\mathcal{K}$ in V_m;
  for $l = 1 \ldots \mathcal{L}$ **do**
    Subsample D'_m using SMOTE;
    for Tuning hyperparameters $j = 1 \ldots 10^p$ **do**
      for Tuning repeats $c = 1 \ldots \mathcal{C}$ **do**
        Generate resamples $R''_{c,m,l}$ and $T''_{c,m,l}$;
        for Tuning resamples $b = 1 \ldots \mathcal{B}$ **do**
          Fit $\widehat{f}_{b,c,m,l,j}(R''_{b,c,m,l}; \theta_{b,c,m,l,j})$;
          Predict $T''_{b,c,m,l,j}$;
          Compute goodness of fit $\kappa_{b,c,m,l,j}$;
        end
      end
      Compute average goodness of fit $\kappa_{m,l,j} = \frac{\sum_b^{\mathcal{B}} \sum_c^{\mathcal{C}} \kappa_{b,c,m,l,j}}{\mathcal{B} \times \mathcal{C}}$;
    end
    Set optimal tuning parameters $\theta^*_{m,l} \in \Theta_{m,l}$ from $\max \kappa_{m,l,j}$;
    Fit the model $f_{m,l}(D''_{m,l}; \theta^*_{m,l})$;
    Extract variable importance $\varphi_{m,l}$ from $\widehat{f}_{m,l}(D''_{m,l}; \theta^*_{m,l})$;
    Fit the model $\widehat{f}_{m,l}(V'_{m,l}; \theta^*_{m,l})$;
    Extract predicted probabilities $\widehat{y}_{m,l}$ from $\widehat{f}_{m,l}(V'_{m,l}; \theta^*_{m,l})$;
    Compute goodness of fit statistics $\mathcal{S}_{m,l}$ from $V'_{m,l}$ and $\widehat{y}_{m,l}$;
  end
end

Outputs: $\mathcal{M} \times \mathcal{L} \times \mathcal{K} = \Phi$ lists of variable importance, $\mathcal{M} \times \mathcal{L} \times 0.25N$ predicted probabilities of deconcentration, and $\mathcal{M} \times \mathcal{L} \times \mathcal{S}$ goodness of fit statistics

Figure 3.4: (Pseudo-)Algorithm for the deconcentration learning task. Notation is described in the main text.

The predictors are represented by $k \in \mathcal{K}$. Allow $'$ to represent data that have been preprocessed, and $''$ preprocessed and subsampled for class imbalance. For the tuning process, allow resamples to be indexed by $b \in \mathcal{B}$ and repeats to be indexed by $c \in \mathcal{C}$, with an individual resample given by $R_{b,c}$ and a holdout from resampling by $T_{b,c}$. The complete set of tuning hyperparameters[17] is given by $\Theta_{l,m}$, and a candidate set of hyperparameters is given by $\theta_{j,l,m}$ for $j \in \mathcal{J}$ where $j = 1 \dots 10^p$ and p is the number of hyperparameters for model l. Allow $\kappa_{j,l,m}$ to represent the mean AUC from cross-validation across the $\mathcal{B} \times \mathcal{C}$ resamples for the j^{th} tuning hyperparameter set candidate. Denote an optimal hyperparameter tuning as $\theta^*_{l,m}$, where optimal is defined as $\max_{j \in \mathcal{J}} \kappa_{j,l,m}$. Finally, fitted models are given by $\widehat{f}$, and goodness-of-fit statistics are denoted $s \in \mathcal{S}$. The learning task can then be described as in Figure 3.4.

3.4 Models

I now describe each of the models used in the learning task. Due to space constraints, I focus on a conceptual introduction and direct readers to the works cited for mathematical exposition, as well as Hastie, Tibshirani, and Friedman (2009) for a general overview. I refer to models by their **method** name in caret, to distinguish among similar models and similar implementations of the same model.

3.4.1 GLMs and extensions

The first model I use is logistic regression. There are no hyperparameters to tune for this model, and variable importance here is just the absolute value of the sign of each

17. The tuning hyperparameters differ across replicates for the same model due to different initialization values (chosen randomly).

β_k (since all predictors are standardized before estimation). I refer to this model as **glm**. Political scientists rarely think of logistic regression as machine learning, yet it is a typical classification algorithm. Logistic regression and linear discriminant analysis (LDA; Fisher 1936), a canonical ML algorithm, produce nearly identical estimates of the log posterior odds between outcome classes—the two methods differ only in how the parameter estimates are computed.[18] Logistic regression also has close ties to naive-Bayes classifiers, perceptrons (a building block in neural networks), and linear support vector machines (see, e.g., Dreiseitl and Ohno-Machado 2002); all of these methods often yield very similar results. And despite an explosion in the variety of ML algorithms available, logistic regression is consistently among the best-performing classifiers (Bagnall and Cawley 2017). Thus, logistic regression can and should be evaluated against the more exotic techniques increasingly receiving political scientists' attention (Hastie, Tibshirani, and Friedman 2009; Hindman 2015).

The second model I use is the elastic net, denoted **glmnet**. The elastic net is a regularization of logistic regression that imposes a penalty on the coefficients consisting of a convex combination of the commonly-used L1 and L2 penalties, known as the Lasso and ridge, respectively (Hoerl 1962; Tibshirani 1996). The elastic net retains the shrinkage and variable-selection properties of the Lasso as well as the predictive performance and grouping effect of the ridge, while overcoming problems associated with each. In addition, the Lasso and ridge regularizations are special cases of the elastic net which can be estimated from the data (Zou and Hastie 2005).

Specifically, the elastic net retains the maximum likelihood objective function for logistic regression but adds the cost $\lambda \left[(1-\eta)\|\beta\|_2^2/2 + \eta\|\beta\|_1 \right]$. Setting $\lambda = 0$

18. As a result of these differences, logistic regression is typically more robust but less efficient (Efron 1975; Hastie, Tibshirani, and Friedman 2009).

produces the objective function for standard logistic regression, while any $\lambda > 0$ acts as a complexity penalty. When $\eta = 1$ the elastic net is equivalent to the lasso, when $\eta = 0$ it is equivalent to the ridge, and $0 < \eta < 1$ it is an optimal tradeoff between the two penalties (Friedman, Hastie, and Tibshirani 2010). It is over these two hyperparameters that the model is tuned. Like with **glm**, variable importance is the absolute value of the sign of the coefficients.

The final GLM variant I study is boosted logit, denoted **LogitBoost**. Boosting is a meta-algorithm that combines a number of "weak" models into a more powerful "committee," with the potential to decrease both bias and variance (Freund and Schapire 1996; Friedman, Hastie, and Tibshirani 2000). Boosted logit works by iteratively fitting weighted logistic regression models on the data, and at each boosting iteration, upweighting observations which have been misclassified by the previous iteration and downweighting those which have been correctly classified. The final classifier is then a combined weighted majority vote of the models fit at each stage (Dettling and Bühlmann 2003). The only tuning hyperparameter for this model is the number of boosting iterations. Variable importance for this model is calculated using ROC curve analysis on each predictor, with variables that provide larger AUCs deemed more important (caret's standard "model-independent" importance metric).

3.4.2 Variations on discriminant analysis

The second family of models I investigate can be broadly construed as extensions and alternatives to classical discriminant analysis. As noted above, canonical LDA produces results nearly identical to those of logistic regression, so I instead focus on support vector machines (SVMs), distance-weighted discrimination, and nearest

shrunken centroids.

Where the outcomes can be perfectly separated in the multivariate distribution of the predictors (the "feature space"), LDA and logistic regression choose the boundary that maximizes the distance (or margin) between the two classes.[19] Support vector machines extend this framework to cases where the classes are not perfectly separable, and points on either or both sides of the boundary are misclassified. SVMs work by choosing a hyperplane that maximizes the margin between classes, while imposing a cost (through "slack variables") that increases both with the number of misclassified cases and with those cases' distance to the decision boundary.

Among the many available SVM implementations, I study **svmRadialWeights**. SVMs rely on kernels to calculate distance between observations in the feature space; radial basis function (RBF) kernels perform well as a default because they generate more flexible boundaries than does the classic linear kernel (Hofmann, Schölkopf, and Smola 2008). This model is tuned over the width of the RBF kernel, as well as a cost hyperparameter which penalizes model complexity. Additionally, this model allows for tuning over weights assigned to each class, rather than assigning deconcentration and no deconcentration equal (or some other fixed) weight. Like with **LogitBoost**, variable importance for this model is the AUC contribution of each predictor.

Under certain conditions, the SVM decision boundary can be sensitive to the training sample, leading to unstable results across resamples (and, here, imputed data replicates). Distance-weighted discrimination addresses this problem by altering the margin-based criterion of the SVM to allow for greater influence of points further

19. Specifically, the margin is typically defined by two times the perpendicular distance to the closest point in the data (D'Orazio et al. 2014).

away from the decision boundary, inducing greater stability (Marron, Todd, and Ahn 2007). I use the sparse distance-weighted discrimination (**sdwd**) implementation because it imposes shrinkage akin to that of the elastic net, tuning over L1 and L2 regularizations in calculating the distance to the decision boundary, improving predictive performance (Wang and Zou 2015). Variable importance here is again just the absolute value of the sign of each coefficient, like with **glm** above.

Nearest centroid classifiers (NCCs) present an alternative to classical discriminant analysis. Instead of finding a boundary that best separates classes, NCCs assign each observation to the class with the closest mean in $\mathcal{K}$-dimensional space. Nearest *shrunken* centroids classifiers go further by first shrinking the classwise mean toward the overall mean for each feature (separately). This shrinkage "de-noises" the centroids for each class, dropping or reducing the importance of features that do not contribute to class predictions (Tibshirani et al. 2002). This model, **pam**, tunes over the shrinkage hyperparameters, and uses the difference between the class centroids and the overall centroid for each feature to measure variable importance. Predictors with class centroids further away from the overall centroid better separate the outcome classes, and so are considered more important.

3.4.3 MARS

The third family of algorithms I study are MARS models. MARS is a form of non-parametric regression where the outcome is regressed on a weighted sum of basis expansions of the predictors, each of which (excluding the intercept) is either a hinge function or a product of hinge functions. A hinge function detects non-linearities in predictors at locations known as knots; although hinge functions remain linear, their interactions can be non-linear. Hinge functions and their interactions are estimated

directly from the data (Friedman 1991).

Conceptually, MARS models are similar to generalized additive models (GAMs), with which political scientists are more familiar (Beck and Jackman 1998). However, GAMs use local approximations such as splines instead of hinge functions, and cannot automatically model interactions. Typically, this limitation leads to worse fit than MARS offers (Hastie, Tibshirani, and Friedman 2009, 295-336). On the other hand, given this extreme flexibility in model-building, MARS runs the risk of overfitting the data. Thus, MARS estimation consists of a first stage in which the hinge functions are built, and a second stage in which the model is trimmed. In this "backward pass," MARS prunes the inputs and basis functions to those that produce the lowest generalized cross-validation (GCV) error. Specifically, MARS approximates the error from leave-one-out cross-validation, adds a complexity penalty to this error, and then removes terms which provide the smallest GCV gains, resulting in an optimal sub-model of the fit obtained during the "forward pass."

The first MARS implementation I study is **gcvEarth**, which tunes over the maximum interaction depth (i.e., how many hinge functions can be multiplied in a single interaction term). Variable importance is a measure of the reduction in GCV error attributable to adding the feature and its interactions to the model; variables never used in a basis function are given an importance value of zero.

The second model I study is **bagEarthGCV**, which supplements **gcvEarth** with bootstrap aggregating, or "bagging" (Breiman 1996). Like boosting, bagging is a meta-algorithm implemented in a number of ML models. The idea behind bagging is to improve predictive accuracy by averaging across sub-models fit to bootstrap replicates sampled from the training data (in this case, sampled from the training data replicate). Bagging is most powerful when the model is unstable, as with MARS and tree-based ensembles. Bagging may also be particularly useful for this task

not only as a way to capture the variation across imputed data replicates, but also because I am primarily interested in variable importance, which is especially unstable across MARS models. Like **gcvEarth**, **bagEarthGCV** tunes over the maximum interaction depth, but calculates variable importance as average GCV error reduction across bootstrap replicates.

3.4.4 Decision trees and forests

Classification and regression trees (CART) and forests are among the most widely studied ML algorithms in political science (Blair, Blattman, and Hartman 2017; Grimmer and Stewart 2013; Muchlinski et al. 2016). These models work by iteratively searching for predictors that best partition the observations into outcome classes. Beginning from the original "node" with the entire replicate training sample, the data are split using whatever predictor minimizes average Gini impurity (a measure of the distribution of classes) in the new nodes created by the split. These nodes are then partitioned further until some stopping criterion is reached, such as the minimum number of observations in a node. I study a simple implementation of this algorithm, **rpart2**, which tunes over the maximum tree depth (i.e., how many splits a branch can have), which acts as a complexity penalty. Variable importance for this model is the summed reduction in Gini impurity contributed by each predictor.

Individual decision trees suffer from instability and overfitting. I therefore also study random forests, which create ensembles from individual trees through bagging (Breiman 2001a). Random forests work by taking samples of the training data and growing trees from each of them, then classifying each observation through a majority vote weighted by the trees' out-of-bag prediction error. Because of the enormous computational costs of building the forest (relative to other models),

predictors are randomly sampled at each node in each tree, and the split is made using the best among the sampled predictors only. The implementation I use, **rf**, tunes over the number of predictors sampled at each node. Variable importance here is "permutation importance," obtained by permuting each predictor randomly and then measuring the average difference in predictive accuracy (see, e.g., Hill and Jones 2014). Larger declines in predictive performance after permutation indicate greater importance.

The second decision forest I estimate is **C5.0**. Instead of minimizing Gini impurity, this model splits nodes by maximizing the normalized information gain (for details see Quinlan 1993). **C5.0** has a few other attractive properties: it incorporates boosting iterations and pre-winnows predictors that contribute little information before building the trees (Hastie, Tibshirani, and Friedman 2009). The number of boosting iterations and the amount of winnowing constitute the tuning hyperparameters. Instead of permutation importance, variable importance here is measured as the proportion of observations in the training sample classified at a split by a predictor, so that predictors which generate branches that affect more observations are more important.

One shortcoming of random forests for model selection is that the splitting procedure identifies *individually* important predictors, not the most important subset. If predictors are strongly correlated, this might lead to redundancy, and a much more compact set of features may be able to produce similar predictive accuracy. Regularized random forests (RRFs) address this issue by splitting nodes based on a regularized information gain criterion, where the regularization is a complexity penalty (Deng and Runger 2013). The model I study, **RRF**, is tuned over the size of the sample of predictors at each node in each tree, as well as the magnitude of the complexity penalty. As with **rf**, variable importance is permutation

importance.

The last random forest I study is **rotationForest**. This approach is motivated by the desire to increase the diversity of trees in the ensemble without sacrificing accuracy (Kuncheva and Rodríguez 2007; Rodríguez, Kuncheva, and Alonso 2006). Diversity is induced by randomly sampling predictors and performing principal components analysis before building each classification tree. The model tunes over the number of predictor subsets and the number of trees comprising the ensemble. Variable importance is again permutation importance, with the added step of filtering the permuted predictors through the factored indices.

3.4.5 Neural networks

The final category of models I estimate are neural networks. Conceptually, the most commonly-studied neural networks are akin to two-stage regression models, where the first stage is a basis expansion of the inputs (Hastie, Tibshirani, and Friedman 2009, 389-416). More formally, the first stage fits a set of expansions to the predictors, producing a set of transformed inputs ("hidden units"), which are then used to predict the outcome. In the second stage, the fitted errors are fed back through the model to the original predictors through a process known as "back-propagation" before the expansions and hidden units are adjusted and the next fitting iteration begins. The parameters of the two sets of functions—the basis expansions and the eventual model fit—are known as "weights" or "bias terms," and are chosen to minimize mean-squared error. This type of neural network, a single hidden layer back-propagation network (or single layer perceptron), has been known to political science for nearly two decades (Beck, King, and Zeng 2000), though its use remains limited due to its perceived difficulty of interpretation.

The first such model I study is **mlp**, which is tuned only over the number of hidden units. I also use **nnet**, which is tuned over both the number of hidden units and the size ("decay") of the weights—the number of basis expansions of the predictors and their estimated effects on deconcentration. Variable importance for both models is given by Garson's algorithm, which tracks the weights assigned to all of the connections branching from a predictor to the output, combines them, and scales them relative to the other predictors. Larger cumulative weights indicate greater variable importance (Garson 1991; Goh 1995). The final model I study is a stacked autoencoder deep neural network (**dnn**). This model builds on **nnet** by adding two more layers of hidden units, effectively passing predictors through three basis expansions instead of just one en route to deconcentration. In addition to tuning over the number of hidden units at each layer and the weights, this model also tunes over how many of the weights to pass upward from each hidden layer toward deconcentration. Garson's algorithm is not available for multiple-layer neural networks, so like with **LogitBoost** and **svmRadialWeights**, variable importance for this model is just the AUC contribution of each predictor.

3.5 Results

3.5.1 Variable importance

The variable importance results across all of these models are presented in Table 3.4. Across the board, the variables identified in the coalition manipulation theory provide more information about deconcentration than do the alternatives. Predictors relating to coalition bargaining have larger importance values and higher importance rankings than those of competing theories, and indeed many of the control variables. These

Table 3.4: Predictor importance across models

		Mean importance		Mean importance rank	
	Variable	Best models	All models	Best models	All models
✓	Number of units	99.52	94.67	1.05	1.70
✓	Mean party age	56.98	54.93	3.46	5.21
	Population	48.70	58.42	4.77	4.66
✓	Leg. independence	36.24	41.36	6.53	7.23
✓	Local elections	35.04	43.72	7.05	7.78
✓	Veto players	34.12	43.13	6.82	7.25
✓	Cabinet members	31.36	44.40	7.72	6.89
	GDP	27.91	49.52	9.77	6.11
	Geographic size	26.41	36.49	10.17	8.60
	Democracy	19.56	30.96	11.51	10.53
	Inequality	19.44	30.92	10.90	10.24
✓	Transition	18.89	14.02	10.70	13.36
✓	Jud. independence	16.14	16.55	11.95	13.23
	Military spending	13.11	23.75	13.15	11.96
	Leg. election	10.40	6.53	14.01	15.89
	Ruggedness	10.09	12.21	14.25	14.86
	Exec. election	8.43	9.03	14.86	15.79
	Ethnic diversity	6.67	19.95	15.47	15.47
	International war	5.58	8.38	15.85	15.85

Mean importance and importance ranks, averaged across all replicate datasets and models. "Best models" refers to the five models with best out-of-sample predictive performance (i.e., the highest mean AUC across replicate validation datasets). The scale is 0-100 for importance and 1-19 for importance rank; larger importance values and smaller ranks indicate more information provided to the models. Checkmarks note the predictors associated with my coalition manipulation theory.

patterns become even more apparent among the five best-performing models—**sdwd**, **nnet**, **glmnet**, **glm**, and **rotationForest**—which all achieve an AUC in out-of-sample prediction of over 0.70, averaged across imputed data replicates, as plotted in Figure 3.5, below. Although the judicial independence and transition variables' performance is underwhelming, all four of the major empirical implications are represented in the top six variables. The number of politicians bargaining over coalitions, the number of veto players, the degree of majoritarianism, and the stability of the extant coalition structure all therefore provide substantial information for predicting deconcentration.

Taken as a whole, these results indicate that the coalition manipulation theory is a useful framework for understanding deconcentration.

In contrast, among the variables identified in other theories, only GDP and democracy average a top-10 variable importance ranking. While the evidence is mixed for alternative theories, the majority-rigging predictors fare particularly poorly, with ethnic diversity and both election indicators all among the five least informative variables. One concern may be that these variables are indicators, have less variation than continuous variables, and so naturally have less information to impart to the model. However, no such clear pattern emerges: legislative independence, local elections, and veto players achieve are all discrete-valued and yet are the 4^{th}-6^{th} most important variables, while ethnic diversity and ruggedness are both continuous but are found near the bottom of the Table.

Interestingly, among the eleven worse-performing models, the converse holds: they assign less importance to the variables associated with the coalition bargaining theory. Figure 3.6 plots variable importance among these two sets of models against each other; predictors that are equally important across models will be placed along the 45° (dashed) line. Predictors relating to the coalition manipulation theory (in blue) are closer to the top right, reflecting the results in Table 3.4. Three of these eight are below the dashed line, indicating that they are more important among the best-performing models than among the others. Variables relating to alternative explanations (in orange) are bunched around the origin, and only one of the eight are below the line. None of the three controls (in bronze) are more important in the best-performing models. This pattern shows up in the average distance to the 45° line among these groups. Predictors relating to coalition bargaining are 4 points less important among the best models, but predictors from alternative

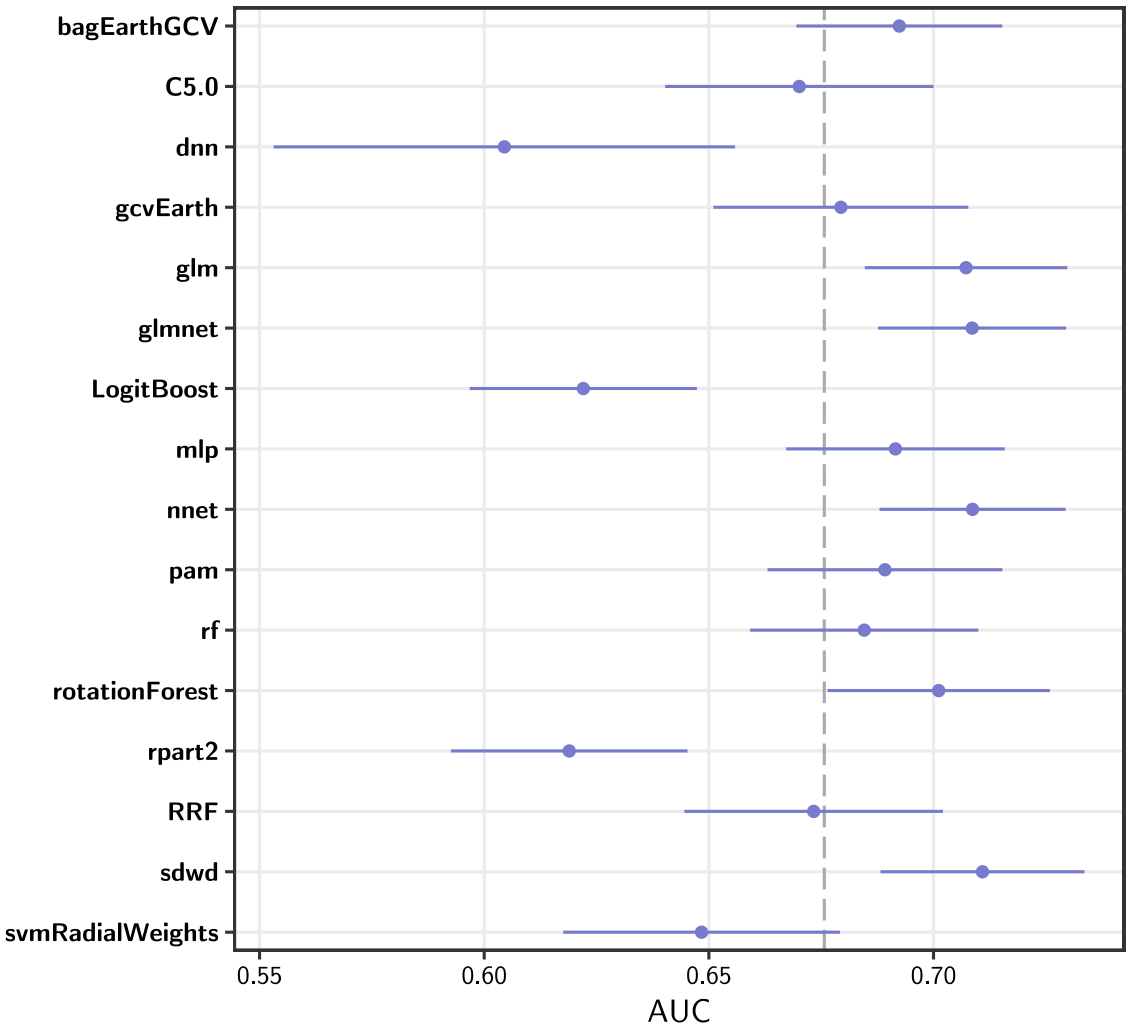


Figure 3.5: Model performance. Mean AUCs for out-of-sample predictions for each of the models, averaged across imputed data replicates, with lines for $+/-$ one standard deviation. The dashed line indicates the mean AUC across all models and replicates. Larger values indicate better performance.

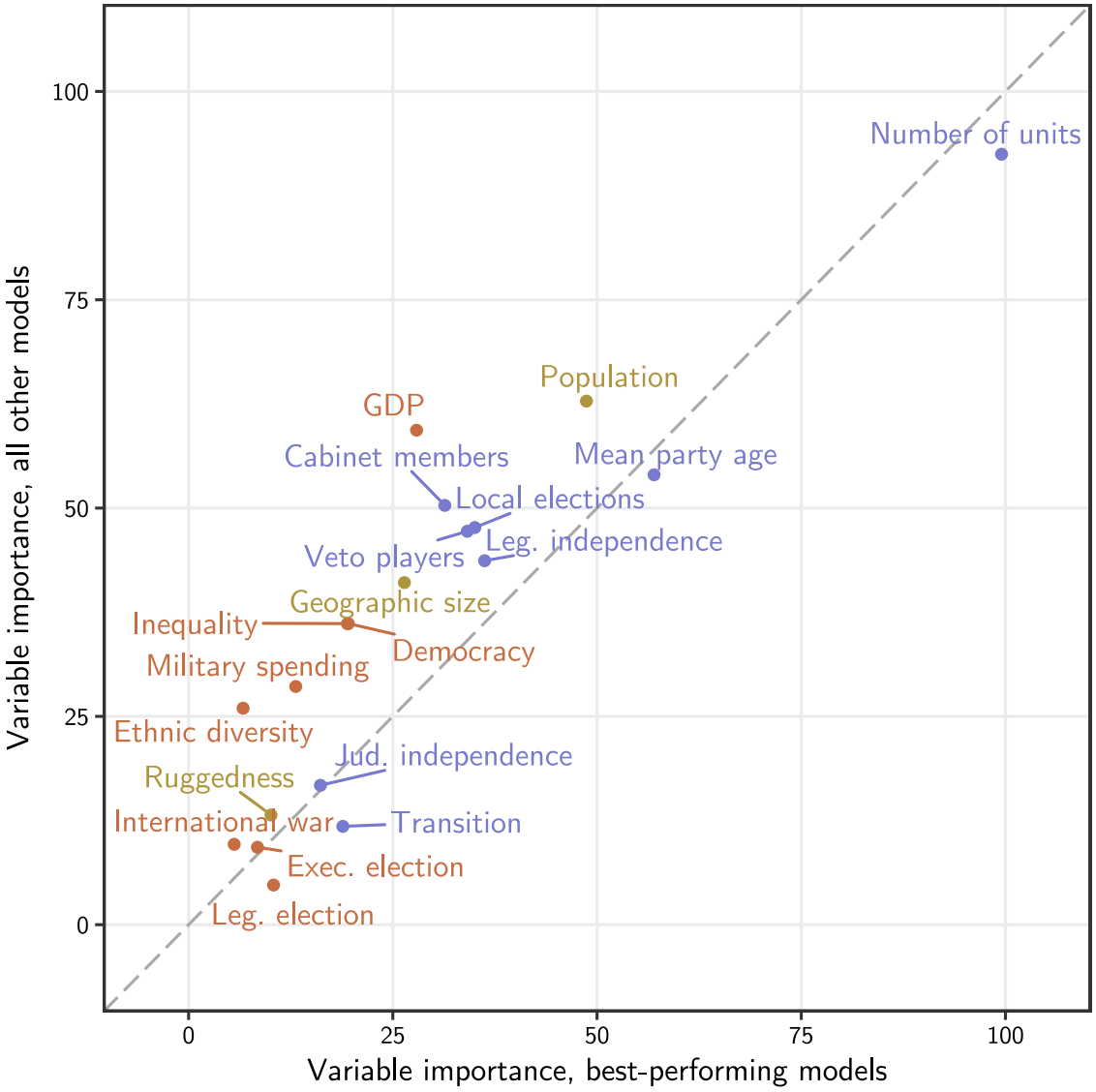


Figure 3.6: Predictor importance by model performance. Mean variable importance is plotted for each of the predictors among the five best-performing models (x-axis) and among the eleven worse-performing models (y-axis). Blue dots indicate variables related to the coalition manipulation theory, orange to alternative explanations, and bronze to control variables.

accounts and controls are 12 and 11 points less important.[20] And the difference between models is significant: while the best models achieve a mean AUC of 0.71, the other models average just 0.66. Taken as a whole, these results suggest that discarding or down-weighting information related to coalitional bargaining causes predictive performance to deteriorate substantially. The coalition manipulation theory provides far more leverage for understanding the incidence of deconcentration than do alternative approaches or meaningful controls.

3.5.2 Predictive performance

In general, the best models' predictive power is high. Figure 3.7 plots ROC curves for each of these five, fit to each of the validation samples. These ROC curves suggest stability across data replicates, confirmed in Table 3.5 by the small AUC standard deviations. These AUCs also compare favorably with many of the classification tasks explored elsewhere in the discipline: the best models here provide about as much predictive power as do prominent theories of local violence, US Supreme Court decisions, and civil war (Blair, Blattman, and Hartman 2017; Montgomery, Hollenbach, and Ward 2012; Ward, Greenhill, and Bakke 2010). Linearly combining these models into an ensemble yields better predictions still, though the gains are very modest.[21]

Table 3.5 also provides performance metrics based on classification accuracy. All of the models provide very similar predictive value across outcome classes. Sensitivity (also known as recall), the ratio of true positives to observed positives,

20. Worse-performing models tend to be unstable and so spread variable importance more evenly, resulting in most predictors being above the 45° line.
21. The mean ensemble AUC is 0.7113 compared to 0.7109 for the best-performing model, **sdwd**, and is slightly more stable, with a standard deviation of 0.0221 versus **sdwd**'s 0.0227. However, these differences are so minute that the results are not meaningfully changed.

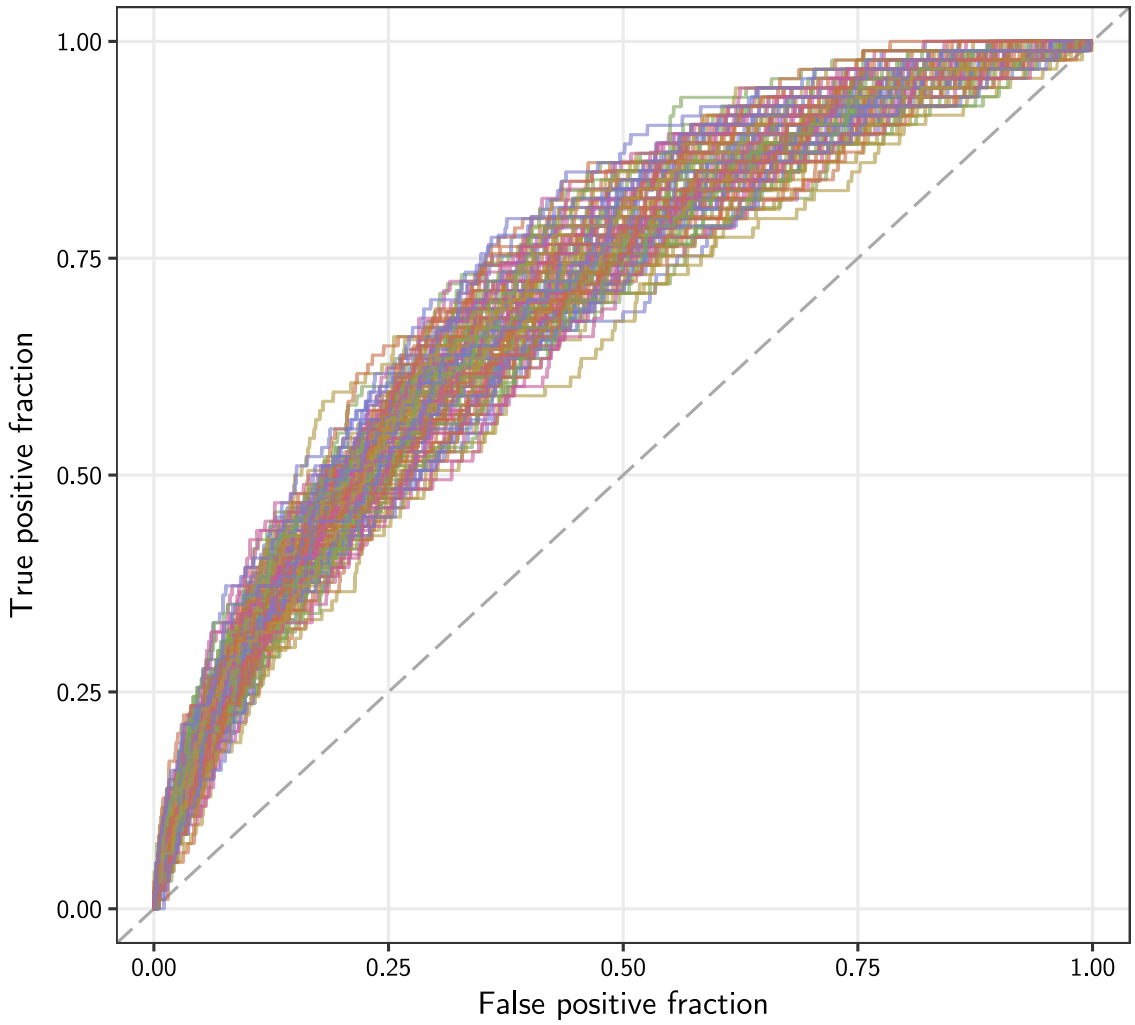


Figure 3.7: ROC curves for the five best-performing models. Each line is a ROC curve from a model fit to a replicate dataset. **sdwd** models are plotted in blue, **nnet** in orange, **glmnet** in purple, **glm** in green, and **rotationForest** in bronze. Curves toward the upper-left indicate better performance.

Table 3.5: Predictive performance for the five best models

	Model				
	sdwd	**nnet**	**glmnet**	**glm**	**rotationForest**
AUC	0.71	0.71	0.71	0.71	0.70
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Brier score	0.17	0.17	0.17	0.17	0.15
	(0.00)	(0.00)	(0.00)	(0.01)	0.00)
Sensitivity (recall)	0.51	0.49	0.48	0.52	0.44
	(0.05)	(0.06)	(0.05)	(0.04)	(0.04)
Specificity	0.77	0.78	0.79	0.76	0.81
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
PPV (precision)	0.09	0.09	0.10	0.09	0.10
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NPV	0.97	0.97	0.97	0.97	0.97
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
F_1 score	0.16	0.16	0.16	0.16	0.16
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
F_2 score	0.27	0.27	0.27	0.27	0.26
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)

Performance metrics are averaged across imputed data replicates. Standard deviations are in parentheses. All metrics are constrained to $[0, 1]$, with higher values indicating better performance for all except Brier Scores.

represents the ability of the model to detect deconcentration. In contrast, specificity is the ratio of true negatives to observed negatives, giving the ability of the model to detect "no deconcentration." Unsurprisingly, all of the models are better able to detect "no deconcentration" than they are deconcentration, a function of the class imbalance in the underlying data. Additionally, the low positive predictive value (PPV, or precision) rates suggest that the models are over-predicting deconcentration. Defined as the ratio of true positives to predicted positives, a PPV of 0.10 suggests that deconcentration occurs only once in every ten cases where it is predicted.[22]

Although these PPVs are low, this does not mean these models are not useful

22. This calculation assumes a classification cutoff of 50%, i.e., that an observation assigned a predicted probability of deconcentration over 0.50 is assigned class 1, and class 0 otherwise.

for predicting deconcentration. Table 3.5 also gives each model's average F_1 score, the harmonic mean of precision and recall, which summarizes its ability to accurately predict deconcentration when it does occur as well as to *not* predict it when it doesn't. F_2 scores are computed similarly but upweight recall relative to precision, penalizing false negatives ("missed" cases of deconcentration) more harshly than false positives, arguably more theoretically relevant here. Both measures are constrained between zero and one, with larger values indicating better performance.[23]

These F scores can be compared to those of a "no information" model, wherein deconcentration is predicted by random guessing, through a small Monte Carlo experiment. Specifically, I generate simulated data with the same class imbalance and sample size as in each of the validation samples. I then predict deconcentration in some proportion π of observations, incrementing π over $(0, 1]$ in steps of 0.01.[24] For each value of π I make 5,000 sets of guesses and F_1 and F_2 score calculations. The results of this experiment indicate that the best F_1 score possible from random guessing is 0.085, approximately half of the F_1 rates reported in Table 3.5. Similarly, the best F_2 score from random guessing is just 0.189. Thus, the models that use data on coalition bargaining to predict deconcentration improve classification accuracy by 50-100% over a "no information" model.

3.5.3 Interpretation

Finally, to demonstrate that the coalition-manipulation variables correlate to deconcentration in the directions predicted by theory, I select a few models for

23. More generally, F_β scores are given by $(1+\beta^2)(\frac{\text{precision}\cdot\text{recall}}{(\beta^2\cdot\text{precision})+\text{recall}})$. Larger values of β correspond to even greater relative weight assigned to recall vis-à-vis precision.
24. Setting $\pi = 0$ generates undefined F scores, since PPV evaluates to $0/0$. By always guessing "no deconcentration," we would achieve a marginally better NPV of approximately 0.985, at the cost of forcing recall to zero and losing all PPV.

interpretation. Supervised ML models are often criticized as "black box" methods for their difficulty of interpretation. However, recent scholarship has introduced new tools for analyzing ML results (Jones and Linder 2016; Muchlinski et al. 2016). I focus on three of the best-performing models for which such methods exist: **glmnet**, **rf**, and **nnet**.[25]

Results from **glmnet** and **rf** are plotted in Figure 3.8 in blue and orange, respectively. Each panel shows the change in predicted probability of deconcentration of moving each variable across its complete range while holding all other variables at their central tendencies. (Each variable is pre-processed, so the scale of this range is not meaningful and is therefore omitted.) For continuous (binary) variables, each line (point) is the loess fit to predictions drawn across replicate datasets, with ribbons (lines) for 95% confidence intervals. Note that the y-axes differ across panels.

Both models uncover relationships predicted by the coalition manipulation theory. All of the effects are in the expected direction, and all of the changes in the predicted probability of deconcentration are substantively large. Unsurprisingly, given the variable importance results in Table 3.4, increasing the number of existing subnational units correlates with the greatest jump in the likelihood of deconcentration, doubling or quadrupling across its observed range, depending on the model.[26] The smallest effects appear to be when a country is experiencing a regime transition, but even here the effect is between 4-10 percentage points, for a 10-30% increase in the probability of deconcentration. The strength of these relationships, and the broad agreement across models, indicate that these variables contribute substantially to understanding variation in deconcentration.

25. Although **rf** is not among the best-predicting models by AUC, it produces a much lower mean Brier score (0.13, with a standard deviation of 0.00) than the models in Table 3.5.
26. Unlike **glmnet**, **rf** allows for nonlinear predicted effects. In line with the argument in Chapter 2 that deconcentration provides diminishing marginal returns in coalition-manipulation value, the effect size tapers off where the number of existing units is highest.

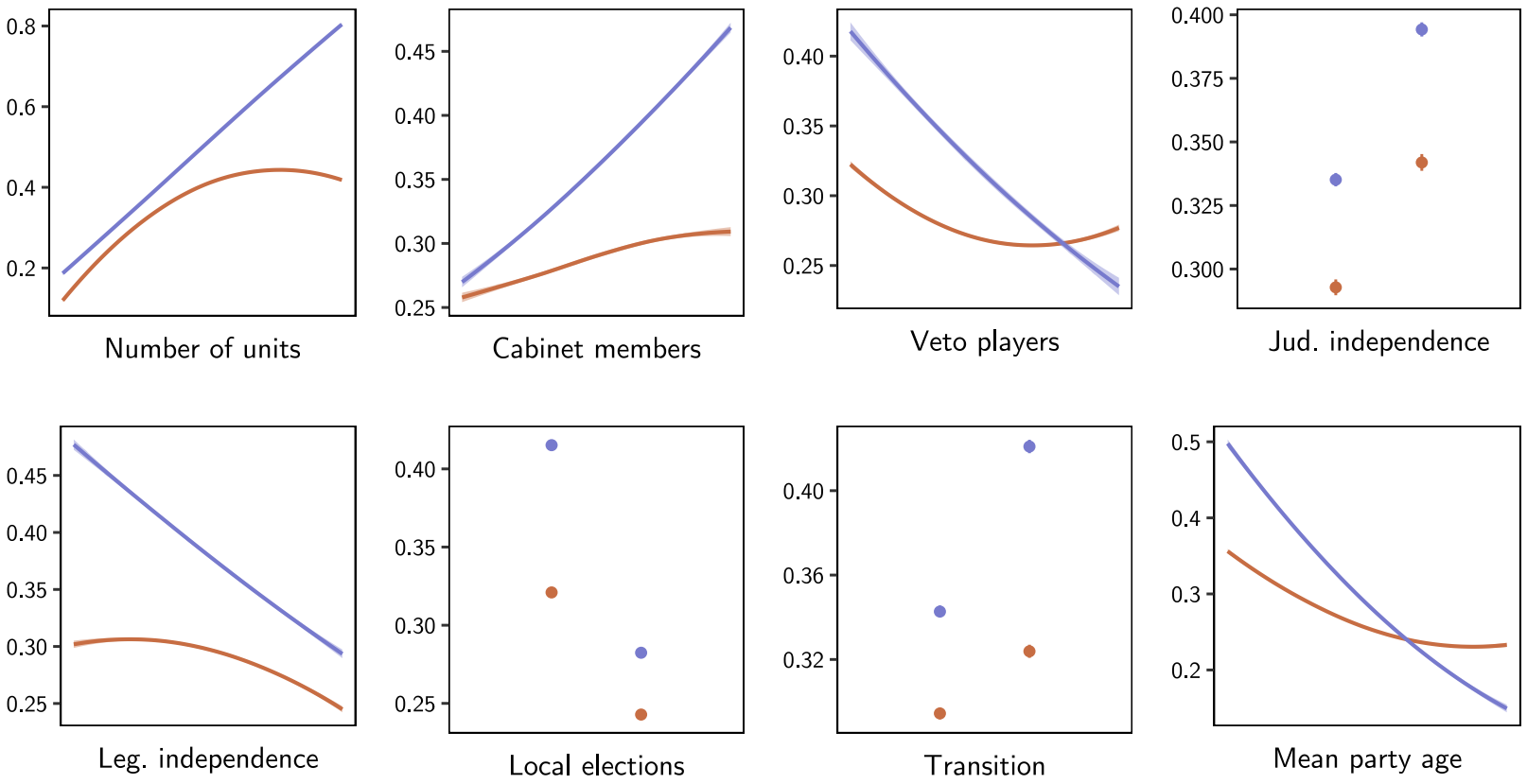


Figure 3.8: Predicted probabilities from the elastic net and random forest. Predicted probabilities of deconcentration from **glmnet** and **rf**, plotted in blue and orange, respectively. Lines (dots) represent mean predictions from loess fits across all imputed data replicates, with ribbons (lines) for 95% confidence intervals, moving each covariate across its range (or from 0 to 1 for binary predictors), while holding all other predictors at their central tendencies. Note that the axes differ across panels.

Similarly, Figure 3.9 plots one of the fitted neural networks. The width of each connecting line is proportional to the strength of the relationship (i.e., the absolute magnitude of the coefficient), and "hidden" nodes refer to the fitted basis expansions of the inputs.[27] Among the inputs, the thickest lines clearly come from the predictors in the top cluster, which relate to my theory, as well as the population control in the bottom cluster. These lines' weight relative to those of the predictors from alternative theories suggests that variable importance for **nnet** mirrors the general findings in Table 3.4. Further, it is possible to compute exact Garson (1991) weights, represented by the dots next to predictors. Larger dots indicate greater magnitude (in absolute value), and thus more importance to the model. That the largest dots are found among the top cluster indicates that this neural network relies most heavily on coalition manipulation variables.

Finally, it is possible to infer the signs of the connections and weights, which are colored orange for positive values and blue for negative. Larger values of predictors colored blue therefore increase the probability of deconcentration, while larger values of those in orange lower it. As expected, the number of units, the number of cabinet members, and transitional periods increase the predicted probability of deconcentration, while the other variables decrease it. Only judicial independence has the "wrong" sign, though its estimate is unstable across **nnet** fits. Similarly, all control variables exert the expected effect on deconcentration. In contrast, variables from alternative accounts are a mixed bag. Ethnic diversity and legislative elections both run in the expected direction, but inequality has the "wrong" sign (and, among alternative theories' variables, is given the greatest weight), as do executive elections and international war. Although this is just one of the 25 fitted **nnet** models and should not be over-interpreted, Figure 3.9 again suggests

27. Intercepts are omitted.

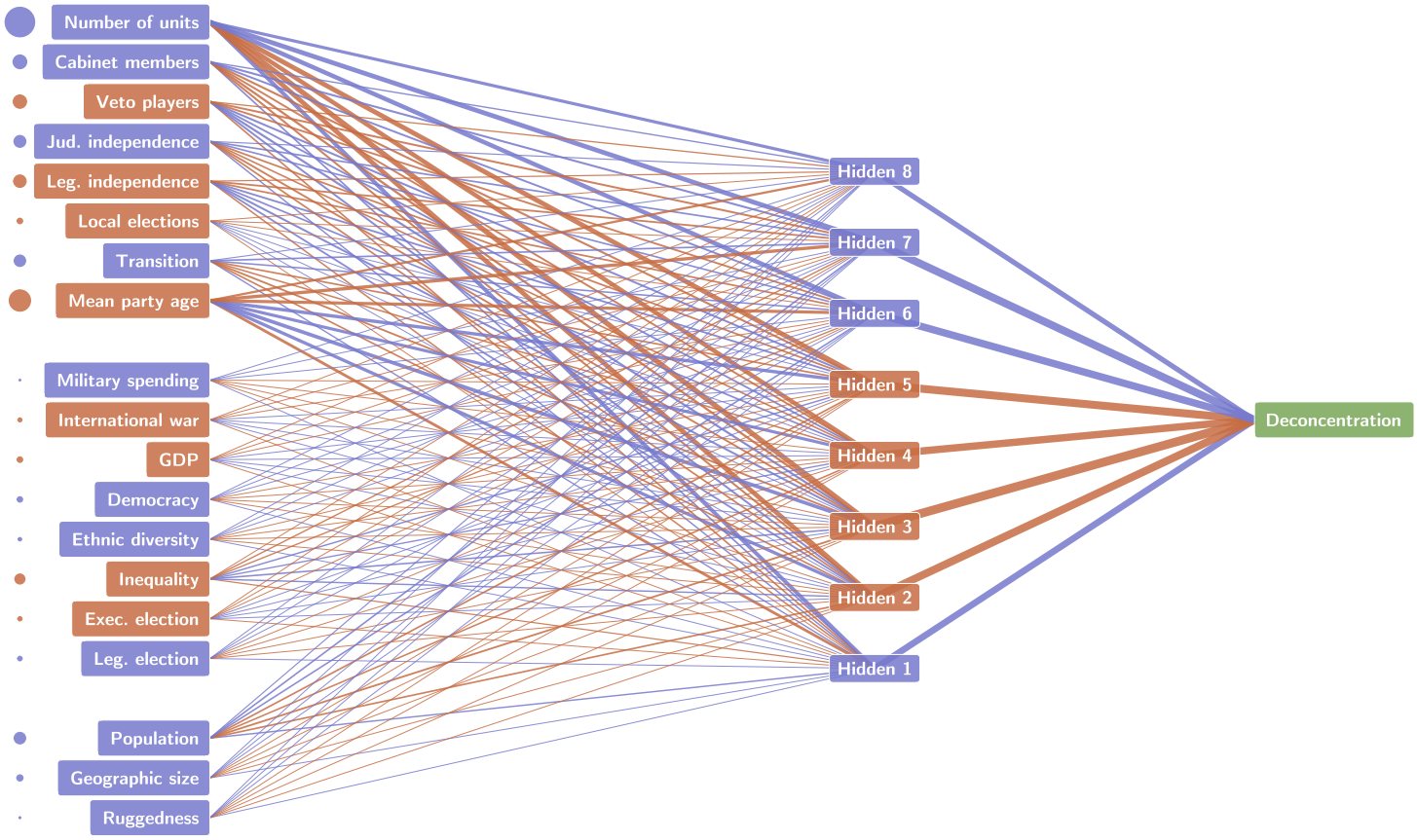


Figure 3.9: An example neural network fit. One of the neural networks fit by **nnet** to an imputed data replicate. "Hidden" nodes represent the basis expansions of the predictors. Positive values are in blue and negative values are in orange, where nodes are colored according to their overall sign. Line thickness and dot size correspond to effect magnitude.

that my theoretical expectations match the patterns in the data better than do alternative accounts.

3.6 Conclusion

Understanding the causes of deconcentration begins with understanding when and where it occurs. Although my theory hypothesizes causal relationships, the first and most direct test of its utility is whether it is able to predict observable patterns in the incidence of deconcentration. After all, prediction is an important goal in social science—perhaps *the* goal of social science (Lakatos 1978).

New data help shed light on this variation. Accounting for nearly 400 cases of deconcentration across 60 years and 189 countries, these data allow me to test the utility of the coalition manipulation model. I find that information about coalition bargaining consistently yields better predictions about the incidence of deconcentration, whether in simple logistic regressions or sophisticated machine learning models. Further, comparable predictive performance implies that the coalition bargaining theory provides as firm a theoretical base for deconcentration as, for instance, the Fearon and Laitin (2003) state weakness theory does for civil war. None of the alternative explanations comes close to providing similar predictive power. And for each variable, the direction of these relationships matches theoretical expectations: more politicians bargaining over coalitions, fewer veto players, less majoritarianism, and an unstable coalition structure all correspond to an increased probability of deconcentration.

This chapter has focused on building models for predictive inference. Yet while causation implies prediction (Muchlinski et al. 2016), prediction is also possible in the absence of a strong causal theory, or even when the causal theory is known

to be false (Breiman 2001b). These results are unable to speak to whether the relationships uncovered in this chapter are causal as theorized, or whether they are endogenous to forces unaccounted for here. It is to this question that the next chapter turns.

4 | Death and disequilibrium

Ahmed Sékou Touré ruled Guinea from independence in 1958 until his unexpected death in office in March 1984. Over 26 years, Touré had maintained a tight grip on power by instituting one-party rule and winning a series of uncontested elections in 1961, 1968, 1974, and 1982. Yet barely a week after his death, his constitutional successor was overthrown in a military coup led by Lansana Conté. A quarter-century of rigid elite power networks were suddenly thrown into disarray as the ruling party was abolished, the National Assembly dissolved, the Constitution suspended, and a thousand political prisoners freed (Kudamatsu 2009). Amid all of this upheaval, Guinea's new rulers pulled a lever that Touré never had: they created five new prefectures. Guinea's administrative divisions had remained essentially fixed over three decades, but instability following Touré's death pushed his successors to deconcentrate within two years.

My theory of deconcentration argues that politicians use deconcentration to manipulate coalitions. One of the key empirical implications of the model is that deconcentration is more likely during periods of coalitional instability. Shocks to the set of politicians bargaining push coalitions out of equilibrium, after which politicians use deconcentration to direct the bargaining to a more favorable new coalition structure. In short, my theory predicts that when politicians like Touré exit coalitional bargaining, the remaining players turn to deconcentration to try and

better their position in the new landscape.

In this chapter, I examine this prediction of the model using exogenously-timed leadership deaths. My data include all deaths of heads of states and governments over the period 1960-2010, built from multiple primary and secondary sources. I restrict attention to deaths that are apolitical, for instance, heart attacks and plane crashes. These deaths are exogenous to the process of coalition bargaining, allowing me to estimate the causal effect of such shocks on the probability of deconcentration. This identification strategy relies on the fact that the *timing* of leadership deaths does not depend on underlying social, political, and economic conditions.

I find that leadership deaths increase the likelihood of deconcentration significantly: the predicted probability of deconcentration within the two years following such a death increases by about 3.2 percentage points, or 71%. These estimates suggest that if deconcentration always occurred as frequently as it does following a shock to coalition bargaining, then we would expect an additional five cases per year. Further, these findings do not just reflect generalized instability following leadership death, as no other major institutional changes (e.g., enacting term limits or amending the constitution) become more likely. Finally, death and deconcentration each ultimately lead to changes in the extant coalition structure, demonstrating direct evidence of the theoretical mechanism. Together, these results provide strong evidence that coalition manipulation is the causal force driving deconcentration around the world over the last 60 years.

4.1 Data and model

4.1.1 Data

Data for the dependent variable, deconcentration, are the same as those employed in Chapter 3. Once again, I use Statoids change histories to code deconcentration as 1 if a country created new layers or top-level administrative units in a given year, and 0 otherwise.

Data for the independent variable, leadership death, come from a variety of sources. Many political science theories emphasize the role of political leadership, and so a number of scholars have compiled datasets on heads of states and governments that include information about how such leaders left office (Ahlquist and Levi 2011). The main source I draw upon is the Archigos project (Goemans, Gleditsch, and Chiozza 2009), which has the widest coverage and best documentation available. I supplement these data with two sources on leaders who died while in office: sudden deaths as identified in Jones and Olken (2005) and assassinations in Jones and Olken (2009), the latter to ensure that such deaths are not coded as exogenous. I also examine reference texts listing dates in office for heads of state, particularly the Central Intelligence Agency (2016) *World Factbook*, da Graça (1985), and Lentz (1994), which identify those who died before they could complete their terms. Finally, since many of these sources do not cover more recent years, I searched major English-language newspapers for mentions of leadership deaths since 2000.[1]

The treatment of interest is a binary indicator for whether a country-year experienced a leadership death due to "natural" causes, including illness, accident, and suicide. I code all assassinations, removals by internal or external forces, and

1. Specifically, I searched the archives of BBC News, the Los Angeles Times, the New York Times, and the Washington Post for country names paired with variations on "death," "died," and "passed away."

battle casualties as 0. I also code any deaths that occur during coups as 0, even if they would otherwise be coded as "natural." For instance, Chilean President Salvador Allende's suicide that came just as his regime was toppled in 1973 is not included as an apolitical leadership death.

The data sources I draw on contain substantial disagreements. Many of these datasets attempt to identify a *single* leader, but in some countries it is difficult to determine who that leader is. In mixed systems with both a Prime Minister and a President, scholars may disagree over which office wields greater authority. Alternatively, in single-party regimes, the head of state or government may often be considered a figurehead, while real power lies with the Secretary or Chairman of the Central Committee of the state party. Further, sudden leadership deaths are almost always followed by rigorous debate over the true cause of death. For instance, when Georgian Prime Minister Zurab Zhvania died from carbon monoxide poisoning during an apparent sexual misadventure, his family and allies maintained that the death was a politically-motivated murder. This uncertainty over true causes of death generates disagreements among datasets on which deaths can be considered apolitical.

I use contemporaneous scholarly sources to resolve coding inconsistencies. Where the disagreement is over who truly holds the most power, I err on the side of including leadership deaths because it constitutes a shock either way. If two politicians' relative power is similar enough that scholars disagree on who was *more* powerful, then the removal of one is essentially indistinguishable from the removal of the other, and should produce the same effect on coalitions. More formally, I want to ensure that the stable unit treatment value assumption is not violated. When two politicians can both be considered a country's "leader," the death of either one can reasonably constitute a consistent application of the treatment.

In contrast, I err on the side of coding deaths for which I am unable to establish an apolitical cause as 0. In doing so, I ensure that my estimates aren't polluted by including some deaths which are endogenous to the same forces driving deconcentration. For instance, it is possible that inter-ethnic competition could lead to a coup in which the head of state is killed, followed closely by deconcentration that serves to redirect centralized resources to an ethnic homeland. In this case, the leader's death and deconcentration would both be endogenous to inequality and ethnic fractionalization. To avoid such bias, I focus only on cases where the cause of death is known to be apolitical.

Finally, I also code as 0 cases where the evidence indicates that leaders were forced into retirement or inactivity due to illness, despite holding formal office until death. For instance, Jones and Olken (2005) include Artur da Costa e Silva, the Brazilian president who died of a heart attack in 1969, and Keizō Obuchi, the Japanese Prime Minister who died of a stroke in 2000. In both cases, Archigos codes these leaders as retiring due to illness, providing evidence that neither held true power by the time of their death. The purpose of this coding decision is to rule out anticipation effects: it is possible that reformers seeking to deconcentrate for reasons such as economic growth or majority-rigging, knowing the ruler was gravely ill, waited until his passing to push for reform. Excluding cases where the ruler was all but incapacitated well in advance of death guards against such bias.

The resulting data include 82 apolitical leadership deaths, presented in Table 4.1. Among these cases are 25 heart attacks, 15 deaths from cancer and another 15 from other long-term diseases, 10 sudden illnesses, 9 accidents, 5 deaths during surgery, 2 strokes, and 1 suicide. Thus, over half of the deaths in my data are very sudden. Although the other half of these cases are the result of cancer and other long-term illnesses, such cases do not threaten my identification strategy: even

Table 4.1: Apolitical leadership deaths

Country	Year	Name	Cause
Albania	1985	Enver Hoxha	Disease
Algeria	1978	Houari Boumédiène	Disease
Angola	1979	Agostinho Neto	Cancer
Argentina	1974	Juan Perón	Disease
Australia	1967	Harold Holt	Accident
Bahrain	1999	Isa bin Salman Al Khalifa	Heart attack
Barbados	1985	Tom Adams	Heart attack
Barbados	1987	Errol Barrow	Sudden illness
Barbados	2010	David Thompson	Cancer
Bhutan	1972	Jigme Dorji Wangchuk	Sudden illness
Bolivia	1969	René Barrientos	Accident
Botswana	1980	Sir Seretse Khama	Cancer
Brazil	1985	Tancredo Neves	Disease
Cambodia	1960	Norodom Suramarit	Sudden illness
Cambodia	1984	Chan Sy	Disease
China	1976	Mao Zedong	Disease
China	1997	Deng Xiaoping	Disease
Comoros	1998	Mohamed Taki Abdoulkarim	Heart attack
Côte d'Ivoire	1993	Félix Houphouët-Boigny	Cancer
Croatia	1999	Franjo Tuđman	Cancer
Denmark	1960	Hans Christian Hansen	Cancer
Dominica	2000	Rosie Douglas	Heart attack
Dominica	2004	Pierre Charles	Heart attack
Dominican Republic	1982	Antonio Guzmán Fernández	Suicide
Ecuador	1981	Jaime Roldós Aguilera	Accident
Egypt	1970	Gamal Abdel Nasser	Heart attack
France	1974	Georges Pompidou	Cancer
Gabon	1967	Léon M'ba	Cancer
Gabon	2009	Omar Bongo	Cancer
Greece	1964	Paul	Cancer
Grenada	1989	Herbert Blaize	Cancer
Guinea	1984	Ahmed Sékou Touré	Heart attack
Guinea	2008	Lansana Conté	Sudden illness
Guyana	1985	Forbes Burnham	Surgery
Guyana	1997	Cheddi Jagan	Heart attack
Haiti	1971	François Duvalier	Disease
Hungary	1993	József Antall	Cancer
Iceland	1970	Bjarni Benediktsson	Accident

Table 4.1 (continued): Apolitical leadership deaths

Country	Year	Name	Cause
India	1964	Jawaharlal Nehru	Stroke
India	1966	Lal Bahadur Shastri	Heart attack
Iran	1989	Ruhollah Khomeini	Surgery
Iraq	1966	Abdul Salam Arif	Accident
Israel	1969	Levi Eshkol	Heart attack
Jamaica	1967	Sir Donald Burns Sangster	Stroke
Japan	1980	Masayoshi Ōhira	Heart attack
Jordan	1999	Hussein bin Talal	Disease
Kenya	1978	Jomo Kenyatta	Sudden illness
Kuwait	1965	Abdullah Al-Salim As-Sabah	Heart attack
Kuwait	1977	Sabah Al-Salim As-Sabah	Cancer
Kuwait	2006	Jaber Al-Ahmad Al-Sabah	Sudden illness
Kyrgyz Republic	1999	Jumabek Ibraimov	Cancer
Laos	1992	Kaysone Phomvihane	Disease
Liberia	1971	William Tubman	Surgery
Malaysia	1976	Abdul Razak Hussein	Disease
Malaysia	2001	Salahuddin	Surgery
Mauritania	1979	Ahmed Ould Bouceif	Accident
Morocco	1961	Mohammed V	Surgery
Morocco	1999	Hassan II	Heart attack
Mozambique	1986	Samora Machel	Accident
Myanmar	2007	Soe Win	Disease
Nepal	1972	Mahendra	Heart attack
New Zealand	1974	Norman Eric Kirk	Heart attack
Nicaragua	1966	René Schick	Heart attack
Niger	1987	Seyni Kountché	Cancer
Nigeria	1998	Sani Abacha	Heart attack
Panama	1981	Omar Torrijos	Accident
Poland	2010	Lech Kaczyński	Accident
Romania	1965	Gheorghe Gheorghiu-Dej	Sudden illness
Saudi Arabia	1982	Khalid	Heart attack
Sierra Leone	1964	Sir Milton Margai	Sudden illness
Spain	1975	Francisco Franco	Disease
Swaziland	1982	Sobhuza II	Sudden illness
Syria	2000	Hafez al-Assad	Heart attack
Thailand	1963	Sarit Thanarat	Heart attack
Togo	2005	Gnassingbé Eyadéma	Heart attack
Trinidad and Tobago	1981	Eric Williams	Disease

Table 4.1 (continued): Apolitical leadership deaths

Country	Year	Name	Cause
Turkmenistan	2006	Saparmurat Niyazov	Heart attack
United Arab Emirates	2004	Zayed bin Sultan Al Nahyan	Disease
Uruguay	1965	Luis Giannattasio	Heart attack
Uruguay	1967	Óscar Diego Gestido	Heart attack
Vietnam	1986	Lê Duẩn	Sudden illness
Vietnam	1988	Phạm Hùng	Heart attack

though they are not sudden, their timing is, creating exogenous variation that allows me to treat my estimates as causal.

4.1.2 Model

My baseline model is a standard logistic regression with the country-year as the unit of analysis:

$$y_{i,t} \sim \mathcal{F}(\pi_{i,t}), \tag{4.1}$$

$$\pi_{i,t} = \text{logit}^{-1}(\alpha + \beta x_{i,t}), \tag{4.2}$$

where y_i is deconcentration, α is the intercept, β is the local average treatment effect (LATE), countries are indexed by $i \in \mathcal{I}$, years are indexed by $t \in \mathcal{T}$, $\mathcal{F}$ is the Bernoulli distribution, and logit^{-1} is the inverse-logit function.[2] The treatment x is an indicator that takes the value 1 if the country has experienced a leadership death in the last two years (inclusive), and 0 otherwise. For example, Mozambican president Samora Machel died in a plane crash in 1986, so this variable is coded as 1 for Mozambique in 1986, 1987, and 1988. I refer to these as "two-year windows." Since

2. Estimating the LATE with panel data and a binary response is non-trivial; Angrist and Pischke (2009) and Wooldridge (2010) both suggest an alternative strategy using linear probability models. Results from these models are nearly identical to those reported here.

both deconcentration and leadership deaths are hand-coded, the data are complete, covering every independent country for each year over the period 1960-2010. (The panel remains unbalanced, however, due to different dates of independence across countries.)

This use of windows in lieu of simple contemporaneous deconcentration has two motivations. The data are noisy, so studying simultaneous deconcentration is likely to obscure more than illuminate. For instance, the second Emir of Kuwait, Sabah Al-Salim As-Sabah, died on December 31, 1977. Restricting the model to contemporaneous deconcentration would unrealistically limit the timeframe for observing the relationship predicted by theory to just a few hours. Using two-year windows ameliorates this measurement problem. Moreover, the theoretical model says little about how long coalitional bargaining is expected to take, but anecdotal evidence suggests that it varies broadly and can take up to 18 months—as in Belgium, where the June 2010 elections did not yield a government until December 2011. Using a post-death window as the main covariate allows time for the bargaining at the heart of the model to unfold. I demonstrate below that my results are robust to choosing alternative window lengths.

Before estimating the model, I check whether treated and untreated observations systematically differ. Figure 4.1 plots the distributions of a range of social, economic, and political variables by treatment condition, with treated (untreated) country-years in blue (orange).[3] The only variable for which there is significant imbalance across treatment conditions is democracy. These results are consistent with

3. Geographic size is given in millions of square kilometers, population in millions, life expectancy in years, unemployment in proportion of the eligible labor force, and GDP in per capita constant US 2010 dollars (imputed)—all from the World Bank (2017). Democracy is given by the revised combined Polity IV score (Marshall, Jaggers, and Gurr 2011), imputed and recoded here into nondemocracies (< -6), democracies (> 6), and hybrid regimes (all others). Geographic size, population and GDP are all logged.

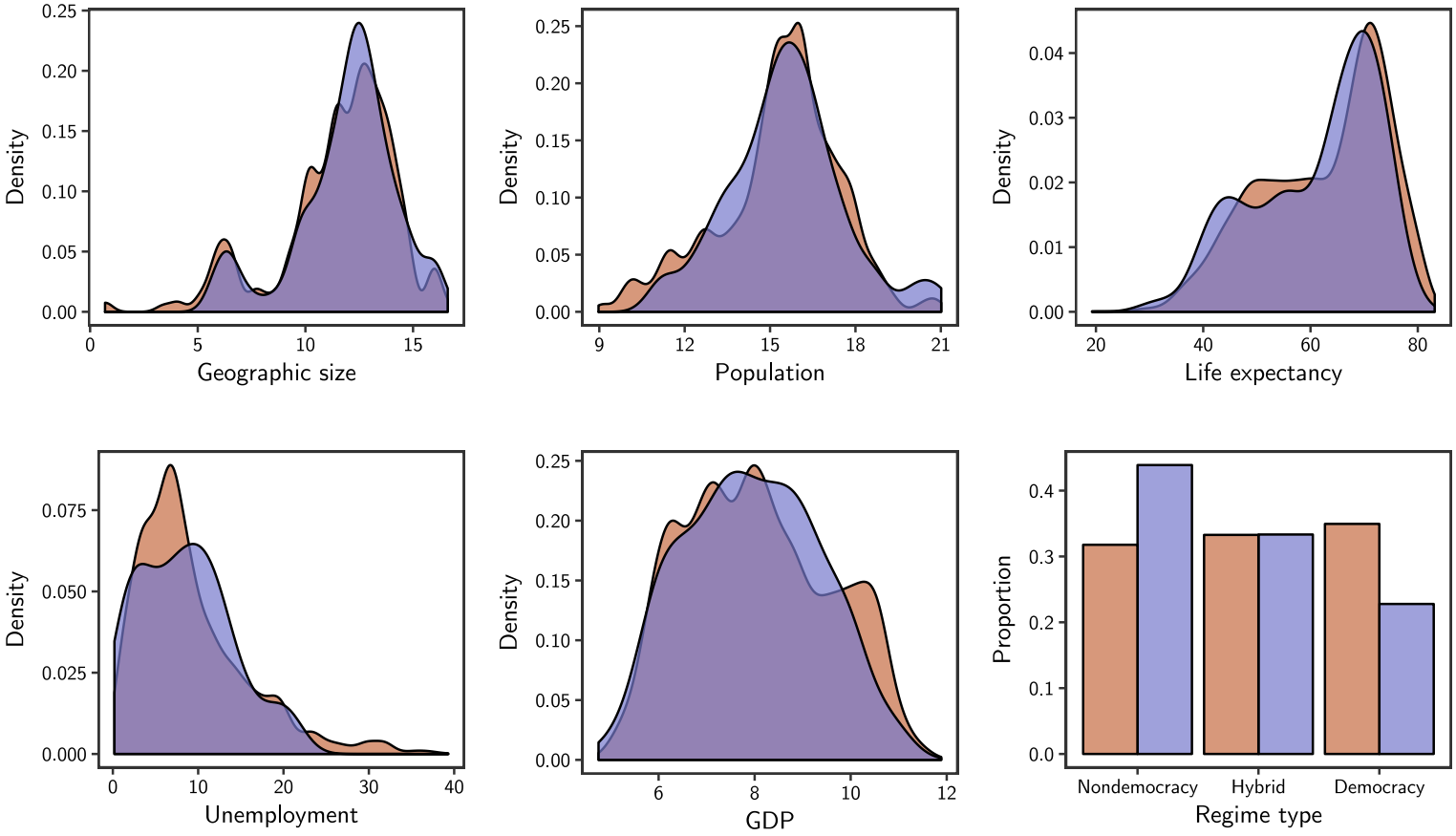


Figure 4.1: Balance across treated and untreated observations. Each panel plots the distribution of a social, economic, or political predictor for treated (blue) and untreated (orange) country-years. The similarity across distributions within each panel suggests that apolitical leadership death is applied as-if randomly by nature. Data sources and definitions are in footnote 3.

previous research connecting leadership death to democratization, but inconsistent with other studies that link it to economic underdevelopment (Treisman 2015). Nevertheless, to ensure that the ignorability assumption holds, I examine models that condition the treatment on both development and democracy.[4]

4.2 Results

4.2.1 Baseline result

The first column in Table 4.2 provides the main result. As expected, a country experiencing an apolitical leadership death is more likely to deconcentrate within the following two years. This effect is also substantively significant. The probability of deconcentration increases from 4.5% to 7.7% following such deaths—a 71% increase.[5] If the baseline frequency of deconcentration were raised to that following a leadership death, then we would expect to observe an additional 265 cases of deconcentration, or more than 5 extra per year, over the period 1960-2010.

To examine whether this result is simply capturing cross-sectional or temporal heterogeneity, the model in column (2) adds country and year fixed effects. Adding these parameters degrades model fit, but the LATE is unchanged. I then condition on development and democracy, following the discussion above. I measure development using gross domestic product (GDP) per capita, in thousands of constant 2010 US dollars (World Bank 2017), while democracy is given by the "revised combined" Polity IV score (Marshall, Jaggers, and Gurr 2011). To avoid post-treatment bias, I

4. In order for me to treat my estimates as causal, I require that $(y^1, y^0) \perp\!\!\!\perp x \mid Z$, where y^1 and y^0 refer to treated and untreated observations, respectively, x is the treatment, and Z are the confounding covariates (here, development and democracy).
5. All effect size calculations in this chapter hold fixed effects at their reference categories and controls at their means.

Table 4.2: The effect of leadership death on deconcentration

	(1)	(2)	(3)	(4)	(5)
Death in last two years	0.56*	0.55*	0.53*	0.55*	0.53†
	(0.25)	(0.28)	(0.27)	(0.28)	(0.27)
Development			−0.06*		−0.04
			(0.03)		(0.03)
Democracy				−0.03	−0.03
				(0.05)	(0.02)
Country and year FE?		✓	✓	✓	✓
BIC	3,025	4,628	4,679	4,628	4,686

*p $<$ .05,† p $<$ .10. Each model includes 8,089 observations across 189 countries.

use the level of economic development or democracy observed immediately prior to the first year for which leadership death data are available in each country. I also use multiple imputation to account for missingness among the controls. These results are reported in columns 3-5 of Table 4.2, with the LATE unchanged. Thus, the effect of leadership death on deconcentration does not appear to be an artifact of underlying conditions.

Returning to the prediction of the model—that a shock to the players bargaining over coalitions increases the probability of deconcentration—scholars may be concerned that leadership deaths constitute a particularly strong treatment. Since leaders are (here, definitionally) the most powerful political actors, we might be tempted to think of their removal as a radical shift with a stronger effect than what we would observe for smaller, more common changes to the set of politicians bargaining over coalitions. However, as the formal example in Chapter 2 demonstrates, coalition structures are exceedingly fragile: even slight shifts can radically change the extant coalition structure. Nothing in my theoretical framework suggests that the likelihood of deconcentration correlates with the relative power of the player removed from coalitional bargaining. Thus, the results here are a faithful estimate of the expected

effect not just from leadership death, but rather of changes to the set of politicians bargaining over coalitions more generally.

Similarly, scholars may worry that countries are exposed to multiple treatments through a series of leadership deaths, biasing my estimate of the LATE. I do not think this is a concern for these data for two reasons. For one, the nature of the treatment is such that its expected effect does not increase with further deaths: because a single death is enough to destabilize coalition structures, another one does not add "more" destabilization. Additionally, receiving the treatment in one year does not affect the probability of receiving it in future years, since the deaths I study here are chosen for their exogeneity to political processes. Thus, I do not expect that multiple treatments will bias my estimate of the LATE.

In sum, these results provide strong evidence that politicians deconcentrate to manipulate coalitions. When the set of players is suddenly shocked by a leadership death, the likelihood of deconcentration in the immediate aftermath nearly doubles. More than just a useful tool for predicting deconcentration, my theory uncovers the causal origins of deconcentration.

4.2.2 Robustness

Scholars may be concerned that these estimates are sensitive to particular modeling decisions. I test a number of alternative specifications and estimators to determine the robustness of these results.

First, to further examine whether cross-sectional and temporal heterogeneity are driving the effect of leadership death on deconcentration, I re-estimate the models with slightly different controls. I study contemporaneous (post-treatment) measures of democracy and development, I use listwise deletion instead of multiple

imputation, I interact both controls with the treatment, and I study models with controls but no fixed effects. All of these specifications yield estimates substantively unchanged from those in Table 4.2. I then re-estimate the fixed effects models but cluster standard errors by country, year, and country and year; the treatment effect remains statistically significant in each model. I then investigate potential "neighborhood effects" which both the fixed effects and groupwise adjustments to standard errors may be unable to capture. While an ideal estimation strategy would model this spatial autocorrelation directly, the performance of such models for panel data with a binary dependent variable and a relatively short time series is unclear (Lee and Yu 2015; Ward and Gleditsch 2008). As a rough approximation, I instead re-estimate the baseline model with added region fixed effects. The estimated LATE remains unchanged.

A second potential concern is stark class imbalance in deconcentration. As discussed in Chapter 3, only about 4.6% of observations are coded as 1, with the rest 0s. On the other hand, despite being *relatively* rare, deconcentration is not absolutely rare, since it is observed in these data nearly 400 times. There is some debate over the circumstances under which absolute and/or relative rareness poses a problem: standard maximum likelihood estimates can suffer from bias not only from the "small sample" problem but also from issues relating to data complexity and small disjuncts (He and Garcia 2009; King and Zeng 2001). To examine whether such bias affects my estimates, I estimate the model with a "rare events correction" (Choirat et al. 2016), as well as a penalized maximum likelihood (PML) approach developed by Firth (1993). In both cases, I find that if anything, class imbalance biases the estimated effect of a leadership death downward. Since the magnitude of this bias is statistically and substantively inconsequential, I prefer standard maximum likelihood over these more computationally intensive estimators.

A related concern is that because there are so few leadership deaths in the data, the estimates may be sensitive to a few particular cases. To study this possibility, I use jackknife estimation, iteratively dropping each observation and re-estimating the model. The jackknifed LATEs all fall in the range $[0.50, 0.56]$, suggesting that the results are not reliant on any individual observations. I then jackknife by country, iteratively dropping all observations for each of the 189 countries in the data and re-estimating the model. This procedure yields estimates ranging over the interval $[0.44, 0.59]$. While this technique produces more varied estimates—as is expected when more observations are dropped simultaneously—nothing deviates substantially from the baseline result. Jackknifing by year also produces very similar estimates. In short, I find very little evidence of high-leverage observations, countries, or years which exert undue influence on the LATE.

The final modeling decision I investigate is whether these results are sensitive to the choice of two-year windows. To study this possibility, I construct three-, four- and five-year windows following leadership deaths using the same coding procedure as in the main analysis. (I do not construct windows shorter than two years, however, because such narrowness would exaggerate measurement error and censor the period of coalitional bargaining during which deconcentration is expected to occur, as discussed above.) Substituting each of the longer windows for the two-year windows used in the main analysis yields estimated effects of leader deaths on the probability of deconcentration of 0.31, 0.52^{*}, and 0.45^{*}, respectively. While there is some variation in these estimates, the overall picture remains the same: leadership deaths increase the probability of deconcentration.

Figure 4.2 presents a selection of these estimates in a "ropeladder" (Adolph 2012). Each point represents the mean difference in predicted probability of deconcentration as estimated by a given model, constructed from 1,000 draws from

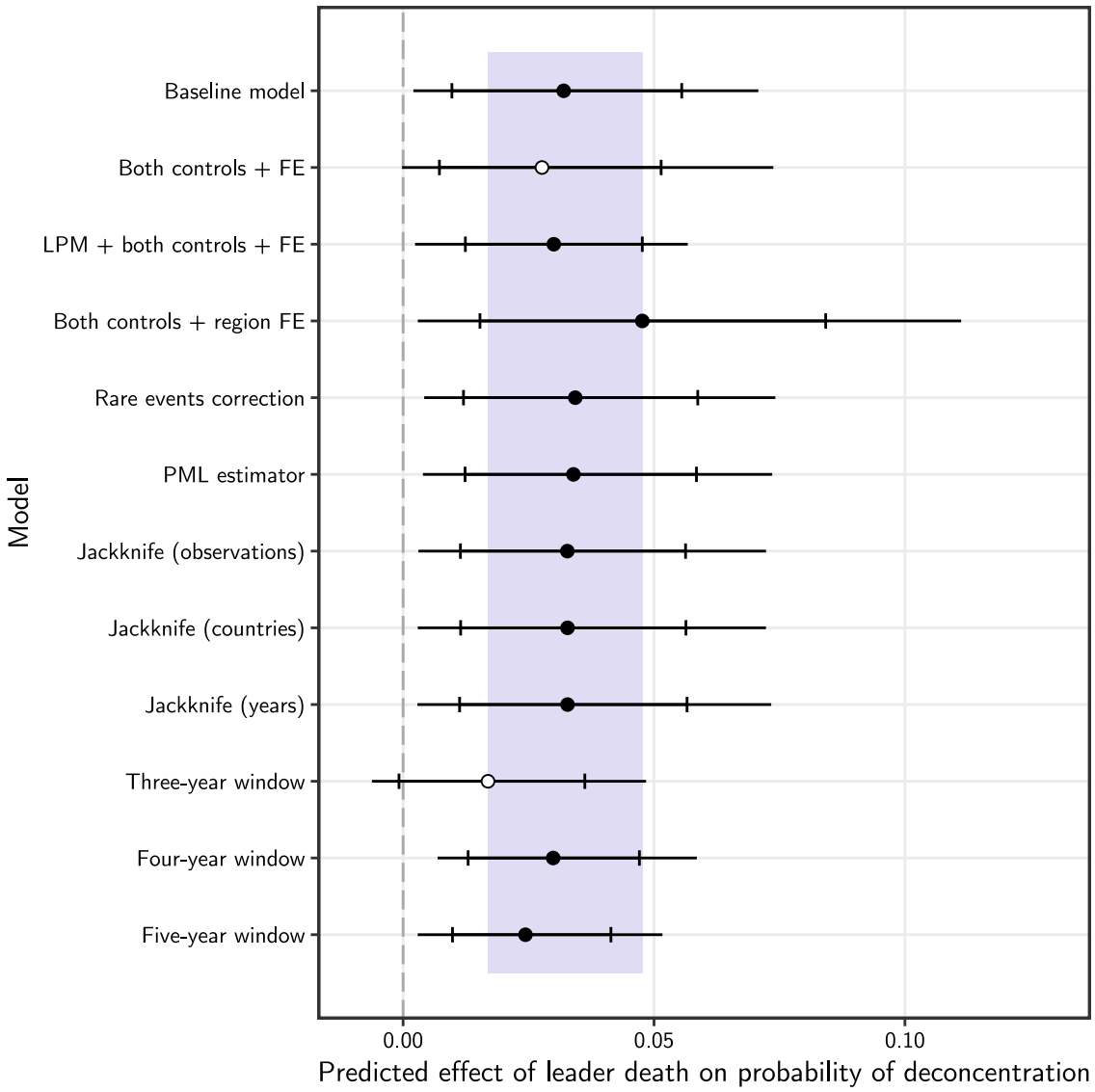


Figure 4.2: The effect of leadership death on deconcentration. Results are only shown for a selection of models. Estimates and standard errors are constructed from 1,000 draws from $\mathcal{N}_{\text{MV}}(\widehat{\boldsymbol{\beta}}, \mathbb{C}[\widehat{\boldsymbol{\beta}}])$. Points represent mean draws, with lines and ticks for 95% and 80% quantile intervals, respectively. Solid dots indicate estimates that are statistically significant at $p < .05$, with hollow dots for those that are not. Jackknife estimates are average means and quantiles. "LPM" refers to linear probability models and "FE" to fixed effects.

$\mathcal{N}_{\text{MV}}(\widehat{\boldsymbol{\beta}}, \mathbb{C}[\widehat{\boldsymbol{\beta}}])$, with lines (ticks) for 95% (80%) quantile intervals.[6] For jackknife specifications, I present means for each set of quantiles across all iterations of the jackknife procedure. The blue box highlights the range of point estimates across these specifications. As is evident, the estimated effect of a leadership death on the probability of deconcentration is stable around the 3 percentage point effect given by the baseline model, ranging over 1.7-4.8 percentage points. Given the observed probability of deconcentration of 4.6%, leadership deaths are expected to increase the probability of deconcentration by approximately 40-100%.[7]

4.2.3 Threats to inference

My identification strategy relies on apolitical leadership death as a plausibly exogenous treatment. The main threats to treating my estimates as causal are therefore (1) non-random assignment of the treatment and (2) anticipation effects (see, e.g., Dunning 2012). First, if leadership deaths are caused by the same forces that lead to deconcentration, then my results could simply be capturing the endogenous relationship between death, deconcentration, and the omitted variable(s). Second, if political elites can anticipate leadership deaths, then they may delay efforts to deconcentrate until after the head of state passes away, biasing upward the estimated effect of the death itself.

There is little evidence to suggest that non-random assignment is a problem. Recall that it is the *timing* of leadership deaths that is exogenous, and not necessarily death itself. A randomization failure would therefore require that the assignment

6. Here, $\mathcal{N}_{\text{MV}}$ indicates the multivariate normal distribution, $\widehat{\boldsymbol{\beta}}$ is the vector of fitted coefficients (including controls and fixed effects), and $\mathbb{C}[\widehat{\boldsymbol{\beta}}]$ is their covariance.
7. Across all models, estimates range over 1.7-10.8 percentage points, for a 37%-235% increase over the baseline probability of deconcentration. Figure 4.2 therefore presents somewhat conservative estimates of the LATE.

of the timing of a leadership death is related to the force driving deconcentration. Very few variables seem likely to affect both. One potential connection could be that more economically developed countries have better health care services and are therefore less likely to experience leadership deaths and, potentially, deconcentration. Yet regressing deaths on GDP per capita yields a null result ($\widehat{\beta} = -0.00, p \approx 0.13$), suggesting that deaths are *not* more common in poorer countries. This may be because rulers typically receive health care far exceeding that available to most of their country's citizens, even in the poorest countries. It is difficult, for instance, to imagine Mao Zedong suffering from China's underdeveloped health infrastructure, with his army of personal physicians (Li 1994). And many of the illness-related deaths in my data occur outside the country which the deceased ruled, such as Gabonese President Léon M'ba's death in Paris.[8] The anecdotal evidence suggests that a null relationship between GDP and the frequency of apolitical leadership death makes sense, and that the state of a country's healthcare infrastructure does not determine when leaders die in office.[9]

Turning now to anticipation effects: if political elites are simply waiting until after heads of state die to initiate reforms, then we should see much of the observed deconcentration essentially simultaneous to deaths. I have already attempted to guard against anticipation effects by excluding leaders coded as "retired due to illness" by Archigos. However, to test whether such anticipation is a problem, I re-code the post-death windows to *exclude* simultaneous deaths. Thus, to revisit the Mozambique example above, instead of coding 1986, 1987, and 1988 as 1 for Samora

8. Similarly, in the Guinean case motivating this chapter, Ahmed Sékou Touré died after traveling to the United States for heart surgery.
9. Nor does there appear to be much difference in leader age across regime types. Nondemocracies have younger leaders, but these differences are marginal (54 years old, on average, against 56 years old in democracies and hybrid regimes). Regressing leadership death on the interaction of leader age and either GDP or democracy yields null results, suggesting that leaders are no more likely to die young in poorer or less democratic countries.

Table 4.3: The effect of leadership death on deconcentration, excluding simultaneous cases

Death in previous two years	0.54†			
	(0.31)			
Death in previous three years		0.21		
		(0.29)		
Death in previous four years			0.50*	
			(0.23)	
Death in previous five years				0.42†
				(0.21)
BIC	3,027	3,029	3,025	3,026

$^{*}p < .05, ^{\dagger} p < .10$. Each model includes 8,089 observations across 189 countries.

Machel's death in 1986, the two-year window is coded as 1 only for 1987 and 1988 in this analysis. I then re-estimate the baseline model with two-year windows as well as the three-, four- and five-year windows considered above. Results from this alternative coding are presented in Table 4.3. The estimates are noisier but again substantively unchanged from the main analysis above, indicating that anticipation effects do not account for my results.

Finally, to strengthen the causal interpretation of the baseline result, I conduct two sets of placebo tests. I first estimate a series of models identical to Equations 4.1-4.2, but where instead of the treatment, deconcentration is regressed on placebo windows of varying lengths occurring prior to a leadership death, as well as after a leadership death but lagged by five years. Table 4.4 reports the results of these alternative windows. In no case is the estimated effect of the placebo window statistically distinguishable from zero. These results reinforce the interpretation that the death of a leader exerts a causal effect on the likelihood of deconcentration.

The second set of placebo tests examines whether leadership deaths lead to manipulation of other political institutions. The results presented so far may just

Table 4.4: Placebo tests for the effect of death on deconcentration

Model	$\widehat{\beta}$
Five year pre-death window	0.20
	(0.23)
Four year pre-death window	0.29
	(0.23)
Three year pre-death window	0.30
	(0.25)
Two year pre-death window	0.37
	(0.27)
Two year window, lagged five years	0.36
	(0.29)
Three year window, lagged five years	0.36
	(0.26)
Four year window, lagged five years	0.19
	(0.25)
Five year window, lagged five years	0.17
	(0.23)

$^{*}p < .05,^{\dagger}\ p < .10$. Each model is estimated separately. N > 7,000 for each regression.

reflect general political maneuvering that occurs during the heightened uncertainty following a leadership death. To investigate whether deconcentration is specifically tied to coalition bargaining, and not just garden-variety institutional engineering, I regress a series of indicators for alternative institutional changes on the two-year post-death window from the baseline model. These indicators are: changes to veto players, the powers of subnational government units, term limits, proportional representation, and presidentialism, all from the Database of Political Institutions (Beck et al. 2001); enacting new constitutions, amending them, or suspending them, all from the Comparative Constitutions Project (Elkins, Ginsburg, and Melton 2009); and large changes (i.e., of three or more points) in the Polity scores used above. Results from these models indicate that leadership deaths do not increase

the probability of any such institutional changes. That all of these effects are null suggests that my estimates are not just picking up institutional instability following leadership death, but rather are capturing the specific role of deconcentration as a means to manipulate coalition bargaining.

4.3 Checking the mechanism

The findings presented in this chapter indicate that sudden shocks to coalition bargaining lead to more deconcentration. If, as I have argued, the causal mechanism is that politicians' preferences over coalitions induce preferences over deconcentration, then we should also observe coalitional realignments following leadership death, and in particular, following deconcentration. Figure 4.3 maps these causal links. The foregoing analysis has focused on establishing (1). I now study linkage (2), the effect of deconcentration on eventual coalitional changes, to verify that coalitional considerations are the driving force. However, since deconcentration and coalition changes are clearly endogenous, I also study the direct link from death to coalitional outcomes in (3), which should not suffer from endogeneity bias due to the exogenous timing of the deaths.

The ideal data to test these links would account for all changes to coalition structures—including factions within dominant parties and cross-party coalitions. Since these data do not exist, I take a narrower approach using changes to the size of the cabinet. Such shifts are definitionally a subset of all cabinet shake-ups, which are in turn a good indicator of changes to the membership of the ruling coalition, since ministerial appointments are used to reward allies and punish defectors (Bäck, Debus, and Dumont 2010). Variation in the ruling coalition is itself a subset of all variation in the coalition structure. Changes to the cabinet size should therefore

correlate to coalitional realignments more broadly.

Data on cabinet size come from the Cross-National Time Series Data Archive (CNTS; Banks and Wilson 2013). As discussed in Chapter 3, CNTS provides a count of the number of cabinet-level members for each country-year. Since my theory does not suggest the direction of this change—whether the number of politicians in the ruling coalition will increase or decrease following deconcentration—I code a binary indicator for any change occurring. I expect that leadership death and deconcentration both increase the probability of a change in the number of cabinet members.

To test this hypothesis, I first estimate a series of models of the form

$$y_{i,t} \sim \mathcal{F}(\pi_{i,t}), \tag{4.3}$$

$$\pi_{i,t} = \text{logit}^{-1}(\alpha + \beta_0 x_{i,t} + \beta_1 z_{i,t}), \tag{4.4}$$

where the notation is the same as in Equations 4.1-4.2, except the outcome y is an indicator for any change in the number of cabinet members and z is a binary indicator for two-year "deconcentration windows" (constructed equivalently to

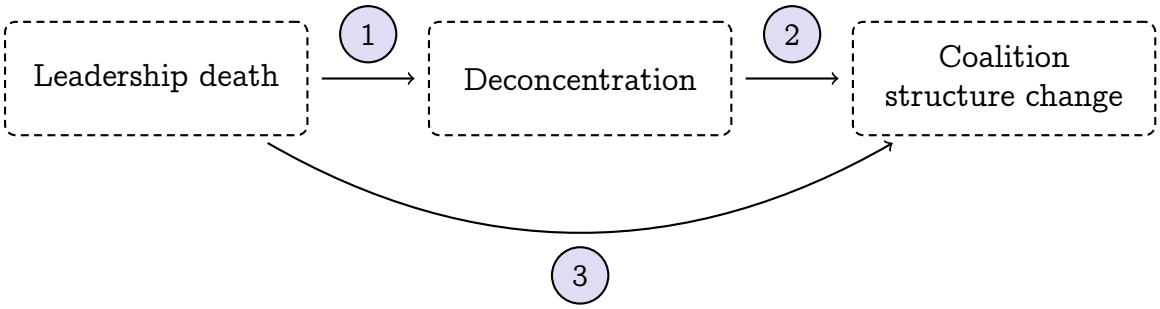


Figure 4.3: Causal structure. Exogenously-timed leadership deaths destabilize extant coalition structures. Politicians then use deconcentration to push coalitional bargaining toward a more favorable new equilibrium.

treatment windows as defined above). Positive values for β_0 and β_1 indicate support for linkages (3) and (2), respectively. Further, instead of effecting change in general, deconcentration and death might be expected to lead to larger changes in the number of cabinet ministers than are observed under normal politics. I therefore estimate a second series of models of the form

$$|\Delta y_{i,t}| \sim \mathcal{F}_{NB}(\mu_{i,t}, \lambda_{i,t}), \tag{4.5}$$

$$\mu_{i,t} = (\alpha + \beta_0 x_{i,t} + \beta_1 z_{i,t}), \tag{4.6}$$

where y is the number of cabinet ministers and $|\Delta y_{i,t}| = |y_{i,t} - y_{i,t-1}|$. Because the dependent variable is positive and integer-valued, I use $\mathcal{F}_{NB}$, the negative binomial distribution with dispersion parameter λ. Finally, it may be possible that changes to the ruling coalition are reflected not in the size of the cabinet but rather in who holds ministerial positions. To study this possibility, I use data on major cabinet changes, which CNTS defines as the simultaneous replacement of more than 50% of cabinet ministers. Since CNTS counts the number of major changes in each country-year, I estimate a second negative binomial model given by

$$y_{i,t} \sim \mathcal{F}_{NB}(\mu_{i,t}, \lambda_{i,t}), \tag{4.7}$$

$$\mu_{i,t} = (\alpha + \beta_0 x_{i,t} + \beta_1 z_{i,t}). \tag{4.8}$$

Again I expect $\beta_0, \beta_1 > 0$, so that both death and deconcentration make changes in the coalition structure more likely.

Results for these models are reported in Table 4.5. These estimates indicate broad support for causal links (2) and (3). Whether modeled separately or together, deconcentration and leadership death notably increase the probability of a change

Table 4.5: The effect of deconcentration and death on coalition structures

	Change in cabinet size (binary)		
	(1)	(2)	(3)
Deconcentration	0.37*		0.37*
	(0.07)		(0.07)
Leader death		0.31*	0.29*
		(0.13)	(0.14)
	Change in cabinet size (count)		
	(4)	(5)	(6)
Deconcentration	0.27*		0.27*
	(0.06)		(0.06)
Leader death		0.21†	0.19†
		(0.11)	(0.11)
	Major cabinet turnover (count)		
	(7)	(8)	(9)
Deconcentration	0.09†		0.09†
	(0.05)		(0.05)
Leader death		0.41*	0.41*
		(0.08)	(0.08)

*$p < .05$,† $p < .10$. $N > 7{,}500$ for each regression.

in cabinet size, generate larger changes in cabinet size, and lead to more frequent major cabinet turnover.

However, because deconcentration and coalitional change are endogenous, these results do not indicate if creating more subnational units is the mechanism driving changes in coalition structures, or if it is just the shock of a leadership death itself doing the work. To examine this question, I re-estimate all of these models with an added multiplicative interaction term, $x_{i,t}z_{i,t}$. If deconcentration is responsible for coalition structure change, then we would expect $\beta_1 \approx 0$, with β_0 and the estimated interaction effect both positive. I also adopt an instrumental

variables strategy, first regressing deconcentration on leadership death and then regressing my measures of coalitional change on the fitted values from the first stage. Given that the instrument, the endogenous regressor, and the outcome are all binary, the canonical two-stage least-squares (2SLS) estimator is unsuited to these data. Yet 2SLS will be inconsistent if the first-stage regression is modeled as a nonlinear function (e.g., a logistic regression, as in the main analysis). Given these challenges, I estimate the first-stage regression as a linear probability model (LPM), which preserves the asymptotic properties of 2SLS while typically recovering estimates quite close to a more realistic, nonlinear first-stage model (Angrist and Pischke 2009; Wooldridge 2010).

The results from these models are presented in Table 4.6. The evidence is somewhat mixed, but broadly supports deconcentration as the mechanism producing change in coalition structures. Deconcentration remains significant in all models, and for two of the dependent variables, the effect of deconcentration is larger than that of death, as is evident in the top and middle panels of Figure 4.4. Including both covariates in a model of cabinet size attenuates the effect of death but leaves that of deconcentration essentially unchanged. This suggests that deconcentration exerts its own effect on coalition structure change, independent from that of leadership death. However, when studying cabinet turnover, the effect of death on turnover is much stronger than that of deconcentration. And across all of the dependent variables, none of the interaction terms are distinguishable from zero. The instrumental-variables strategy produces more encouraging results, with both the first- and second-stage results all indicating that deconcentration generates greater cabinet change.

Interpreting these results at face value is difficult for three reasons: all of these models are non-linear, the treatment and mediator are both binary, and all covariates exhibit stark class imbalance—all of which trouble causal mediation analysis (Imai,

Table 4.6: Testing the deconcentration mechanism

	Change in cabinet size (binary)		
	(1)	(2SLS)	
Deconcentration	0.36*		2.91*
	(0.07)		(1.24)
Leader death	0.24	0.07*	
	(0.15)	(0.02)	
Deconcentration × death	0.29		
	(0.37)		
	Change in cabinet size (count)		
	(3)	(2SLS)	
Deconcentration	0.27*		3.05†
	(0.06)		(1.59)
Leader death	0.21†	0.07*	
	(0.12)	(0.02)	
Deconcentration × death	−0.10		
	(0.27)		
	Major cabinet turnover (count)		
	(5)	(2SLS)	
Deconcentration	0.10*		6.12*
	(0.05)		(1.24)
Leader death	0.46*	0.07*	
	(0.09)	(0.02)	
Deconcentration × death	−0.31		
	(0.23)		

*$p < .05$, †$p < .10$. $N > 7{,}500$ for each regression. The first-stage regression, in the middle column, is a linear probability model. The second-stage regressions are probit and negative binomial models for the binary and integer-valued outcomes, respectively. The F-statistic for the first-stage regression is 9.67, indicating a potentially weak instrument—as expected when both the instrument and endogenous regressor are binary.

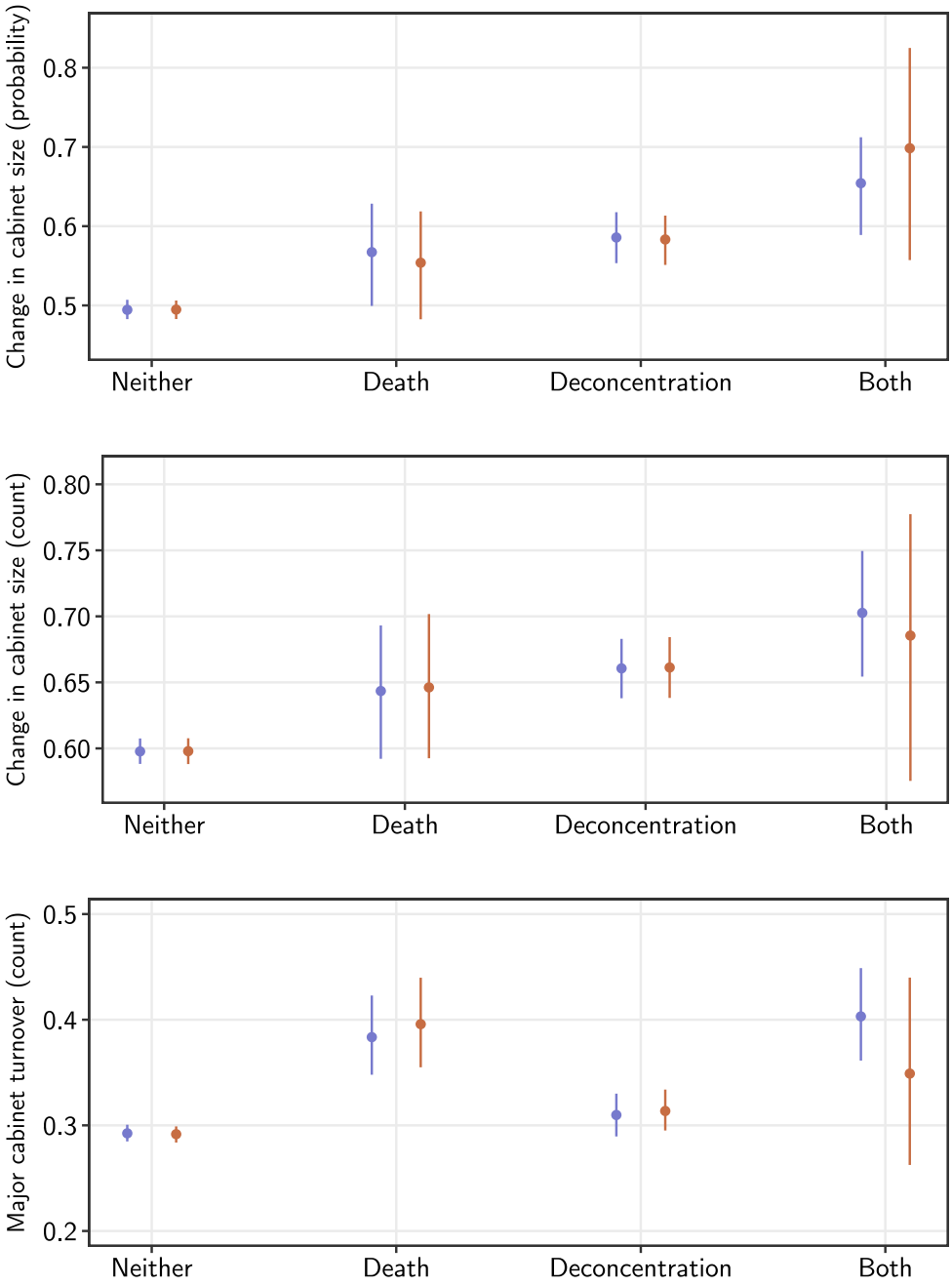


Figure 4.4: The effect of death and deconcentration on coalition structures. Each panel plots the predicted effects of death and deconcentration on changes to cabinet size or composition. Blue indicates predictions from models with both covariates combined additively, while orange indicates models interacting the two.

Keele, and Tingley 2010; Imai, Keele, and Yamamoto 2010). Given these challenges, it is difficult to adjudicate the relative contributions of death and deconcentration to coalition structure change. However, the broad picture is one in which both covariates make such changes more likely. Taken together with the main results above, it is clear that leadership death causes a notable increase in deconcentration, which then carries forward through to changes in coalition structures, as proxied by changes to the cabinet. While not dispositive, the evidence indicates that coalitional bargaining is the mechanism driving deconcentration.

4.4 Conclusion

A key implication of my theory of deconcentration is that shocks to the coalition bargaining environment can precipitate deconcentration by pushing coalitions out of equilibrium. When a politician is suddenly removed from the bargaining process, the existing coalition structure is destabilized, providing an opportunity for the remaining politicians to use deconcentration to push the bargaining toward a more favorable new coalition structure. Apolitical leadership deaths provide clear evidence of this implication of the model: when heads of state and government die unexpectedly, the probability of deconcentration over the following two years increases by 3.2 percentage points, or 71%. By using deaths for which the timing is random, I ensure that this relationship is causal and not just the artifact of underlying socioeconomic conditions.

Measurement error and class imbalance in both the dependent and independent variables present difficulties for an otherwise straightforward model. Nevertheless, these estimates are robust to a variety of alternative modeling strategies. In addition, my identification strategy does not appear to be threatened by non-random

assignment or anticipation effects, and a battery of placebo tests indicates that they are not just artifacts of noisy data. Finally, both deconcentration and leadership death both lead to more changes in coalition structures, providing further support for the mechanism identified in my theoretical framework.

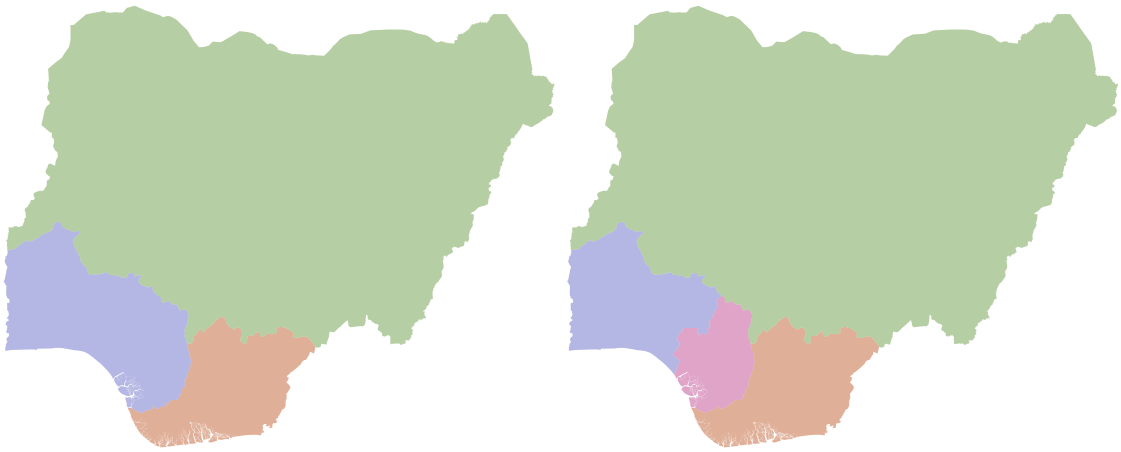
This analysis provides compelling evidence of the link between deconcentration and coalition manipulation. It is clear that an exogenous shock to coalitional bargaining increases the probability of deconcentration. Yet most deconcentration occurs in the absence of such a shock. Do these results hold outside the narrow context of the instability induced by leadership death? Do coalitional considerations determine deconcentration even in periods of "politics as usual"? To answer these questions, I now turn to a close comparison of three statehood movements in early post-independence Nigeria.

5 | Deconcentration in Nigeria

> It seems to me wrong that when Britain has done so much to create Nigeria as a country, she should, just before giving us independence, cut the country into little bits. There are countless tribes, and it is impossible to find any area which is homogeneous. If there is to be a new Region, this would be a very difficult task over the setting of boundaries. There would always be groups which would be a minority. There would be great trouble and bloodshed and it would postpone independence by 7 years. I am all against additional Regions.
>
> —Nigerian Prime Minister Sir Abubakar Tafawa Balewa, in a private meeting with Sir Henry Willink, Chairman of the Minorities Commission of the United Kingdom, 26 November 1957 (TNA, CO 957/1, 11)

My theory suggests that deconcentration is the result of politicians trying to manipulate political coalitions. In the foregoing chapters, I have demonstrated the strategic logic underpinning deconcentration as coalition manipulation, and shown how the empirical implications of this theory are consistent with the patterns evident in a global dataset. In this chapter, I trace the coalition manipulation mechanism through a case study to further test this argument.

My focus is on Nigeria over the 1946-1966 period. These two decades encapsulate the beginning of regional party politics under the "Richards" Constitution of 1946, the end of British colonial rule in October 1960, and the entire life of the First Republic, which fell in the military coup of January 1966. Over these years the Eastern, Northern, and Western Regions grew up as units of government alongside the parties that called them home, generating significant inter-regional conflicts. As successive constitutions in 1951, 1954, and 1957 entrenched the three



1960-1963 1963-1967

Figure 5.1: Regions in Nigeria's First Republic. Maps of Nigeria's regional boundaries before and after deconcentration in the First Republic. Orange is the Eastern Region, green is the Northern Region, blue is the Western Region, and purple is the Mid-West Region.

regions' dominance in soon-to-be-independent Nigeria, minority communities in each region lobbied the colonial administration for deconcentration. These "statehood movements" were so large as to prompt the British to form an official Commission, led by Sir Henry Willink, to "Enquire into the Fears of Minorities and the Means of Allaying Them" (Colonial Office 1958), yet the tripartite federation remained at independence. Only in 1963 did a new region emerge, when the Mid-West was carved out of the Western Region. It was, and remains, the only episode of deconcentration under civilian government in independent Nigeria.

In this chapter, I conduct a close case comparison of the three most prominent statehood movements during this period: those of the proposed Mid-West, Calabar-Ogoja-Rivers (COR), and Middle Belt regions.[1] (Following the convention in

1. Other statehood movements were significantly less prominent. For instance, the Rivers State movement was led by Rivers Chiefs who argued that treaties signed between 1884 and 1888 granted

Nigerian political writings of the era, I use "state" and "region" interchangeably throughout this chapter.) Unlike the Mid-West, neither the COR nor the Middle Belt region ever came to be.[2] Focusing on three movements across this twenty-year period allows me to make two important comparisons. First, by studying one successful and two "failed" movements in the same institutional setting, I am able to hold contextual factors constant and isolate the role of coalitional considerations in determining deconcentration outcomes. Second, by studying these movements' prospects before and after independence, I am able to examine change in the institutional context, and particularly those rules governing coalitional bargaining, as identified in the theoretical exposition in Chapter 2.

Though this case comparison provides a close look at the causes of state creation in Nigeria, it can only capture one of the many shapes which deconcentration can take. As I demonstrate below, Nigeria's two largest parties—the Northern People's Congress and the National Council of Nigeria and the Cameroons, led by Sir Abubakar Tafawa Balewa and Nnamdi Azikiwe, respectively—created the Mid-West to divide the country's largest opposition party, Obafemi Awolowo's Action Group. This episode indicates how politicians can use deconcentration to split opposition parties—how they can "divide to rule." However, the lessons of this chapter may be less instructive for understanding deconcentration that reshapes the governing coalition, either by broadening it or by trimming it down closer to minimum-winning

---

them legal rights to self-government, which the British summarily denied. Similarly, the Lagos Separate State Movement rejected both the establishment of Lagos as a federal territory and its reintegration into the Western Region at constitutional talks in 1957, but its membership was small and short-lived. Also under constant debate was the Ilorin irredentism movement, seeking to "return" Ilorin and Kabba provinces to the Western Region from the North (TNA, CO 879/164, 49-50, 132-138).

2. Since 1966 Nigeria has undergone successive deconcentration. The country now consists of 36 states and the Federal Capital Territory of Abuja. While neither the COR nor the Middle Belt regions (as proposed over the period I study here) were ever created, the minority elements that advocated these regions have won deconcentration in various forms (Suberu 2001).

size. While later Nigerian deconcentration (including the Abacha-era state creation highlighted in Chapter 1) resonates with these other coalitional realignments, such cases are beyond the scope of this chapter.

The data for this case study is drawn from four months of archival research at the National Archives (TNA) of the United Kingdom in 2017 and 2018. I examined all available records relating to Nigerian politics over this period, drawing from the files of the Prime Minister, Cabinet, Foreign, and what were the Colonial and Commonwealth Relations Offices.[3] These records provide incredibly detailed insights into partisan and institutional struggles over the entire period in question, particularly as the officers writing them were eyewitnesses to many of the key discussions among British and Nigerian actors, even after independence. Although there are some threats to inference from relying on British sources that must be addressed—discussed in greater detail below—these primary source provide a nuanced view of all three statehood movements' fates.[4]

5.1 Parties-in-regions, parties-as-regions

5.1.1 Historical setting

By the end of the nineteenth century, Nigeria was comprised of two distinct regions. In the north, the *jihad* of the Shehu Usman dan Fodio had established rigidly hierarchical rule based around the twin pillars of a conservative ideology steeped in Islamic revival and a ruling aristocracy drawn from Hausa-speaking Fulani.[5]

3. Following the TNA citation guidelines, these offices are abbreviated as PREM, CAB, FCO, and DO (for Dominions Office, containing records for both the Colonial and Commonwealth Relations Offices).
4. Photographs of all original documents cited in this chapter are available upon request.
5. "Shehu" is an honorific in the Hausa language, equivalent to "Sheikh" in Arabic, and is the name by which dan Fodio is commonly known in northern Nigeria (Kendhammer 2010).

At its height in the 1830s, the Sokoto Caliphate—named after the city which dan Fodio made its capital—was one of the largest states in West African history (Falola and Heaton 2008, 65). However, its size and the heterogeneity of its subjects generated immense administrative challenges. After dan Fodio's death in 1817, his son transformed the Caliphate into a fully-fledged system of decentralized government: emirs drawn from the aristocratic *sarakuna* class ruled over emirates that paid tribute to, and took instruction from, Sokoto (Kendhammer 2010; Paden 1973; Vaughan 2016). In contrast, in the south, the collapse of the Oyo Empire in 1833 sparked 60 years of civil war and fragmented rule among the Yoruba. When British merchants arrived in mid-century, they found willing trading partners in the Yoruba *obas* (kings or "traditional rulers") seeking an advantage in the many localized conflicts. To the east, in the delta of the Niger River, government was similarly fragmented as local leaders vied to control the lucrative palm oil trade (Falola and Heaton 2008).

British conquest accelerated when European powers awarded the United Kingdom the sole claim to these areas at the Berlin Conference in 1885. The Protectorates of Northern and Southern Nigeria were established in name in 1900, and in reality when the British forces marched on Sokoto in 1903, killing the Emir and ending the Caliphate. By 1906, the Colony of Lagos was folded into the Southern Protectorate, and in 1914 the entire region was merged into the single Colony and Protectorate of Nigeria, with the capital in Lagos and Lord Frederick Lugard as Governor-General.

Lugard had earned his spurs by leading the conquest of Sokoto as Governor of the North (1900-1906), and admired the decentralized stability of the Caliphate. To shore up control over the fragmented south, he transplanted the administrative structure he found in the old Caliphate, empowering Yoruba *obas* in the south-west

and creating new traditional rulers to govern the (mostly) Igbo communities of the south-east. This patchwork system of local authority became known as "indirect rule." The nascent colonial government underwent a major overhaul in 1939, when the provinces of Southern Nigeria were divided into two groups that reflected their divergent institutional legacies: the West, covering most of Yorubaland, and the East. For the first time, Nigeria was officially organized into Northern, Western, and Eastern Regions.

5.1.2 Regionalized parties

These administrative divisions became political entities in the 1946 Richards Constitution (named for the Governor-General), which created regional legislatures. Although non-elective and still dominated by British officials, these assemblies represented the first sites of power for nascent Nigerian nationalist movements to contest. By making partisan competition regional, the British spurred the development of regional parties.

To be sure, political parties existed well before the Richards Constitution, in spirit if not in name. Nationalist activity can be traced to 1908, and when the "Clifford" Constitution of 1922 granted the first semblance of voting rights to residents of Lagos, partisan organization began in earnest. Proto-nationalist associations emerged in the mid-1920s, consisting of large coalitions of market traders, young radicals, the growing professional class, and *obas* not content with their middle-management role in the system of indirect rule. But it was not until the 1940s that leading Lagosian politicians called for a National Council to form a united front in fighting for independence. These demands led, in 1945, to the formation of the National Council of Nigeria and the Cameroons (NCNC).[6] Veteran

6. "The Cameroons" were administered as part of Nigeria out of administrative convenience. When

operative Nnamdi Azikiwe ("Zik"), a Northern-born and Western-educated Igbo, was elected Secretary-General. Under Zik's stewardship the NCNC immediately set to pressuring the British for greater devolution of power to Nigerians (Gould 2013; Sklar 1963).

Zik's primary objective was independence, but since that possibility seemed remote, in 1946 the NCNC set its sights on a much more attainable goal: direct elections to regional legislatures which had real authority. This demand grew directly out of the shortcomings of the newly-promulgated Richards Constitution, which met neither condition. Members of the regional legislatures were delegates from the unelected Native Administrations, the key local government bodies in the indirect rule system. Further, the legislatures' primary purpose was not to legislate directly, but rather to send delegates to a central Legislative Council, itself dominated by the colonial administrators. Far from modernizing indirect rule, the regional legislatures of the Richards Constitution entrenched it, and so served as a focal point of NCNC attacks (Sklar 1963; Vaughan 2000). This agitation proved decisive, kick-starting a constitutional review in 1948.

Yet as the NCNC came to dominate the independence movement, Yoruba leaders grew wary of Zik's rising star and the corresponding influence of Igbo-interest organizations on national politics (TNA, CO 879/192, 7-8). In 1948, Lagosian politician (and former Zik ally) Obafemi Awolowo formally inaugurated a pan-Yoruba alliance, *Egbe Omo Oduduwa* (Society for the Descendants of Oduduwa, the mythical progenitor of the Yoruba). Espousing Yoruba unity through shared cultural heritage, *Egbe* quickly won much of the support of the *obas* in the West that the NCNC and its local allies had previously enjoyed. In 1949, Zik unsuccessfully used the levers of

---

Nigeria attained independence, a plebiscite was held to determine the status of the territory. Northern Cameroonians voted to remain in Nigeria as part of the Northern Region, while Southern Cameroonians voted to join the new country called the Cameroon Republic.

NCNC influence to attempt to block the accession of the new *Oba* of Lagos, an old foe from Lagosian youth politics in the 1930s. The affair stank of Igbo influence in Yoruba internal affairs, and the resentful (and very influential) *Oba* quietly set about working against Zik. NCNC influence in the West ebbed and never truly recovered (Sklar 1963).

What had been a unified national front in the NCNC quickly evolved into a ethno-regional rivalry, with *Egbe* speaking for the West and the NCNC for the East. That the NCNC's member organizations continued to add Igbo-interest groups (such as the Igbo Federal Union) only accelerated this regionalization. By the time the Action Group (AG) emerged as the official "political wing" of *Egbe* in 1951, regionalism constituted a core tenet of its ideology (Akinola 2014, 12-13).

With the constitutional review underway and the tide of regionalism decidedly rising in Nigeria, Northern leaders saw the writing on the wall. Emirate rulers became wary of a political transition led by "southern radicals," while the educated middle class of the North, small though it was, feared permanent second-class status in a federation dominated by the wealthier south. In 1948, this shared fear of southern overrule led an admixture of emirate leaders and conservative public servants to announce the first broad-based political party in the North. After a few early rebrandings, the party took the name Northern People's Congress (NPC, or *Jam'iyyar Mutanen Arewa* in Hausa).

The NPC's central mission was to reconstruct and extend the conservative Islamic rule of the Sokoto Caliphate, and in particular, to resist southern attempts to erode this tradition in the name of national unity. Notably, the foremost patron of the NPC at its founding was Alhaji Sir Ahmadu Bello, the Sardauna (literally, "War Chief") of Sokoto. A great-great-grandson of Usman dan Fodio, the Sardauna appeared to all observers as intent on building on the legacy of the Shehu's *jihad* by

re-establishing the core features of Caliphate rule in an independent Nigeria. For instance, Whitaker (1970, 348-349) writes that the Sardauna

> chose March 15th, anniversary of the fall of the Sokoto forces to Lugard's troops in 1903, as the day on which self-government was granted to Northern Nigeria... [His] speeches frequently touched on the past glories of the Fulani empire, the desirability of preserving tradition, and the religious sanctification and duties underlying political office.

While the Sardauna spoke for the the *sarakuna*, his lieutenant Abubakar Tafawa Balewa, a former teacher whose father had been born a slave, represented the *talakawa* (commoner) class. Under their leadership, the NPC reasserted a form of Northern rule that matched the North's historical identity. The NPC motto was "One North: One People Irrespective of Religion, Rank, or Tribe," but it was clear from the first days of the party that this meant One North in the image of the Caliphate.

5.1.3 "Virtually made to collapse"

The constitutional review that had begun in 1948 reached its conclusion in 1950. Nigerian representatives to the conference—which would produce the 1951 "Macpherson" Constitution—were firmly organized into ethno-regional blocs led by the AG, NCNC, and NPC. Negotiations centered on the extent of federalism: AG and NCNC leaders pushed for greater regional autonomy for fear of domination by the North, while advocating distribution of resource rents on a "derivation" basis to ensure that the south received the fruits from its profitable cocoa and palm production. In contrast, the Northern leaders pushed for a strong central government due to their belief in their ability to dominate it, the result of the North's much larger population.

The document that emerged was a compromise between these interests, introducing the first steps toward federal government in Nigeria. Four central ministers were to be appointed by the Governor from among the representatives of each region, while members of the central House of Representatives were elected from the regional Houses of Assembly. The Regions themselves were governed by elected, majority-Nigerian executive committees. This arrangement generated extreme centrifugal forces: the regions were formally subservient to the center, but central representatives were chosen by the regions. These incentives moved the most powerful party leaders to stay in their regions and send deputies to the center, preferring regional policy autonomy over national compromise (Akinola 2014; TNA, CO 879/192, 158-163). As Sklar (1963, 118) writes, it was an arrangement "of the kind that invites trouble and is virtually made to collapse."

These centrifugal forces brought the North to the brink of secession in 1953. Representing the comparatively more developed southern regions, the AG and NCNC together tabled a motion for self-government within three years. The Sardauna offered an amendment to replace this timetable with the phrase "as soon as practicable," leading to a walkout and the resignation of AG and NCNC ministers. The Sardauna rose to adjourn the legislative session and spoke only: "The mistake of 1914 has come to light and I should like to go no further" (Sklar 1963, 128). Over the next few days, Northern leaders were heckled in the streets of Lagos. Incensed, they released a public plan of action amounting to an intent to secede. Rioting in Kano followed, and the country nearly disintegrated (TNA, CO 879/192, 26).

The British immediately scheduled another constitutional renegotiation, signaling an end of the Macpherson experiment. In particular, outgoing Secretary of State for the Colonies, Oliver Lyttleton, believed that NPC demands would have to be met in order to preserve the unity of Nigeria. In a memo drafted for the Cabinet

in 1953 (TNA, CAB 129/61/4), he concluded:

> I am clear that the present cumbersome constitution will have to be radically revised and that the best hope of preserving the unity of the territory will lie in seeking to reach agreement on some modified and looser form of association at the Centre... such as to satisfy the North that their interests were effectively safeguarded.

Beyond their interest in Nigerian unity, however, the British also tended toward Northern appeasement due to the Region's numerical superiority, as well as the commitment to conservativism its leaders shared.[7] For example, in a second memo (TNA, CAB 129/62/35), Lyttleton added: "We cannot let the North down. They are more than half the population, more attached to the British and more trustful of the Colonial Service than the other two" regions. Finally, Lyttleton and other British administrators enjoyed better interpersonal relationships with the Northern leaders, the vast majority of whom had backgrounds in the civil service or colonial government (Sklar 1963). This predisposition is evident in Lyttleton's comparison of the major party leaders to his Cabinet colleagues (TNA, CAB 129/62/35):

> The Northern spokesman is Abubakar [Tafawa] Balewa—his soubriquet "the golden voice of the North"—a natural orator with real eloquence, nearly perfect English and all the arts of modulation at his command. The East is represented by Dr. Azikiwe, an American-educated demagogue, who makes few interventions but is able, persuasive, apparently moderate and wholly unreliable; the West by Mr. Awolowo with a flood of words always on the same wave length, a lawyer with considerable legal

7. British Conservatives in the Colonial Office were eager to draw comparisons between the "arrogance" of southern nationalists and their Labour opponents in Westminster. For example, according to the minutes of a Cabinet meeting on 27 May 1953, Lyttleton told an approving Prime Minister—Winston Churchill—that "the collapse of the cumbrous constitution under the auspices of the Labour Government had provided an opportunity of according a larger measure of autonomy to the 14 million Moslem [sic] inhabitants of the Northern provinces who were more favourably disposed to this country than the Southern Nigerians" who espoused radical nationalism (TNA, CAB 128/26).

> equipment and outstanding ambition. He is no more reliable than his Eastern colleague.

These interpersonal attachments almost certainly colored British receptiveness to each party's demands and furthered the shift toward loose federation in the late colonial period.

5.1.4 The last days of empire

The document that emerged from the 1953 talks, the Lyttleton Constitution of 1954, would carry Nigeria through to independence (after amendment in 1957-1958). In part due to British sympathy for Northern interests, the regions reached their peak vis-à-vis the center. Regions were given complete authority to devise their own legislative and executive institutions. In consequence, the AG, NCNC, and NPC were granted unprecedented autonomy to rule the regions as their fiefdoms, building permanent majorities for themselves. This feature did not go unnoticed by the few actors at the 1953 talks not allied to these parties: before withdrawing from the talks, the delegation from the National Independence Party (NIP) declared that

> [to] give the Regions carte blanche authority in respect of these matters is in effect the creation of Regions with complete autonomy, which is bound to lead to the division of Nigeria into three separate countries with no protection for minority groups (TNA, CO 879/159, 95).

The regions were so prominent in the Lyttleton Constitution that the East and West attained internal self-government in 1957, and the North in 1959—well before Nigeria obtained independence on 1 October 1960 (TNA, CO 879/192, 158-161). Slight revisions in 1957 and 1958 produced an Independence Constitution, enacted as an Order in Council of the United Kingdom, that left regional supremacy essentially

unchallenged. Balewa's lament in 1953 remained apt seven years later (TNA, CO 879/159, 27):

> There is not one genuine political party in Nigeria today. All parties are regional and, to a certain extent, tribal in character. This statement does not need any clarification as 95% of the membership of the Action Group, NCNC and NPC are Yorubas, Ibos [sic] and Hausas respectively, and their activities are mainly confined to the Western, Eastern and Northern Regions respectively.

At independence, Nigeria had a ceremonial President as head of state, and a Prime Minister chosen from the federal House of Representatives who acted as head of government. The regions had separate constitutions and so legislative structures varied, but each was run by a Premier. Beyond their responsibilities as regional executives, Premiers appointed officials to a federal Senate, operating much like the British House of Lords. The federal government was given enumerated powers covering such matters as external affairs, immigration, defense, and trade. Concurrent powers, shared by the federal and regional governments, included welfare, higher education, and prison management, with federal law supreme in the event of legal conflicts. All residual powers were left to the regions, including health, socioeconomic development, and agriculture, constituting the majority of Nigerian economic activity (Sklar 1963; Suberu 2001). Below the regions were provinces, districts, and villages, variously administered according to the regional constitutions and prone to frequent revision in the fluid political environment of the First Republic.

Although equal on paper, the three regions were considerably mismatched. The North accounted for over 75% of the country's geographic area and 50% of its population—over a quarter of British Africa's 60 million subjects (TNA, CAB 129/62/35). Yet the more densely-populated southern regions accounted for the majority of economic activity: while the principal exports in all regions were

Table 5.1: Nigeria's regions at independence

	East	North	West
Seats in House of Representatives	73	174	62
Area (thousands sq. kilometers)	78.11	721.11	111.82
Population (millions)	7.22	16.84	6.09
English literacy rate (%)	15.60	2.10	17.30
Revenue (£, millions)	4.57	4.56	7.30

Geographic area is approximate. Population data are from the last colonial census, published in 1953. Revenue data are from 1955, the last year available prior to independence. Lagos data are omitted. Data are from GADM (2018), Sklar (1963), Suberu (2001), and TNA (CO 879/164).

agricultural, those of the south tended toward high-value cash crops in cocoa and palm, and those in the North were the less-profitable groundnuts and hides (TNA, CO 879/192, 24). As a result of this disparity of wealth, the majority of Nigeria's mercantile and administrative class was drawn from the East and the West.

The Constitution specified an arduous and lengthy process for deconcentration. Creating another region would require an act approved in both houses of the federal legislature, each with a two-thirds majority. Identical acts would then have to be passed by a simple majority in each house of a majority of regions, among which the "rump" region that was to lose territory must be included. A plebiscite would then be held in the proposed region, which needed to pass with at least 60% of the vote. Finally, the result of the plebiscite would trigger a ratification vote in the regional legislatures, among which a majority needed to pass. The Constitution's voting rule was thus very strict, engineered to promote regional stability and discourage deconcentration.

5.2 "Fissiparous tendencies"

The evolution of Nigerian government into a three-region federation was the result of a long campaign to entrench subnational partisan enclaves. Yet this tripartite structure had always masked immense ethnic, linguistic, and social diversity. The 1953 census revealed approximately 250 distinct languages spoken by the country's then-30 million citizens—with ten languages being spoken by at least 350,000 people. No region was homogenous; each consisted of a "nucleus" and a "peripheral zone," distinct both geographically and culturally (Coleman 1958). Despite the best efforts of the Sokoto *jihadists* and the old kings of Oyo, there remained sizeable minorities in all regions which had never been folded into the political or cultural majority by which they found themselves governed in the early 1950s (Sklar 1963; Vaughan 2000, 2016). This feature of the three-region structure was by then well-known: British colonialist E. D. Morel had proposed an alternative four-region structure in 1911, and Senior Resident in Northern Nigeria Charles Temple had proposed a 7-province structure in 1912. Lugard rejected both.

When the Richards Constitution transferred political responsibilities to the Regions, minority groups around the country responded by organizing political parties and action groups—organizations broadly termed "statehood movements" by their contemporaries. By 1955, to the extent that multipartism could be said to exist in any of the regions, the opposition parties were almost exclusively organized around minority community agitation for separate regions (Sklar 1963). These "fissiparous tendencies," as the colonial government's official euphemism termed them (Whitaker 1970, 116), had previously sought to advance their interests from within the folds of the major parties. But with regions acquiring the formal trappings of government, and the major parties entrenching their hold on regional executive and

legislative bodies, it became clear that these communities were destined to remain permanent minorities in a tripartite federation (Alapiki 2005).

Three statehood movements are particularly noteworthy: those organizations advocating new regions in the Mid-West, Calabar-Ogoja-Rivers, and Middle Belt zones. These "Big Three" or "Major Minorities" represented the largest statehood movements in the Western, Eastern, and Northern regions, respectively (Vickers 2010). Together, they constituted the greatest challenges to tripartism in the period with which this chapter is concerned.

5.2.1 The Mid-West

The West as it emerged from the 1939 reorganization consisted of nine provinces. All of these were dominated by Yoruba except for the two southeastern provinces, Benin and Delta. Here most inhabitants were drawn from the Edo and Urhobo peoples (both of whom speak Edo and not Yoruba), and in the far reaches bordering the Niger River, Igbo-, Ijaw-, and Itsekiri-speakers. Together, these minority elements accounted for 25% of the Region's inhabitants in the 1953 census (TNA, CO 879/192, 17). Given their similarity of experience under Western, Yoruba rule, and their position in the center-south of the country, these provinces came to be known as the Mid-West.

More than just ethno-linguistic differences, the Mid-West experience differed from that of the "Yoruba West" in local administrative structure. Traditional forms of local government in the Mid-West tended to be more conciliatory than those of Yorubaland (Sklar 1963; Vaughan 2000), while the Mid-Western institutions with Yoruba equivalents differed in subtle ways. For instance, the Benin royal family can trace its lineage to the *One* of Ife, establishing a direct familial tie to one of the most

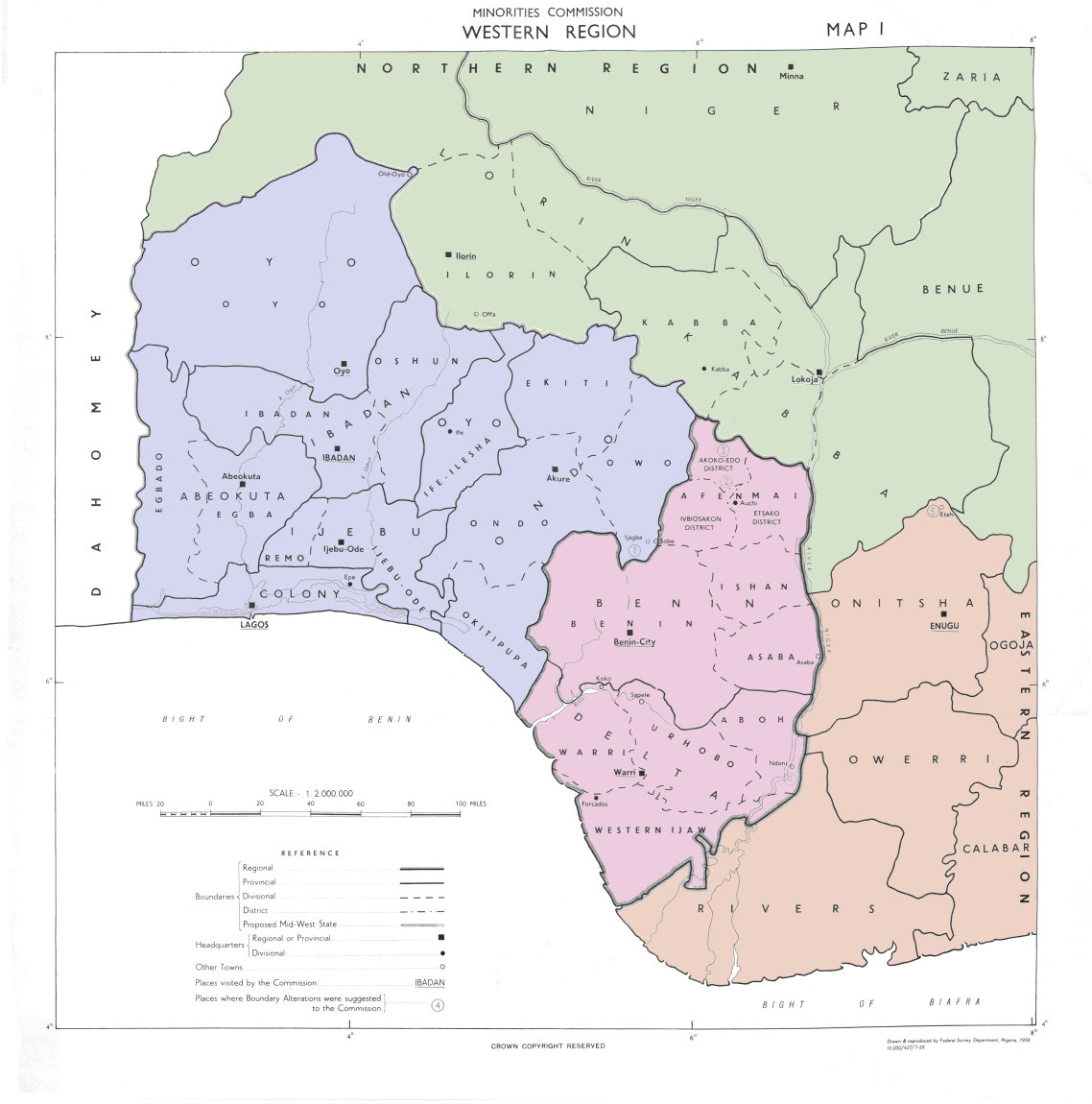


Figure 5.2: The Mid-West Region. A map of the Mid-West Region, as presented in the Willink Commission Report of 1958 and retouched (TNA, DO 177/42). Orange is the Eastern Region, green is the Northern Region, blue is the rump Western Region, and purple is the Mid-West Region as it was then proposed. See Figure 5.6 in the Appendix (section 5.A) for the original map.

important Yoruba *obas*.[8] Yet while the Yoruba *obas* were customarily constrained by the assent of their subjects, the *Oba* of Benin required no such approval, allowing him to rule more absolutely. The equal application of indirect rule across these various institutions subsequently yielded different places for traditional rulers in the emergence of Western parties, with the Yoruba *obas*' role far more advisory, if not muted.

The Mid-West statehood movement's formal organization can be traced to August 1948, when the *Oba* of Benin, Akenzua II, created the Reformed Benin Community (RBC), guided by Chief Humphrey Omo-Osagie. It is undoubtedly true that part of the impetus for the initial formation of the RBC was personal incentive. The *Oba*'s vision of the Mid-West was a "family reunion" of local communities sharing common myths of origin, cultural heritage, and ancestral parentage—in which the *Oba* is the central figure. Part of this project appeared to be an attempt to re-establish the ancient empire of Benin within the modern administrative framework (TNA, CO 879/164, 49), much like the project of the NPC in the North. On a less grand scale, the RBC also appears to have been conceived as a vehicle for Akenzua and Omo-Osagie to counter the aspirations of local political enemies, and in particular, the leaders of the Benin Tax-Payers' Association (Vickers 2000).

However much elite incentives motivated the enunciation a Mid-Western identity, Akenzua and Omo-Osagie found a large audience with whom the message resonated. To many, Western rule meant benign neglect at the hands of AG politicians keen to divert funds to Yoruba provinces. Despite being home to one of the country's most prominent ports in Warri and sizeable export earnings from palm oil, the Mid-West did not receive the same investment as elsewhere in the

8. The *One* 's position in Yoruba society is in large part due to Ife's position as the spiritual home and "cradle" of the Yoruba.

Region. For instance, in 1952, the Western government spent £169,000 in Benin province and £225,000 in Ijebu, a province in the Yoruba heartland, despite the latter having only 60% of the population of the former. Nor did Mid-Westerners feel they received their share of public sector jobs, with the civil service, Marketing and Production Boards, and regional investment vehicles going largely to Yoruba. Mid-Westerners were effectively shut out of major economic decisions of the region, which, under the Lyttleton Constitution, were substantial.

AG corruption and electioneering were also notorious. A 1956 letter from Governor of the Western Region, Sir John Rankine, to the Secretary of State for the Colonies, Alan Lennox-Boyd, includes an account of the leading AG figure Anthony Enahoro at the 1956 Regional election (TNA, CO 879/192, 93):

> It is reported that when he left for his constituency in the Benin Province prior to the election he was heard to say 'It's every man for himself', and that he took with him a considerable sum, two to three thousand pounds, with which to seek golden opinions.

The same Enahoro would later become Deputy Leader of the Federal Opposition in 1962, earning an estimated £60,000-70,000 annually by "facilitating contacts" (TNA, PREM 11/4383). Vote-buying and corruption of this sort was hardly the exception for this period in Nigerian politics; it was Mid-Westerners general inaccess to the spoils, not anger over financial mismanagement, that chafed.

Elsewhere, AG government took the form less of neglect and more of coercion. Nowhere was this more evident than in the years-long effort to gain partisan control over traditional rulers in the West. Although the AG had always enjoyed some measure of kingly support from its close links to *Egbe*, the party used its control of Western government to obtain a monopoly. The turning point came in 1956, when the AG succeeded in deposing the *Alafin* of Oyo, one of the most prominent

Yoruba *obas*, over his political allegiances (TNA, CO 879/192, 49-50). As this battle played out, Western kings—including the *Oba* of Benin—crossed the aisle to support the AG, until only one NCNC supporter remained among the 54 members of the Western House of Chiefs in 1958. Besides just carrying enormous social importance, the (enforced) loyalty of the chiefs was then used to engineer AG majorities in sub-regional government. For example, in Benin province, the AG had never won an outright majority in any of the local council elections. Yet in 1958, acting on the AG's instructions, the *Oba* of Benin injected 26 chiefs on these councils, giving AG majorities in three (of six) and a 56% majority on the provincial council—despite winning just 35% of the provincial vote, compared to the NCNC's 56% (Sklar 1963).

AG subversion of institutions to partisan ends was not limited to traditional rulers. In one letter from January, 1955, Rankine reported to Lennox-Boyd no fewer than three such attempts in the space of the 6 months since the former's appointment. The AG had attempted with varying degrees of success to: organize a regional police accountable to the AG, against express provisions in the Lyttleton Constitution leaving the Police under federal control; reorganize all provincial authorities so that they reported directly to the Minister for Local Government, effectively turning district and village officials into political appointees; and stack the Production Development Board with party men, who were able to distribute contracts so to shore up AG support (TNA, CO 879/192, 1-7). Later in 1955, the AG also engineered an investigation into the NCNC-dominated Lagos Town Council under the charge of financial mismanagement. This justification did not wash, given its comparatively efficient administration in regional context, and the fact that the allegations had emerged from AG-owned newspapers. As Rankine noted, the real object appeared to be to dissolve the Council and replace it with an AG patsy (TNA, CO 879/192, 49, 62).

Together, these factors combined to make a separate region attractive to Mid-Westerners. The RBC soon grew into a series of associated movements, including the Benin Delta Peoples' Party (BDPP, 1952-1955), Mid-West State Movement (MWSM, 1956), and the Mid-West Peoples Congress (MWPC, 1962-1963), among many smaller organizations. Throughout the 15-year period from the foundation of the RBC to the eventual success of the movement with the creation of the Mid-West state in 1963, the movement enjoyed a majority (but not uniformity) of support among people of the Benin and Delta provinces (Vickers 2000).[9]

5.2.2 Calabar-Ogoja-Rivers

Just as the Mid-West was culturally, politically, and economically distinct from the Yoruba core of the Western region, so too were the Calabar, Ogoja, and Rivers provinces from the Igbo core of the Eastern region. While 60% of Eastern citizens were Igbo in 1953, they were largely concentrated in the region's interior, mostly Onitsha and Owerri Provinces, with non-Igbo occupying the southern and eastern periphery. Almost all of the 265,000 Ijaw clustered in the coastal creeks of the region's southern districts. In the southeast, the Efik-speaking peoples together counted well over a million citizens—including 750,000 Ibibio alone—occupying the basin of the Cross River, which flows up from Cameroon and creates an inland delta. In the far east of the Region was Ogoja Province, which Sklar (1963, 13) describes as "a linguistic medley which is yet undocumented in its entirety."

Like their Mid-Western counterparts, residents of the COR provinces had a political and cultural trajectory that diverged substantially from that of the regional

9. On 18 April 1955 the Lieutenant Governor of the Western Region, Sir Hugo Marshall, wrote to Permanent Under-Secretary of the State for the Colonies, Sir Thomas Lloyd, that in Rankine's opinion, it was a "cold truth that it would be political suicide for any Western politician in present circumstances to come out openly against the separate State now desired by the people of Benin and Delta provinces" (TNA, CO 879/192, 17).

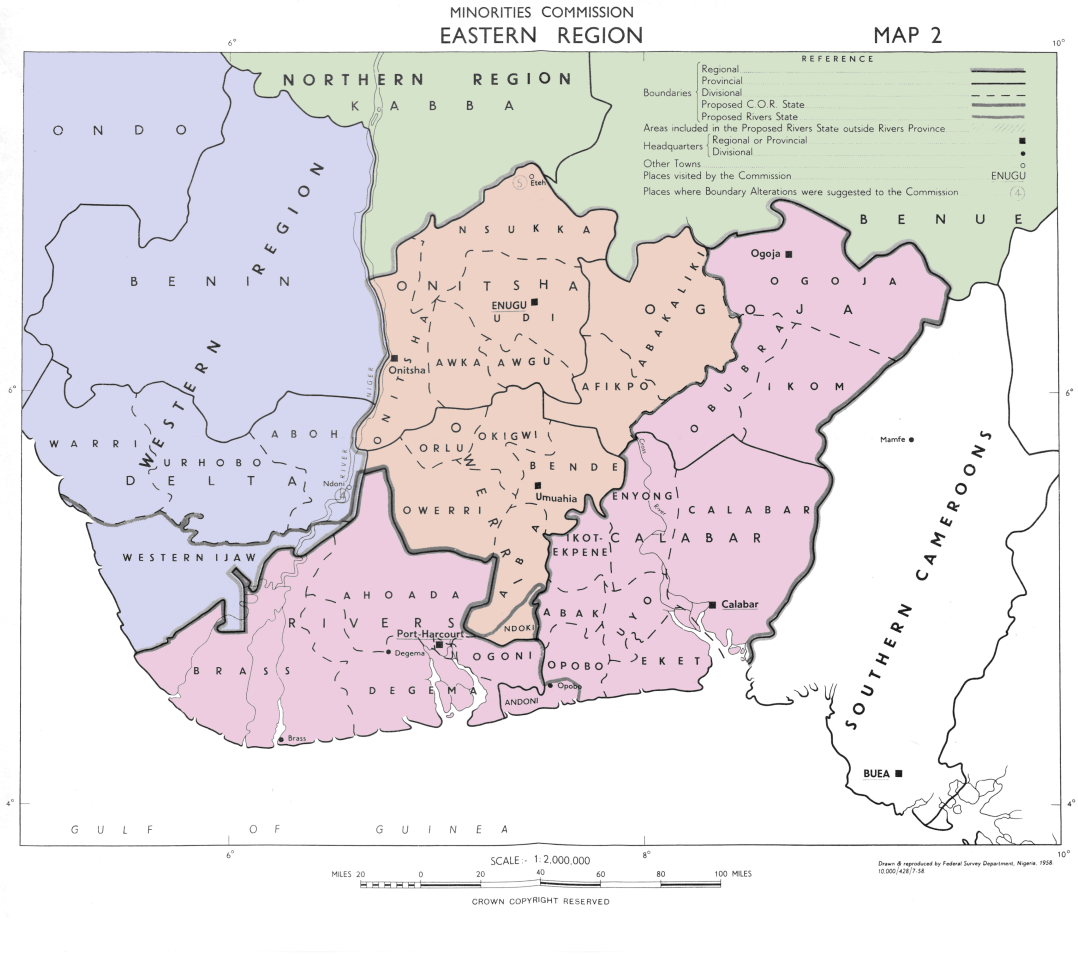


Figure 5.3: The COR Region. A map of the COR Region, as presented in the Willink Commission Report of 1958 and retouched (TNA, DO 177/42). Orange is the rump Eastern Region, green is the Northern Region, blue is the Western Region, and purple is the COR Region as it was then proposed. See Figure 5.7 in the Appendix (section 5.A) for the original map.

majority. Ijaw separatist leaders presented British officials with evidence of the singularity of the Ijaw language, and argued that as "water people" they could not be adequately governed by the Igbo "land people" who did not understand their way of life (Vickers 2010, 144-151). The residents of Calabar province, meanwhile, had experienced centuries of European contact, accounting for a large proportion of the British slave (and later, palm oil) trade in West Africa (Falola and Heaton 2008). In contrast, the Igbo hinterland had been almost untouched by either global trade. Thus, while the coastal belt had signed treaties of protection over the course of the 19th century, it was not until 1910 that the Regional capital, Enugu, was brought under effective British control (TNA, CO 879/192, 30).

The diverse peoples of the COR provinces were mainly united in their discontent with Igbo-oriented NCNC rule. While perhaps not as coercive as AG government in the West,[10] the East was undoubtedly the most inept of the three regional governments (TNA, CAB 129/87/20). In 1955, the Governor of the Eastern Region, Sir Clem Pleass, wrote in a characteristic letter to Lennox-Boyd that (TNA, CO 879/192, 33-34):

> [Zik's] main failing has always been that he cannot tolerate competition and in consequence he surrounds himself with incompetents and non-entities, in order that his own light may shine the brighter... Mr. Akpabio, the Minister of Education... is certainly not competent to be Minister of Education. He is stupid and obstinate and so incompetent in Council that he is becoming the butt of his colleagues... The other Ministers may be discussed more briefly. The only one of any genuine ability is Dr. Esin, the Minister of Welfare... Dr. Emole, the Minister of Agriculture...

10. This coercive aspect would later grow amid heightened ethnic tensions after the coups of 1966 and the civil war of 1967-1970. British Prime Minister Harold Wilson, returning from a trip to Nigeria in 1969, told his Cabinet that the non-Igbo minorities in the East had suffered atrocities under the breakaway Biafran regime (TNA, CAB 128/44/15). Minutes from a similar meeting in November 1969 note that "the rebellion was essentially a tribal movement, based on Ibo [sic] domination. A Biafran State could not be viable without the inclusion in it of large numbers of non-Ibos who detested Ibo rule [sic]" (TNA, CAB 128/44/57).

> [is a] sadly ineffectual little man. Mr. Imeh, the Minister of Trade... is a self-opinionated individual of doubtful honesty and little ability; he finds the problems of the Marketing Board—his former employers—quite behind him. Messrs. Ururuka and Okoya, Ministers of Industries and Transport respectively, are ineffectual cyphers.

The weakness of the Eastern government was such that Pleass considered it an "excusable exaggeration" to call the House of Assembly a rubber stamp for the NCNC Executive, which was, like the Assembly itself, an organ directly controlled by Zik and his "incompetents and nonentities." It seemed to Lennox-Boyd that the party remained in power not through good governance but rather on the strength of the party's Igbo associations (TNA, CAB 129/94/21).

An indicative example of this ineptitude is the near-bankruptcy of the East. Governor-General James Robertson wrote to Lennox-Boyd on 23 December 1957 that the Regional government had overspent so badly over the previous two years, largely due to the exorbitant cost of Zik's universal primary education scheme, that it would not be solvent just three months later. Robertson, Lennox-Boyd, the recently-retired Macpherson, and the new Governor of the East, Sir Robert Stapledon, mulled washing their hands of the crisis "so that the bankruptcy of the Region may expose the bankruptcy of Zik and the NCNC as a political party" (TNA, CO 879/192, 84). In the end, fear of souring Anglo-Nigerian public relations forced the British to intervene, whereupon they informed the Eastern government that they would declare a state of emergency if the books were not immediately balanced. Despite Zik's conviction that there was no need to change course, the regional executive voted to cut education spending and raise revenues, partially by imposing school fees. Large-scale demonstrations followed, and the NCNC suffered a loss of public confidence, but bankruptcy was narrowly avoided (Sklar 1963).

Though they may have struggled to govern, NCNC ministers excelled in

corruption (TNA, CO 879/192, 30-32). Local government bodies were so badly mismanaged that three Urban District Councils had to be replaced by caretaker bodies in the mid-1950s. Over the 1954-1955 period, the government established a series of state-run Corporations for various industries—including finance, development, tourism film, printing, and pharmaceuticals—which played little discernible role except to direct public contracts to allies. In 1956, the NCNC Chief Whip resigned and organized a revolt against Zik, alleging that the regional government had funneled £2,000,000 into African Continental Bank Limited, in which Zik owned a substantial private stake (TNA, CAB 128/30/48; CO 879/192, 79-80). The Secretary of State then appointed a commission, headed by the Chief Justice of the Federation, which eventually found that while Zik may not have attempted to personally enrich himself through graft, he had "fallen short of the expectations of honest, reasonable people" (quoted in Sklar 1963, 185). Particularly damning was the fact that these investments had probably kept Zik's main source of personal wealth, the Zik Group of real estate and newspaper holdings, solvent.

It was in reaction to this chronically inept and corrupt government that COR separatism emerged. Demands for a COR state were first voiced in 1953, but the movement was not formally inaugurated until 1954, whereupon it quickly allied with the United National Independence Party (UNIP, successor to the NIP). Yet COR organization began and remained much weaker than elsewhere. So effectively were minority elements shut out of Eastern elections that although COR state proponents were motivated by resentment of of Igbo rule, they depended on non-NCNC Igbo politicians to advance their cause (TNA, CO 879/192, 17). In addition, the NCNC was by far the most effective of the three major parties in poisoning the well for statehood movements. On the one hand the NCNC effectively pitted COR elements against each other by supporting individual factions, e.g., a Rivers or Calabar State

to stand alone. On the other, the NCNC also effectively flooded the political space with statehood demands, including supporting an Annang State movement, to reduce the prominence of the COR movement (TNA, CO 879/192, 55). Despite these efforts, demands for the COR state appeared at constitutional talks in 1953 and 1957 (TNA, CO 879/159, CO 879/164), and became a focal point into late-colonial inquiries into minority rights (Colonial Office 1958).

5.2.3 The Middle Belt

Of the three major statehood movements, perhaps none had a stronger case for self-rule than that of the Middle Belt. The ethno-linguistic communities of the Middle Belt were tightly grouped in the southernmost provinces of the Northern Region: the Nupe in southern Niger and eastern Ilorin, the Tiv in Benue, the Yoruba in southern Ilorin and western Kabba (bordering the heart of Yorubaland in the Western province), and some 220 smaller groups together accounting for 5 million citizens in these provinces as well as Adamawa and Plateau. These minority groups were also much larger than their Mid-West and COR counterparts, with 1.3 million Kanuri, 800,000 Tiv, 500,000 Yoruba, and 350,000 Nupe counted among the North's 17 million inhabitants. And whereas the "core North" was almost exclusively Muslim, nearly all of the Region's Christians and adherents to indigenous faiths were concentrated in the Middle Belt. As Sklar (1963, 345) notes, almost everyone in the Middle Belt was either non-Hausa-speaking, non-Muslim, or both.

The Middle Belt also formed a political periphery in a much more meaningful way than did the Mid-West or the COR provinces. It was from the northernmost reaches of the Region that the Shehu's rebellion had emerged, and it was there that the emirate system continued to structure social and political life in the colonial era.

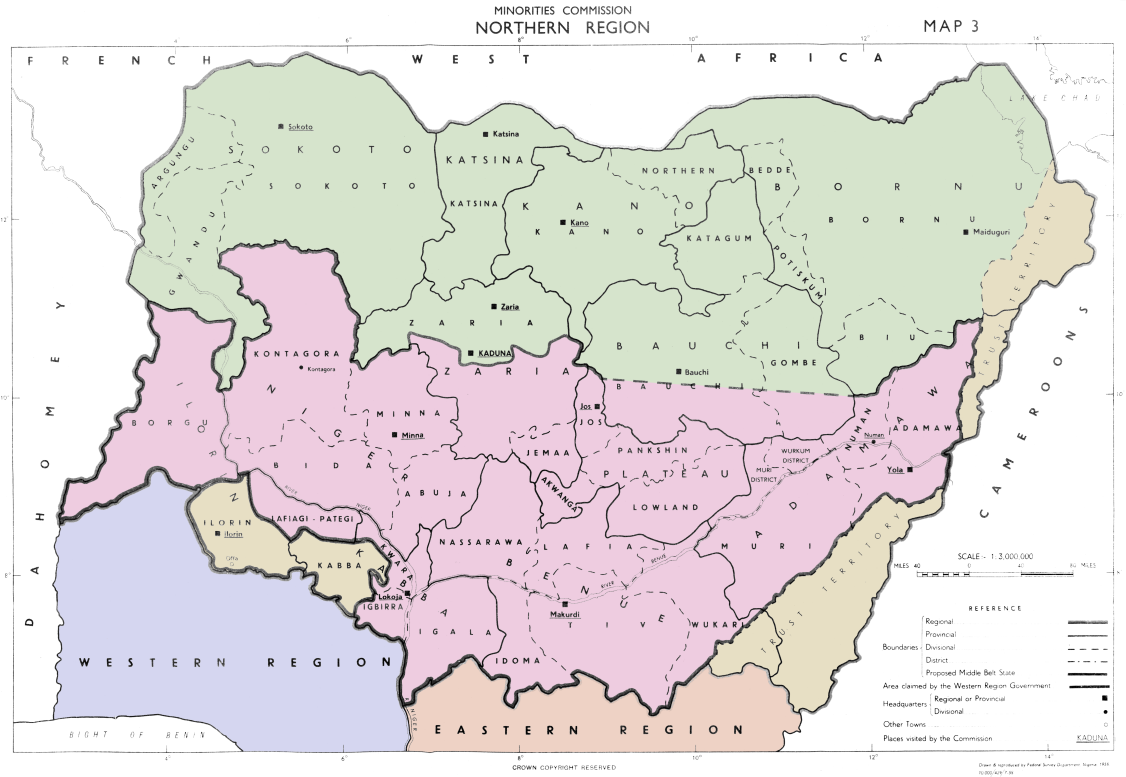


Figure 5.4: The Middle Belt Region. A map of the Middle Belt Region, as presented in the Willink Commission Report of 1958 and retouched (TNA, DO 177/42). Orange is the Eastern Region, green is the rump Northern Region, blue is the Western Region, purple is the Middle Belt Region as it was then proposed, and tan are Northern territories expected to leave the Region but not to join the Middle Belt under the movement's proposals. See Figure 5.8 in the Appendix (section 5.A) for the original map.

Yet the Caliphate had never been able to "radiate power" through satellite emirates in the Middle Belt (Boone 2003; Herbst 2000). The conservative political order that Lugard had "preserved" in the Middle Belt was as new and alien to these provinces as British rule was elsewhere in Nigeria—and scarcely distinguishable from colonial rule itself (Harnischfeger 2004; Kendhammer 2010; Vaughan 2016). Macpherson summarized this experience in a February 1955 letter to Lennox-Boyd (TNA, CO 879/192, 9):

> there is a very considerable non-Moslem population, mainly classed as 'Pagans' and concentrated in the more inaccessible areas where their forefathers successfully resisted or escaped from the Fulani invaders who brought Islam with fire, sword and the slave-gang to the North.

What few pre-colonial governance structures remained when the British arrived were dismissed as "mystico-religious rather than political or military" (TNA, CO 879/164, 53), and obliterated through the Lugardian expansion of emirate governance structures throughout the North.

Moreover, nowhere in Nigeria was the will of the regional majority imposed on the minority with such vigor as in the North. As with the AG in the West, traditional rulers served as the primary vehicle for enforcing the will of the NPC. However, unlike the AG, the NPC had at its disposal Islamic courts which were staffed by *alkali* (Islamic judges), who were more vulnerable to manipulation within the emirate system than their common law equivalents were in the southern Regions. From the mid-1950s onward, Native Administration officials noted increasing pressure to administer Islamic justice as the NPC defined it. "Whispers" of "unpleasant consequences" for failing to follow the party line became open discussion (TNA, CO 879/192, 14). In 1958, Tafawa Balewa admitted ruefully in private to Governor of the North, Sir Gawain Bell, that the Northern emirs had stacked Islamic courts with

amenable judges, and then intervened in individual cases for political and personal reasons (TNA, CO 879/192, 76). This judicial capture coincided with religious discrimination, including the promotion of Islamic institutions of governance over indigenous structures, discrimination against Christians in public hiring, and the banning of Christian missionary activities while providing state funding for Islamic equivalents, including a conversion tour by the Sardauna himself (Harnischfeger 2008; Logams 1985).

Middle Belt representatives to the 1953 constitutional conference told the British that "most of these places are being neglected by the Moslem rulers whose interest is not in the education and progress of the people who differ with them in outlook, culture, religion, and social patterns" (TNA, CO 879/159, 285). They cited a litany of offenses. Among these were items such as imbalance in the appointment of ministers to the Northern Executive Council, discrimination against Middle Belters in awarding government scholarships, the concentration of higher education institutions exclusively in the core North, and the complete absence of the Middle Belt provinces from the Agricultural Development plan. Tafawa Balewa's reply, characteristic of the NPC stance, was that "the boundary of the Northern Region was based on history and tradition and suggestions that there was such a thing as a Middle Belt were fairy stories" (TNA, CO 879/159, 77). To the NPC, the Middle Belt could not suffer oppression because the Middle Belt did not exist.

The vigor with which the NPC attempted to enforce political assimilation into the emirate hierarchy created a movement for Middle Belt statehood that began earlier, grew faster, and reached greater strength than either the Mid-West or COR movements. As early as 1908, traditional rulers in Adamawa lobbied the British to be removed from emirate rulership, with similar demands also being heard in Bauchi, Plateau, and Zaria over the next two decades (Logams 1985).

In 1938, the Tiv Progressive Union became the first Middle Belt political party, joining the NCNC when the latter formed officially in 1944. In 1945 the first formal demand for a Middle Belt region came from the Birom Progressive Union, and in 1949 the broader Middle Belt statehood movement was born with the Northern Nigerian Non-Muslim league, renamed the Middle Zone League (MZL) the next year. So large was the Middle Belt movement that when the MZL agreed upon a policy of working cooperatively with the NPC in 1953, the more radical Middle Belt People's Party (MBPP) formed in defiance, and both parties returned members to the federal House of Representatives in 1954. These factions merged in 1955 as the United Middle Belt Congress (UMBC; TNA, CO 879/192, 52). In the 1956 regional elections, pro-Middle Belt state elements won 15 of the North's 131 seats, and therefore formed a very real legislative bloc (TNA, CO 879/192, 99-100). Thus, despite its "congenital lack of cohesion," the movement for a Middle Belt region gained increasing strength in the decade leading up to independence (Sklar 1963, 348), with the UMBC emerging as an important political force in the last years of decolonization.

The revolt against NPC rule was as much in the streets as at the polls. Riots in the Tiv heartland claimed victims in 1929, 1939, 1945, 1948, 1960, and 1964 (Tamuno 1970). Even Northern Governor Sir Bryan Sharwood-Smith, who was otherwise very sympathetic to the NPC leadership and called Middle Belt claims of neglect "without justification" (TNA, CO 879/192, 27), noted in 1955 that

> [i]t has, on the other hand, been far too true that many Northern leaders for a long time tended to look down their noses at all who live below the 11th parallel and, in this respect, the tendency to separatism in these areas is a fitting reward for excessive race and class consciousness.

The cycle of protest and crackdown only increased the sense of victimization of the

Middle Belt at the hands of the NPC, strengthening the statehood movement well into the independence era (TNA, PREM 11/4963).

5.2.4 The Willink Commission and independence

In 1957, the colonial government convened a final round of constitutional talks in London with self-government on the horizon. Noting the fever pitch that minority claims had reached, the conference decided "after bitter discussion" to appoint a Commission to enquire into the plight of these "region-locked" minorities (TNA, CAB 129/94/21). This "Commission to Enquire into the Fears of Minorities and the Means of Allaying Them," chaired by Sir Henry Willink, was tasked with assessing whether minority fears were justifiable, what safeguards could be enacted to prevent majority domination, and whether the creation of new regions was necessary to protect minorities. However, as the terms of reference made clear, state creation was a "last resort" that would only be countenanced if "no other solution" could prevent regional majority domination (TNA, CO 879/192, 155). Further, the burden of proof on minorities would be high: Lennox-Boyd informed the Governors that minority movements' evidence must be "based on specific happenings in the past... A general statement of fears for the future unrelated to the established facts of the past would carry little weight" (TNA, CO 879/192, 156). With constitutional talks paused to await the results of the inquiry, the Commission quickly went to work. Willink and his associates hosted 61 public hearings, including meetings in all regions, Lagos, and each of the proposed Mid-West, COR, and Middle Belt regions.

After six months of gathering evidence, the final report was released in July 1958. Although the Commission recognized that "there remained a Body [sic] of genuine fears and that the future was regarded with real apprehension" by the

minority movements, it rejected the call for new regions. To the commissioners, there were four grounds to dismiss statehood movements: new regions would create new minorities, exacerbating the problem it was supposed to solve; none of the proposed new regions would be fiscally viable; new regions were not necessary to solve minority discrimination, as the existing regions could instead just transfer greater powers over taxation and spending to minority areas, and constitutional safeguards could be added to provide minorities legal protections; and (somewhat implausibly) that the "tribalism" that generated minority discrimination would wane with independence (Colonial Office 1958).

The problem of subdividing minorities *ad infinitum* had long occupied British thinking on minority fears. A briefing memo for the Secretary of State for the Colonies ahead of the 1957 constitutional convention described the Middle Belt thus (TNA, CO 879/164, 49; see also TNA, CO 879/192, 11):

> The concept of a separate 'Middle Belt' Region is a nebulous one... There are many differences of culture and language between the 'Middle Belt' tribes, and it would be impossible to draw any clear boundary line between the Muslim and non-Muslim areas. The only unifying influence in any 'Middle Belt' Region which might be established would be fear of Islam. Once that fear died through separation from the Muslim North, the new Region would tend to break up into its constituent tribal and provincial units.

Nor was this concern limited to the Middle Belt. The Lieutenant Governor of the Western Region, Sir Hugo Marshall, wrote in 1955 that "there would be old rivalries and old jealousies within the proposed 'Mid-West' State just as bitter as there are at present between the 'Mid-West' and the 'Yoruba West'" (TNA, CO 879/192, 18). While this theory has been somewhat borne out by independent Nigeria's repeated deconcentration, it did minimize the very deep cultural and institutional differences separating the Middle Belt from the core North, and the Mid-West from the Yoruba

West. Any such differences would be greatly dampened within either new region, and the "old rivalries" were likely to be interpersonal more than institutional.

Perhaps no concern held greater weight with the colonial government than the fear of non-viability. In another brief to the Secretary of State for the Colonies for the 1957 constitutional talks, the subject of fragmentation drew the following commentary (TNA, CO 879/164, 51-52; see also TNA, CO 879/192, 11, 40-41; DO 195/82/66):

> The machine of Government in Nigeria has undergone two major reorganizations in the last six years, the second of which is still barely complete. A further upheaval through the creation of more Regions would probably prove more than it could stand. The machine might well break down and there would then *inter alia* be an end for an indeterminate period to much hope of further orderly economic development. Even if fresh governmental machines could be created, the cost of this machinery would substantially increase, although there would be no increase in resources available, so that an undue proportion of the financial resources of the country would go simply into turning the wheels of government.

Less convincing was the argument that devolving authority could allay minority fears. The major parties, looking for ways to clamp down on threats to their regional hegemony, had already sought ways to bring statehood activists into their fold. Decentralization played a critical role in their courtship. In the North, the Hudson Commission had recommended in 1957 the reorganization and strengthening of provincial authorities. Although the eventual bill to establish provincial councils passed by the Regional legislature in 1959 bore little resemblance to the Commission's recommendations, the effect was to give greater autonomy to local actors, particularly in the Middle Belt (TNA, CO 879/164, 52; CO 879/192, 29; Whitaker 1970). In the West, the AG had created in 1955 an Advisory Council and a Cabinet position, the Minister of Mid-West Affairs, to represent the Benin and Delta

provinces' interests in policy discussions—what colonial administrators referred to the "Welsh solution." Similar proposals were mulled in the East (TNA, CO 879/164, 50-52). Though piecemeal, these reforms left the minorities' representatives with as much authority as they had enjoyed at any point in the previous two decades—yet the statehood movements had only grown in strength. Decentralization seemed an unlikely solution.

With the Willink Commission Report delivered, constitutional talks resumed in 1958. The leaders of the major parties negotiated the final changes to the Constitution and prepared for the 1959 elections under the assumption that each would sweep its respective region, which indeed they all did. Thus, when independence came on October 1, 1960, deconcentration looked a bleak prospect. The three regions had significant powers, each was governed by a dominant party with a vested interest in protecting its fiefdom (as well as the institutional resources to do so), and party leaders had the Commission's findings to fall back on when pressed on statehood movements.

5.3 Statehood politics as coalition politics

5.3.1 Policy by circumstance

Major party leaders' attitudes toward minority statehood, and particularly those of Awolowo and Zik, bedeviled colonial administrators. For decades, Awolowo had presented himself publicly as a fierce advocate of state creation (Vickers 2000, 3-4), calling it an "article of faith" in his 1960 autobiography (Awolowo 1960). Yet it was the AG who stymied Mid-Western statehood aspirations at successive constitutional conferences and in the early days of the First Republic. Similarly, the NCNC

"Freedom Charter" of 1948 had proposed a commonwealth of states organized on a linguistic basis, and its 1954 election manifesto had advocated the division of all regions into smaller states (TNA, CO 879/192, 35, 63). But by 1955, the NCNC had revoked its approval of any attempt to break up the East, and by the 1957-1958 conferences in London, the NCNC appeared to drop all claims for state creation. Exasperated colonial officials echoed Marshall, who wrote in 1955 that the parties' positions "are apt to change so rapidly that it is dangerous" to assume any stability even six months into the future (TNA, CO 879/192, 17).

With the benefit of sixty years' hindsight, it is apparent that these positions fluctuated with the coalitional considerations of the day. Nowhere is this clearer than in the parties' statehood strategies heading into the 1957 talks. By 1955 fragmentation was considered a "major objective" of both the AG and NCNC, due to their inability to compete with the NPC (TNA, CO 879/192, 18). It was therefore unsurprising that the AG was active in supporting the COR movement in 1955, while the NCNC supported the Mid-West movement. But to the bemusement of the British, both parties also appeared willing to bow to the claims of the statehood movements within their own territories if it meant the creation of the Middle Belt and the break-up of the North (TNA, CO 879/192, 63). Both the AG and the NCNC recognized that they would not be able to lead a majority federal government so long as the North remained united, and so were willing to trade away some of their own authority in order to drive a wedge into the monolithic NPC bloc. Governor-General Robertson wrote to Lennox-Boyd that "[t]he NCNC's policy of more States is based solely upon Dr. Azikiwe's desire to break up the overwhelming strength of the North, and to enable him and his party to dominate Nigeria, which he knows he can never do while the North remains one" (TNA, CO 879/192, 63), a sentiment just as easily applied to Awolowo's AG. Despite grand claims of ideological commitment to

state creation, both leaders' support for new Regions was entirely dependent on the coalitional context: should the North remain as one, both the AG and NCNC would resist any attempt to create the Mid-West or COR regions (TNA, CO 879/192, 35).

The same pattern emerged in the run-up to the 1959 elections, which would determine the government at independence. By the constitutional talks of 1958, it was apparent that the only potential governments that could command a majority would be an AG-NPC or NCNC-NPC coalition. Based on a mutual distrust of Awolowo, Tafawa Balewa and Zik entered into a gentlemen's agreement to form a government, and the prospect of NCNC-NPC rule became public knowledge well in advance of the election (TNA, CO 879/192, 180). The Secretary of State for the Colonies summarized the maneuvering in a Cabinet memo about the "stormy" 1958 conference (TNA, CAB 129/95/13):

> It is apparent that for the time being at least the [NCNC] is lining up with the North. Consequently Chief Awolowo, the Action Group leader, has every reason to fear that on independence the West may be faced with a North-East alliance. Partly, if not mainly, for this reason he has pressed with great vigour for his major objective—the weakening of the North by the creation of a Middle Belt State and for changes in the Northern system of government. In so doing he has aroused bitter opposition from the North.

When Awolowo's efforts at the conference were thwarted by NPC and NCNC intransigence, the AG was left with few options, representing the smallest Region (in both size and population) and looking down the barrel of life in the federal opposition. Awolowo decided to ratchet up his advocacy of minority rights. The AG built a national coalition, with statehood movements such as the UMBC, UNIP, the Borno Youth Movement (BYM), and the Ilorin-based *Talaka Parapo* in an attempt to chip away at regional majorities in the North and East.[11] As Acting Governor of

11. While the UMBC and BYM favored state creation, the *Parapo* was mostly Ilorin Yoruba who wished

the Western Region George Mooring wrote to Assistant Under-Secretary of State Christopher Eastwood in August 1958, "Awolowo is convinced that a new states issue would make a very good plank in the platform for the Federal election campaign next year" (TNA, CO 879/192, 181).[12] And so it did: Tafawa Balewa would later complain to British Prime Minister Harold Macmillan that the AG had picked up seats in the East by "exploiting inter-tribal differences and had also gained ground in the North, where they had exploited religious differences" (TNA, PREM 11/3048).[13]

5.3.2 Equilibrium emerges

The AG's efforts to break up the NCNC-NPC coalition were for naught. As in Table 5.2, the NPC's dominance of the North turned a 61% majority into 134 seats, 16 short of an outright majority after the 7 independents who won seats in the North declared for the NPC. Rather than attempt to form a minority government, the NPC honored the gentlemen's agreement and entered into a coalition government with the NCNC. Tafawa Balewa became Nigeria's first Prime Minister, while Zik took the ceremonial role of Governor-General (later President, when Nigeria became a Republic under the 1963 "Independence" Constitution). Awolowo led the official opposition, comprised of the AG and its minority-interest allies.

The formation of the NCNC-NPC government sounded the death knell for minority aspirations in the North and East. Three considerations determined the NPC's obstinate resistance to Middle Belt statehood. First, despite a *talakawa* Prime Minister in Tafawa Balewa, the NPC remained emirate-dominated. For

to "re-join" the West, citing ethnic and cultural similarity. Ilorin had fallen to the Caliphate in the 1830s, and so had always been considered part of Northern Nigeria by the British, despite belonging culturally to Yorubaland.

12. This strategy is also reflected in the AG's caustic reaction to the Willink Commission Report's recommendation against new states (TNA, CO 879/192, 180).
13. Awolowo may also have been motivated by the desire to counter the AG's image as a Yoruba party, given its *Egbe* roots.

Table 5.2: The result of the 1959 Federal House of Representatives election in Nigeria

Votes

	East	North	West	Lagos (FCT)	Total
NCNC	**1,246,984**	525,575	758,462	61,608	2,592,629
	(64.62)	(16.13)	(40.19)	(55.97)	(36.08)
NPC	0	**1,994,045**	32,960	189	2,027,194
	(0.00)	**(61.19)**	(1.75)	(0.17)	(28.21)
AG	445,144	559,878	**933,680**	48,137	1,986,839
	(23.07)	(17.18)	**(49.47)**	(43.73)	(27.65)
Other	237,626	179,022	162,107	138	578,893
	(12.31)	(5.49)	(8.59)	(0.13)	(8.06)
Total	1,929,754	3,258,520	1,887,209	110,072	7,185,555

Seats

	East	North	West	Lagos (FCT)	Total
NCNC	**58**	8	21	2	89
	(79.45)	(4.60)	(33.87)	(66.67)	(28.53)
NPC	0	**134**	0	0	134
	(0.00)	**(77.01)**	(0.00)	(0.00)	(42.95)
AG	14	25	**33**	1	73
	(19.18)	(14.37)	**(53.23)**	(33.33)	(23.40)
Other	1	7	8	0	16
	(1.37)	(4.02)	(12.90)	(0.00)	(5.13)
Total	73	174	62	3	312

Vote and seat percentages by region are given in parentheses. Bold corresponds to each party's "home" region. Blank and invalid votes are excluded. Note that figures include votes and seats won by junior allies, and that percentages may not sum to 100 due to rounding. Data are from Nunley (2012) and Sklar (1963).

traditional rulers like the Sardauna—then the Northern Premier—the NPC was the vehicle for the restoration of the Caliphate. The emirs' commitment to retaining the Middle Belt provinces under the influence of the Sokoto elite was literally a matter of religious conviction, and any effort by the NPC's moderate wing to appease separatist elements would very likely have driven apart the party itself. Second, the NPC feared it could not compete in a Middle Belt region. In Adamawa, Benue,

Ilorin, and Plateau provinces, the UMBC/AG alliance contested 51 seats and won an average of 32% of the vote. Contesting all of the North's 123 other constituencies, the UMBC won just 8% elsewhere, earning single-digit returns in exactly 100 races (Dudley 1968). Given the UMBC's performance across the Middle Belt, the NPC had no reason to believe it could win a governing majority in a Middle Belt state, despite its dominance in the North more broadly. Finally, to the NPC, the retention of the North's numerical superiority was a matter of life and death. Any rebalancing that allowed a majority-Southern coalition to govern would reverse decades of "Northernization" and the effort to erase the gap in socioeconomic development that had prevailed since the days of the Richards Constitution (TNA, CO 879/192, 125). Emirate leaders' strategic calculus was largely the same as it had been for decades: carving the Middle Belt out of the North could only result in southern domination of the North and a return to second-class status within the Federation. "Not an inch of Northern soil" would be willingly given up (TNA, CO 879/192, 174).

The NCNC's position, fluid throughout the colonial era, suddenly crystallized. Now assured of its place in government, and with no prospect of Northern fragmentation, the NCNC categorically rejected COR statehood as being mere anti-Igbo agitation (Sklar 1963; TNA, CO 879/192, 55). Dropping its advocacy for the commonwealth of states, the NCNC sought instead to build its base in the West back up—having emerged from Lagosian politics in the 1940s—so as to gain control over both southern regions and rival the NPC's numerical advantage in the North. Since the entrenchment of the AG in the West meant such a strategy was unlikely to succeed in the short run, the Mid-West presented an intermediate outcome: a foothold for expansion westward, a means of granting the NCNC control of two regional executives, and thus effectively a unilateral veto in inter-regional bargaining.

Beyond just shoring up their own positions, the NPC and NCNC saw in

Mid-West statehood an opportunity to destroy the AG. For one, the AG's electoral strategy had included loud and frequent declarations in favor of state creation. By passing bills in the federal legislature and the Northern and Eastern Assemblies, the NPC-NCNC leaders could force the AG-led Western government to formally vote down an equivalent bill in the Western legislature, decimating its support in the Mid-West and likely breaking its alliances with minority groups in the North and East. If instead the AG allowed a Mid-West bill and a Mid-Western plebiscite to pass, then the AG's home region would be torn in half. The result of such a division would be to render the AG toothless in federal politics. Already the region with the fewest seats, the West sans the Mid-West would leave the AG a base of no more than 40 seats in the House of Representatives, leaving it utterly dependent on the NPC or NCNC. As the Private Secretary to the Secretary of State for the Colonies, Reginald Maudling, concluded in a note to British Prime Minister Harold Macmillan in November, 1961: "The NPC and the NCNC [saw] the splitting of Western Nigeria as an opportunity to reduce to the Action Group Government there to comparative insignificance" (TNA, PREM 11/3444).

Nor did the NPC or NCNC fear an AG successor. To Zik, the party most likely to take the reins in the West was the NCNC itself. Zik was a master of Lagosian and Western politics from the early days of the NCNC, had broad support in Yorubaland, and as Nigeria's foremost independence activist, was by then the most popular political figure in the country (TNA, PREM 11/3443). After three years above the fray as Nigeria's ceremonial Governor-General, Zik appeared ready to return to battle and form independent Nigeria's first southern-led government. The NPC, in contrast, expected to be able to ally with enough anti-NCNC Yoruba to form an almost exclusively Northern government. The NPC might even have been able to pick up some support through allies whose southern campaigns it had

recently begun funding, most notably the nascent Mid-Western Democratic Front (MWDF; TNA, PREM 11/4963). Although the Mid-West would give the NCNC control of an extra regional government, the North's numerical superiority was likely to be an effective safeguard against any attempt by the NCNC to encroach on NPC territory—as it had been since 1946.

5.3.3 "Fire on the mountain"

The challenge for NPC and NCNC leaders was how to overcome the Western veto that attempts to create a Mid-West region would surely induce. Opportunity arrived in 1962. After the requisite bills were passed in the federal, Eastern, and Northern legislatures, the AG stalled a vote in the Western legislature while lodging legal challenges over the interpretation of the Constitution's provisions for state creation. Before these lawsuits could work their way through the courts, however, simmering factional resentments within the AG boiled over. A letter from the Deputy High Commissioner in Ibadan, Roger Barltrop, to High Commissioner Lord Antony Head, provides an inside account of this episode.

At the AG's annual convention in February 1962, the Western Premier, Samuel Akintola, corralled a majority of party delegates into open revolt against Awolowo's leadership. Among the many points of contention were alleged financial improprieties involving banks in which Awolowo owned a large stake—not dissimilar to Zik's bank scandal—as well as substantive disagreements on agricultural, education, and taxation policies. Most importantly, Akintola resented Awolowo's attempts to rule the West from his position in the federal government, and sought a vote to curtail Awolowo's ability to intrude on the duties of the Western Premier. However, the power struggle was inconclusive, and moved into the background over the following

months as Awolowo maneuvered to regain the support of top AG officials.

On May 19, after party elders' peace talks collapsed, Awolowo convened a joint meeting of the parties' regional Executive Committees and conducted a ten-hour "trial" of Akintola, after which those in attendance voted 81-29 to "convict" Akintola on charges of "maladministration, anti-party activities and indiscipline" and order that he resign immediately. That decision was ratified on May 20 by the AG Federal Executive Committee. Akintola fought back by instructing the Speaker of the Western House to convene the Assembly, with the intent to force a vote of no confidence which he felt he could defeat. He also asked the Governor—a ceremonial position, filled by royal appointment—to dissolve the legislature and thereby force a regional election, so that he could take his case to the people. Bowing to pressure from Awolowo, but within their constitutional rights, both rejected his requests. On May 21, on advice from AG officials, the Governor dismissed Akintola. Akintola then lobbied a sympathetic Tafawa Balewa, and wrote an urgent letter to the Queen, asking both to remove the Governor. With their help not forthcoming, and the Awolowo faction having already appointed his replacement (Dauda Adegbenro), Akintola's fate was sealed.

On May 25, the House of Assembly convened. David Hunt of the High Commissioner's Office relayed to John Chadwick of the Commonwealth Relations Office (CRO) the climax of the conflict (TNA, PREM 11/3893):

> All was peaceful as the members took their seats at 9 o'clock and a pious hush was observed during the customary Prayers for the peace and good Government of Western Nigeria. The Speaker then called on [Jonathan] Odebiyi to move the suspension of Standing Orders to allow the sitting to go on beyond noon. As Odebiyi rose, and before he could open his mouth, a member sitting on the back benches on the Government side leapt on to the desk in front shouting "Fire on the mountain"...
>
> This was a pre-arranged signal; a member sitting on the opposition

> benches immediately made a dash at Odebiyi and his Premier, Adegbenro, lashing out with chairs. Another seized the Mace and attempt to belabour the Speaker with it. A number of members were injured, clothes were torn... and the Minister of Trade and Industry, Mr. Kessington Momoh, was stabbed in the neck by a former Parliamentary Secretary. After this had gone on for [three minutes] the Police[,] summoned by the Commissioner, who was in the official box, threw tear gas into the Chamber. The House then dissolved in tears.
>
> A second meeting was held at 11:50 am. It followed almost exactly the same course except that this time Police in steel helmets with batons and riot shields were actually present on the floor of the House... Fighting started in exactly the same way, and the Police had the pleasure of using their batons before closing the session once more with tear gas. The press and spectators outside the House had the edifying spectacle of Honourable Members picking up their petticoats to speed their flight and some of them squeezed through the ventilation louvres, which provided a tight fit since most prosperous Yorubas, on a high starch diet richly laced with palm oil, are men of full figure.

Within days Tafawa Balewa declared a state of emergency in the Western Region and appointed a caretaker administration accountable to the federal government. It was obvious to all that the AG was in grave danger. Lord Head wrote just five days later that, although the NPC and NCNC had long sought to break up the AG, the suddenness with which the opportunity had arisen "[took] the Action Group's breath away."

5.3.4 Three become four

To the surprise of no one, the caretaker administration dropped the legal challenges to Mid-Western statehood that the former Western government had lodged, and gave the necessary approval for the plebiscite to move forward. In July, voters in Benin and Delta provinces returned an overwhelming victory for the Mid-West

referendum,[14] which was quickly ratified. An interim government was appointed, and regional elections for the new Mid-West House of Assembly were scheduled for 1964. After two decades of struggle, the Mid-West statehood movement had gone from futility to victory in the space of fourteen months. The intentionally cumbersome constitutional procedure for deconcentration had been designed to prevent political manipulation of regional boundaries, and yet that is precisely how the Mid-West became Nigeria's fourth region in 1963.

With half of its territory gone, the AG could no longer just patch up its factional conflict and retake its place as one of Nigeria's major parties. Some officials followed Awolowo as he sought to reshape the AG into a national opposition, while others stuck with Akintola and sought to form a smaller party that could join the NPC and NCNC in a national coalition. When the Western House reopened six months later, in January 1963, Akintola leveraged NCNC support to win reinstatement as Premier under the banner of his new United People's Party (UPP). No longer could a single Western party marshal Yoruba support in the way that the AG had been able to, drawing on its ties to *Egbe*. And by breaking up the AG's vicegrip on local institutions, the national coalition had ensured that no AG successor would emerge. As Sklar (1963, 501) noted on the eve of the AG crisis,

> [a]t bottom, the stability of the three-party system depends upon each party's effective control of a regional pattern of chieftaincy, local government, and commercial patronage. The loss of such control by a major party would drastically impair its capacity to compete under existing conditions.

Neither the remains of the AG nor the UPP were able to re-establish control over these institutions after the six-month interregnum under federal control.

14. The vote was widely rigged, with many polling stations returning more votes than there were registered voters. However, the referendum probably would have passed without such fraud (TNA, DO 195/82/66, DO 195/83/91).

Split into warring factions, restricted to competition over the rump of the West, and stripped of its traditional fiefdom, the AG was no longer able to compete nationally. The NPC and NCNC had effectively transformed a three-party coalition structure into a two-party coalition structure. Although Tafawa Balewa and Zik remained split over how Yoruba politicians would align themselves once the dust settled, they shared in the spoils from eliminating the AG. "The greatest conspiracy in Nigeria's 'trilateral' relations" had been a rousing success (Akinola 2014, 48).

The effect of the Mid-West on coalitional politics was evident in the 1964 elections, which took place just ten months after the first Mid-West government was elected. Akintola's UPP realigned with other Western elements to form the Nigerian National Democratic Party (NNDP), which entered into alliance with the NPC, in keeping with Akintola's commitment to joining a national coalition and ensuring Yorubaland got its "turn to eat" (Bayart 1993). Awolowo and the remnants of the AG allied with the NCNC and a group of Northern parties, together known as the Northern Progressive Front (NPF). The NNDP-NPC alliance contested the election as the Nigerian National Alliance (NNA), while the AG-NCNC-NPF bloc took the name United Progressive Grand Alliance (UPGA; TNA, PREM 13/451). As the results in Table 5.3 show, the NPC increased its seat share to an outright majority, the NCNC's nominal seat share declined marginally but that of its UPGA coalition increased. The Action Group—"the best organized, best financed, and most efficiently run political party in Nigeria" (Sklar 1963, 422)—was decimated. As Figure 5.5 indicates, though the dust from deconcentration had barely settled, a new two-party structure was already crystallizing.

Unfortunately, it is impossible to give a full account of the coalitional shift that took place after the creation of the Mid-West, as the 1964 results were widely considered unreliable. The 1964 constituencies reflected a redistricting which, in turn,

Table 5.3: The result of the 1964 Federal House of Representatives election in Nigeria

	East	Mid-West	North	West	Lagos (FCT)	Total
NPC	0	0	162	0	0	162
NNDP	0	0	0	36	0	36
NNA Total	0	0	162	36	0	198
NCNC	64	14	0	5	1	84
AG	4	0	0	15	2	21
NPF	0	0	4	0	0	4
UPGA Total	68	14	4	20	3	109
Others	2	0	1	1	1	5
Grand Total	70	14	167	57	4	312

Voting data are unavailable. Seat totals correspond to results after the supplementary elections of March 1965. Data are from Nunley (2012) and Sklar (1963).

was based on the census of 1962-1963. By all accounts, this census was completely fraudulent: provisional tallies in 1962 showed an outlandish increase in the Eastern population, giving the southern regions an outright majority. After these figures were "corrected" by the (Northern) minister in charge of the census to restore the numerical superiority of the North, Tafawa Balewa canceled the count and ordered a new census. The recount in November 1963 showed an even wider margin for the North, at 29.8 million residents to the East's 12.4 million, which southerners viewed as impossible (TNA, DO 195/280/15, PREM 11/4963).

Further, the election itself was considered fraudulent. Throughout the country ballot papers failed to arrive in time for the vote (TNA, PREM 13/451). Dozens of Northern opposition candidates were unable to file their papers, including in Tafawa Balewa's constituency (TNA, DO 195/282/239). British estimates placed the amount spent on illicit electioneering at £3 million over the summer leading up to

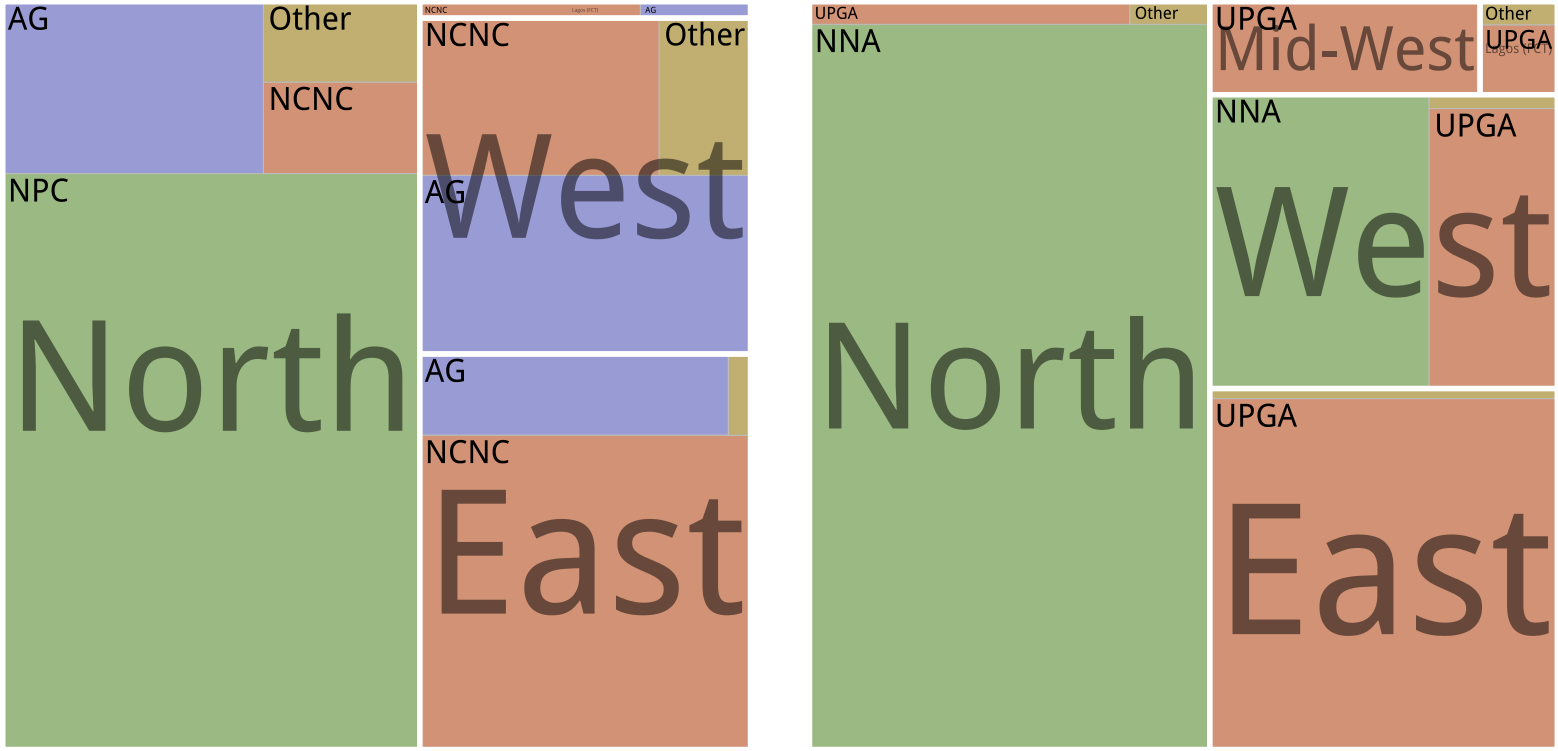


Figure 5.5: Coalition structure change in Nigeria. Blocks are sized proportionally to seats in the Federal House of Representatives after each election, divided by region, and colored according to party. Nigeria's three dominant coalitions in 1959 became just two in 1964. The creation of the Mid-West splintered the AG and caused the remaining parties to reorganize into the NNA and UPGA.

the election (TNA, DO 195/280/50A). Of the Western Region, High Commissioner Sir Francis Cumming-Bruce wrote in November 1964 (TNA, DO 195/281/71):

> Pressures and thuggery are widespread; bribery is on a lavish scale; and the recently revised constituency boundaries have been redrawn on lines favourable to the Akintola Government. Thus the constituencies likely to return NNDP members have proliferated, while the number of those in Action Group areas has been reduced. There is also clear evidence of malpractice in the compilation of the electoral rolls. These Tammany methods are likely to win a quarter of the seats.

In reaction to these electoral malpractices, the UPGA instructed its supporters to boycott the vote.[15] After some 75% of constituencies in Lagos, the Mid-West, and the East failed to return a candidate, the government was forced to hold supplementary elections in March 1965. Seeking stability, Tafawa Balewa formed a national unity government. With the major party leaders assuaged, he remarked that "though he was not out of the woods yet, he believed that the crisis and its aftermath were now 70% over" (TNA, DO 195/282/209A).

5.3.5 The fall of the Republic

In the small hours of January 15, 1966, just as this new two-party system had settled, a group of mid-ranking Igbo officers staged a military coup. The Sardauna, Akintola, Tafawa Balewa, and eighteen others were assassinated. Overnight, virtually every major player was removed from coalitional bargaining.[16]

The January plot was almost entirely unforeseen. British officials had long been convinced that Nigeria's greatest security threats were communist-inspired

15. When Akintola was assassinated in the January 1966 coup, Lagosians were reportedly jubilant, as he had been seen as the architect of the rigged elections and subsequent security problems in the West (TNA, PREM 13/1040).
16. Awolowo and Zik both escaped harm. Awolowo was in prison, on trial for offenses relating to an alleged plan to overthrow Tafawa Balewa by force, while Zik was traveling in the United Kingdom, with which the plotters likely timed their attack to coincide.

revolution and long-term administrative decline (TNA, DO 195/280/46A, DO 195/281/81, DO 195/282/209A).[17] The possibility of a military coup had been consistently discounted (e.g., TNA, DO 195/281/106). Even when the 1964 elections brought sectarian violence, the main concern was secession of the Eastern Region, and not an attempt to overthrow civilian government (TNA, DO 195/280/48, DO 195/281/115B, DO 195/282/155A).[18] Thus it was on 8 December 1965 that Colonel Tom Hunt, the Defence Advisor to the High Commissioner and former Chief of Staff of the Nigerian Army headquarters, circulated a memo on the "Possibility of an Army Coup d'Etat" (TNA, DO 195/283/251). He concluded:

> Most African countries are preoccupied with the possibility of a military coup d'etat, and no doubt with good reason... I have reported on a number of occasions about the prospects of this happening and I am still firmly of the opinion that this is unlikely at the present moment... The senior ranks of the Nigerian Army at present consist of officers around the age of forty, many of whom have seen long service in their youth... They are not particularly politically minded and most of them are well satisfied with their new status and their high rate of pay (by African standards). I do not believe that this type of more senior officer has any compelling urge to change the present political regime, however rotten it may appear to be to them. Below the GOC [General Officer Commanding], Brigadiers and full Colonels, there are some [Lieutenant] Colonels and a number of Majors and Captains who, in my opinion, are likely to become increasingly restive... [but] I think the pressures have not yet sufficiently built up within the Army to generate any immediate threat of political action... I have discussed this topic with a number of people, including the Commodore of the Nigerian Navy yesterday, and there seems to be a consensus of opinion that there is no imminent danger at the present.

17. In the days following the coup, much of the intelligence passed through Lagos was concerned with the possibility that the culprits were leftists in cahoots with Ghanaian President Kwame Nkrumah (TNA, DO 195/294/23, DO 195/294/110).
18. Even if the East did secede, it did not seem likely to provoke war. For example, the former secretary to the Sardauna, Bruce Greatbatch, informed John Moreton of the High Commissioner's office that the NPC and Northern emirs would not resist Eastern secession (TNA, DO 195/280/45).

Six weeks later, seven Majors and a Captain ended the First Republic (TNA, DO 195/294/98A).

Like the British, Nigerian political elites had failed to recognize the extent of mid-ranking officers' dissatisfaction. For example, Archbishop Makarios III, the first President of Cyprus, was in Enugu as a guest of the Eastern Government. By chance, he was with both the Governor and the Premier of the East as the coup unfolded, leaving just hours before an attempt was made on the latter's life (TNA, DO 195/290/48A). Further, the day before the coup, the Attorney General of the Federation, Taslim Owalawe Elias, was anonymously tipped off, and specifically about the threat to Tafawa Balewa's life. But when he informed Tafawa Balewa, the Prime Minister brushed him off and returned to work (TNA, DO 195/294/10, DO 195/294/84).

Even senior Nigerian military officials failed to foresee the January plot, despite it being planned and executed by their subordinates, and directly targeted at them. As the coup was primarily inspired by the frustration of junior officers with their superiors, *all* of the top brass were on the "list for liquidation," including every Lieutenant Colonel and above. Dozens of them were killed or wounded in the coup (TNA, DO 195/294/16, DO 195/294/17, DO 195/294/51A, DO 195/294/64). The two most senior surviving officers were Brigadier Babafemi Ogundipe, on assignment in London, and the GOC, Major-General Johnson Aguiyi-Ironsi. After being phoned by a Lieutenant Colonel under duress (himself later killed), Ironsi raced to the headquarters of the federal guard, quickly established control, and at the prodding of a high-ranking British expatriate officer, assumed unified command over the Army and Police. In the vacuum the coup had created, he found himself a reluctant head of state (TNA, DO 195/284/49, DO 195/294/58A, DO 195/294/93A, DO

195/294/109A, DO 195/294/110).[19] The revolt was so sudden, and so unexpected, that Ironsi's only recourse was to mount the horse that had bolted and begin governing.

With both civilian and military elite networks in complete disarray, Ironsi's new military government suspended the Constitution, dissolved the federal legislature, and abolished the regions entirely. Riots immediately spread through the North, encouraged by the remaining emirate leaders (TNA, CAB 128/41/29). Refusing to accept a centralized, unitary government, Northern soldiers staged a counter-coup in July, "returning effective power to the Regions" due to the reality that the new head of government, Colonel Yakubu Gowon, had "little more than nominal control of either the country or the Army" (TNA, PREM 13/1040).[20] Again riots spread through the North as thousands were massacred in anti-Igbo pogroms, sponsored, aided, and abetted by Northern elements in the Army and Police (TNA, CAB 128/41/48, CAB 128/41/50, PREM 13/1041, PREM 13/1662, PREM 13/2259).[21] Expectations of Eastern secession slowly grew over the following six months, and by April 1967, the East was siphoning oil royalties rather than sending them to the center as stipulated in the Constitution, in clear preparation for civil war (TNA, CAB 128/41/52, CAB 128/42/17). Trying to hold the country together, Gowon declared the division of the country into twelve states on May 27, 1967 (TNA, PREM 13/1661), as well as a state of emergency in the East so as to enforce this decree.

19. Ironsi's eventual successor, Colonel Yakubu Gowon, himself only narrowly survived the January plot. He was on the list for liquidation, but his would-be assassin was stopped by an alert security guard (TNA, DO 195/294/109A).
20. On 4 October, Cumming-Bruce informed Prime Minister Harold Wilson that "Gowon and Central authorities are no longer able to exercise any effective control in North," noting that their sole means of communication was a single morse radio used for police dispatches (TNA, PREM 13/1662).
21. After Gowon refused to accept eyewitness reports presented to him directly by Cumming-Bruce, the latter wrote to Sir Morrice James of the CRO (TNA, PREM 13/1662), "I am afraid that my main impression of [Gowon] is that he simply is not up to his responsibilities... He refuses to face up to the stark facts of the scale of brutalities in the North, and the extent of the Army's positive responsibilities for them."

On May 30, 1967, Military Governor Colonel Chukwuemeka Ojukwu announced the secession of the "Republic of Biafra"—the former Eastern region—and the Nigerian Civil War began. Over the next three years, 100,000 soldiers and at least 500,000 civilians died, the latter overwhelmingly from starvation (TNA, CAB 128/43/46, CAB 128/43/49, CAB 128/44/57, PREM 13/2258, PREM 13/2259, PREM 13/3380).

The breakdown of the First Republic and the Biafran tragedy indicate the power that deconcentration has to destabilize coalition structures. At independence, Nigeria's problems were by no means limited to inter-regional struggle. But the three-region, three-party system that had held for the better part of two decades was, ironically, perhaps the strongest centripetal force in the First Republic. By forcing the parties to develop extraregional allies, the tripartite coalition structure imposed stability on national political competition. In creating the Mid-West, the leaders of the NPC and NCNC achieved their desired goals of reshaping political competition into a two-party contest, but they also upended the "three-legged stool" that had kept Nigeria stable for two decades. Ultimately, millions of Nigerians paid the price for the coalitional realignment effected by Mid-West statehood.

5.4 Analysis

5.4.1 The role of institutions

The Mid-West was the only new region created in the First Republic. Although coalitional considerations were the decisive factor in determining the fate of the Big Three statehood movements, variation in the institutional context also played an important role. The constitutional provisions governing both deconcentration and the procedures for forming a ruling coalition—in the language of Chapter 2, the

voting and majoritarianism rules—contributed to these different trajectories.

Under British rule, the minimal requirements for deconcentration were fairly informal: the Secretary of State for the Colonies had an effective veto over any proposal, as well as the power to enact new regions unilaterally, so deconcentration simply required convincing him that it was good policy. However, no such proposals were suitable to the British. So assured of the British veto was the AG that, in 1954, it passed a resolution in the regional House of Assembly backing the Mid-West in an attempt to solidify Mid-Western support before a regional election. Marshall had summarized Rankine's reports of this calculus (TNA, CO 879/192, 18):

> Although his conversations with Mr. Awolowo on the proposal for a Mid-West State had not shown clearly what Mr. Awolowo would do if that idea were rejected by Her Majesty's Government, it might well be that he and his colleagues would not be ill content with such an apparently adverse decision. They must support the proposal now; but if they could claim credit for supporting it and blame Her Majesty's Government for making it impossible of fulfilment, they might not be unwilling to accept the resulting situation.

Three years later, inviting Willink to chair the Minorities Commission, Lennox-Boyd wrote that he "felt bound to warn the Conference [of 1957] that there could be no easy solution of minority problems through the liberal creation of new Regions or States" (TNA, CO 879/192, 154). Lennox-Boyd reiterated this stance in 1958 when he told participants at the resumed talks that "there can be no question of new states being created before independence" (TNA, CAB 129/95/13). With his approval both necessary and sufficient for deconcentration, Nigerian party leaders could propose a wide variety of statehood plans, safe in the knowledge that the British would block all of them equally.

Once Nigeria attained independence, the bar for deconcentration was set by the constitutional provisions for creating new regions. Neither Middle Belt nor COR

statehood were feasible because they both failed to satisfy this very high bar. A Middle Belt region would inevitably diminish the NPC's overall share of seats in the federal House of Representatives and eliminate the numerical superiority protecting it from southern domination. In effect, deconcentration in the form of a Middle Belt region seemed guaranteed to remove the NPC from the ruling coalition, and so the NPC never gave its assent to this deconcentration proposal. Equally, a COR state would have split the NCNC in half, resulting in two smaller parties, neither of which would be able to challenge the NPC or AG on a national stage, guaranteeing a much smaller share in the spoils of joining the ruling coalition. Again the effect was to deny the COR movement the required assent of Eastern politicians.

Equally, the sole impediment to the Mid-West was the majority assent of the Western legislature. But the AG crisis of 1962 provided an opening during which this Western veto was effectively suspended. The federally-appointed administrator dropped the Western lawsuits holding up the plebiscite, advancing Mid-Western statehood efforts beyond the point where Middle Belt and COR movements had gotten stuck. As if it had been "pre-ordained" (Akinola 2014, 50), the AG crisis allowed the NPC and NCNC leaders to overcome the last constitutional hurdle, satisfying the voting rule and paving the way for deconcentration.

5.4.2 Alternative explanations

Beyond demonstrating the logic of deconcentration as coalition manipulation, the statehood movements of Nigeria's First Republic provide a useful test of alternative theories of deconcentration. First, it is possible to imagine that the Mid-West was created because its geography, demographics, or historical context differed sharply from that of the Western region. In this telling, creating the Mid-West

served to "correct" the mistaken grouping of Mid-Westerners under Yoruba rule. As the foregoing discussion suggests, this explanation is unpersuasive because, while containing some truth, it cannot account for the different trajectories of the Big Three statehood movements. Among these, the Middle Belt's cultural and institutional landscape diverged more sharply from the core North's than did that of the Mid-West from the Yoruba West's. Nigeria's fourth region was the Mid-West, and not the Middle Belt (or Lagos State, Calabar State, or any other group with a more distinct cultural and institutional identity) suggesting that this lens is less useful for understanding Nigerian statehood politics.

Second, in a similar vein, theories of federalism and decentralization suggest that new subnational units of government are created to further goals like economic development or improved public services. Such explanations make little sense in the context of the First Republic. The Mid-West was known to be non-viable, as recognized in the Willink Commission and admitted even among Mid-West advocates (Vickers 2000). The natural resources for which the Mid-West is now known, and in particular the petrochemical industry, did not exist at this stage of Nigerian history. It was not until 1965 that commercial oil production appeared on the country's horizon, when Tafawa Balewa predicted it would double gross national product over the next five years (TNA, PREM 13/451). Even then, production was concentrated in Calabar province (TNA, CO 879/192, 35). Thus, a Mid-West shorn of Western subsidies would barely be in a position to fund a regional government, much less invest the capital required to propel economic development. If viability determined movements' success, the COR region certainly would have made a better candidate than the Mid-West did in 1963.

Third, scholars may be concerned that the Mid-West's success merely reflects the better organizational strength of its proponents, particularly activists based in

Benin. However, the Middle Belt statehood movement was larger and stronger than the Mid-West movement ever was. Joseph Tarka, founder of the Tiv Progressive Union and leader of the UMBC, won more votes than any politician in Nigeria in the 1959 federal election. While Mid-West parties struggled to win seats in national, regional, or local elections throughout the early independence period, the UMBC won 25 seats in 1959—fully one-third of the seats in the lower North—including all 7 Tiv seats. Similarly, in the 1957 Eastern Regional elections, COR elements won 44% of the vote in the proposed region against the NCNC's 41%—good for 18 seats and a majority (TNA, CO 879/164, 49). If organizational strength determined statehood outcomes, then by all accounts deconcentration in Nigeria would have produced a Middle Belt or COR region before it did a Mid-West.

A final alternative perspective might suggest that deconcentration was an attempt to manipulate electoral outcomes by rigging a majority in the soon-to-be-held 1964 federal elections. However, it is unclear how creating the Mid-West would help further this goal. Considering its relatively small population, the Mid-West did not provide enough winnable seats to meaningfully advance NPC or NCNC electoral strategies. Further, as discussed above, census manipulation and electoral maladministration provided much cheaper and more effective means of manufacturing a victory. It is difficult to explain the emergence of the Mid-West from this perspective.

5.4.3 Threats to inference

The evidence presented above suggests two primary threats to inference. The first is that it is possible for the NPC and NCNC to have engineered the AG crisis, making deconcentration epiphenomenal to the broader partisan conflict playing out.

This interpretation is unconvincing for two reasons. First, even if the AG crisis was the work of Tafawa Balewa, Zik, or their allies, deconcentration still played the decisive role in permanently fracturing the AG and realigning coalition structures. Without the Mid-West, the Awolowo-led AG would have likely recaptured its Western majority at the end of the period of emergency in January 1963 (TNA, DO 195/281/71). Second, the archival materials suggest that the AG crisis arose naturally from forces unrelated to coalition bargaining among the major actors. As discussed above, the main factors driving the conflict were policy disagreements and interpersonal differences among AG elites. The Awolowo faction had won the power struggle the previous week, and had everything to gain from the peaceful installation of Adegbenro (TNA, PREM 11/3893). Responsibility for the violence was clear: the man who had given the signal was the representative from Akintola's home town of Ogbomosho, and a close ally. British observers noted that "but for one man's stupidity it could have been avoided": if Akintola had just accepted the decision of the AG to remove him, then it never would have come to pass (TNA, PREM 11/3893). And while Tafawa Balewa had initially been pleased by the AG's internal strife, he quickly soured on it. On the day violence broke out on the floor of the Western House, David Hunt wrote to the CRO that Tafawa Balewa was "now deeply incensed by damage to Nigeria's good name caused by today's disgraceful scenes... He is aware responsibility for them lies entirely with [the] Akintola faction plus NCNC Opposition [sic], and has now lost sympathy with Akintola (TNA, PREM 11/3893). The best interpretation of the available evidence is that, however fortunate for Mid-West activists the AG crisis may have been, it was not deliberately manufactured to produce deconcentration.

A second threat to inference lies with the reliability and accuracy of the archival materials themselves. It may be that the story of the First Republic is

obscured by inaccuracies or biases in the British perspective. While it is impossible to determine with certainty whether such bias is a problem, a few factors suggest that these documents present a reasonable approximation of the truth.

For one, the majority of documents studied here are factual reports from on-the-ground observers, reporting up the chain to the High Commissioner, the Secretary of State for the Colonies (or Commonwealth Relations), and ultimately the Prime Minister. The express purpose of these reports was to provide unvarnished information, not to recommend policy or promote particular actors' interests. The incentives for the actors producing these data are therefore structured to reward accurate inferences about local intelligence reports, and while the "fog of war" at times generated substantial confusion, the authors of these reports faithfully recorded their uncertainty and corrected prior mistakes.

Further, the authors of these reports were themselves skilled civil servants with vast in-country experience and excellent sources. Examples of British officials' ability to get very close to the inner workings of Nigerian politics abound. For instance, Barltrop had as a confidant the aide to the Western Governor who personally delivered the letter of dismissal to Premier Akintola in May of 1962 (TNA, PREM 11/3893). During the crisis following the 1964 federal election, Tafawa Balewa confided in the GOC of the Nigerian Army, Christopher Welby-Everard (a British expatriate), who then relayed the maneuverings at nightly meetings at the High Commissioner's private residence (TNA, PREM 13/451). This level of access was particularly apparent during the coup of January 1966. The High Commissioner was notified just hours after the revolt began, by Deputy Inspector-General of the Police Leslie Marsden. The latter was then the "moving force and guiding hand" in getting Ironsi to organize a military government as the coup was still unfolding, providing an insider's view of the power vacuum effected by the mutiny (TNA, DO 195/294/45,

DO 195/294/109A). Given Nigeria's strategic importance to the United Kingdom both before and after independence, the British government invested considerable resources in its intelligence and diplomatic network in the country, leaving a reliable record of Nigerian politics of the day.

Finally, all indications suggest that very little of the intelligence gathered during this period has been redacted or destroyed. British laws and regulations governing archival records leave very few exceptions to making these documents public, particularly given their age. In my experience, only records relating to the royal family's finances were removed, with codeword-classified documents freely available. On the rare occasion intelligence was redacted for security reasons, it was done in-line, with the rest of the document still legible (e.g., TNA, PREM 13/3380). Together, these factors suggest that the data-generating process for these archival materials is one that produces a complete picture of Nigerian statehood politics from 1946 to 1966.

5.5 Summing up

Awolowo opened the 1953 constitutional conference by telling leaders from the North, East, and West that "we have all come here, like surgeons, to perform a major operation on the ailing corpus of the Macpherson Constitution." To the AG and the NCNC, the problem was clear: Nigeria's tripartite federation was destined to fail so long as the North remained united. It is a fact "which cannot be gainsaid," he continued, "that the Northern Region as a bloc is large and powerful enough to override the wishes of the other two regions combined, or at the very best to make the wishes of the other Regions barren and unfulfilled" (TNA, CO 879/159, 6-13). Yet it was not Nigeria's sprawling North or densely populated East that was split in

half in the First Republic, but rather the West, the smallest and least populous of the three.

Statehood politics in Nigeria over the 1946-1966 period is instructive for scholars of deconcentration. The central force driving the success or failure of statehood proposals was coalitional politics, conditioned by the institutional context. The new region that emerged only did so because it transformed Nigeria's three-party system into a two-party system, and it was only possible when an important veto player was temporarily sidelined. In contrast, neither the Middle Belt nor the COR regions were likely to produce coalition structures that made the remaining parties better off, and at no time did either movement win the necessary consent of the Regional actors in the North or the East. Comparing these statehood movements before and after independence provides a clear view of the centrality of coalitional bargaining, and the institutions structuring that bargaining, for deconcentration.

This analysis provides strong evidence for the mechanisms identified in my theoretical framework. The archival materials studied in this chapter uncover a micro-logic of state creation in Nigeria that corroborates the broad, cross-national patterns presented in earlier chapters. Taken together, this evidence provides robust empirical support for the theory of deconcentration as coalition manipulation.

5.A Appendix

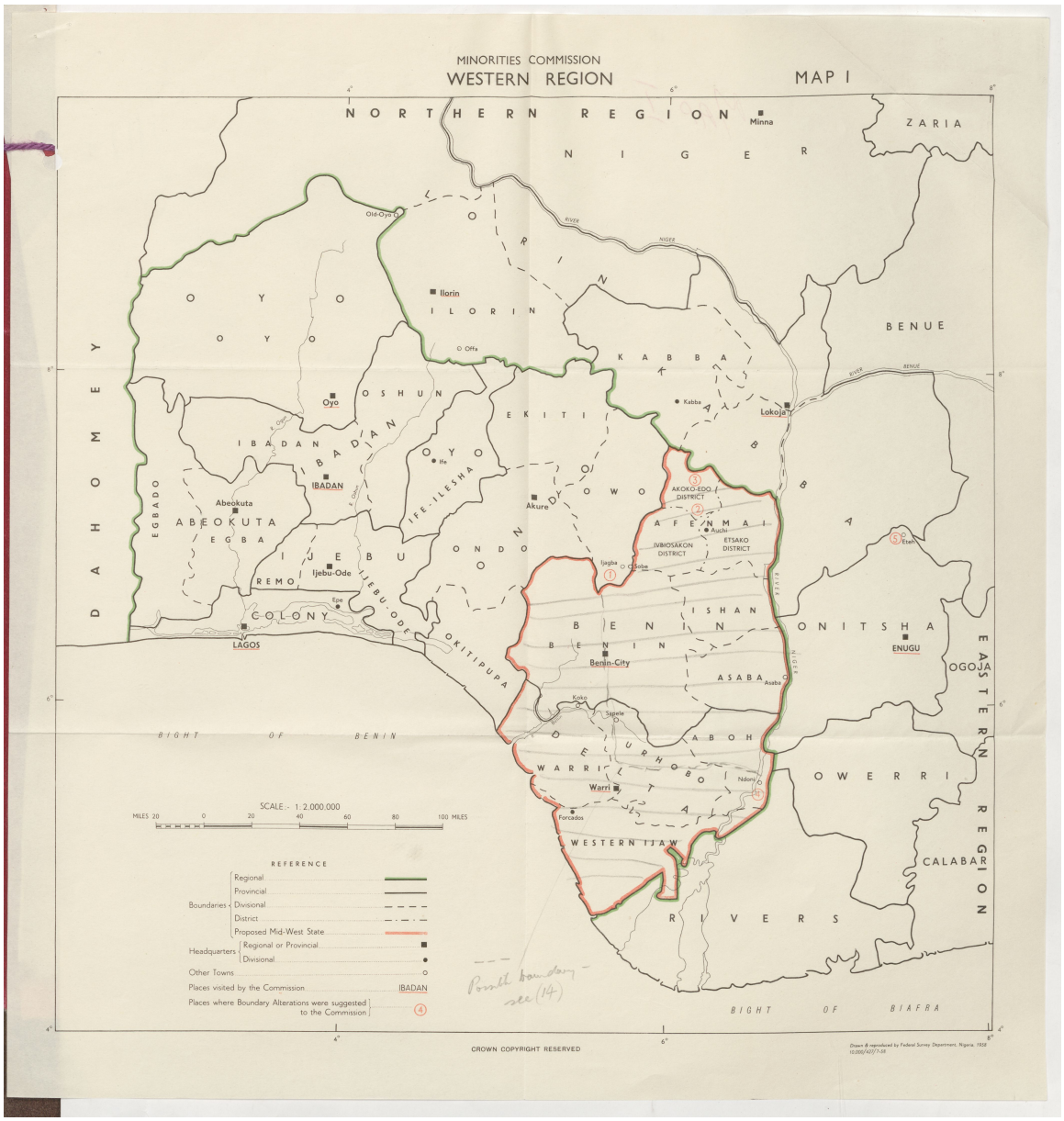


Figure 5.6: The Mid-West Region. A map of the proposed Mid-West Region, outlined in orange, as presented in the Willink Commission Report of 1958 (TNA, DO 177/42). See Figure 5.2 for a retouched version.

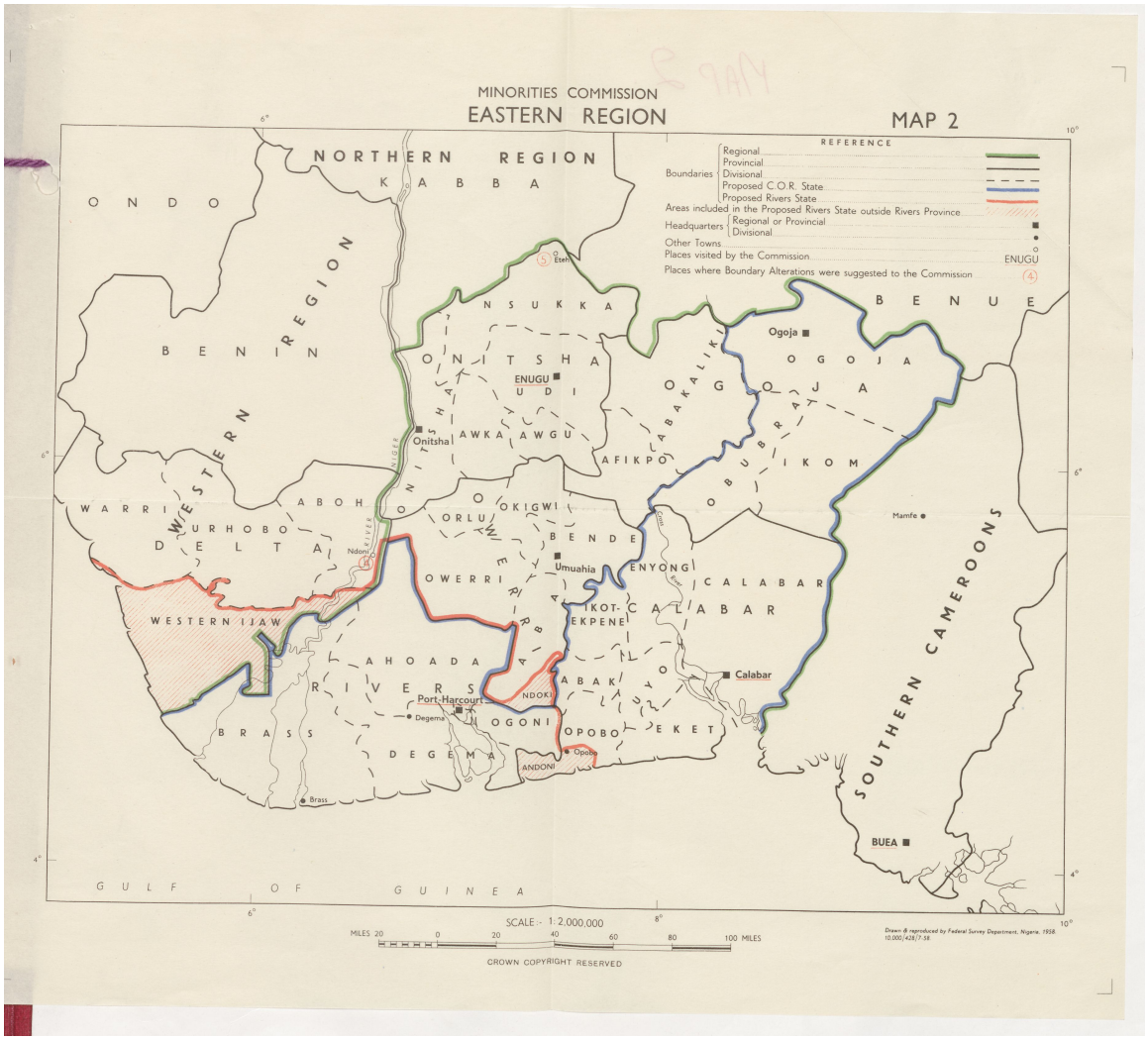


Figure 5.7: The COR Region. A map of the proposed COR Region, outlined in blue, as presented in the Willink Commission Report of 1958 (TNA, DO 177/42). See Figure 5.3 for a retouched version.

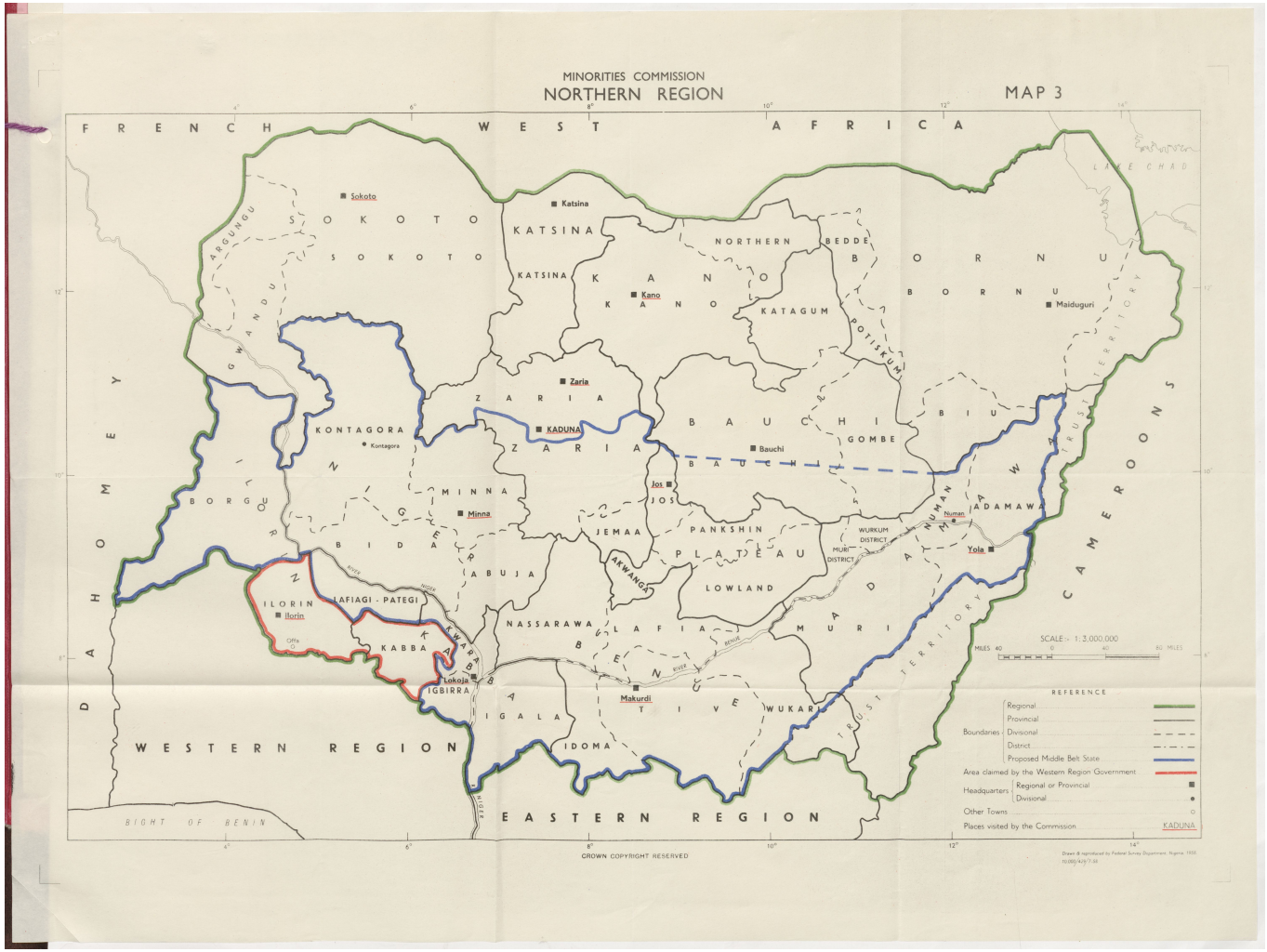


Figure 5.8: The Middle Belt Region. A map of the proposed Middle Belt Region, outlined in blue, as presented in the Willink Commission Report of 1958 (TNA, DO 177/42). See Figure 5.4 for a retouched version.

6 | Conclusions

> The old schism of federal & republican, [sic] threatened nothing because it existed in every state, and united them together by the fraternism of party. But the coincidence of a marked principle, moral & political with a geographical line, once concieved, [sic] I feared would never more be obliterated from the mind; that it would be recurring on every occasion & renewing irritations until it would kindle such mutual & mortal hatred, as to render separation preferable to eternal discord. I have been among the most sanguine in believing that our Union would be of long duration. I now doubt it much, and see the event at no great distance, and the direct consequence of this question.
>
> —Thomas Jefferson, in a letter to William Short, 13 April 1820 (LC, MTJ/BIB/023789)

The spring of 1820 saw American democracy in peril. Victory in the War of 1812,[1] the fall of the Federalists, and their amalgamation into the dominant Democrat-Republican party had created a brief Era of Good Feelings (Dangerfield 1952). But factional resentments had quickly spoiled the pervasive sense of national unity by pushing statehood politics to the fore. When a bill to add Missouri to the Union came up for debate in 1819, Representative James Tallmadge, Jr. proposed an amendment that stipulated its status as a "free state," wherein slavery would be banned. Intraparty warfare immediately broke out. Southern, Jeffersonian Republicans teamed up to resist what they saw as unconstitutional overreach of Northern Republicans, many of whom had previously been Federalists. After raging for more than a year, the conflict was finally resolved in February 1820 when leaders

1. Historians disagree on which participants (if any) could claim to have won the war, which ended in a military stalemate, but agree that Americans believed themselves victorious in the aftermath (Hickey 2012).

in the House and Senate negotiated the simultaneous passage of statehood for Maine, as well as the partial outlawing of slavery in what remained of the Louisiana Territory. The Missouri Compromise may have temporarily saved the Union, but the Era of Good Feelings was over.

This episode is a well-known precursor to the slave state-free state rivalry that would culminate in civil war forty years later. Lost in this narrative, however, is the role of the Missouri Crisis in realigning the party system. That factional interests were at the root of the conflict was obvious to contemporaries. President James Monroe wrote to former President James Madison in February 1820 that he feared the entire affair was a plot hatched by erstwhile Federalists to take down the Democrat-Republican party from within (LC, MJM/022697).[2] Madison shared his suspicion (LC, MJM/018701):

> It appears to me as it does to you, that a coupling of Missouri with Maine, in order to force the entrance of the former [through] the door voluntarily opened to the latter is, to say the least, a very doubtful policy. Those who regard the claims of both as similar & equal, and distrust the views of such as wish to disjoin them may be strongly tempted to resort to the expedient... I find the idea is fast spreading that the zeal [with] which the extension, so called, of slavery is opposed, has, with the coalesced leaders, an object very different from the welfare of the slaves, of the check to their increase; and that their real object is, as you intimate, to form a new state of parties founded on local instead of political distinctions; thereby dividing the Republicans of the North from those of the South, and making the former instrumental in giving to the opponents of both an ascendancy over the whole.

In particular, Southern Republicans saw in Rufus King the architect of the Missouri Crisis. The Senator for New York, a signer of the Constitution and the last nationally prominent Federalist, was assumed to be pulling the strings behind Tallmadge's amendment and the subsequent organization of Northern Republicans. As Jefferson

2. "LC" refers to archival materials available through the Library of Congress.

concluded in a letter to Monroe, "King is ready to risk the union [sic] for any chance of restoring his party to power and wriggling himself to the head of it" (LC, MTJ/BIB/023743).[3]

The realignment kickstarted by the Crisis would take four years to unfold. Monroe's conciliatory efforts kept factional warfare contained within the party for the rest of his term, but these fault lines finally burst open during the presidential election of 1824. Four candidates ran, all drawn from the Democrat-Republicans, and each representing a different party caucus. When none of them managed to obtain a majority in the Electoral College, an 1825 contingent election in the House of Representatives installed former Federalist John Quincy Adams as President. The party splintered, and by 1828, the factions had reorganized into two new parties. Andrew Jackson, buoyed by a perceived "corrupt bargain" that ushered Adams into office despite Jackson's plurality of electoral votes, became the *de facto* leader of the Democratic Party. Philosophical descendants of the Federalists coalesced around the National Republican (later Whig) Party. This new party system was organized along much the same lines as it had been two decades earlier, with conflict centering on issues such as the extent of federalism, regionalism, and slavery. The "corrupt bargain" finished the coalitional realignment that the Missouri Crisis had started.

6.1 Divide to rule

Episodes like the Missouri Crisis are common around the world. Democracies and nondemocracies alike have split districts and carved out new regions nearly 400 times

3. Jefferson's dismay was apparent throughout the spring of 1820. On 22 April, he wrote that the "Missouri question" had, "like a fire bell in the night, awakened and filled me with terror. I considered it at once as the knell of the Union... I regret that I am now to die in the belief that the useless sacrifice of themselves, by the generation of [1776] to acquire self government and happiness to their country, is to be thrown away by the unwise and unworthy passions of their sons, and that my only consolation is to be that I live not to weep over it" (LC, MTJ/BIB/023795).

across 126 countries since 1960. Deconcentration has frequently brought renegotiation of the intergovernmental balance of power, intraparty conflict, and constitutional crisis. And it has also often brought civil war. The fear that deconcentration would "be renewing irritations until it would kindle such mutual & mortal hatred, as to render separation preferable to eternal discord" was just as surely felt by Nigerians in 1963 as it was by Thomas Jefferson in 1820.

Scholars have only recently taken notice of deconcentration. The emerging consensus, if one exists, is that countries create new subnational units to distribute patronage and rig majorities. However, this explanation overpredicts deconcentration and fails to explain the intraparty conflict that it generates. In this dissertation, I have offered an alternative theory: deconcentration is used to manipulate political coalitions. Creating new units empowers actors who can change the outcome of coalitional bargaining. Politicians' preferences over deconcentration are induced by their preferences over alternative coalition structures. Whether deconcentration occurs depends both on these induced preferences and the institutions that underpin the bargaining.

I tested the implications of this theory using novel data and a variety of methods. A large machine learning task in Chapter 3 uncovered cross-national patterns broadly consistent with my predictions: that deconcentration is more likely where there are more existing subnational units, fewer veto players, higher majoritarian thresholds, and during periods of coalitional transition. In contrast, this analysis found little evidence to support alternative explanations. In Chapter 4, I examined whether these relationships were causal. Using sudden leadership deaths as exogenous shocks, I showed that coalitional instability leads politicians to pursue deconcentration as a means of directing bargaining toward more favorable new coalition structures. Finally, tracing Nigerian statehood movements over

the 1946-1966 period, I found significant evidence of the coalition-manipulation mechanism at work. This close case comparison indicated that coalitional politics drove variation not only in *which* movement was ultimately successful, but also *when* it was, in Nigeria's First Republic.

6.2 Implications

This study has offered an explanation for deconcentration that accounts both for the puzzling alliances that emerge in conflicts over subnational units, and for the institutions that determine the outcomes of those conflicts. Along the way, the arguments and evidence have suggested contributions to a number of scholarly debates. Most immediately, my theory is the first explanation of deconcentration as a general phenomenon. A number of case studies have proposed answers limited to the context from which they are drawn, but by taking a more comparative approach, I have provided new ways to think about deconcentration's "hard cases" which do not fit previous explanations. My model also offers an extension of selectorate theory, in which both the governing coalition and those who choose it are negotiated simultaneously. Further, focusing on the endogenous evolution of institutions and coalition structures generates new ways to think about the origins of party systems. Finally, my theory of deconcentration suggests that coalitional considerations can motivate politicians to trade away their own power, helping answer a critical question in political economy.

6.2.1 Deconcentration's hard cases

Previous studies have proposed explanations for individual episodes of deconcentration in countries such as Indonesia, Uganda, the US, and Vietnam. This dissertation took

a broader perspective, offering a comparative analysis of global deconcentration over the 1960-2010 period. My theory provides the first explanation for deconcentration as a general phenomenon.

This broader focus facilitated more reliable inferences about the causes of deconcentration. I demonstrated theoretically that the coalition manipulation mechanism is applicable across regime types, and in particular, is not dependent on electoral competition. I also outlined the institutional constraints on deconcentration, clarifying why strong statehood movements like that of Nigeria's Middle Belt routinely fail. Empirically, by introducing new data on global deconcentration, I was able to draw on a larger sample, and one in which the observations had not been selected according to the outcome. This strategy allowed me to rigorously test the coalition manipulation theory against alternative hypotheses, which a narrower study would not have allowed. Together, these contributions provide a better vantage point from which to assess the many other "hard cases" of deconcentration that scholars have yet to take up.

6.2.2 Selectorate theory

My model demonstrates how deconcentration occurs endogenously as part of broader coalition bargaining. In the language of selectorate theory, politicians can simultaneously negotiate winning coalitions as well as membership in the selectorate, the broader political elite from which these coalitions are drawn. While it is well known that altering the selectorate implies contestation over the winning coalition, my model suggests that the reverse also holds: altering the winning coalition implies contestation over the selectorate. Renegotiating one entails renegotiating the other.

Previous studies have taken the selectorate as exogenously-given (Bausch 2017; Bueno de Mesquita et al. 2003; Shirk 1993). This approach leaves rulers and politicians unrealistically limited in the coalition structures they can credibly offer. In contrast, my model demonstrates how scholars can better account for coalitional possibilities by endogenizing the selectorate itself. Allowing for more flexibility in *who decides* the shape of the ruling coalition will enrich our understanding of how these coalitions emerge.

More generally, in treating the selectorate as fixed, scholars have artificially constrained politicians' ability to manipulate institutions in pursuit of more favorable winning coalitions. This focus on coalition bargaining divorced from institutional choice has invited criticism that selectorate theory ignores important contextual variation, and is therefore a "blunt instrument" (Gallagher and Hanson 2015). My model suggests one way to sharpen these analyses, by allowing coalition bargaining to include renegotiation of the institutions structuring that bargaining. Most immediately, my theory suggests that scholars need to expand the two-dimensional institutional space of selectorate theory to include a third: the voting rule, which determines deconcentration veto players and therefore constrains the shape of the selectorate itself. Similarly endogenizing other institutional details in a broader model of coalition bargaining can improve selectorate theory's ability to account for the logic of political survival across diverse contexts.

6.2.3 Party systems and electoral institutions

My theory also recasts the relationship between electoral institutions and party systems. According to strict institutionalists, parties organize in response to electoral and constitutional rules, including the number of representatives elected in each

district, the degree of legislative proportionality, the extent of decentralization, and the mechanism for executive selection (Chhibber and Kollman 2004; Cox 1997; Duverger 1954; Shugart and Carey 1992). More recent scholarship has flipped this argument on its head, documenting parties' strategic selection of these same institutions (Boix 1999; Colomer 2005; Robinson and Torvik 2015). In this account, parties emerge from a mix of economic, historical, and sociological factors (Kitschelt 2000; Lipset and Rokkan 1967). They then choose institutions that further their private interests, manipulating such rules so as to entrench themselves in office and extend their hold on the levers of government.

Ample evidence exists to support both sides of the debate. It is clear that parties organize and coordinate in response to the institutional context, but also seek to alter these rules in any way that might improve their odds of winning elections. The challenge for scholars is developing a theory of party systems and electoral institutions that account for both processes.

My model provides a blueprint for addressing this problem. I demonstrated that coalitions such as parties and factions can be negotiated alongside the institutions of subnational government. By endogenizing deconcentration to coalition bargaining, I showed how institutional change can coincide with shifts in the composition of the party system. And in focusing attention on the coalition manipulation mechanism, I provided an approach that can help scholars reconcile divergent perspectives on the origins of party systems. Future research can follow in this vein by modeling simultaneous bargaining over the shape of electoral institutions and party organizations.

6.2.4 Giving away power

A final contribution of this study is in demonstrating one condition under which politicians can be persuaded to give up power. Among the key insights of the twentieth century's "new institutionalism" is the argument that political institutions are constructed by self-interested actors who benefit from their distributive implications. These actors endlessly accumulate power so as to maintain their control over these institutions, as well as to stave off competitors who seek to redirect rents to themselves (Bates 1989; Knight 1992; Moe 2005; North 1990). In consequence, few political actors can be persuaded to give up or trade away their power. Scholars have discovered few general mechanisms for convincing actors to cede this authority.

My model highlights one under-studied mechanism: politicians may be willing to cede power if it improves their coalitional position. Deconcentration may entail giving up territory, centrally-controlled rents, patron-client networks, constituents, and a share of legislative or executive authority. Yet for some politicians, these sacrifices may be worthwhile if the coalitional incentives are strong enough. In particular, politicians may be willing to trade off their power within the polity as a whole if they stand to become more powerful within their respective coalitions.

This coalitional perspective provides one answer to the puzzle raised by new institutionalists, one which can be deployed in a number of areas of scholarly inquiry. Studies of post-conflict disarmament, demobilization, and reintegration struggle to find mechanisms by which combatants can be convinced to give up arms (Autesserre 2010; Walter 2015). Theories of transition from military to civilian rule get stuck on what makes such transfers of power credible (Acemoğlu, Ticchi, and Vindigni 2010; Barany 2015). And although recent scholarship has provided nuanced insights into the mechanisms of nondemocratic rule, we still lack an understanding of why

actors cede power to *particular* institutions such as parties or legislatures, and when they do (Pepinsky 2014). The coalition-manipulation mechanism developed here may shed new light on some or all of these different manifestations of the same general puzzle. Much more research is needed in order to understand how and when coalitional considerations can motivate actors to give power away.

6.3 Directions for future research

This study also suggests four questions for further investigation. First, does the coalition manipulation mechanism provide any leverage over deconcentration at the lowest levels of subnational government? Data constraints have limited the empirical evidence presented here to top-level units such as states or regions. But politicians often reorganize lower-level subdivisions. For instance, the number of local government areas in Nigeria has tripled since independence, while Brazil's 1,574 municipalities in 1940 ballooned to 3,952 in 1970 and 5,560 in 2000 (Suberu 2001; Tomio 2002). Such units frequently provide valuable services such as maintaining roads, building schools, and distributing grants tied to social programs (Crawford and Hartmann 2008; Hassan and Sheely 2017; Soares, Ribas, and Osório 2010). Studying lower-level deconcentration is therefore just as important as is studying top-level deconcentration for understanding citizens' lived experience of government.

I have implicitly claimed that coalition manipulation can account for deconcentration among top- and lower-level units. In Chapter 2, I argued that the institutional and partisan ties binding politicians together work across these levels. Even if municipal officials are relatively weak, their empowerment seems just as likely to upset coalitional bargaining as does the creation of top-level units, and so the causal mechanism remains the same. However, this implication of the theory

has gone largely untested in this study. Anecdotal evidence suggests that local boundaries are chosen with the same coalitional logic in mind (e.g., Eaton 2004), but only with better data will it be possible to establish whether my model is useful for explaining deconcentration at the lowest levels of government more generally. Future research is needed to determine if my model travels downward to the more complex institutional environment in which lower-level deconcentration is situated.

Similarly, can coalitional politics tell us anything about why countries join supranational governments such as the European Union (EU)? Although both involve extending the number of government units, EU accession differs qualitatively from deconcentration, not the least because individual countries can only join or leave—and even then, only through protracted negotiations with foreign governments. Scholars have typically focused on accession as being driven by such forces as corporate activism, ideational and cultural factors, economic and monetary interests, and politicians' electoral incentives (Featherstone and Radaelli 2003; Schimmelfennig and Sedelmeier 2002; Schneider 2017). Yet joining the EU has clearly shifted coalition structures in countries such as Denmark, Finland, and Sweden (Johansson and Raunio 2001; Vachudova 2005). Coalitional considerations may be another factor that enters into the political calculus of accession. Regional integration involves enormous complexity, and thus substantial uncertainty, so coalition manipulation is unlikely to be the only (or even the main) causal force. Still, greater emphasis on coalition bargaining may shed new light on why countries join the EU, and the distributional implications thereof.

Third, what insights can the coalition manipulation mechanism provide into *concentration*, the amalgamation or elimination of existing subnational units of government? Scholars have documented a wave of "municipal mergers" across the developed world in the last fifty years (Baldersheim and Rose 2010; Blom-Hansen,

Houlberg, and Serritzlew 2014; Forde 2005). Echoing the outcomes theory of decentralization, studies in this literature have generally assumed that such (re-)concentration is driven by the desire to make public service delivery more efficient through economies of scale. However, here too theory and evidence are both mixed. While larger units may deliver services more efficiently, they might also induce problems of scale, particularly communicating and controlling a larger administrative bureaucracy, as well as problems arising from serving citizens with more heterogeneous preferences (Oates 1972; Williamson 1967). It is unsurprising that scholars have found only highly contingent and inconsistent effects of such reforms, prompting one review to conclude that "the quest for an optimal jurisdiction size is futile" (Blom-Hansen et al. 2016, 813).

My theory suggests an alternative reason to concentrate subnational government. Politicians may be able improve their position in the coalition structure by eliminating subnational units, *removing* some of their peers from bargaining entirely. We would expect that, in order for such concentration proposals to succeed, they would need to overcome the unanimous opposition of those who would be disempowered. Thus, (re-)concentration would likely be more common where the voting and majoritarianism rules are both considerably lax—where there are few veto players and small governments can form. The patterns in municipal amalgamation seem to support this prediction, as such reforms are clustered in parliamentary democracies where units can be created or destroyed through normal legislative procedure, including Denmark, France, Iceland, Japan, and the UK. Such episodes of concentration provide considerable scope for extending this study of deconcentration into a more general theory of coalitional politics driving the structure of subnational government.

Finally, what can studying coalitional politics tell us about the organization

of representative institutions across regime types? The last two decades have seen an explosion in the study of parties, elections, and legislatures in nondemocracies. The lessons of this literature—broadly, that rulers build institutions to solve problems of asymmetric information and credible commitments (Gehlbach, Sonin, and Svolik 2016)—have not been easily translated to democratic contexts, in which these problems are less dire. But rulers everywhere struggle to build and maintain coalitions. In drawing institutional change into coalitional bargaining, scholars can shine new light on points of convergence across democracies and nondemocracies. Studying coalition manipulation may help make sense of how representative institutions emerge and evolve in democracies and nondemocracies alike.

Archival Sources

LC. MJM/018701. Letter from James Madison to James Monroe, 10 February 1820. The Library of Congress, Washington, D.C.

——. MJM/022697. Letter from James Monroe to James Madison, 5 February 1820. The Library of Congress, Washington, D.C.

——. MTJ/BIB/023743. Letter from Thomas Jefferson to James Monroe, 3 March 1820. The Library of Congress, Washington, D.C.

——. MTJ/BIB/023789. Letter from Thomas Jefferson to William Short, 13 April 1820. The Library of Congress, Washington, D.C.

——. MTJ/BIB/023795. Letter from Thomas Jefferson to John Holmes, 22 April 1820. The Library of Congress, Washington, D.C.

TNA. CAB 128/26. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Wednesday, 27th May, 1953, at 3:00 pm. The National Archives, Kew, London, United Kingdom.

——. CAB 128/30/48. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Wednesday, 11th July, 1956, at 11:00 am. The National Archives, Kew, London, United Kingdom.

TNA. CAB 128/41/29. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 16th June, 1966, at 10:30 am. The National Archives, Kew, London, United Kingdom.

——. CAB 128/41/48. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 29th September, 1966, at 10:00 am. The National Archives, Kew, London, United Kingdom.

——. CAB 128/41/50. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 13th October, 1966, at 10:00 am. The National Archives, Kew, London, United Kingdom.

——. CAB 128/41/52. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 27th October, 1966, at 10:00 am. The National Archives, Kew, London, United Kingdom.

——. CAB 128/42/17. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 6th April, 1967, at 10:00 am. The National Archives, Kew, London, United Kingdom.

——. CAB 128/43/46. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Tuesday, 12th November, 1968, at 11:00 am. The National Archives, Kew, London, United Kingdom.

——. CAB 128/43/49. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 5th December, 1968, at 10:00 am. The National Archives, Kew, London, United Kingdom.

TNA. CAB 128/44/15. Conclusions of a meeting of the Cabinet held in the Prime Minister's Room, House of Commons, SW1, on Thursday, 3rd April, 1969, at 10:00 am. The National Archives, Kew, London, United Kingdom.

———. CAB 128/44/57. Conclusions of a meeting of the Cabinet held at 10 Downing Street, SW1, on Thursday, 27th November, 1969, at 10:00 am. The National Archives, Kew, London, United Kingdom.

———. CAB 129/61/4. "Constitutional Developments in the Gold Coast and Nigeria," memorandum by the Secretary of State for the Colonies, 13 May 1953. The National Archives, Kew, London, United Kingdom.

———. CAB 129/62/35. "The Nigerian Constitution," memorandum by the Secretary of State for the Colonies, 17 August 1953. The National Archives, Kew, London, United Kingdom.

———. CAB 129/87/20. "Nigeria," memorandum by the Secretary of State for the Colonies, 7 May 1957. The National Archives, Kew, London, United Kingdom.

———. CAB 129/94/21. "Nigeria," memorandum by the Secretary of State for the Colonies, 31 July 1958. The National Archives, Kew, London, United Kingdom.

———. CAB 129/95/13. "Nigeria," memorandum by the Secretary of State for the Colonies, 20 October 1958. The National Archives, Kew, London, United Kingdom.

———. CO 879/159. Record of proceedings of the Nigeria Constitution Conference held in London in July and August 1953. The National Archives, Kew, London, United Kingdom.

TNA. CO 879/164. Nigeria Constitutional Conference May–June, 1957: Briefs for the Secretary of State. The National Archives, Kew, London, United Kingdom.

———. CO 879/192. Nigeria: Constitutional and Political Correspondence 1955-1958. The National Archives, Kew, London, United Kingdom.

———. CO 957/1. Chairman's Meeting with Prime Minister, 26 November 1957. The National Archives, Kew, London, United Kingdom.

———. DO 177/42. Minorities Commission: Western Region. The National Archives, Kew, London, United Kingdom.

———. DO 177/42. Minorities Commission: Eastern Region. The National Archives, Kew, London, United Kingdom.

———. DO 177/42. Minorities Commission: Northern Region. The National Archives, Kew, London, United Kingdom.

———. DO 195/82/66. "Federation of Nigeria: The Mid-West Before the Referendum," letter from Robin Jasper to Lord Antony Head, 17 April 1963. The National Archives, Kew, London, United Kingdom.

———. DO 195/83/91. Letter from R. L. D. Jasper to R. W. D. Fowler, 16 July 1963, enclosed in a letter from J. O. Moreton to V. C. Martin, 19 July 1963. The National Archives, Kew, London, United Kingdom.

———. DO 195/280/15. Letter from R. L. D. Jasper to Duncan Sandys, 26 February 1964. The National Archives, Kew, London, United Kingdom.

TNA. DO 195/280/45. Letter from John Moreton to Victor Martin, 11 August 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/280/46A. Minutes of a meeting on 18 August 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/280/48. Letter from Martin Ewans to Victor Martin, 21 August 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/280/50A. Letter from John Moreton to Victor Martin, 2 September 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/281/71. Letter from Sir Francis Cumming-Bruce to Arthur Bottomley, 9 November 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/281/81. "Some Impressions of Nigeria's Future," letter from Sir Francis Cumming-Bruce to Arthur Bottomley, 5 December 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/281/106. Telegram from Sir Francis Cumming-Bruce to John Chadwick, 31 December 1964. The National Archives, Kew, London, United Kingdom.

———. DO 195/281/115B. Letter from Sir Francis Cumming-Bruce to Sir Saville Garner, 2 January 1965. The National Archives, Kew, London, United Kingdom.

———. DO 195/282/155A. Letter from Sir Francis Cumming-Bruce to Sir Saville Garner, 16 January 1965. The National Archives, Kew, London, United Kingdom.

TNA. DO 195/282/209A. Letter from E. N. Larmour to John Chadwick, 2 June 1965. The National Archives, Kew, London, United Kingdom.

———. DO 195/282/239. Record of a meeting between Michael I. Okpara and C. J. Treadwell, 18 October 1965. The National Archives, Kew, London, United Kingdom.

———. DO 195/283/251. Letter from P. D. McEntee to P. C. Duff, 9 December 1965. The National Archives, Kew, London, United Kingdom.

———. DO 195/284/49. Letter from Dennis Pepper to D. F. Hawley, 3 February 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/290/48A. Enugu Fortnightly Summary No. 2 (66), 5 January 1966 to 18 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/10. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 15 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/16. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 15 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/17. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 15 January 1966. The National Archives, Kew, London, United Kingdom.

TNA. DO 195/294/23. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 15 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/45. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 17 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/51A. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 17 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/58A. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 18 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/64. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 19 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/84. Telegram from Sir Francis Cumming-Bruce to Commonwealth Relations Office, 20 January 1966. The National Archives, Kew, London, United Kingdom.

———. DO 195/294/93A. “Situation Report: Nigeria,” 21 January 1966. The National Archives, Kew, London, United Kingdom.

TNA. DO 195/294/98A. Telegram from John Chadwick to Sir Francis Cumming-Bruce, 22 January 1966. The National Archives, Kew, London, United Kingdom.

——. DO 195/294/109A. Record of a meeting between Mr. D. F. Hawley and Mr. L. Marsden, 22 January 1966. The National Archives, Kew, London, United Kingdom.

——. DO 195/294/110. "The Nigerian Coup d'Etat," despatch from Sir Francis Cumming-Bruce to Arthur Bottomley, 22 January 1966. The National Archives, Kew, London, United Kingdom.

——. PREM 11/3048. 1960: Nigeria. The National Archives, Kew, London, United Kingdom.

——. PREM 11/3443. 1961: Nigeria. The National Archives, Kew, London, United Kingdom.

——. PREM 11/3444. 1961: Nigeria. The National Archives, Kew, London, United Kingdom.

——. PREM 11/3893. 1962: Nigeria. The National Archives, Kew, London, United Kingdom.

——. PREM 11/4383. 1962-1963: Nigeria. The National Archives, Kew, London, United Kingdom.

——. PREM 11/4963. 1964: Nigeria. The National Archives, Kew, London, United Kingdom.

TNA. PREM 13/451. 1964-1965: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/1040. 1966: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/1041. 1966: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/1661. 1966-1967: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/1662. 1967: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/2258. 1968: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/2259. 1968: Nigeria. The National Archives, Kew, London, United Kingdom.

———. PREM 13/3380. 1970: Nigeria. The National Archives, Kew, London, United Kingdom.

Bibliography

Acemoğlu, Daron, Georgy Egorov, and Konstantin Sonin. 2008. "Coalition Formation in Non-Democracies." *Review of Economic Studies* 75 (4): 987–1010.

Acemoğlu, Daron, Davide Ticchi, and Andrea Vindigni. 2010. "A Theory of Military Dictatorships." *American Economic Journal: Macroeconomics* 2 (1): 1–42.

Adolph, Chris. 2012. *tile: A fully featured R graphics package built on the grid graphics environment.* R package version 0.4.13.

Ahlquist, John S., and Margaret Levi. 2011. "Leadership: What It Means, What It Does, and What We Want to Know About It." *Annual Review of Political Science* 14:1–24.

Akinola, Anthony. 2014. *Party Coalitions in Nigeria: History, Trends and Prospects.* Ibadan, Nigeria: Safari Books Ltd.

Alapiki, Henry E. 2005. "State Creation in Nigeria: Failed Approaches to National Integration and Local Autonomy." *African Studies Review* 48 (3): 49–65.

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Political Parties in America.* Chicago: University of Chicago Press.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43 (3): 762–800.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

Arjomand, Saïd Amir. 2009. *After Khomeini: Iran Under His Successors.* Oxford: Oxford University Press.

Arzaghi, Mohammad, and J. Vernon Henderson. 2005. "Why Countries are Fiscally Decentralizing." *Journal of Public Economics* 89 (7): 1157–1189.

Austen-Smith, David, and Jeffrey S. Banks. 1999. *Positive Political Theory: Collective Preference.* Vol. 1. Ann Arbor, MI: University of Michigan Press.

Autesserre, Séverine. 2010. *The Trouble with the Congo: Local Violence and the Failure of International Peacebuilding.* Cambridge: Cambridge University Press.

Awolowo, Obafemi. 1960. *Awo: The Autobiography of Chief Obafemi Awolowo.* Cambridge: Cambridge University Press.

Ayee, Joseph R. A. 2012. "The Political Economy of the Creation of Districts in Ghana." *Journal of Asian and African Studies* 48 (5): 623–645.

Bäck, Hanna, Marc Debus, and Patrick Dumont. 2010. "Who Gets What in Coalition Governments? Predictors of Portfolio Allocation in Parliamentary Democracies." *European Journal of Political Research* 50 (4): 441–478.

Bagnall, Anthony, and Gavin C. Cawley. 2017. "On the Use of Default Parameter Settings in the Empirical Evaluation of Classification Algorithms." ArXiv: 1703.06777v1, March.

Bakke, Kristin M., and Erik Wibbels. 2006. "Diversity, Disparity, and Civil Conflict in Federal States." *World Politics* 59 (1): 1–50.

Baldersheim, Harald, and Lawrence E. Rose, eds. 2010. *Territorial Choice: The Politics of Boundaries and Borders.* Houndmills: Palgrave Macmillan.

Banks, Arthur S., and Kenneth A. Wilson. 2013. "Cross-National Time-Series Data Archive." Jerusalem: Databanks International.

Barany, Zoltan. 2015. "Exits from Military Rule: Lessons for Burma." *Journal of Democracy* 26 (2): 86–100.

Baron, David P., and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83 (4): 1181–1206.

Bates, Robert H. 1989. *Beyond the Miracle of the Market: The Political Economy of Agrarian Development in Kenya.* Cambridge: Cambridge University Press.

Bausch, Andrew W. 2017. "Coalition Formation and Selectorate Theory: An Experiment." *Political Science Research and Methods* 5 (2): 261–275.

Bayart, Jean-François. 1993. *The State in Africa: The Politics of the Belly.* London: Longman.

Beck, Nathaniel, and Simon Jackman. 1998. "Beyond Linearity by Default: Generalized Additive Models." *American Journal of Political Science* 42 (2): 596–627.

Beck, Nathaniel, Gary King, and Langche Zeng. 2000. "Improving Quantitative Studies of International Conflict: A Conjecture." *American Political Science Review* 94 (1): 21–35.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15 (1): 165–176.

Bell, E. T. 1938. "The Iterated Exponential Numbers." *Annals of Mathematics* 39 (3): 539–557.

Benoit, Kenneth. 1996. "Democracies Really Are More Pacific (in General)." *Journal of Conflict Resolution* 40 (4): 636–657.

Beramendi, Pablo. 2007. "Federalism." In *The Oxford Handbook of Comparative Politics,* edited by Carles Boix and Susan C. Stokes, 752–781. Oxford: Oxford University Press.

———. 2012. *The Political Geography of Inequality: Regions and Redistribution.* Cambridge: Cambridge University Press.

Beramendi, Pablo, and Sandra León. 2015. "Federalism." In *The Routledge Handbook of Comparative Political Institutions,* edited by Jennifer Gandhi and Rubén Ruiz-Rufino, 209–225. New York: Routledge.

Berry, Christopher. 2008. "Piling On: Multilevel Government and the Fiscal Common-Pool." *American Journal of Political Science* 52 (4): 802–820.

Besley, Timothy, and Stephen Coate. 2003. "Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach." *Journal of Public Economics* 87 (12): 2611–2637.

Bhavnani, Rikhil R. 2015. "The Effects of Malapportionment on Cabinet Formation: Subnational Evidence from India." *British Journal of Political Science* 48 (1): 69–89.

Blair, Robert A., Christopher Blattman, and Alexandra C. Hartman. 2017. "Predicting Local Violence: Evidence from a Panel Survey in Liberia." *Journal of Peace Research* 54 (2): 298–312.

Blaydes, Lisa, and Mark Andreas Kayser. 2011. "Counting Calories: Democracy and Distribution in the Developing World." *International Studies Quarterly* 55 (4): 887–908.

Blom-Hansen, Jens, Kurt Houlberg, and Søren Serritzlew. 2014. "Size, Democracy, and the Economic Costs of Running the Political System." *American Journal of Political Science* 58 (4): 790–803.

Blom-Hansen, Jens, Kurt Houlberg, Søren Serritzlew, and Daniel Treisman. 2016. "Jurisdiction Size and Local Government Policy Expenditure: Assessing the Effect of Municipal Amalgamation." *American Political Science Review* 110 (4): 812–831.

Bodner, Todd E. 2008. "What Improves with Increased Missing Data Imputations." *Structural Equation Modeling* 15 (4): 651–675.

Boix, Carles. 1999. "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review* 93 (3): 609–624.

Boix, Carles, and Milan W. Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-Sharing in Dictatorships." *Journal of Politics* 75 (2): 300–316.

Boone, Catherine. 2003. *Political Topographies of the African State: Territorial Authority and Institutional Choice.* Cambridge: Cambridge University Press.

Brancati, Dawn. 2014. "Democratic Authoritarianism: Origins and Effects." *Annual Review of Political Science* 17:313–326.

Brand, J. A. 1965. "The Mid-West State Movement in Nigerian Politics: A Study in Party Formation." *Political Studies* 13 (3): 346–365.

Breiman, Leo. 1996. "Bagging Predictors." *Machine Learning* 24 (2): 123–140.

——. 2001a. "Random Forests." *Machine Learning* 45 (1): 5–32.

——. 2001b. "Statistical Modeling: The Two Cultures." *Statistical Science* 16 (3): 199–231.

Brier, Glenn W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78 (1): 1–3.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival.* Cambridge, MA: MIT Press.

Central Intelligence Agency. 2016. "The World Factbook." Washington, D.C.

Ceron, Andrea. 2012. "Bounded Oligarchy: How and When Factions Constrain Leaders in Party Position-Taking." *Electoral Studies* 31 (4): 689–701.

Chatterjee, Kalyan, Bhaskar Dutta, Debraj Ray, and Kunal Sengupta. 1993. "A Noncooperative Theory of Coalitional Bargaining." *Review of Economic Studies* 60 (2): 463–477.

Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. "SMOTE: Synthetic Minority Over-sampling Technique." *Journal of Artificial Intelligence Research* 16:321–357.

Cheeseman, Nic, Gabrielle Lynch, and Justin Willis. 2016. "Decentralisation in Kenya: The Governance of Governors." *Journal of Modern African Studies* 54 (1): 1–35.

Chhibber, Pradeep, and Ken Kollman. 2004. *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States.* Princeton, NJ: Princeton University Press.

Choirat, Christine, James Honaker, Kosuke Imai, Gary King, and Olivia Lau. 2016. "relogit: Rare Events Logistic Regression for Dichotomous Dependent Variables." In *Zelig: Everyone's Statistical Software.* http://zeligproject.org/.

Coase, Ronald. 1937. "The Nature of the Firm." *Economica* 4 (16): 386–405.

Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3 (1): 1–31.

Cohen, Jacob. 1960. "A Coefficient of Agreement for Nominal Scales." *Educational and Psychological Measurement* 20 (1): 37–46.

Coleman, James S. 1958. *Nigeria: Background to Nationalism.* Berkeley, CA: University of California Press.

Colomer, Josep M. 2005. "It's Parties That Choose Electoral Systems (or Duverger's Law Upside Down)." *Political Studies* 53 (1): 1–21.

Colonial Office. 1958. *Commission to Enquire into the Fears of Minorities and the Means of Allaying Them.*

Cox, Gary W. 1987. *The Efficient Secret: The Cabinet and the Development of Political Parties in Victorian England.* Cambridge: Cambridge University Press.

———. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems.* Cambridge: Cambridge University Press.

Crawford, Gordon, and Christof Hartmann, eds. 2008. *Decentralisation in Africa: A Pathway out of Poverty and Conflict?* Amsterdam: Amsterdam University Press.

Crémer, Jacques, and Thomas R. Palfrey. 1999. "Political Confederation." *American Political Science Review* 93 (1): 69–83.

D'Orazio, Vito, Steven T. Landis, Glenn Palmer, and Philip Schrodt. 2014. "Separating the Wheat from the Chaff: Applications of Automated Document Classification Using Support Vector Machines." *Political Analysis* 22 (2): 224–242.

da Graça, John V. 1985. *Heads of State and Government.* London: Macmillan.

Dangerfield, George. 1952. *The Era of Good Feelings.* New York: Harcourt, Brace & Co.

Deng, Houtao, and George Runger. 2013. "Gene Selection with Guided Regularized Random Forest." *Pattern Recognition* 46 (12): 3483–3489.

Dettling, Marcel, and Peter Bühlmann. 2003. "Boosting for Tumor Classification with Gene Expression Data." *Bioinformatics* 19 (9): 1061–1069.

Deutsch, Karl W., Sidney A. Burrell, Robert A. Kann, and Maurice Lee Jr. 1957. *Political Community in the North Atlantic Area: International Organization in the Light of Historical Experience.* Princeton, NJ: Princeton University Press.

Diaz-Cayeros, Alberto. 2006. *Federalism, Fiscal Authority, and Centralization in Latin America.* New York: Cambridge University Press.

Dickovick, J. Tyler. 2011. *Decentralization and Recentralization in the Developing World.* University Park, PA: Pennsylvania State University Press.

——. 2014. "Federalism in Africa: Origins, Operation and (In)Significance." *Regional and Federal Studies* 24 (5): 553–570.

Dreiseitl, Stephan, and Lucila Ohno-Machado. 2002. "Logistic Regression and Artificial Neural Network Classification Models: A Methodology Review." *Journal of Biomedical Informatics* 35 (5): 352–359.

Dudley, B. J. 1968. *Parties and Politics in Northern Nigeria.* London: Frank Cass & Co. Ltd.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences.* New York: Cambridge University Press.

Duverger, Maurice. 1954. *Political Parties.* New York: Wiley.

Eaton, Kent. 2004. *Politics Beyond the Capital: The Design of Subnational Institutions in South America.* Stanford, CA: Stanford University Press.

Eaton, Kent, and Tyler Dickovick. 2004. "The Politics of Re-centralization in Argentina and Brazil." *Latin American Research Review* 39 (1): 90–122.

Efron, Bradley. 1975. "The Efficiency of Logistic Regression Compared to Normal Discriminant Analysis." *Journal of the American Statistical Association* 70 (352): 892–898.

Elkins, Zachary, Tom Ginsburg, and James Melton. 2009. *The Endurance of National Constitutions.* New York: Cambridge University Press.

Engstrom, Erik J. 2006. "Stacking the States, Stacking the House: The Partisan Consequences of Congressional Redistricting in the 19th Century." *American Political Science Review* 100 (3): 419–427.

Enweremadu, David U. 2013. "Nigeria's Quest to Recover Looted Assets: The Abacha Affair." *Africa Spectrum* 48 (2): 51–70.

Falleti, Tulia G. 2005. "A Sequential Theory of Decentralization: Latin American Cases in Comparative Perspective." *American Political Science Review* 99 (3): 327–346.

Falola, Toyin, and Matthew M. Heaton. 2008. *A History of Nigeria.* Cambridge: Cambridge University Press.

FAO. 2017. "FAOSTAT Statistics Database." Available at `http://www.fao.org/faostat/en/`. Accessed 10 April 2017.

Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8 (2): 195–222.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75–90.

Featherstone, Kevin, and Claudio M. Radaelli, eds. 2003. *The Politics of Europeanization.* Oxford: Oxford University Press.

Fernández-Delgado, Manuel, Evan Cernadas, and Senén Barro. 2014. "Do We Need Hundreds of Classifiers to Solve Real World Classification Problems?" *Journal of Machine Learning Research* 15 (1): 3133–3181.

Filippov, Mikhail, Peter C. Ordeshook, and Olga Shvetsova. 2004. *Designing Federalism: A Theory of Self-Sustainable Federal Institutions.* New York: Cambridge University Press.

Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80 (1): 27–38.

Fisher, R. A. 1936. "The Use of Multiple Measurements in Taxonomic Problems." *Annals of Eugenics* 7 (2): 179–188.

Fitrani, Fitria, Bert Hofman, and Kai Kaiser. 2005. "Unity in Diversity? The Creation of New Local Governments in Decentralising Indonesia." *Bulletin of Indonesian Economic Studies* 41 (1): 57–79.

Forde, Catherine. 2005. "Participatory Democracy or Pseudo-Participation? Local Government Reform in Ireland." *Local Government Studies* 31 (2): 137–148.

Freund, Yoav, and Robert E. Schapire. 1996. "Experiments with a New Boosting Algorithm." *Machine Learning: Proceedings of the Thirteenth International Conference:* 148–156.

Friedman, Jerome H. 1991. "Multivariate Adaptive Regression Splines." *The Annals of Statistics* 19 (1): 1–67.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. 2000. "Additive Logistic Regression: A Statistical View of Boosting." *The Annals of Statistics* 28 (2): 337–407.

———. 2010. "Regularization Paths for Generalized Linear Models via Coordinate Descent." *Journal of Statistical Software* 33 (1): 1–22.

Friedrich, Carl J., and Zbigniew K. Brzezinski. 1956. *Totalitarian Dictatorship and Autocracy.* Cambridge, MA: Harvard University Press.

GADM. 2018. "Global Administrative Areas." Available at `http://gadm.org`. Accessed 13 February 2018.

Gallagher, Mary E., and Jonathan K. Hanson. 2015. "Power Tool or Dull Blade? Selectorate Theory for Autocracies." *Annual Review of Political Science* 18:67–85.

Gandhi, Jennifer, and Ellen Lust-Okar. 2009. "Elections Under Authoritarianism." *Annual Review of Political Science* 12:403–422.

Garson, G. David. 1991. "Interpreting Neural Network Connection Weights." *Artificial Intelligence Expert* 6 (4): 46–51.

Gehlbach, Scott, Konstantin Sonin, and Milan W. Svolik. 2016. "Formal Models of Nondemocratic Politics." *Annual Review of Political Science* 19:565–584.

Gill, Jeff. 1999. "The Insignificance of Null Hypothesis Significance Testing." *Political Research Quarterly* 52 (3): 647–674.

Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102 (477): 359–378.

Goemans, Henk E., Kristian Skrede Gleditsch, and Giacomo Chiozza. 2009. "Introducing Archigos: A Dataset of Political Leaders." *Journal of Peace Research* 46 (2): 269–283.

Goh, A. T. C. 1995. "Back-Propagation Neural Networks for Modeling Complex Systems." *Artificial Intelligence in Engineering* 9 (3): 143–151.

Golder, Matt, Sona N. Golder, and David A. Siegel. 2012. "Modeling the Institutional Foundation of Parliamentary Government Formation." *Journal of Politics* 74 (2): 427–445.

Gould, Michael. 2013. *The Biafran War: The Struggle for Modern Nigeria.* London: I. B. Tauris.

Green, Elliott. 2010. "Patronage, District Creation, and Reform in Uganda." *Studies in Comparative International Development* 45 (1): 83–103.

Greif, Avner, and David D. Laitin. 2004. "A Theory of Endogenous Institutional Change." *American Political Science Review* 98 (4): 633–652.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267–297.

Grossman, Guy, and Janet I. Lewis. 2014. "Administrative Unit Proliferation." *American Political Science Review* 108 (1): 196–217.

Grossman, Guy, Jan H. Pierskalla, and Emma Boswell Dean. 2017. "Government Fragmentation and Public Goods Provision." *Journal of Politics* 79 (3): 823–840.

Hainmueller, Jens, and Chad Hazlett. 2014. "Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach." *Political Analysis* 22 (2): 143–168.

Hale, Henry E. 2004. "Divided We Stand: Institutional Sources of Ethnofederal State Survival and Collapse." *World Politics* 56 (2): 165–193.

Hamilton, Alexander, John Jay, and James Madison. 1892 [1787-1788]. *The Federalist: A Commentary on the Constitution of the United States, Being a Collection of Essays Written in Support of the Constitution Agreed upon September 17, 1787.* Edited by Henry Cabot Lodge. New York: G. P. Putnam's Sons.

Harnischfeger, Johannes. 2004. "Sharia and Control over Territory: Conflicts between 'Settlers' and 'Indigenes' in Nigeria." *African Affairs* 103 (412): 431–452.

———. 2008. *Democratization and Islamic Law: The Sharia Conflict in Nigeria.* Frankfurt: Campus Verlag.

Harriss, John, Janet Hunter, and Colin M. Lewis, eds. 1995. *The New Institutional Economics and Third World Development.* London: Routledge.

Hassan, Mai. 2014. "A State of Change: The Kenyan State after Multi-Party Elections." PhD diss., Harvard University.

———. 2016. "A State of Change: District Creation in Kenya After the Beginning of Multi-Party Elections." *Political Research Quarterly* 69 (3): 510–521.

Hassan, Mai, and Ryan Sheely. 2017. "Executive-Legislative Relations, Party Defections, and Lower Level Administrative Unit Proliferation: Evidence from Kenya." *Comparative Political Studies* 50 (12): 1595–1631.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd ed. Berlin: Springer.

Hayek, Friedrich A. 1944. *The Road to Serfdom.* Chicago: University of Chicago Press.

He, Haibo, and Edwardo A. Garcia. 2009. "Learning from Imbalanced Data." *IEEE Transactions on Knowledge and Data Engineering* 21 (9): 1263–1284.

Hegre, Håvard. 2014. "Democracy and Armed Conflict." *Journal of Peace Research* 51 (2): 159–172.

Henisz, Witold J. 2002. "The Political Constraint (POLCON) Dataset." Available at `https://mgmt.wharton.upenn.edu/profile/henisz/`. Accessed via the Quality of Government Institute.

Herbst, Jeffrey Ira. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control.* Princeton, NJ: Princeton University Press.

Hickey, Donald R. 2012. *The War of 1812: A Forgotten Conflict.* Bicentennial Edition. University of Illinois Press.

Hill, Daniel W., Jr., and Zachary M. Jones. 2014. "An Empirical Evaluation of Explanations for State Repression." *American Political Science Review* 108 (3): 661–687.

Hindman, Matthew. 2015. "Building Better Models: Prediction, Replication, and Machine Learning in the Social Sciences." *The Annals of the American Academy of Political and Social Science* 659 (1): 48–62.

Hoerl, A. E. 1962. "Application of Ridge Analysis to Regression Problems." *Chemical Engineering Progress* 58 (3): 54–59.

Hofmann, Thomas, Bernhard Schölkopf, and Alexander J. Smola. 2008. "Kernel Methods in Machine Learning." *The Annals of Statistics* 36 (3): 1171–1220.

Hooghe, Liesbet, Gary Marks, Arjan H. Schakel, Sara Niedzwiecki, Sandra Chapman Osterkatz, and Sarah Shair-Rosenfield. 2016. *Measuring Regional Authority: A Postfunctionalist Theory of Governance.* Vol. 1. Oxford: Oxford University Press.

Horowitz, Donald. 1985. *Ethnic Groups in Conflict.* Berkeley, CA: University of California Press.

Humphreys, Macartan, and Jeremy M. Weinstein. 2007. "Demobilization and Reintegration." *Journal of Conflict Resolution* 51 (4): 531–567.

Ichino, Nahomi. 2008. "Essays on Ethnic Diversity and Political Instability in Sub-Saharan Africa." PhD diss., Stanford University.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15 (4): 309–334.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25 (1): 51–71.

Johansson, Karl Magnus, and Tapio Raunio. 2001. "Partisan Responses to Europe: Comparing Finnish and Swedish Parties." *European Journal of Political Research* 39 (2): 225–249.

Johnson, Boris. 2014. "Oral Evidence: Fiscal Devolution to Cities and City Regions." HC 1018 Q241. Given 3 March. Available at `http://www.parliament.uk/business/committees/`. Accessed 16 March 2018.

Jones, Benjamin F., and Benjamin A. Olken. 2005. "Do Leaders Matter? National Leadership and Growth Since World War II." *Quarterly Journal of Economics* 120 (3): 835–864.

——. 2009. "Hit or Miss? The Effect of Assassinations on Institutions and War." *American Economic Journal: Macroeconomics* 1 (2): 55–87.

Jones, Zachary M., and Fridolin J. Linder. 2016. "edarf: Exploratory Data Analysis using Random Forests." *Journal of Open Source Software* 1 (6).

Kasara, Kimuli. 2006. "Ethnic Beachheads and Vote Buying: The Creation of New Administrative Districts in Kenya, 1992-2001." Mimeo.

Kendhammer, Brandon. 2010. "Muslims Talking Politics: Framing Islam and Democracy in Northern Nigeria." PhD diss., University of Wisconsin-Madison.

Khemani, Stuti. 2010. "Political Capture of Decentralization: Vote-Buying through Grants-Financed Local Jurisdictions." World Bank Policy Research Working Paper 5350. June.

Kimura, Ehito. 2013. *Political Change and Territoriality in Indonesia: Provincial Proliferation.* New York: Routledge.

King, Gary, and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9 (1): 137–163.

Kitschelt, Herbert. 2000. "Linkages Between Citizens and Politicians in Democratic Polities." *Comparative Political Studies* 33 (6-7): 845–879.

Kitschelt, Herbert, and Steven Wilkinson, eds. 2007. *Patrons, Clients, and Policies: Patterns of Democratic Accountability and Political Competition.* Cambridge: Cambridge University Press.

Knight, Jack. 1992. *Institutions and Social Conflict.* New York: Cambridge University Press.

Kraxberger, Brandon. 2004. "The Geography of Regime Survival: Abacha's Nigeria." *African Affairs* 103 (412): 413–430.

Kropko, Jonathan, Ben Goodrich, Andrew Gelman, and Jennifer Hill. 2014. "Multiple Imputation for Continuous and Categorical Data: Comparing Joint Multivariate Normal and Conditional Approaches." *Political Analysis* 22 (4): 497–519.

Kudamatsu, Masayuki. 2009. "Ethnic Favoritism: Micro Evidence from Guinea." Unpublished manuscript. July.

Kuhn, Max. 2008. "Building Predictive Models in R using the caret Package." *Journal of Statistical Software* 28 (5): 1–26.

——. 2014. "Futility Analysis in the Cross-Validation of Machine Learning Models." ArXiv: 1405.6974v1, May.

Kuhn, Max, and Kjell Johnson. 2013. *Applied Predictive Modeling.* Berlin: Springer.

Kuncheva, Ludmila I., and Juan J. Rodríguez. 2007. "An Experimental Study on Rotation Forest Ensembles." In *Multiple Classifier Systems: 7th International Workshop,* edited by Michal Haindl, Josef Kittler, and Fabio Roli, 459–468. Berlin: Springer.

Lakatos, Imre. 1978. *The Methodology of Scientific Research Programmes: Philosophical Papers.* Edited by John Worrall and Gregory Currie. Vol. 1. Cambridge: Cambridge University Press.

Lall, Ranjit. 2016. "How Multiple Imputation Makes a Difference." *Political Analysis* 24 (4): 414–433.

Lasswell, Harold D. 1936. *Politics: Who Gets What, When, How.* New York: Whittlesey House.

Laver, Michael. 1998. "Models of Government Formation." *Annual Review of Political Science* 1:1–25.

Laver, Michael, and Kenneth A. Shepsle. 1996. *Making and Breaking Governments: Cabinets and Legislatures in Parliamentary Democracies.* New York: Cambridge University Press.

Law, Gwillim. 2016. "Statoids: Administrative Divisions of Countries." Available at `http://www.statoids.com`. Accessed 30 November 2016.

Lee, Lung-fei, and Jihai Yu. 2015. "Spatial Panel Data Models." In *The Oxford Handbook of Panel Data,* edited by Badi H. Baltagi, 363–401. Oxford: Oxford University Press.

Lentz, Harris M., III. 1994. *Heads of States and Governments: A Worldwide Encyclopedia of Over 2,300 Leaders, 1945 through 1992.* London: Routledge.

Lewis, Janet I. 2014. "When Decentralization Leads to Recentralization: Subnational State Transformation in Uganda." *Regional and Federal Studies* 24 (5): 571–588.

Li, Zhisui. 1994. *The Private Life of Chairman Mao: The Memoirs of Mao's Personal Physician.* New York: Random House.

Lijphart, Arend. 1984. *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries.* New Haven: Yale University Press.

Lipset, Seymour, and Stein Rokkan. 1967. "Cleavage Structures, Party Systems, and Voter Alignments: An Introduction." In *The West European Party System,* edited by P. Mair, 1–64. New York: Oxford University Press.

Logams, P. Chunun. 1985. "The Middle-Belt Movement in Nigerian Political Development: A Study in Political Identity 1949-1967." PhD diss., University of Keele.

Lupu, Noam. 2016. *Party Brands in Crisis: Partisanship, Brand Dilution, and the Breakdown of Political Parties in Latin America.* Cambridge: Cambridge University Press.

Mackintosh, John P. 1962. "Federalism in Nigeria." *Political Studies* 10 (3): 223–247.

Malesky, Edmund. 2009. "Gerrymandering—Vietnamese Style: Escaping the Partial Reform Equilibrium in a Nondemocratic Regime." *Journal of Politics* 71 (1): 132–159.

Manor, James. 1999. *The Political Economy of Democratic Decentralization.* Washington, DC: The World Bank.

Marron, J. S., Michael J. Todd, and Jeongyoun Ahn. 2007. "Distance-Weighted Discrimination." *Journal of the American Statistical Association* 102 (480): 1267–1271.

Marshall, Monty G. 2015. "Major Episodes of Political Violence (MEPV) and Conflict Regions, 1946-2015." Center for Systemic Peace.

Marshall, Monty G., Keith Jaggers, and Ted Robert Gurr. 2011. "Polity IV Project." Center for Systemic Peace.

Mawdsley, Emma. 2002. "Redrawing the Body Politic: Federalism, Regionalism and the Creation of New States in India." *Journal of Commonwealth and Comparative Politics* 40 (3): 34–54.

McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal. 1998. "Congress and the Territorial Expansion of the United States." Paper prepared for the Stanford-UCSD conference on the History of Congress, January 15-16, 1999. December.

Merkle, Edgar C., and Mark Steyvers. 2013. "Choosing a Strictly Proper Scoring Rule." *Decision Analysis* 10 (4): 292–304.

Moe, Terry M. 2005. "Power and Political Institutions." *Perspectives on Politics* 3 (2): 215–233.

Montesquieu, Charles de Secondat, Baron de. 1989 [1748]. *L'espirit des Lois*. Edited and translated by Anne Cohler, Basia Miller, and Harold Stone. New York: Cambridge University Press.

Montgomery, Jacob M., Florian M. Hollenbach, and Michael D. Ward. 2012. "Improving Predictions Using Ensemble Bayesian Model Averaging." *Political Analysis* 20 (3): 271–291.

Muchlinski, David, David Siroky, Jingrui He, and Matthew Kocker. 2016. "Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data." *Political Analysis* 24 (1): 83–103.

Negretto, Gabriel L. 2013. *Making Constitutions: Presidents, Parties and Institutional Design in Latin America.* Cambridge: Cambridge University Press.

Norris, Pippa. 2008. *Driving Demcoracy: Do Power-Sharing Institutions Work?* Cambridge: Cambridge University Press.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance.* New York: Cambridge University Press.

North, Douglass, and Barry Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England." *The Journal of Economic History* 49 (4): 803–832.

Nunley, Albert C. 2012. "African Elections Database." Available at `http://africanelections.tripod.com/index.html`. Accessed 13 February 2018.

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94 (1): 20–36.

O'Dwyer, Conor. 2006. *Runaway State-Building: Patronage Politics and Democratic Development.* Baltimore, MD: Johns Hopkins University Press.

O'Neill, Kathleen. 2005. *Decentralizing the State: Elections, Parties, and Local Power in the Andes.* New York: Cambridge University Press.

Oates, Wallace E. 1972. *Fiscal Federalism.* New York: Harcourt Brace Jovanovich.

Okafor, S. I. 1987. "Jurisdictional Partitioning, Distribution Policies and the Spatial Structure of Health-Care Provision in Nigeria." *Political Geography Quarterly* 6 (4): 335–346.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge: Cambridge University Press.

Paden, John N. 1973. *Religion and Political Culture in Kano.* Berkeley, CA: University of California Press.

Pepinsky, Thomas. 2014. "The Institutional Turn in Comparative Authoritarianism." *British Journal of Political Science* 44 (3): 631–653.

Persico, Nicola, José C. R. Pueblita, and Dan Silverman. 2011. "Factions and Political Competition." *Journal of Political Economy* 119 (2): 242–288.

Persson, Torsten, and Guido Tabellini. 1996. "Federal Fiscal Constitutions: Risk Sharing and Redistribution." *Journal of Political Economy* 104 (5): 979–1009.

Pierskalla, Jan H. 2016. "Splitting the Difference? The Politics of District Creation in Indonesia." *Comparative Politics* 48 (2): 249–268.

Quinlan, Ross. 1993. *C4.5: Programs for Machine Learning.* Burlington, MA: Morgan Kaufmann.

Ray, Debraj. 2007. *A Game-Theoretic Perspective on Coalition Formation.* Oxford: Oxford University Press.

Ray, Debraj, and Rajiv Vohra. 2014. "Coalition Formation." In *The Handbook of Game Theory with Economic Applications,* edited by H. Peyton Young and Shmuel Zamir, 4:239–326. Amsterdam: Elsevier B.V.

Riedl, Rachel Beatty, and J. Tyler Dickovick. 2014. "Party Systems and Decentralization in Africa." *Studies in Comparative International Development* 49 (3): 321–342.

Riker, William H. 1962. *The Theory of Political Coalitions.* New Haven: Yale University Press.

———. 1964. *Federalism: Origins, Operation, Significance.* New York: Little, Brown & Company.

———. 1980. "Implications from the Disequilibrium of Majority Rule for the Study of Institutions." *American Political Science Review* 74 (2): 432–446.

Ringe, Nils, and Jennifer Nicoll Victor. 2013. *Bridging the Information Gap: Legislative Member Organizations as Social Networks in the United States and the European Union.* Ann Arbor, MI: University of Michigan Press (with Christopher J. Carman).

Robinson, James A., and Ragnar Torvik. 2015. "Endogenous Presidentialism." *Journal of the European Economic Association* 14 (4): 907–942.

Rodden, Jonathan. 2002. "The Dilemmas of Fiscal Federalism: Grants and Fiscal Performance around the World." *American Journal of Political Science* 46 (3): 670–687.

Rodden, Jonathan. 2005. *Hamilton's Paradox: The Promise and Peril of Fiscal Federalism.* New York: Cambridge University Press.

Rodríguez, Juan J., Ludmila I. Kuncheva, and Carlos J. Alonso. 2006. "Rotation Forest: A New Classifier Ensemble Method." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (10): 1619–1630.

Roland, Gérard. 2000. *Transition and Economics: Politics, Markets, and Firms.* Cambridge, MA: MIT Press.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley & Sons.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50 (1): 97–109.

Schattschneider, Elmer Eric. 1942. *Party Government.* New York: Farrar & Rinehart.

Schimmelfennig, Frank, and Ulrich Sedelmeier. 2002. "Theorizing EU Enlargement: Research Focus, Hypotheses, and the State of Research." *Journal of European Public Policy* 9 (4): 500–528.

Schneider, Aaron. 2003. "Decentralization: Conceptualization and Measurement." *Studies in Comparative International Development* 38 (3): 32–56.

Schneider, Christina J. 2017. "The Political Economy of Regional Integration." *Annual Review of Political Science* 20:229–248.

Schomaker, Michael, and Christian Heumann. 2014. "Model Selection and Model Averaging After Multiple Imputation." *Computational Statistics and Data Analysis* 71:758–770.

Serneels, Sven, Evert De Nolf, and Pierre J. Van Espen. 2006. "Spatial Sign Preprocessing: A Simple Way to Impart Moderate Robustness to Multivariate Estimators." *Journal of Chemical Information and Modeling* 46 (3): 1402–1409.

Shepsle, Kenneth A. 1989. "Studying Institutions: Some Lessons from the Rational Choice Approach." *Journal of Theoretical Politics* 1 (2): 131–148.

Shirk, Susan L. 1993. *The Political Logic of Economic Reform in China.* Berkeley, CA: University of California Press.

Shugart, Matthew Søberg, and John M. Carey. 1992. *Presidents and Assemblies: Constitutional Design and Electoral Dynamics.* Cambridge: Cambridge University Press.

Sklar, Richard L. 1963. *Nigerian Political Parties: Power in an Emergent African Nation.* Princeton, NJ: Princeton University Press.

Snyder, Richard. 1992. "Explaining Transitions from Neopatrimonial Dictatorships." *Comparative Politics* 24 (4): 379–399.

Soares, Fábio Veras, Rafael Perez Ribas, and Rafael Guerreiro Osório. 2010. "Evaluating the Impact of Brazil's Bolsa Família: Cash Transfer Programs in Comparative Perspective." *Latin American Research Review* 45 (2): 173–190.

Ståhl, Ingolf. 1977. "An N-Person Bargaining Game in the Extensive Form." In *Mathematical Economics and Game Theory,* edited by Rudolf Henn and Otto Moeschlin, 1:156–172. Berlin: Springer.

Stepan, Alfred C. 1999. "Federalism and Democracy: Beyond the U.S. Model." *Journal of Democracy* 10 (4): 19–34.

Stephanopoulos, Nicholas O., and Eric M. McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." *University of Chicago Law Review* 82 (2): 831–900.

Stewart, Charles, III, and Barry R. Weingast. 1992. "Stacking the Senate, Changing the Nation: Republican Rotten Boroughs, Statehood Politics, and American Political Development." *Studies in American Political Development* 6 (2): 223–271.

Stokes, Susan, Thad Dunning, Marcelo Nazareno, and Valeria Brusco. 2014. *Brokers, Voters, and Clientelism.* Cambridge: Cambridge University Press.

Suberu, Rotimi T. 2001. *Federalism and Ethnic Conflict in Nigeria.* Washington, DC: United States Institute of Peace.

Svolik, Milan. 2012. *The Politics of Authoritarian Rule.* Cambridge: Cambridge University Press.

Swianiewicz, Paweł. 2010. "If Territorial Fragmentation is a Problem, is Amalgamation a Solution? An East European Perspective." *Local Government Studies* 36 (2): 183–203.

Tamuno, Tekena N. 1970. "Separatist Agitations in Nigeria since 1914." *Journal of Modern African Studies* 8 (4): 563–584.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society Series B* 58 (1): 267–288.

Tibshirani, Robert, Trevor Hastie, Balasubramanian Narasimhan, and Gilbert Chu. 2002. "Diagnosis of Multiple Cancer Types by Shrunken Centroids of Gene Expression." *Proceedings of the National Academy of Sciences* 99 (10): 6567–6572.

Tiebout, Charles. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (5): 416–424.

Tocqueville, Alexis de. 1964 [1835]. *Democracy in America.* New York: Washington Square Press.

Tomio, Fabricio Ricardo de Limas. 2002. "A criação de municípios após a Constituição de 1988." *Revista Brasileira de Ciências Sociais* 17 (48): 61–89.

Treisman, Daniel. 2007. *The Architecture of Government: Rethinking Political Decentralization.* Cambridge: Cambridge University Press.

——. 2015. "Income, Democracy, and Leader Turnover." *American Journal of Political Science* 59 (4): 927–942.

Tsebelis, George. 2002. *Veto Players: How Political Institutions Work.* Princeton, NJ: Princeton University Press.

Vachudova, Milada Anna. 2005. *Europe Undivided: Democracy, Leverage, and Integration after Communism.* Oxford: Oxford University Press.

van Buuren, Stef, J. P. L. Brand, C. G. M. Groothuis-Oudshoorn, and D. B. Rubin. 2006. "Fully Conditional Specification in Multivariate Imputation." *Journal of Statistical Computation and Simulation* 76 (12): 1049–1064.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45 (3): 1–67.

Vaughan, Olufemi. 2000. *Nigerian Chiefs: Traditional Power in Modern Politics, 1890s-1990s.* Rochester, NY: University of Rochester Press.

——. 2016. *Religion and the Making of Nigeria.* Durham, NC: Duke University Press.

Vickers, Michael. 2000. *Ethnicity and Sub-Nationalism in Nigeria: Movement for a Mid-West State.* Oxford: WorldView Publishing.

——. 2010. *A Nation Betrayed: Nigeria and the Minorities Commission of 1957.* Trenton, NJ: Africa World Press.

von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Walter, Barbara F. 2006a. "Building Reputation: Why Governments Fight Some Separatists but Not Others." *American Journal of Political Science* 50 (2): 313–330.

——. 2006b. "Information, Uncertainty, and the Decision to Secede." *International Organization* 60 (1): 105–135.

Walter, Barbara F. 2015. "Why Bad Governance Leads to Repeat Civil War." *Journal of Conflict Resolution* 59 (7): 1242–1272.

Wang, Boxiang, and Hui Zou. 2015. "Sparse Distance Weighted Discrimination." ArXiv: 1501.06066v1, January.

Ward, Michael D., and John S. Ahlquist. 2018. *Maximum Likelihood Strategies for Social Science.* Cambridge: Cambridge University Press, forthcoming.

Ward, Michael D., and Kristian Skrede Gleditsch. 2008. *Spatial Regression Models.* New York: SAGE Publications.

Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "The Perils of Policy by p-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47 (4): 363–375.

Watts, Ronald L. 1998. "Federalism, Federal Political Systems, and Federations." *Annual Review of Political Science* 1:117–137.

Weingast, Barry R. 1995. "The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Growth." *Journal of Law, Economics, & Organization* 11 (1): 1–31.

Wheare, Kenneth C. 1946. *Federal Government.* London: Oxford University Press.

Whitaker, C. Sylvester, Jr. 1970. *The Politics of Tradition: Continuity and Change in Northern Nigeria, 1946-1966.* Princeton, NJ: Princeton University Press.

White, Ian R., Patrick Royston, and Angela M. Wood. 2011. "Multiple Imputation Using Chained Equations: Issues and Guidance for Practice." *Statistics in Medicine* 30 (4): 377–399.

Wibbels, Erik. 2005. *Federalism and the Market: Intergovernmental Conflict and Economic Reform in the Developing World.* New York: Cambridge University Press.

Williamson, Oliver E. 1967. "Hierarchical Control and Optimum Firm Size." *Journal of Political Economy* 75 (2): 123–138.

Wooldridge, Jeffrey. 2010. *Econometric Analysis of Cross Section and Panel Data.* 2nd ed. Cambridge, MA: MIT Press.

World Bank. 1997. *World Development Report.* New York: Oxford University Press for the World Bank.

———. 2017. "World Development Indicators." Available at `http://data.worldbank.org/data-catalog/world-development-indicators`. Accessed 1 April 2017.

Young, Crawford, and Thomas Turner. 1985. *The Rise and Decline of the Zairian State.* Madison, WI: University of Wisconsin Press.

Zanker, Franzisca, Claudia Simons, and Andreas Mehler. 2015. "Power, Peace, and Space in Africa: Revisiting Territorial Power Sharing." *African Affairs* 114 (454): 72–91.

Ziblatt, Daniel. 2004. "Rethinking The Origins of Federalism: Puzzle, Theory, and Evidence from Nineteenth-Century Europe." *World Politics* 57 (1): 70–98.

Ziblatt, Daniel. 2006. *Structuring the State: The Formation of Italy and Germany and the Puzzle of Federalism.* Princeton, NJ: Princeton University Press.

Zou, Hui, and Trevor Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society Series B* 67 (2): 301–320.